

Drug Discovery on Pentameric Ligand-Gated Ion Channels

by

Marta M. Wells

B.S., Johns Hopkins University, 2013

Submitted to the Graduate Faculty of  
the School of Medicine in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH  
SCHOOL OF MEDICINE

This dissertation was presented

by

Marta M. Wells

It was defended on

August 3, 2018

and approved by

Ivet Bahar, Professor, Computational & Systems Biology

Maria Kurnikova, Associate Professor, Chemistry, Carnegie Mellon University

Yan Xu, Professor, Anesthesiology

Dissertation Director: Pei Tang, Professor, Anesthesiology

Copyright © by Marta M. Wells

2018

## **Drug Discovery on Pentameric Ligand-Gated Ion Channels**

Marta M. Wells, PhD

University of Pittsburgh, 2018

Pentameric ligand-gated ion channels (pLGICs) are extensively involved in fast synaptic transmission and play a key role in many different neurological processes, such as pain sensation, memory, and addiction. These receptors are known targets for a variety of established pharmacological agents with diverse clinical uses, but recent major progress in our understanding of pLGIC structure-function relationships allows for the enhanced discovery of new drugs with improved therapeutic potential and reduced adverse effects. This dissertation presents the integrated computational and experimental techniques employed towards this end for two different pLGICs, the  $\alpha 3$  glycine receptor (GlyR) and the  $\alpha 7$  nicotinic acetylcholine receptor (nAChR). Three primary accomplishments resulted from this work. First, an ensemble-based virtual screening protocol was developed to target the GlyR transmembrane domain (TMD). Small-scale screening calculations were performed as a pilot study and validated by functional electrophysiology measurements, resulting in a hit rate of over 90% success. Second, this screening protocol was expanded to identify compounds specifically targeting the cannabinoid-binding site in  $\alpha 3$ GlyR. The lead compound from *in vitro* functional validation experiments progressed to *in vivo* behavioral tests in mice and was found to be a potent analgesic for inflammatory pain. Third, experimental restraints from nuclear magnetic resonance (NMR) and

electron paramagnetic resonance (EPR) spectroscopy were applied to determine the first structures of the  $\alpha 7$ nAChR intracellular domain (ICD). Although many different pLGIC structures have been solved in recent years, no structure obtained so far contains a complete ICD. This domain is an important target for therapeutic interventions for a variety of neurological disorders associated with intracellular signal transduction, but structure-based drug discovery is impossible without accurate structural information. The  $\alpha 7$ nAChR ICD structures determined here lay the foundation for future screening efforts in a previously unexplored area of drug discovery.

## TABLE OF CONTENTS

<b>PREFACE.....</b>	<b>XV</b>
<b>1.0 INTRODUCTION.....</b>	<b>1</b>
<b>1.1 BACKGROUND AND SIGNIFICANCE.....</b>	<b>1</b>
<b>1.2 THESIS CONTRIBUTIONS.....</b>	<b>3</b>
<b>2.0 ENSEMBLE-BASED VIRTUAL SCREENING AT THE CANNABINOID- BINDING SITE IN GLYCINE RECEPTORS .....</b>	<b>5</b>
<b>2.1 INTRODUCTION .....</b>	<b>5</b>
<b>2.2 METHODS.....</b>	<b>8</b>
<b>2.2.1 System Preparation.....</b>	<b>9</b>
<b>2.2.2 Virtual Screening .....</b>	<b>11</b>
<b>2.2.3 Electrophysiology with <i>Xenopus</i> Oocytes .....</b>	<b>12</b>
<b>2.3 RESULTS.....</b>	<b>13</b>
<b>2.3.1 Effects of Lipid Composition on Conformations of <math>\alpha</math>1GlyR TMD .....</b>	<b>13</b>
<b>2.3.2 Virtual Screening on an Ensemble of <math>\alpha</math>1GlyR TMD Structures .....</b>	<b>16</b>
<b>2.3.3 Rank and Selection of Screened Compounds.....</b>	<b>19</b>
<b>2.3.4 Functional Validation of Screened Compounds.....</b>	<b>22</b>
<b>2.4 DISCUSSION.....</b>	<b>26</b>
<b>3.0 A NEW ANALGESIC POTENTIATOR OF A3 GLYCINE RECEPTORS .....</b>	<b>29</b>

3.1	<b>INTRODUCTION</b> .....	29
3.2	<b>METHODS</b> .....	31
3.2.1	<b>Molecular Dynamics Simulations to Generate an Ensemble of <math>\alpha</math>3GlyR Structures</b> .....	31
3.2.2	<b>Ensemble-Based Virtual Screening</b> .....	32
3.2.3	<b>Molecular Dynamics Simulations to Elucidate the Binding Mode of the Lead Compound</b> .....	34
3.2.4	<b>Electrophysiology with <i>Xenopus</i> Oocytes</b> .....	35
3.2.5	<b>Behavioral Evaluation of Analgesia</b> .....	36
3.3	<b>RESULTS</b> .....	38
3.3.1	<b>Virtual Screening on an Ensemble of <math>\alpha</math>3GlyR TMD Structures</b> .....	38
3.3.2	<b>Functional Validation of Screened Compounds</b> .....	41
3.3.3	<b>pLGIC Selectivity and Off-Target Effects of the Strongest <math>\alpha</math>3GlyR Modulators</b> .....	43
3.3.4	<b>Behavioral Evaluation of Lead Compound Analgesia</b> .....	46
3.3.5	<b>Compound 8 Binds to the Cannabinoid Site in <math>\alpha</math>3GlyR</b> .....	49
3.4	<b>DISCUSSION</b> .....	56
4.0	<b>STRUCTURES OF THE A7 NICOTINIC ACETYLCHOLINE RECEPTOR INTRACELLULAR DOMAIN</b> .....	60
4.1	<b>INTRODUCTION</b> .....	60
4.2	<b>METHODS</b> .....	63
4.2.1	<b>Experimental Methods</b> .....	64
4.2.1.1	<b>Sample Preparations</b> .....	64

4.2.1.2	Electrophysiology with <i>Xenopus</i> Oocytes .....	66
4.2.1.3	NMR Spectroscopy .....	67
4.2.1.4	EPR Spectroscopy.....	71
4.2.2	Computational Methods.....	72
4.2.2.1	Rosetta Input Preparation .....	72
4.2.2.2	TMD/ICD Structure Folding Calculations.....	73
4.2.2.3	TMD+ICD Structure Refinement .....	76
4.2.2.4	Structure Validation, Quality Assessment, and Analysis.....	78
4.3	RESULTS .....	79
4.3.1	Functional Characterization of $\alpha 7$ nAChR Constructs for Structural Studies .....	79
4.3.2	Structural Restraints from NMR Experiments .....	81
4.3.3	Distance Restraints from EPR Experiments .....	84
4.3.4	Weight Optimization of Experimental Distance Restraints.....	85
4.3.5	Structure Determination of the $\alpha 7$ nAChR ICD and TMD .....	86
4.3.6	Refined $\alpha 7$ nAChR TMD+ICD Structures .....	89
4.3.7	Pore Conformation and Ion Permeation Pathway .....	94
4.3.8	Structural Basis for Metabotropic Signaling .....	96
4.4	DISCUSSION.....	99
	BIBLIOGRAPHY .....	103



## LIST OF TABLES

Table 1. Top ranked compounds from the virtual screening. ....	20
Table 2. Additional compounds selected for <i>in vitro</i> functional validation. ....	21
Table 3. Potentiation at 1 $\mu$ M, EC <sub>50</sub> , and maximum potentiation of the screened compounds. ...	24
Table 4. Top-ranked compounds from virtual screening. ....	39
Table 5. Off-target profiling for compound 8. ....	46
Table 6. Summary of Lennard-Jones and electrostatic parameters for compound 8. ....	52
Table 7. Interaction energies and geometries for compound 8-water complexes. ....	53
Table 8. Molecular and quantum mechanical dipole moments for compound 8. ....	53
Table 9. Structural restraints from NMR and EPR experiments. ....	82
Table 10. Structure statistics for the $\alpha$ 7nAChR TMD+ICD. ....	91

## LIST OF FIGURES

Figure 1. Side view of the $\alpha$ 1GlyR TMD NMR structure. ....	7
Figure 2. Flowchart highlighting the screening protocol. ....	9
Figure 3. Backbone RMSD trajectories of the $\alpha$ 1GlyR TMD in MD simulations. ....	10
Figure 4. Lipid effects on the $\alpha$ 1GlyR TMD structure during MD simulations. ....	15
Figure 5. Lipids increase virtual screening selectivity. ....	17
Figure 6. Representative examples of lipid effects on virtual screening. ....	18
Figure 7. Functional validation of virtual screening. ....	23
Figure 8. First tier $\alpha$ 1GlyR modulators do not affect $\alpha$ 7nAChR. ....	25
Figure 9. Correlation between virtual screening predictions and in vitro functional validation. .	26
Figure 10. Functional validation of virtual screening on $\alpha$ 3GlyR. ....	42
Figure 11. Modulation of off-target pLGICs. ....	44
Figure 12. Compound 8 does not act on cannabinoid receptors. ....	45
Figure 13. Analgesic effects of compound 8 <i>in vivo</i> . ....	48
Figure 14. Compound 8 binds at the cannabinoid site in $\alpha$ 3GlyR. ....	51
Figure 15. Compound 8 binds an intra-subunit site between TM3 and TM4 helices. ....	55
Figure 16. Current structural knowledge of the pLGIC ICD. ....	61
Figure 17. Overview of the $\alpha$ 7nAChR TMD+ICD structure determination procedures. ....	63

Figure 18. Sequence alignment of $\alpha 7$ nAChR with input template structures. ....	76
Figure 19. Functional characterization of $\alpha 7$ nAChR constructs.....	80
Figure 20. Structural restraints from NMR experiments. ....	83
Figure 21. Distance restraints from EPR experiments.....	84
Figure 22. Optimization of experimental distance restraint weights in Rosetta. ....	86
Figure 23. Convergence of the iterative $\alpha 7$ nAChR ICD folding calculations.....	88
Figure 24. Refined structures of the $\alpha 7$ nAChR TMD+ICD. ....	90
Figure 25. Salt bridges stabilize the $\alpha 7$ nAChR ICD structure.....	93
Figure 26. Pore conformation and ion permeation pathway.....	96
Figure 27. Protein binding sites in the $\alpha 7$ nAChR ICD for cytoplasmic signaling. ....	98

## LIST OF EQUATIONS

Equation 1. Maximum possible effect in the Hargreaves test. ....	38
Equation 2. PRE NMR distance restraint. ....	70
Equation 3. PRE NMR intensity ratio.....	70
Equation 4. NMR transverse relaxation rate.....	70
Equation 5. RDC NMR angles.....	70
Equation 6. Total score for Rosetta structures .....	74
Equation 7. Q-factors for calculated structures.....	78

## LIST OF ABBREVIATIONS

5-HT	5-hydroxytryptamine
CB	Cannabinoid
CFA	Complete Freund's Adjuvant
CNS	Central nervous system
CW	Continuous wave
DEER	Double electron-electron resonance
DMPC	1,2-Dimyristoyl-sn-glycero-3-phosphocholine
DMSO	Dimethyl sulfoxide
DSS	2,2-Dimethyl-2-silapentane-5-sulfonate
DTT	Dithiotreitol
EC	Effective concentration
ECD	Extracellular domain
ED	Effective dose
EDTA	Ethylenediaminetetraacetic acid
ELIC	Erwinia chrysanthemi ligand-gated ion channel
EM	Electron microscopy
EPR	Electron paramagnetic resonance
GABA	$\gamma$ -Aminobutyric acid
GLIC	Gloeobacter violaceus ligand-gated ion channel
GlyR	Glycine receptor
HPLC	High-performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
HTS	High-throughput screening
ICD	Intracellular domain
INEPT	Insensitive nuclei enhanced by polarization transfer
IP <sub>3</sub>	Inositol trisphosphate
IPAP	In-Phase Anti-Phase
LDAO	Lauryl dimethylamine oxide
LINCS	Linear constraint solver
MD	Molecular dynamics
MPE	Maximum possible effect
MTSL	S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
nAChR	Nicotinic acetylcholine receptor
NiNTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance

NOE	Nuclear Overhauser effect
PAINS	Pan Assay Interference Compounds
PBS	Phosphate-buffered saline
PDB	Protein Data Bank
PDSP	Psychoactive Drug Screening Program
PGE <sub>2</sub>	Prostaglandin E2
pLGIC	Pentameric ligand-gated ion channel
POPC	1-Palmitoyl-2-oleoylphosphatidylcholine
PRE	Paramagnetic relaxation enhancement
PWL	Paw withdrawal latency
RDC	Residual dipolar coupling
RIC-3	Resistance to inhibitors of cholinesterase 3
RLU	Relative luminescence unit
RMSD	Root mean squared deviation
RMSF	Root mean squared fluctuation
S <sup>2</sup>	Order parameter
SU	Screening unit
TCEP	Tris(2-carboxyethyl)phosphine
TET	Trifluoroethylthio
THC	$\Delta^9$ -tetrahydrocannabinol
TLC	Thin layer chromatography
TMD	Transmembrane domain
TROSY	Transverse relaxation optimized spectroscopy

## PREFACE

I first want to thank my dissertation advisor, Prof. Pei Tang. I would not have made it to this point in my graduate studies without her support, guidance, and patience. I learned so much from working closely with her, not just about conducting scientific research, but how to think critically and creatively to solve a problem. She's been a great mentor to me over the past five years and an inspiration as a scientist that I'm sure will influence me throughout the rest of my career. I would also like to thank Prof. Yan Xu for his advice and direction on all my different projects over the past few years. I'm so grateful to have been welcomed into your lab groups.

I would like to thank the other members of my dissertation committee, Prof. Ivet Bahar and Prof. Maria Kurnikova, for providing helpful feedback that helped shape this work. Additionally, I will always appreciate the constructive recommendations I received from Prof. Jeffry Madura, a former committee member who unfortunately passed away before my dissertation defense. I would also like to thank our collaborators, Prof. Erik Lindahl and Prof. Sunil Saxena, for their insightful discussions on the structure of  $\alpha 7nAChR$ .

I also want to thank all the current and former members of the Tang and Xu labs who have given me invaluable advice and taught me many of the computational and experimental techniques used in this work, as well as Ms. Sandra Hirsch for her editorial assistance. So many of you collected so much of the concrete biological data used in my theoretical computations; this dissertation would not be possible without your contributions. Joel, Tommy, Charles, Vasyl,

Nicole, Sharon, Nathan, Yali, Arjunan, Jessica, Bogdan, Monica, Brandon, Sherry, Lijie, Edom, and David, I'm so happy to have been able to work with all of you.

Thanks to everyone in the CPCB program, both past and present students, faculty, and staff, for providing a great environment for graduate school. Special thanks to everyone in my year, Nick, Dan, Cihan, Brad, and Kelvin. I would never have survived the beginning of grad school without you guys.

Finally, huge thanks to my family: Mom, Dad, Spencer, and Mitchell. My parents have given me the greatest support throughout my entire education and my brothers have always been there for me. I can't thank you all enough!



A portion of this dissertation has been published:

**Wells, M. M.**, Tillman, T. S., Mowrey, D. D., Sun, T., Xu, Y., and Tang, P. (2015) Ensemble-based virtual screening for cannabinoid-like potentiators of the human glycine receptor alpha1 for the treatment of pain. *Journal of medicinal chemistry* **58**, 2958-2966.

Additional publications from my graduate work that are not contained in this thesis:

Chen, Q., Kinde, M. N., Arjunan, P., **Wells, M. M.**, Cohen, A. E., Xu, Y., and Tang, P. (2015) Direct Pore Binding as a Mechanism for Isoflurane Inhibition of the Pentameric Ligand-gated Ion Channel ELIC. *Sci Rep* **5**, 13833.

Chen, Q., **Wells, M. M.**, Tillman, T. S., Kinde, M. N., Cohen, A., Xu, Y., and Tang, P. (2017) Structural Basis of Alcohol Inhibition of the Pentameric Ligand-Gated Ion Channel ELIC. *Structure* **25**, 180-187.

Ion, B. F., **Wells, M. M.**, Chen, Q., Xu, Y., and Tang, P. (2017) Ketamine Inhibition of the Pentameric Ligand-Gated Ion Channel GLIC. *Biophys J* **113**, 605-612.

Bondarenko, V., **Wells, M. M.**, Xu, Y., and Tang, P. (2018) Solution NMR Studies of Anesthetic Interactions with Ion Channels. *Methods Enzymol* **603**, 49-66.

Wang, Y., Yang, E., **Wells, M. M.**, Bondarenko, V., Woll, K., Carnevale, V., Granata, D., Klein, M. L., Eckenhoff, R. G., Dailey, W. P., Covarrubias, M., Tang, P., and Xu, Y. (2018) Propofol inhibits the voltage-gated sodium channel NaChBac at multiple sites. *J Gen Physiol* (in press).

## 1.0 INTRODUCTION

### 1.1 BACKGROUND AND SIGNIFICANCE

Pentameric ligand-gated ion channels (pLGICs) mediate synaptic transmission in the central and peripheral nervous systems, converting chemical signals from presynaptic neurotransmitters into postsynaptic electrical signals. Eukaryotic members of this superfamily include inhibitory anion-conducting glycine receptors (GlyRs) and  $\gamma$ -aminobutyric acid type-A receptors (GABA<sub>A</sub>Rs) as well as excitatory cation-conducting nicotinic acetylcholine receptors (nAChRs) and serotonin type-3 receptors (5-HT<sub>3</sub>Rs). These channels are found throughout the nervous system and play a critical role in many neurological processes, including cognition, memory, muscle control, sleep, and pain processing. Disruption of normal pLGIC signaling can lead to a variety of different neurological disorders, such as schizophrenia, Alzheimer's disease, hyperekplexia, Parkinson's disease, epilepsy, depression, anxiety, and addiction. Accordingly, these receptors are known targets for a number of major classes of pharmaceutical drugs, including anesthetics, analgesics, antipsychotics, anticonvulsants, hypnotics, antidepressants, anxiolytics and smoking cessation aids (1-5). Despite their prevalence in clinical use, many of these medications are plagued by adverse side effects, high abuse liability, and/or relative inefficacy (6-8), suggesting that there are still opportunities for improvement. Recent significant advances in our understanding of pLGIC

structures and functions open a new pathway for the rational discovery of target-specific drugs with greater clinical outcomes (9-11).

Traditional drug discovery efforts typically rely on high-throughput screening (HTS), which requires expensive infrastructure and extensive physical compound libraries, as well as systematic assay development and optimization efforts (12). In particular, HTS on pLGICs is further complicated by all the standard difficulties associated with studying membrane proteins, including recombinant expression, solubilization, and purification (13,14). Furthermore, HTS assumes no prior knowledge of drug binding sites on the target protein, failing to take advantage of available structural information to direct discovery efforts (12,15,16). Structure-based virtual screening is a cost-effective and efficient alternative to traditional HTS, providing a knowledge-driven approach to drug discovery that narrows the extensive reach of chemical space (17). Virtual screening identifies the most likely lead compounds for a target protein (18-20), while simultaneously revealing the molecular mechanisms behind drug-target interactions to provide a basis for rational lead optimization and design (21,22). The recent surge in pLGIC structural knowledge provides an excellent foundation for *in silico* screening, as evidenced by a number of recent studies that successfully employed structure-based virtual screening to identify novel pLGIC ligands (23-30).

Despite the wide-ranging pharmacology of the pLGIC superfamily, these channels all share a common structural framework consisting of five identical or homologous subunits arranged symmetrically around a central channel pore (9,31). Each subunit contains an N-terminal extracellular domain (ECD) folded in an immunoglobulin-like  $\beta$ -sandwich and a transmembrane domain (TMD) containing four transmembrane  $\alpha$ -helices (TM1-TM4) where TM2 forms the ion conducting pore. Eukaryotic pLGICs also contain an additional intracellular domain (ICD)

connecting TM3 and TM4 that is more diverse in size and composition compared to the highly conserved ECDs and TMDs. Although a number of high-resolution pLGIC structures have been published in the past decade, no structure obtained so far contains a complete ICD, severely limiting our ability to identify compounds that can influence pLGIC trafficking, localization, assembly, and, most importantly, interactions with intracellular protein partners (32-41).

In addition to sharing a common structural topology and organization, all pLGICs undergo conformational changes critical for normal channel function, transitioning between the closed, open, and desensitized states for ion conduction. Typically, signals from a neurotransmitter binding to the orthosteric site in the ECD are propagated to the TMD by small connecting loops between the conserved  $\beta$ -sheets and  $\alpha$ -helices with variable length and structure among pLGICs (5,9). However, this archetypal ionotropic activity can also be modulated by ligand binding outside the orthosteric site to produce allosteric effects (2,9,42,43). In fact, allosteric modulators often achieve higher selectivity for specific pLGIC targets than orthosteric agonists due to the conserved nature of the orthosteric binding site among pLGICs (44). Thus, allosteric modulators are the most promising direction for selective pLGIC drug discovery.

## **1.2 THESIS CONTRIBUTIONS**

The scope of this dissertation covers our drug discovery efforts to identify allosteric modulators of two different pLGICs:  $\alpha$ 3GlyR and  $\alpha$ 7nAChR. Specifically, three major accomplishments resulted from this work: 1) an effective virtual screening protocol for the GlyR TMD was established; 2) a novel analgesic compound acting at the cannabinoid-binding site in  $\alpha$ 3GlyR TMD was identified using this protocol; and 3) structures of the  $\alpha$ 7nAChR ICD were determined as a basis for future

structure-based screening efforts. The combined computational and experimental approaches employed to achieve each aim are summarized below and detailed in the subsequent chapters.

Chapter 2 covers the development of an ensemble-based virtual screening strategy for lipid-facing binding sites in the GlyR TMD. A small database of ~1,500 FDA-approved drugs was screened on an ensemble of  $\alpha 1$ GlyR TMD structures generated from molecular dynamics simulations (MD) in varied lipid environments. The computational protocol was validated by electrophysiology to confirm glycinergic modulation of the top ranked compounds *in vitro*. In chapter 3, we expanded this protocol to screen a large database of over 2 million compounds on an ensemble of  $\alpha 3$ GlyR TMD structures from MD simulations. Functional validation of the top ranked compounds was performed using *in vitro* electrophysiology experiments and the *in vivo* analgesic activity of the lead compound was verified using behavioral tests in mice. Lastly, chapter 4 details our structure determination calculations for the  $\alpha 7$ nAChR ICD. Experimental restraints from a variety of NMR and EPR spectra were used to guide iterative Rosetta calculations and validate the final  $\alpha 7$ nAChR ICD structures, which provide a first look at the complete structure of any pLGIC ICD.

## 2.0 ENSEMBLE-BASED VIRTUAL SCREENING AT THE CANNABINOID-BINDING SITE IN GLYCINE RECEPTORS

This chapter was adapted with permission from the [\*Journal of Medicinal Chemistry\* 58 \(7\): 2958-2966](#). Copyright 2015 American Chemical Society.

### 2.1 INTRODUCTION

Glycine receptors (GlyRs) are members of the superfamily of pentameric ligand-gated ion channels (pLGICs) that are activated by the neurotransmitter glycine. GlyRs are chloride-selective channels that mediate inhibitory neurotransmission in the central nervous system and regulate a variety of behaviors, including pain sensation, neuromotor activity, muscle relaxation, and anxiety (45-47). As such, these receptors are promising targets for therapeutic intervention.

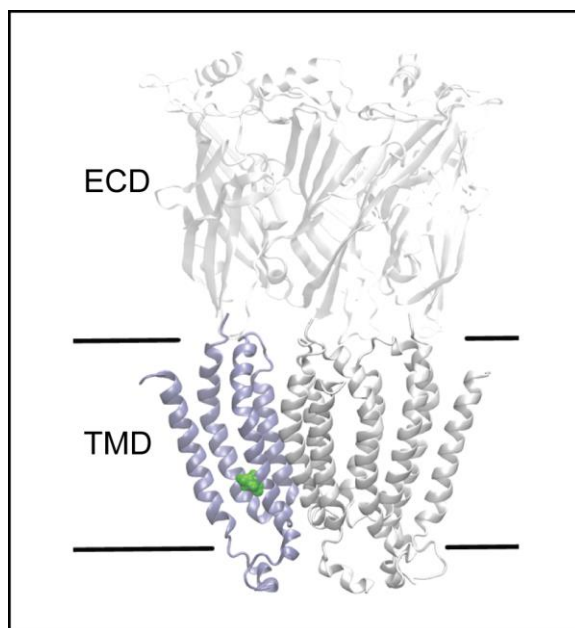
Most of the principal effects of  $\Delta^9$ -tetrahydrocannabinol (THC), the primary psychoactive component in cannabis (48), are mediated through the activation of the cannabinoid receptors (CB1 and CB2) (49). THC has serious negative effects on human health, including motor impairment and psychosis, but it has been found to provide relief in alleviating chronic pain (50). Previous evidence indicates that some effects of THC are independent of CB1 receptors (51). In particular, the analgesic effects of THC are mediated through THC potentiation of GlyRs, since

cannabinoid-induced analgesia is absent in mice lacking  $\alpha 3$ GlyR, but not in those lacking CB1 and CB2 receptors (52). THC potentiates GlyRs with high affinity (53) and residue S296 in TM3 of  $\alpha 1$ GlyR and  $\alpha 3$ GlyR has been shown to be critical for THC potentiation: the S296A mutation abolished THC potentiation of  $\alpha 1$ GlyR and  $\alpha 3$ GlyR and THC binding to the S296 site of  $\alpha 1$ GlyR was confirmed using high resolution NMR (52). Thus, the THC binding site at residue S296 of  $\alpha 1$ GlyR offers an ideal structural template for virtual screening to identify novel positive modulators acting at the same binding site as THC but without its negative effects.

Structurally, each GlyR subunit consists of an extracellular domain (ECD), a transmembrane domain (TMD) with four transmembrane helices (TM1 to TM4), and an intracellular domain (ICD). Agonist binding to the ECD transiently opens the channel allowing  $\text{Cl}^-$  ions to pass through the pore formed by the TM2 helices (54,55). At the time of this study, the only available structure of any GlyR was the NMR structure of the open-channel  $\alpha 1$ GlyR TMD, which can spontaneously form channels permeable to chloride ions in the absence of the ECD or ICD (56). This NMR structure provided a structural basis for screening and designing drugs that act on  $\alpha 1$ GlyR.

It is well documented that membrane composition affects the activity of GlyRs and other pLGICs (57-59). Particularly, cholesterol has been shown to affect channel functions (60,61). In addition, the NMR structure of the  $\alpha 1$ GlyR TMD shows that residue S296 faces away from the intra-subunit helical bundle (**Figure 1**) and is exposed to lipids. The conformations of lipids around the site may affect the binding of THC and other modulators. Therefore, the potential involvement of lipids and/or cholesterol in the binding of THC and other modulators should be considered in virtual screening.

Virtual screening allows for fast and economical evaluation of compounds in a large chemical library, but the accuracy of virtual screening strongly depends on conformations of the target receptor. Most virtual screening programs allow ligands to move flexibly around a rigid receptor. In reality, receptor motion often plays a critical role in the dynamic process of drug binding. A dynamic ensemble of receptor structures can reveal binding pockets that could not be found in a single static structure. More accurate results can be obtained by performing virtual screening on an ensemble of receptor structures generated through molecular dynamics (MD) simulations instead of a single structure (62,63).



**Figure 1. Side view of the  $\alpha$ 1GlyR TMD NMR structure.**

Residue S296 is highlighted in green in one subunit. A model of the ECD is positioned over the TMD for context.

In this study, we performed virtual screening with 1,549 FDA-approved drugs in the DrugBank database at the S296 site in the  $\alpha$ 1GlyR TMD. The promiscuous nature of small molecules and the interconnectedness of cell signaling pathways offers the potential for drug



repurposing, which enjoys the benefit of readily available pharmacokinetic and toxicological data obtained in previous clinical trials. Drug repurposing also significantly reduces the risks and timeline associated with clinical development. To increase the success of our virtual screening, we generated an ensemble of protein conformations through MD simulations with different lipid compositions and screened drugs in the presence and absence of lipids. The study served two purposes: 1) identify approved drugs that have potential analgesic effects by potentiating GlyRs and 2) establish a protocol for screening lipid-exposed druggable sites. More than a dozen compounds identified from the virtual screening were validated experimentally for their abilities to potentiate glycine-activated currents in *Xenopus laevis* oocytes expressing  $\alpha 1$ GlyR. The study has not only revealed drugs acting on the  $\alpha 1$ GlyR TMD with a mechanism similar to that of THC, but also paved a path for discovering new analgesic agents.

## 2.2 METHODS

An overview of the step-by-step protocol used in this study is shown in **Figure 2**.

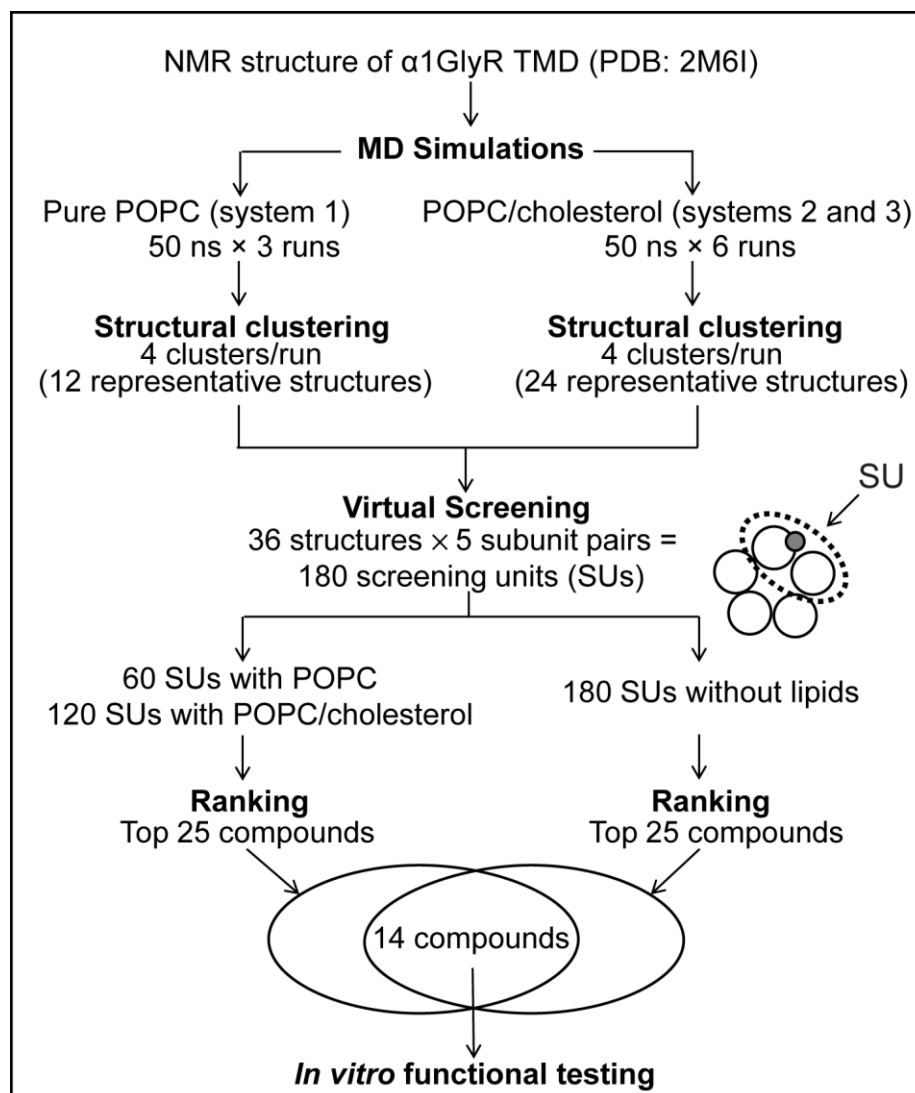
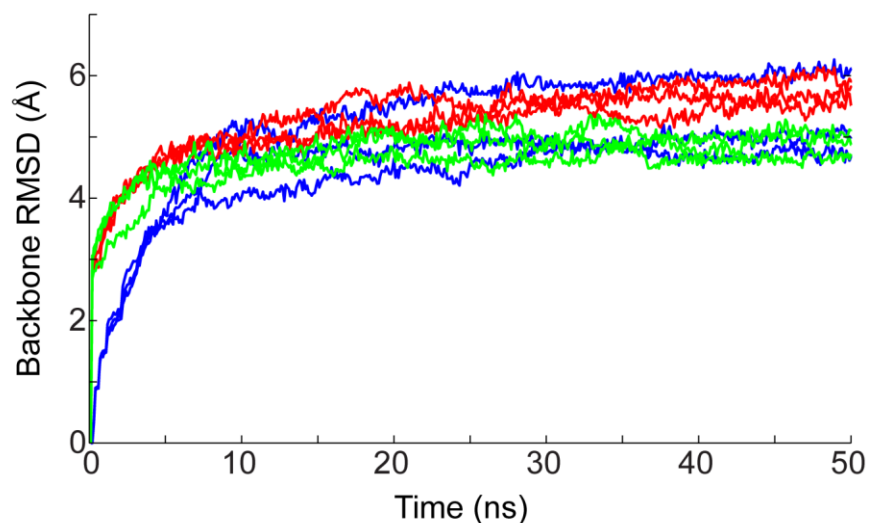


Figure 2. Flowchart highlighting the screening protocol.

### 2.2.1 System Preparation

The NMR-determined structure of the human open-channel  $\alpha 1$ GlyR TMD (PDB: 2M6I) (56) was used for the initial coordinates of the protein in MD simulations. The protein was embedded into three different pre-equilibrated lipid bilayers: 1) pure POPC, 2) POPC/cholesterol in a 4/1 molar ratio with five cholesterol molecules initially within 4 Å of TM1 to TM3, and 3) identical to system 2, except with eight cholesterol molecules initially within 4 Å of TM1 to TM3. The 4/1 molar ratio

of POPC/cholesterol was chosen because eukaryotic cell membranes typically are composed of approximately 20% cholesterol (64). The GROMACS `g_membed` tool (65) was used to insert the protein into POPC or POPC/cholesterol lipid bilayers (66,67). Each system was solvated in TIP3P water, ionized with 25mM NaCl, and contained ~90K atoms.



**Figure 3. Backbone RMSD trajectories of the  $\alpha$ 1GlyR TMD in MD simulations.**

Three  $\alpha$ 1GlyR TMD systems with different lipid environments are colored as follows: blue - pure POPC; red - 5 cholesterol molecules initially within 4 Å of the  $\alpha$ 1GlyR TMD; green – 8 cholesterol molecules initially within 4 Å of the  $\alpha$ 1GlyR TMD. Three replicate simulations were calculated for each system.

MD simulations of the three systems were run to obtain ensemble conformations of the  $\alpha$ 1GlyR TMD equilibrated in lipids. All simulations were performed using GROMACS 4.6.1 (68,69) with the AMBER03 (70) force field and additional parameters for POPC and cholesterol (71-73). The same simulation protocol was applied to each system. Energy minimization was performed for 10,000 steps with harmonic position restraints of 10,000 kJ/mol/nm<sup>2</sup> on the protein backbone atoms followed by 3 ns of equilibration, during which position restraints on the protein

backbone were gradually reduced from 10,000 to 0 kJ/mol/nm<sup>2</sup>. NMR-derived distance restraints) (56) were included over the entire trajectory with a force constant of 1000 kJ/mol/nm<sup>2</sup>. For each system, three 50-ns replicate simulations were performed at a constant pressure and temperature (NPT) of 1 atm and 310 K. After 10-20 ns, the backbone RMSD of the  $\alpha$ 1GlyR TMD in each simulation converged to  $\leq 6$  Å (**Figure 3**). All bonds were constrained using the LINCS algorithm (74). The integration time step used for all simulations was 2 fs. Particle mesh Ewald was used for long-range electrostatic interactions. A 12-Å cutoff was used for nonbonded interactions. Full electrostatic and nonbonded interactions were evaluated every five steps. Systems were simulated with periodic boundary conditions in three dimensions.

### 2.2.2 Virtual Screening

To identify representative protein conformations from each system for virtual screening, we performed structural RMSD clustering using the Quality Threshold Clustering Algorithm (75) implemented in VMD (76) with a threshold distance of 1.3 Å. This threshold distance divided the structures into similarly sized clusters with a smaller intra-cluster RMSD and larger inter-cluster RMSD. RMSD clustering was performed over the backbone atoms in TM1 to TM3 (residues 210 to 300) for each replicate simulation. TM4 was not included as it was furthest from the pore and most susceptible to random fluctuations. A representative structure was selected from each of the four major clusters. Each representative structure is the pentamer that has the minimum average RMSD between itself and all other structures in a given cluster. These representative pentameric structures were each divided into five sets of screening units, two adjacent subunits (**Figure 2**), in order to include potential intra- and inter-subunit binding sites in virtual screening. Four major

clusters were identified for each simulation. This yielded a total of 180 structures (4 representative structures  $\times$  3 systems  $\times$  3 replicates  $\times$  5 subunit pairs) for docking calculations.

Virtual screening was performed on the DrugBank database of 1549 FDA-approved compounds. This database is already a small subset of compounds and thus no further filtering was done prior to screening. To account for the potential effects of lipids in drug binding, all drug compounds were screened on all 180  $\alpha$ 1GlyR TMD structures twice: in the presence and absence of all lipid molecules within 7 Å of residue S296. The 7-Å threshold was selected to include only those lipid molecules that enclose the binding pocket. Docking calculations were performed using AutoDock Vina (77), where the protein and lipids remained rigid while drug compounds were flexible. The search space was set to a 20 Å square box centered on residue S296, encompassing both potential intra- and inter-subunit binding sites. Hits were defined as any docked conformation of a compound with binding energy corresponding to a disassociation constant  $K_d \leq 1 \mu\text{M}$  (binding energy  $\Delta G = RT \ln K_d$ , where  $T = 298 \text{ K}$  and  $R$  is the ideal gas constant) on the 180 screening units. The screened compounds were ranked based on the total predicted binding energy of their hits among 180 screening units in the presence or absence of lipids. Using the total predicted binding energy in the ranking allowed us to account for both the number of hits and the binding energy of the screened compounds. Because of the small search space for molecular docking, high-scoring compounds were, by design, likely to be close to S296.

### **2.2.3 Electrophysiology with *Xenopus* Oocytes**

RNA encoding  $\alpha$ 1GlyR or  $\alpha$ 7nAChR (25-50 ng) was injected into *Xenopus laevis* oocytes (stages 5-6) and channel activity was measured using two-electrode voltage clamp experiments as described previously (78). Oocytes were maintained at 17 °C in a modified Barth's solution (79).

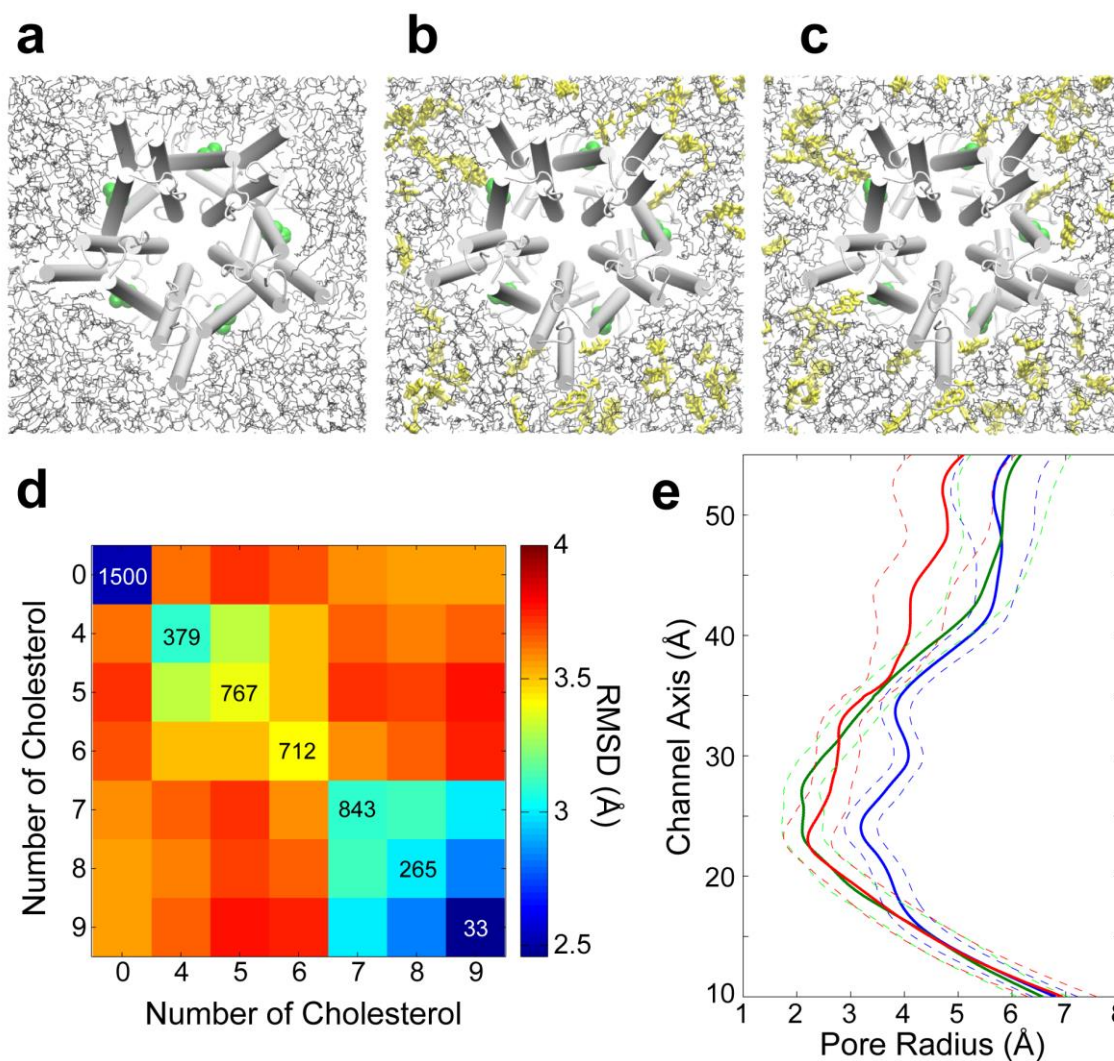
Functional measurements were performed 1-4 days after injection with an OC-725C Amplifier (Warner Instruments) in a 20- $\mu$ l oocyte recording chamber (Automate Scientific). Oocytes were clamped at a holding potential of -60 mV. The recording solutions contained 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub> and 5 mM HEPES at pH 7.0. Channel modulation was measured by the simultaneous co-application of the desired compounds with an effective glycine concentration equivalent to ~5% of the maximal current (EC<sub>5</sub>). Because  $\alpha$ 7nAChR has fast channel desensitization, an effective acetylcholine concentration equivalent to 50% of the maximal current (EC<sub>50</sub>) was used to ensure the accuracy of modulation assessments for the selected drugs. THC was obtained from the National Institute on Drug Abuse. PNU-120596 was purchased from R&D Systems (MN). Conivaptan was purchased from Toronto Research Chemicals, cinacalcet and regorafenib from MedChem Express, and others from Sigma-Aldrich. The vendors have verified the compounds' purity  $\geq$  98% by HPLC or TLC experiments. All compounds were first dissolved in DMSO and then diluted into buffer, with the final concentration of DMSO  $\leq$  0.05%. Data were collected and processed using Clampex 10 software (Molecular Devices).

## 2.3 RESULTS

### 2.3.1 Effects of Lipid Composition on Conformations of $\alpha$ 1GlyR TMD

We performed MD simulations of the  $\alpha$ 1GlyR TMD NMR structure in three systems with different lipid environments (**Figure 4**): (1) pure 1-palmitoyl-2-oleoylphosphatidylcholine (POPC); (2) POPC/cholesterol with five cholesterol molecules initially within 4 Å of TM1 to TM3; and (3) POPC/cholesterol with eight cholesterol molecules initially within 4 Å of TM1 to TM3. Each

system was subjected to three parallel 50-ns simulations. In all simulations, the RMSD of the  $\alpha$ 1GlyR TMD converged or stabilized after 10-20 ns (**Figure 3**). To examine whether the lipid composition affected the structure of the  $\alpha$ 1GlyR TMD, we calculated the pairwise RMSD values of the protein over the course of the MD simulations for each system. From each of the nine 50-ns trajectories, frames were extracted every 100 ps for analysis: 1,500 snapshots from simulations in pure POPC and 3,000 snapshots from simulations in POPC/cholesterol were collected. The presence of cholesterol did not greatly affect the tertiary structure of individual subunits. The mean pairwise RMSD for individual  $\alpha$ 1GlyR TMD monomers was 2.2 Å. In contrast, the mean pairwise RMSD calculated on pentamer structures was 3.5 Å, indicating more significant variations in quaternary structures. **Figure 4d** shows the pairwise RMSDs calculated from pentamers simulated in the absence and presence of cholesterol molecules. Snapshots were classified by the number of cholesterol molecules within 4 Å of TM1 to TM3. Note that during the simulation, the number of cholesterol close to the  $\alpha$ 1GlyR TMD varied due to the migration of cholesterol molecules. Overall, structures within the same group exhibit less structural variations than those belonging to groups with different numbers of cholesterols, demonstrating that MD simulations in the presence of different lipid compositions generate distinct ensembles of protein structures. The effect of cholesterol on the quaternary structure of  $\alpha$ 1GlyR TMD was also evidenced by the change in the pore radius (**Figure 4e**). Without cholesterol, the average minimum pore radius is  $3.1 \pm 0.3$  Å, whereas the average minimum pore radius from simulations with cholesterol is  $1.8 \pm 0.4$  Å. The penetration of cholesterol molecules into the intra-subunit space between TM3 and TM4 may have caused the conformational change of the channel and reduced the pore radius.



**Figure 4. Lipid effects on the  $\alpha$ 1GlyR TMD structure during MD simulations.**

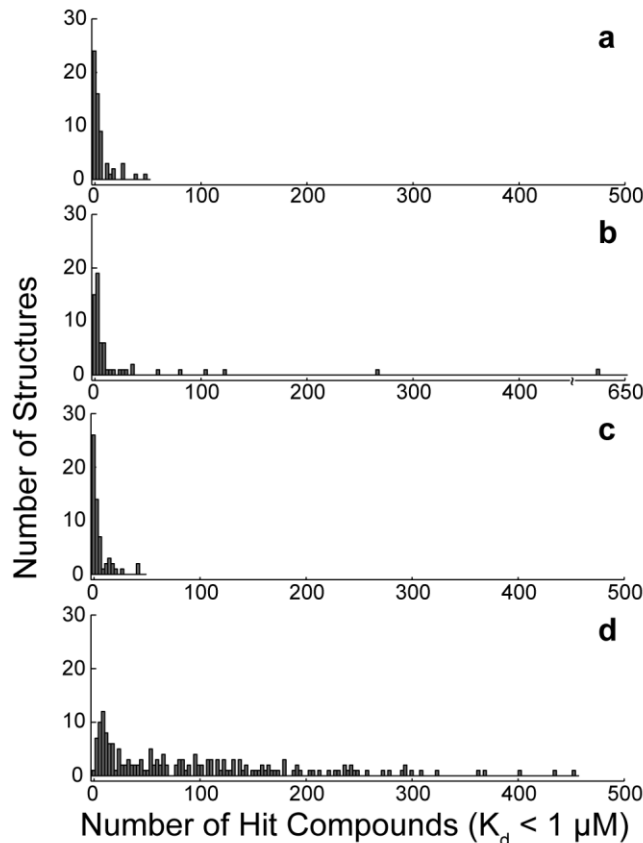
Top views of three  $\alpha$ 1GlyR TMD systems with adjacent lipid molecules from MD simulations. **(a)** System 1 with POPC; **(b)** system 2 with POPC/cholesterol; **(c)** system 3, identical to system 2 except with three more cholesterol molecules within 4 Å of TM1 to TM3 in the initial simulation set up. Residue S296 is highlighted in green. POPC and cholesterol molecules are shown as black and yellow sticks, respectively. **(d)** Mean pairwise RMSD of the  $\alpha$ 1GlyR TMD structures grouped by the number of cholesterol molecules within 4 Å of TM1-TM3 throughout MD simulations. The number of structures in each group is noted along the diagonal. **(e)** The presence of cholesterol in systems 2 (green) and 3 (red) in MD simulations changed the pore profile of the  $\alpha$ 1GlyR TMD compared to system 1 (blue). The pore radii were calculated using the HOLE program (80). The presented mean (solid line) and standard deviation (dashed line) of the pore radii were sampled from 500 structures from three replicate simulations for each system.



### 2.3.2 Virtual Screening on an Ensemble of $\alpha$ 1GlyR TMD Structures

Representative protein conformations were generated through structural RMSD clustering. We examined the major structural clusters from each simulation using the Quality Threshold Clustering Algorithm (75) that provides clusters within a given size threshold and optimizes the intra- and inter-cluster RMSD. Through this process, we found that most of the simulations contained four major structural clusters. For a given cluster, the pentamer with the minimum average RMSD between itself and all other structures was selected as a representative structure. Each representative pentameric structure was divided into five screening units comprised of two adjacent subunits to include potential intra- and inter-subunit binding sites in virtual screening. This yielded a total of 180 screening units for docking calculations. In order to consider direct lipid involvement in the S296 binding site, we screened all compounds against the same 180 structures in the presence and absence of the surrounding lipids (POPC or POPC/cholesterol) from the MD simulations. The lipid molecules may help to form the binding pocket, but they could also potentially block the binding site. Nevertheless, the inclusion of a variety of conformations of lipid molecules is biologically relevant. We reasoned that compounds scoring well in both the absence and presence of lipids might increase the success of *in vitro* functional validation.

Virtual screening in the presence of lipids showed that most of the  $\alpha$ 1GlyR TMD structures bound no more than 25 hit compounds, defined as those with predicted binding disassociation constants ( $K_d$ )  $\leq 1$   $\mu$ M (**Figure 5a-c**). However, when lipids were excluded from the virtual screening, the number of hit compounds on individual structures over all three systems increased significantly ( $p < 0.001$  in the two-sample Kolmogorov-Smirnov test) (**Figure 5d**). The data suggest a higher degree of selectivity when screening in the presence of lipids.

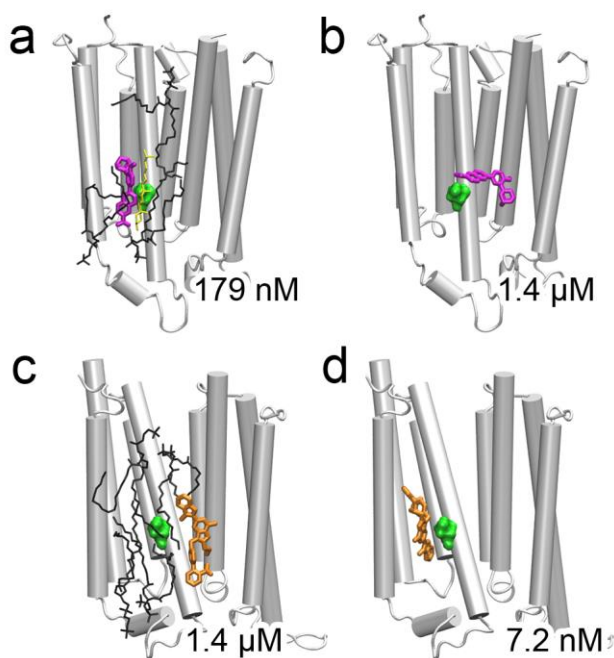


**Figure 5. Lipids increase virtual screening selectivity.**

Distributions of the number of hit compounds ( $K_d \leq 1 \mu\text{M}$ ) from the virtual screening on  $\alpha 1\text{GlyR}$  TMD structures from **(a)** system 1, simulated and screened with POPC; **(b)** system 2, simulated and screened with POPC/cholesterol; **(c)** system 3, identical to system 2 except with three more cholesterol molecule within  $4 \text{ \AA}$  of TM1 to TM3 in the initial simulation set up; and **(d)** screened all structures from three systems, but with neither POPC nor cholesterol in the virtual screening.

For a given  $\alpha 1\text{GlyR}$  TMD structure, we examined how lipid molecules affected drug docking to the S296 site. The presence of lipids effectively reduced the number of hit compounds. Moreover, the lowest predicted binding energies of each compound in the presence and absence of lipids were not necessarily obtained on the same  $\alpha 1\text{GlyR}$  TMD structures, demonstrating the necessity of screening with an ensemble of structures. For the same  $\alpha 1\text{GlyR}$  TMD structures, the

presence of lipids could enhance or reduce the binding affinity of certain drugs. **Figure 6** shows examples of both cases. For a given structure, we observed that the predicted binding affinity of adapalene improved when screened with lipids (**Figure 6a-b**). We also observed the opposite situation with a different drug: lipids blocked telmisartan binding to residue S296. The predicted  $K_d$  of telmisartan on this  $\alpha$ 1GlyR TMD significantly decreased in the absence of lipids (**Figure 6c-d**).



**Figure 6. Representative examples of lipid effects on virtual screening.**

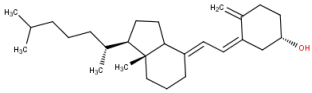
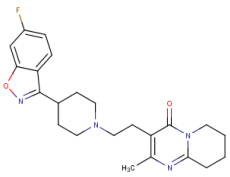
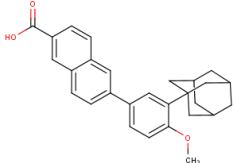
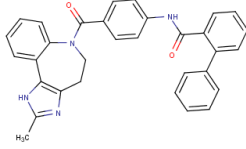
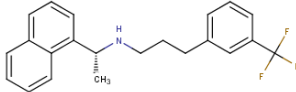
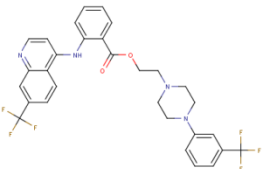
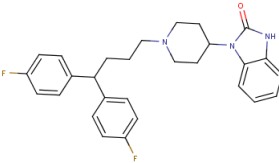
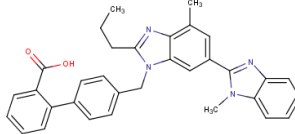
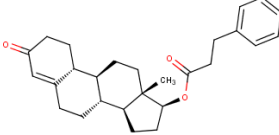
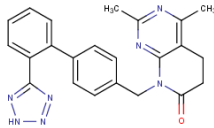
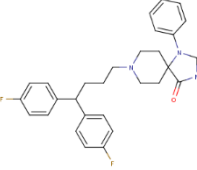
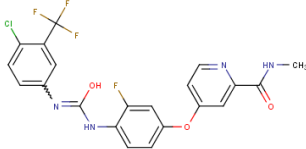
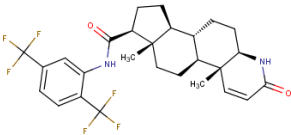
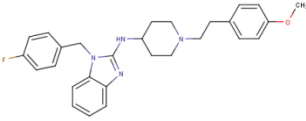
The presence of lipids improved adapalene (magenta) binding to the S296 site in the same structure and reduced the predicted  $K_d$  to **(a)** 179 nM from **(b)** 1.4  $\mu$ M in the absence of lipids. Lipids weakened telmisartan (orange) binding to the S296 site and increased the predicted  $K_d$  to **(c)** 1.4  $\mu$ M from **(d)** 7.2 nM.

### 2.3.3 Rank and Selection of Screened Compounds

The 1549 FDA-approved screened compounds were separately ranked, with and without lipids, based on their predicted binding energies  $\leq -8.18$  kcal/mol (equivalent to  $K_d \leq 1 \mu\text{M}$ ) summed over the 180 screening units. Among the top 25 hits from the screenings with and without lipids, there were 14 overlapping compounds (**Table 1**). These 14 compounds were considered for further experimental testing.

Cholesterol had a profound effect on the screening results; of the 14 top ranking compounds, six would not be highly ranked if screened only with POPC and without cholesterol. Since cholesterol is a component of native cell membranes, the inclusion of cholesterol in the screening is biologically relevant. Some of these top 14 compounds have been previously suggested to interact with receptors other than GlyRs, such as the drug pimozone with the dopamine D2 receptors in the central nervous system (81). None of these compounds, however, have been previously indicated to modulate GlyRs. All 14 compounds were filtered for Pan Assay Interference Compounds (PAINS) (82) and none were found to contain sub-structural features that would label them as “frequent hitters” in high throughput screens. In addition, none of these compounds are structurally similar to THC, as measured by low Tanimoto coefficients (83). A simple 2D structural similarity search would not have yielded the same leading compounds.

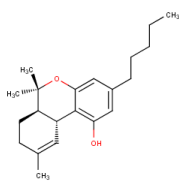
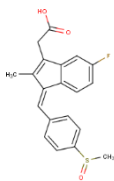
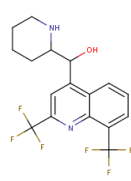
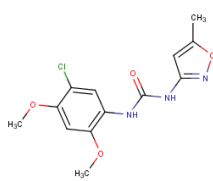
**Table 1. Top ranked compounds from the virtual screening.**

Compound	Structure	Compound	Structure
<b>Cholecalciferol</b> 1 (6) <sup>a</sup> 423 Å <sup>3</sup> <sup>b</sup> 0.275 <sup>c</sup>		<b>Risperidone</b> 12 (22) 372 Å <sup>3</sup> 0.186	
<b>Adapalene</b> 2 (13) 387 Å <sup>3</sup> 0.173		<b>Conivaptan</b> 14 (3) 446 Å <sup>3</sup> 0.072	
<b>Cinacalcet</b> 3 (24) 325 Å <sup>3</sup> 0.158		<b>Antrafenine</b> 16 (1) 483 Å <sup>3</sup> 0.089	
<b>Pimozide</b> 4 (7) 419 Å <sup>3</sup> 0.120		<b>Telmisartan</b> 21 (11) 470 Å <sup>3</sup> 0.121	
<b>Nandrolone phenpropionate</b> 6 (17) 401 Å <sup>3</sup> 0.207		<b>Tasosartan</b> 22 (25) 360 Å <sup>3</sup> 0.113	
<b>Fluspirilene</b> 7 (5) 439 Å <sup>3</sup> 0.113		<b>Regorafenib</b> 24 (21) 369 Å <sup>3</sup> 0.055	
<b>Dutasteride</b> 9 (9) 445 Å <sup>3</sup> 0.152		<b>Astemizole</b> 25 (19) 428 Å <sup>3</sup> 0.134	

<sup>a</sup> Ranks from virtual screening with lipids (without lipids). <sup>b</sup> Molecular volume. <sup>c</sup> Tanimoto coefficient as compared to

THC.

**Table 2. Additional compounds selected for *in vitro* functional validation.**

Compound	Structure
<b>THC</b> 452 (134) <sup>a</sup> 325 Å <sup>3</sup> <sup>b</sup> 1.0 <sup>c</sup>	
<b>Sulindac</b> 78 (114) 304 Å <sup>3</sup> 0.140	
<b>Mefloquine</b> 172 (124) 297 Å <sup>3</sup> 0.182	
<b>PNU-120596</b> 255 Å <sup>3</sup>	

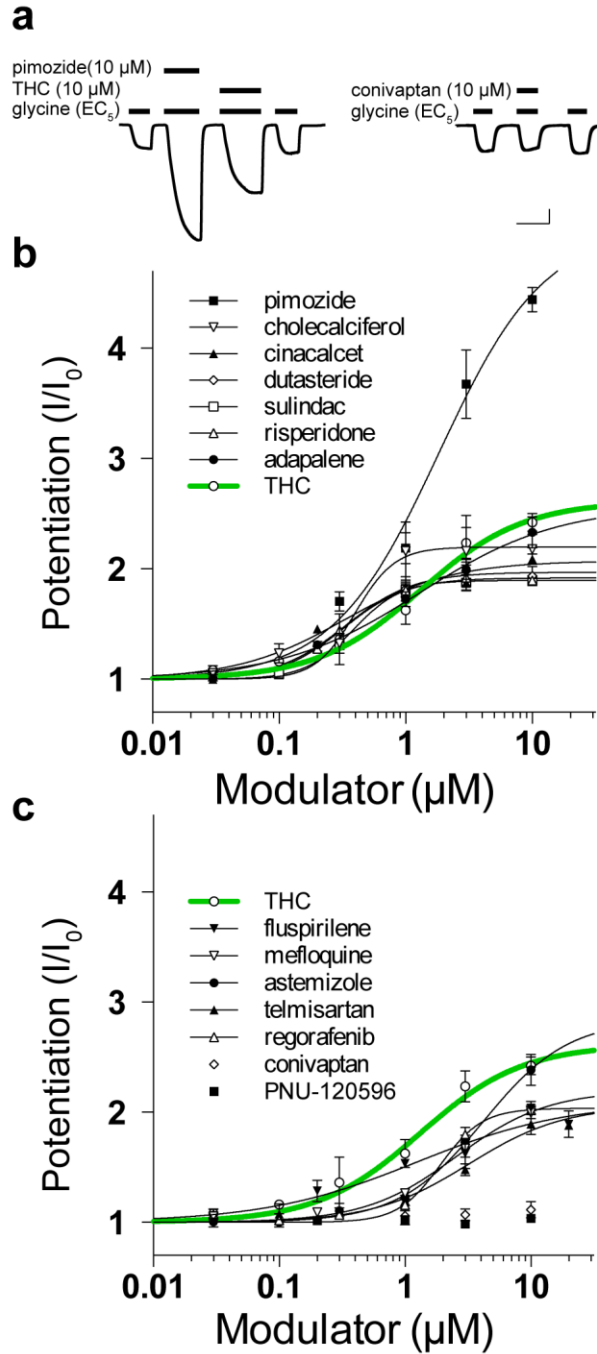
<sup>a</sup> Ranks from virtual screening with lipids (without lipids). <sup>b</sup> Molecular volume. <sup>c</sup> Tanimoto coefficient as compared to THC.

Sulindac and mefloquine, two compounds that ranked below the top 25 but higher than THC, were also selected for *in vitro* functional analysis (**Table 2**). Sulindac is a nonsteroidal anti-inflammatory agent with analgesic and antipyretic effects, but its mechanism of action was unknown prior to this study (84). Mefloquine is an anti-malarial agent that binds to adenosine receptors in the central nervous system (85). Both of these compounds have a hydroxyl group, which was shown to be critical for THC binding to  $\alpha 1$ GlyR (52). However, neither of them are structurally similar to THC, as measured by low Tanimoto coefficients (83). PNU-120596 (**Table 2**), a known highly specific positive allosteric modulator for the  $\alpha 7$ nAChR (86), was chosen as a

negative control in functional assays. This negative control was also predicted to bind the  $\alpha$ 1GlyR TMD with low affinity in virtual screening calculations. For all 180  $\alpha$ 1GlyR TMD structures screened in the presence of surrounding lipids, PNU-120596 never bound with  $K_d \leq 1 \mu\text{M}$ . When lipids were excluded from the screening, one  $\alpha$ 1GlyR TMD structure bound PNU-120596 with  $K_d \leq 1 \mu\text{M}$ .

### 2.3.4 Functional Validation of Screened Compounds

We were able to purchase 11 of the 14 top ranked compounds shown in **Table 1** for functional validation. Antrafenine, tasosartan and nandrolone phenpropionate were not commercially available at the time of our experiments. These 11 top-ranked compounds and two additional compounds (sulindac and mefloquine) from **Table 2**, along with the positive control THC and negative control PNU-120596, were tested on oocytes expressing  $\alpha$ 1GlyR. Representative functional traces in **Figure 7a** show pimozone is a stronger potentiator than THC. In contrast, conivaptan shows no potentiation for  $\alpha$ 1GlyR. Among all thirteen tested compounds, seven (first tier) were found to potentiate the glycine receptor better than THC at  $1 \mu\text{M}$  (**Figure 7b**), the  $K_d$  used as a threshold to define hits in the virtual screening. Another five compounds (second tier) also significantly potentiated the glycine receptor, but less effectively than THC (**Figure 7c**). Conivaptan and the negative control, PNU-120596, did not show potentiation up to the maximum tested concentration,  $10 \mu\text{M}$  (**Figure 7c**). The measured potentiation and  $EC_{50}$  values of all tested compounds are summarized in **Table 3**. The compounds from the first tier give a hit rate of 54%, but the combined results from all tested compounds gives an overall hit rate of 92%.



**Figure 7. Functional validation of virtual screening.**

(a) Representative current traces of *Xenopus laevis* oocytes expressing  $\alpha 1\text{GlyR}$  activated by glycine at 5% maximal effective concentration ( $\text{EC}_5$ ) and modulated by 10  $\mu\text{M}$  pimoizide, THC, or conivaptan. Black bars over the traces depict application of the indicated agonist or modulator. Scale bars indicate 30 seconds and 0.1  $\mu\text{A}$  current. Potentiation of  $\alpha 1\text{GlyR}$  by the indicated concentrations of (b) first tier and (c) second tier modulators was normalized



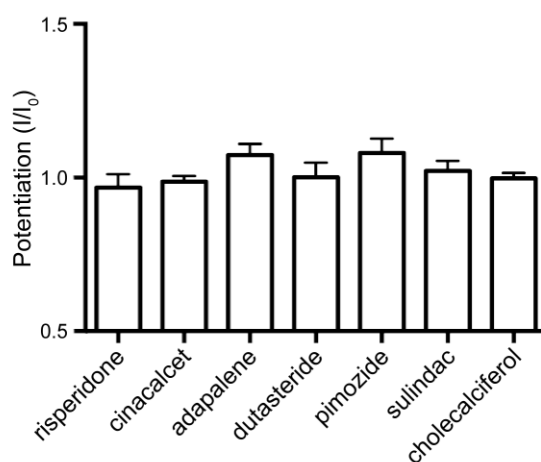
to the glycine-elicited current without modulators. Data are fit to Hill equations with the parameters shown in **Table**

**3.** Error bars designate the standard error of the mean ( $n \geq 4$  oocytes).

**Table 3. Potentiation at 1  $\mu$ M, EC<sub>50</sub>, and maximum potentiation of the screened compounds.**

<b>Compound</b>	<b>Fold Potentiation at 1 <math>\mu</math>M</b>	<b>EC<sub>50</sub> (<math>\mu</math>M)</b>	<b>Maximum Fold Potentiation</b>
pimozide	2.19 $\pm$ 14	1.7 $\pm$ 0.8	5.10 $\pm$ 70
cholecalciferol	2.16 $\pm$ 26	0.4 $\pm$ 0.1	2.20 $\pm$ 14
cinacalcet	1.82 $\pm$ 3	0.32 $\pm$ 0.07	2.10 $\pm$ 6
dutasteride	1.82 $\pm$ 11	0.33 $\pm$ 0.09	2.00 $\pm$ 8
sulindac	1.81 $\pm$ 7	0.38 $\pm$ 0.05	1.90 $\pm$ 4
risperidone	1.80 $\pm$ 6	0.32 $\pm$ 0.06	1.90 $\pm$ 5
adapalene	1.72 $\pm$ 7	1.3 $\pm$ 0.8	2.60 $\pm$ 27
<b>THC</b>	<b>1.62 <math>\pm</math> 12</b>	<b>1.3 <math>\pm</math> 0.6</b>	<b>2.60 <math>\pm</math> 30</b>
fluspirilene	1.53 $\pm$ 2	1.2 $\pm$ 0.7	2.10 $\pm$ 15
mefloquine	1.26 $\pm$ 4	2.5 $\pm$ 0.9	2.20 $\pm$ 19
astemizole	1.22 $\pm$ 2	4.2 $\pm$ 1.8	2.80 $\pm$ 39
telmisartan	1.20 $\pm$ 5	3.2 $\pm$ 1.5	2.10 $\pm$ 18
regorafenib	1.16 $\pm$ 5	1.8 $\pm$ 0.2	2.00 $\pm$ 5
conivaptan	1.06 $\pm$ 3	—	1.07 $\pm$ 5
PNU-120596	1.02 $\pm$ 4	—	1.03 $\pm$ 2

To explore whether these drugs also act as modulators for other members of the pLGIC superfamily, we additionally tested each of the seven first tier compounds in **Figure 7b** on oocytes expressing  $\alpha 7$ nAChR. None showed significant modulation of acetylcholine-activated channel currents (**Figure 8**). This is consistent with the previous finding that THC has little effect on  $\alpha 7$ nAChR function, though some endogenous cannabinoids have been found to inhibit  $\alpha 7$ nAChR (87).

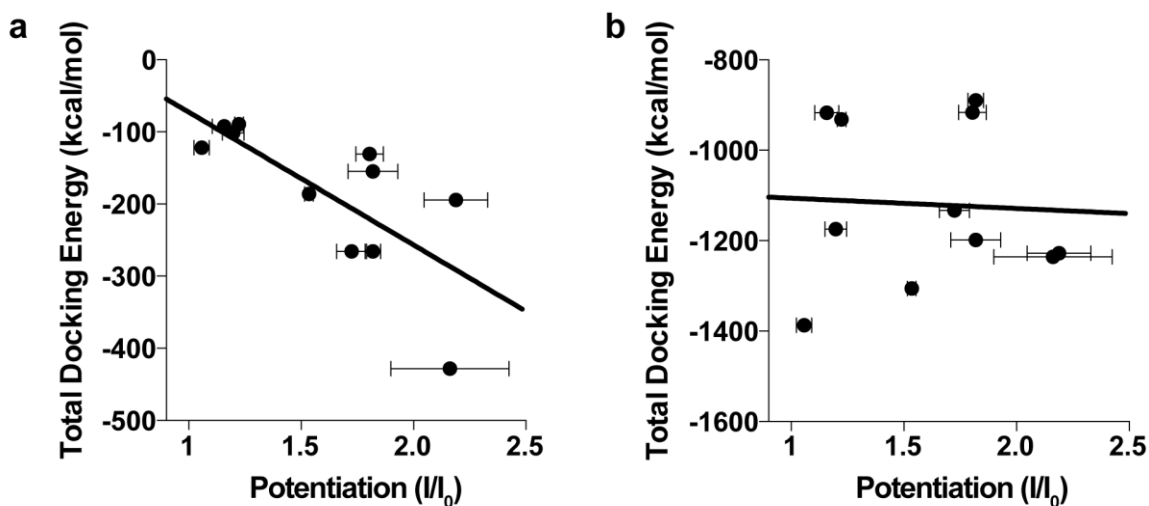


**Figure 8. First tier  $\alpha 1$ GlyR modulators do not affect  $\alpha 7$ nAChR.**

Modulation of  $\alpha 7$ nAChR expressed in *Xenopus* oocytes was measured at the EC<sub>50</sub> for acetylcholine. Bar graphs show modulation of agonist-induced current in the presence of 10  $\mu$ M of each of the indicated modulators. Error bars indicate standard error ( $n \geq 4$  oocytes).

The measured potentiation of the top ranked compounds shows a moderate correlation (Pearson correlation coefficient,  $r = -0.72$ ) with their total binding energies predicted by the virtual screening in the presence of lipids (**Figure 9a**). In general, compounds with a higher rank in the virtual screening in the presence of lipids provided a higher degree of potentiation of the  $\alpha 1$ GlyR current. However, the ranking generated from the virtual screening in the absence of lipids showed

no correlation ( $r = -0.052$ ) with the experimental potentiation results (**Figure 9b**). The experimental data further highlights the importance of including lipids in the virtual screening.



**Figure 9. Correlation between virtual screening predictions and in vitro functional validation.**

Potentiation was measured at 1  $\mu$ M for each of the top ranked tested compounds vs. the total docking energy obtained in (a) the presence of lipids or (b) the absence of lipids. Error bars show the standard error of the measured potentiation,  $n \geq 4$  oocytes. The difference in the total binding energy between (a) and (b) reflects a higher degree of binding selectivity when the virtual screening was done in the presence of lipids.

## 2.4 DISCUSSION

Using an ensemble of structures in the presence and absence of lipids for virtual screening is effective in identifying drugs that are positive allosteric modulators of GlyRs. Ten of the fourteen top-ranked compounds from the virtual screening (**Table 1**), as well as sulindac and mefloquine, have patent applications related to the treatment of pain; however, prior to this study, none of these compounds were known to potentiate  $\alpha$ 1GlyR. Such a high hit rate at the  $\alpha$ 1GlyR S296 site for

compounds relevant to analgesia is not likely a coincidence. Our results suggest that potentiation of  $\alpha 1$ GlyR by these compounds may provide similar analgesic effects to the action on GlyRs as exhibited by THC (52). These findings support the relevance of GlyRs as a therapeutic target in pain pathways.

Virtual screening in the presence of lipids produced a subset of structurally diverse hit compounds distinct from those obtained in the absence of lipids. The presence of lipids affected the screening results at two stages: generating protein structures with the desired lipid composition over the course of MD simulations (**Figure 4a-c**) and virtual screening with lipids as a part of the binding pocket. The NMR structure of the  $\alpha 1$ GlyR TMD provided a reliable starting point for generating an ensemble of protein conformations in both POPC and POPC/cholesterol environments. Using a large ensemble of conformations for virtual screening instead of a single conformer reproduces the dynamic nature of channel proteins and reduces bias in the screening process. The presence of a varied lipid composition during MD simulations increased the diversity of the ensemble (**Figure 4**). Limiting the number of protein structures used in the virtual screening certainly saves computation time. However, the “best” structure for one pre-selected compound is not necessarily the most suitable structure for other compounds in the screened library. In fact, among the 14 compounds shown in **Table 1**, only six share the “best” docking structure with THC. Therefore, the inclusion of an ensemble of protein conformations in the virtual screening is highly recommended if computational resources permit.

For an interior drug-binding site, it would be sufficient to directly use the ensemble of protein structures for screening without including lipids as part of the binding pocket. For the  $\alpha 1$ GlyR S296 site or any similar lipid-facing binding sites, however, lipid molecules may help form the binding pocket and should be considered for their influence on drug binding. Molecular

volumes for the top hits in **Table 1** and **Table 2** varied from below 300 to close to 500 Å<sup>3</sup>. Such flexibility in the molecular volume of the binding compound is likely a characteristic shared by binding sites at the protein-lipid interface in transmembrane proteins. Given the wide range of molecular volumes, it is not surprising that the presence of lipids during screening influenced the results. As shown in **Figure 6**, lipids can have a profound influence on the predicted binding affinity by forming additional stabilizing contacts at the binding site. In this way, lipids help accommodate compounds with widely varying molecular volumes at lipid-exposed binding sites. The vast majority of the compounds identified from our screening in the presence of lipids overlap with the compounds identified from the screening without lipids, but the total number of hit compounds from screening in the presence of lipids is less. It appears that screening in the presence of lipids can effectively filter the results. Furthermore, the results in **Figure 9** clearly demonstrate the importance of including lipids in the virtual screening. Thus, for lipid-facing binding sites, including the surrounding lipid molecules in the virtual screening is highly recommended.

In conclusion, our virtual screening on an ensemble of  $\alpha$ 1GlyR structures in different lipid environments has proved efficient to identify FDA-approved drugs that potentiate  $\alpha$ 1GlyR. Among the identified drug candidates, several were known or suspected to have analgesic effects; potentiation of  $\alpha$ 1GlyRs provides a plausible mechanism for the analgesic effects of these drugs. The screening results provide a valuable basis to further evaluate their analgesic properties *in vitro*. The protocol established in the current study can readily be transferred for discovering novel compounds of higher efficacy on GlyRs by screening larger chemical libraries. Finally, this screening methodology can be applied to other membrane proteins with lipid-exposed binding sites.

### 3.0 A NEW ANALGESIC POTENTIATOR OF $\alpha$ 3 GLYCINE RECEPTORS

#### 3.1 INTRODUCTION

In previous chapter, we developed a virtual screening protocol to identify compounds acting at the cannabinoid-binding site in GlyRs (88). At the time of this pilot study, the only available GlyR structure was the NMR structure of the human open-channel  $\alpha$ 1GlyR TMD (56). As such, we focused our initial investigations on the cannabinoid-binding site in  $\alpha$ 1GlyR. Although potentiation of both  $\alpha$ 1 and  $\alpha$ 3GlyR contributes to cannabinoid-induced analgesia (52,89),  $\alpha$ 1GlyR is widely expressed throughout the central nervous system and mediates the majority of glycinergic neurotransmission. Altering the native function of  $\alpha$ 1GlyR can have severe neurological consequences, such as hyperekplexia or complete muscle rigidity (45). In contrast,  $\alpha$ 3GlyR expression throughout most of the CNS is relatively lower than that of  $\alpha$ 1GlyR (45);  $\alpha$ 3GlyR knockout mice exhibit no obvious abnormal behaviors and respond normally to tactile stimuli (90). However,  $\alpha$ 3GlyR is distinctly expressed in the superficial laminae of the dorsal horn over  $\alpha$ 1GlyR and inhibition of  $\alpha$ 3GlyR in this region is directly involved with the sensation of inflammatory pain (90,91). Consequently, selective potentiation of  $\alpha$ 3GlyR in nociceptive neurons may provide analgesia by reversing this inflammation-induced inhibition of  $\alpha$ 3GlyRs with fewer negative side effects (92-94).

Since our initial drug discovery efforts (88), additional structures of both  $\alpha 1$  and  $\alpha 3$ GlyR have been published (95-98); however, all but one of these structures captures a nonconductive (closed or desensitized) channel state. The x-ray structure of human  $\alpha 3$ GlyR bound with the antagonist strychnine (95) offers an ideal foundation for screening calculations on the closed-state of  $\alpha 3$ GlyR. There is no experimental structure for the open-channel  $\alpha 3$ GlyR. The cryo-EM structure of the zebrafish  $\alpha 1$ GlyR (98) is the only new open-channel GlyR structure published since the development of our virtual screening protocol; however, a later study determined that this structure is likely physically unrealistic and does not accurately represent the open-channel conformation (99). Thus, the NMR structure of the open-channel human  $\alpha 1$ GlyR TMD, which shares ~90% sequence identity with  $\alpha 3$ GlyR, is the best template for screening calculations on the open-state  $\alpha 3$ GlyR. Overall, these structures, combined with our prior knowledge that cannabinoid potentiation involves direct hydrogen bond interactions with S296 in  $\alpha 3$ GlyR (52,100), provide a robust basis for virtual screening calculations to identify novel modulators of  $\alpha 3$ GlyR.

Our initial screening efforts with a small library of ~1,500 FDA-approved drugs resulted in the *in vitro* validation of ~90% of the top-ranked drugs as novel potentiators of  $\alpha 1$ GlyR (88), laying a solid foundation for the  $\alpha 3$ GlyR screening calculations performed in here. We expanded our computational approach to a targeted library of over 2 million compounds to identify compounds acting on  $\alpha 3$ GlyR. Thirteen top-ranked compounds from virtual screening on an ensemble of  $\alpha 3$ GlyR structures were tested *in vitro* using two-electrode voltage clamp electrophysiology, and ten significantly modulated (potentiated and/or inhibited)  $\alpha 3$ GlyR function. Of these confirmed  $\alpha 3$ GlyR modulators, the strongest potentiator showed selectivity for  $\alpha 3$ GlyR and was successful in treating inflammatory pain *in vivo*. Discovery of this novel glycinergic

analgesic opens a new avenue in the development of effective painkillers with fewer negative side effects.

## 3.2 METHODS

### 3.2.1 Molecular Dynamics Simulations to Generate an Ensemble of $\alpha$ 3GlyR Structures

Protein motion plays a critical role in the dynamic process of drug binding and protein flexibility is inherently linked to drug specificity (63,101,102). Ensemble docking, where drugs are screened across multiple varied static protein conformations, is an efficient way to reproduce the essential modes of protein dynamics in drug binding (18,22,88,103,104). Molecular dynamics (MD) simulations are an effective means of generating this structural ensemble for a target protein (63,101,104,105). Thus, to reflect the dynamic nature of  $\alpha$ 3GlyR in our virtual screening calculations, MD simulations were run for both open- and closed-channel conformations of  $\alpha$ 3GlyR. A homology model of the open-state  $\alpha$ 3GlyR TMD was generated based on the NMR structure of the human open-channel  $\alpha$ 1GlyR TMD (PDB: 2M6I) (56) using Modeller 9.14 (106,107) (~90% sequence identity between the  $\alpha$ 3 and  $\alpha$ 1GlyR TMDs). The TMD of the x-ray structure of  $\alpha$ 3GlyR bound to the competitive antagonist strychnine (PDB: 5CFB) (95) was used as initial coordinates for the closed-state simulations. The open and closed  $\alpha$ 3GlyR structures were each embedded into a pre-equilibrated lipid bilayer consisting of POPC/cholesterol in a 4/1 molar ratio (66,67). Our previous efforts with  $\alpha$ 1GlyR demonstrated that the accuracy of docking calculations at a lipid-facing binding site like S296 in  $\alpha$ 3GlyR is increased by the inclusion of cholesterol (88). The 4/1 molar ratio of POPC/cholesterol was chosen because eukaryotic cell



membranes typically are composed of approximately 20% cholesterol (64). Both the open- and closed-channel systems were solvated in TIP4P water, ionized with 25mM NaCl, and contained ~130K atoms.

MD simulations were run using GROMACS 5.0 (108) with the AMBER03 (70) force field and additional parameters for POPC and cholesterol (71-73). The same simulation protocol was applied to both the open- and closed-channel systems. Energy minimization was performed with 1,000 kJ/mol/nm<sup>2</sup> harmonic position restraints on all non-hydrogen atoms until the maximum force on any single atom was not greater than 250 kJ/mol/nm<sup>2</sup> (~10,000 steps), followed by 3 ns of equilibration, during which harmonic position restraints on all non-C $\alpha$  atoms were gradually reduced from 1,000 to 0 kJ/mol/nm<sup>2</sup>. Three 50-ns replicate simulations were calculated for each system at a constant pressure and temperature (NPT) of 1 atm and 310 K. All bonds were constrained using the LINCS algorithm (74). The integration time step used for all simulations was 2 fs. Particle mesh Ewald was used for long-range electrostatic interactions (109). A 12-Å cutoff was used for nonbonded interactions. Full electrostatic and nonbonded interactions were evaluated every five steps. Systems were simulated with periodic boundary conditions in three dimensions. MD simulations were run using XSEDE resources through the NSF (MCB040002) at the Pittsburgh Supercomputing Center (Bridges) and Texas Advanced Computing Center (Stampede).

### **3.2.2 Ensemble-Based Virtual Screening**

To identify representative  $\alpha$ 3GlyR conformations for virtual screening, structures from MD simulations were clustered using the Quality Threshold Clustering Algorithm (75) implemented in VMD (76). Replicate simulations from both the open- and closed-channel systems were individually clustered using an RMSD cutoff that optimally divided the structures into clusters

with a smaller intra-cluster RMSD ( $< \sim 1.3 \text{ \AA}$ ) and larger inter-cluster RMSD ( $> \sim 2.0 \text{ \AA}$ ). RMSD was measured over the backbone atoms in TM1 to TM3 (residues 210 to 300). TM4 was not included as it was furthest from the pore and most susceptible to random fluctuations. A representative structure was selected from each major cluster as the pentamer closest to the cluster center. Ten major clusters were identified for each system (open- and closed-channel), providing a total of 20 representative pentameric structures. Each pentamer was then divided into five sets of screening units consisting of two adjacent subunits in order to include potential intra- and inter-subunit binding sites in virtual screening. Ultimately, 100 screening units were used for virtual screening.

THC was screened across these 50 screening units from each system as a positive control compound using AutoDock Vina (77), where the protein and lipids remained rigid while THC was flexible. The search space was set to a  $20 \text{ \AA}$  cube centered on residue S296, encompassing both potential intra- and inter-subunit binding sites. All surrounding lipids from MD simulations were included in all docking calculations since S296 is at the protein-lipid interface facing the lipid bilayer. The five  $\alpha 3\text{GlyR}$  screening units from each system that produced the lowest THC binding energy were selected for docking calculations on a large library of over 2 million compounds. This screening compound library was generated by filtering the ZINC database of over 35 million small molecules (110) for compounds with drug-like properties (111) and additional physicochemical features required for compounds to penetrate the blood-brain barrier and act on the central nervous system (112). Specifically, the filtering criteria included molecule size restrictions (molecular weight between 150 and 450 g/mol), chemical properties (lipophilicity  $< 3.0$ , hydrogen bond acceptors  $\leq 5$ , hydrogen bond donors  $\leq 5$ , and rotatable bonds  $\leq 5$ ), and pharmacophoric properties (net charge  $\geq 0$  and polar surface area  $< 90 \text{ \AA}^2$ ). Virtual screening with this focused library was

executed using AutoDock Vina (77) with the same parameters as described above for THC docking. All docking calculations were performed using resources from Open Science Grid (113,114), which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. The screened compounds were ranked based on the total predicted binding energy across all five  $\alpha$ 3GlyR screening units in the open- and closed-states, respectively.

### **3.2.3 Molecular Dynamics Simulations to Elucidate the Binding Mode of the Lead**

#### **Compound**

Geometry and parameters for the lead compound (compound 8) were assigned by analogy using the CHARMM General Force Field (CGenFF) for drug-like molecules (115-117) and refined using the Force Field Toolkit (ffTK) protocol (118). Initial coordinates of  $\alpha$ 3GlyR for ligand-bound MD simulations were selected from the representative structures identified in apo MD simulations. The closed- and open-channel structures that provided the highest number of screening units for docking calculations on the large compound library described above (ranked by THC binding energy) were used as the initial coordinates for ligand-bound MD simulations, including the surrounding POPC/cholesterol lipid bilayer. Compound 8 was docked to each subunit in the open- and closed-state pentamer structures using AutoDock Vina (77). The search space was set to a 20 Å cube centered on residue S296. Both systems were solvated in TIP4P water, ionized with 25mM NaCl, and contained ~130K atoms. MD simulations were run using GROMACS 5.1 (108), the CHARMM36 force field (115,119-122), and the same simulation protocol as described above for the apo systems. Simulations were calculated using XSEDE resources through the NSF (MCB040002) at the Pittsburgh Supercomputing Center (Bridges) and Texas Advanced Computing Center (Stampede).

### 3.2.4 Electrophysiology with *Xenopus* Oocytes

RNA encoding  $\alpha 3$ GlyR,  $\alpha 3\beta$ GlyR,  $\alpha 1$ GlyR,  $\alpha 1\beta$ GlyR,  $\alpha 1\beta 3$ GABA<sub>A</sub>R,  $\alpha 7$ nAChR, ELIC, or the ELIC- $\alpha 3$ GlyR chimera (10-50 ng) was injected into *Xenopus laevis* oocytes (stages 5-6) and channel activity was measured using two-electrode voltage clamp experiments as described previously (78). Oocytes were maintained at 17°C in a modified Barth's solution (79). Functional measurements were performed 1-4 days after injection with an OC-725C Amplifier (Warner Instruments) and Digidata 1440A digitizer (Axon Instruments) in a 20- $\mu$ l oocyte recording chamber (Automate Scientific). Data were collected and processed using Clampex 10 software (Molecular Devices). Oocytes were clamped at a holding potential of -60 mV. The recording solutions for oocytes expressing  $\alpha 3$ GlyR,  $\alpha 3\beta$ GlyR,  $\alpha 1$ GlyR,  $\alpha 1\beta$ GlyR,  $\alpha 1\beta 3$ GABA<sub>A</sub>R, and  $\alpha 7$ nAChR contained 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub> and 5 mM HEPES at pH 7.0. The recording solutions for oocytes expressing ELIC and ELIC- $\alpha 3$ GlyR contained 130 mM NaCl and 10 mM HEPES at pH 7.0. All recording solutions for oocytes expressing heteromeric  $\alpha 3\beta$ GlyR or  $\alpha 1\beta$ GlyR also contained 0.1 mM picrotoxin in order to mask any residual currents from expression of homomeric  $\alpha 3$ GlyR or  $\alpha 1$ GlyR, respectively (123). With the exception of  $\alpha 7$ nAChR, modulation was measured by the simultaneous co-application of the desired compounds with an effective agonist concentration equivalent to ~2% of the maximal current (EC<sub>2</sub>) or ~20% of the maximal current (EC<sub>20</sub>). Due to the extremely fast desensitization of  $\alpha 7$ nAChR, an effective acetylcholine concentration equivalent to ~50% of the maximal current (EC<sub>50</sub>) was used to ensure the accuracy of modulation measurements.

Site-directed mutagenesis of  $\alpha 3$ GlyR was introduced with the QuickChange Lightning Kit (Agilent). The ELIC- $\alpha 3$ GlyR chimera was constructed using overlapping polymerase reaction by fusing the ECD of ELIC ending at R199 with the TMD of  $\alpha 3$ GlyR starting at Q219. The large

intracellular loop between TM3 and TM4 in  $\alpha$ 3GlyR was replaced with a short linker (310QPAAK427). The C-terminal of TM4 starting from I450 was replaced with the C-terminal of ELIC (318RGITL322) and additional mutations were introduced to the ELIC ECD (R91A, F188Y) to increase sensitivity to the orthosteric agonist propylamine.

Compounds 1, 3, 9, 11, and 13 were obtained from the National Cancer Institute Developmental Therapeutics Program Open Chemical Repository. Compounds 2, 4, 5, 6, 10, and 12 were purchased from Sigma-Aldrich. Compound 7 was purchased from ChemBridge and compound 8 was purchased from Enamine. All compounds were first dissolved in DMSO and then diluted into buffer, with the final concentration of DMSO  $\leq$  0.1%. Channel activity in the presence of agonist alone was measured with an equivalent concentration of DMSO to that used during the simultaneous co-application of both agonist and compound.

### **3.2.5 Behavioral Evaluation of Analgesia**

This work is a collaborative effort and Dr. Joel Caporoso, a postdoctoral scholar in Professor Yan Xu's lab, carried out all animal experiments. The methods are included here for completeness. All animal experimental procedures were performed with approval from the Institutional Animal Care and Use Committee (IACUC) at the University of Pittsburgh and carried out in accordance with the approved guidelines. CD1 strain mice (25-35 g) were purchased from Charles River Laboratories and housed in a temperature- and humidity-controlled room with a 12-hour light/dark cycle. Food and water were available *ad libitum*. Mice were randomly divided into vehicle and test groups and were handled identically, with the exception of including screened compounds in the treatment for the latter group.

Inflammation was induced by injection of 5  $\mu$ L of 1 mg/mL Complete Freund's Adjuvant (CFA) obtained from Millipore Sigma. Mice were anesthetized with 4% isoflurane and CFA was administered via intraplantar injection to the left hind paw using a Hamilton syringe with a 29-gauge needle. Inflammation of the paw was incubated for 24 hours to maximize CFA-induced hyperalgesia. Mice were then habituated individually in plastic boxes for 30 minutes prior to behavioral assessment using the Hargreaves Test (124). Two different apparatuses were used to measure paw withdrawal latency (PWL) in response to painful stimuli. Initial behavior tests were performed using a radiant heat plantar stimulator meter (IITC Life Science Inc.). The active light intensity was set to 15% and applied to the midplantar region of the CFA-inflamed footpad. A cutoff time of 23 seconds was employed to prevent tissue damage. A 980 nm infrared laser with a 1.6 mm beam (BWT Beijing Ltd.) was later acquired for PWL measurements. The beam was applied to the midplantar region of the CFA-inflamed footpad with laser power set to 1.5 W for male mice and 1.2 W for female mice due to inherent differences in pain sensitivity. In both experimental setups, the baseline PWL was measured from the average of 3 tests, separated by at least 1 minute. Following these baseline measurements, mice were removed from the apparatus and given an intraperitoneal injection of either vehicle or one of the test compounds. The experimenter was blinded to the identity of the injection solutions. After another 30-minute habituation period, PWL was measured every 15 minutes until mice no longer exhibited any analgesic effects. The maximum possible effect (MPE) of each treatment was calculated at each time point by **Equation 1**, where  $t_{\text{treatment}}$  is the average PWL for a given treatment (vehicle or test compound),  $t_{\text{CFA}}$  is the average baseline PWL before treatment, and  $t_{\text{naive}}$  is the average PWL for naïve mice.

**Equation 1. Maximum possible effect in the Hargreaves test.**

$$MPE(\%) = \frac{t_{treatment} - t_{CFA}}{t_{naive} - t_{CFA}} * 100$$

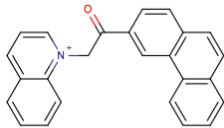
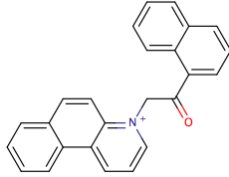
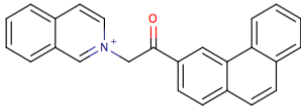
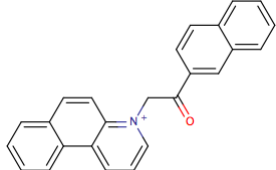
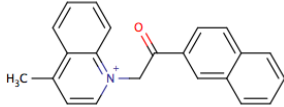
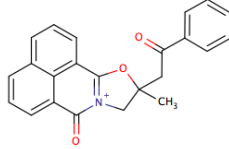
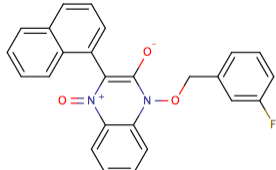
### 3.3 RESULTS

#### 3.3.1 Virtual Screening on an Ensemble of $\alpha$ 3GlyR TMD Structures

To generate a diverse ensemble of  $\alpha$ 3GlyR structures for virtual screening, MD simulations were run for both the open- and closed-state of the channel. Ten representative structures in each channel state across all three replicate simulations were identified by structural RMSD clustering, yielding a total of 100 screening units for the initial docking calculations with THC, the positive control compound. The average THC binding energy was not significantly different for the open- and closed-state  $\alpha$ 3GlyR structures ( $-5.17 \pm 3.60$  and  $-5.15 \pm 3.32$  kcal/mol, respectively), suggesting that both channel conformations are accessible for THC and likely other potentiators. For each channel state, the top five screening units that provided the lowest THC binding energy were used for screening calculations on a large library of over 2 million compounds, selected from the ZINC database (110) based on drug-like properties (111) and additional physicochemical features required for compounds to penetrate the blood-brain barrier and act on the central nervous system (112). Screened compounds were ranked based on their average predicted binding energies across the ensemble of open- and closed-state structures, respectively. Between the two groups of top-10 hits from the respective open- and closed-channel screening calculations, there were 5 overlapping compounds (**Table 4**), resulting in 15 unique hits. Among these 15 compounds, 13 were

commercially available for functional validation; compounds 14 and 15 could not be purchased at the time of our experiments.

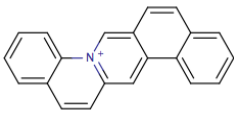
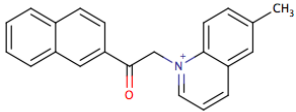
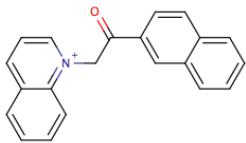
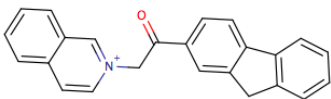
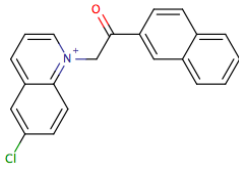
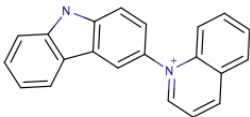
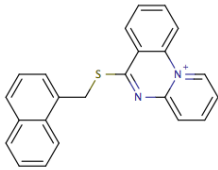
**Table 4. Top-ranked compounds from virtual screening.**

Compound	ZINC ID	Rank*	Structure
1	01668173	1, 1	
2	02381395	2, 5	
3	01668172	3, 3	
4	02519816	5, 2	
5	05681142	7, 7	
6	00640036	4, 1513	
7	20322444	8, 4788	

\*Ranks from virtual screening on open-channel, closed-channel conformations of  $\alpha$ 3GlyR.



Table 4 (continued)

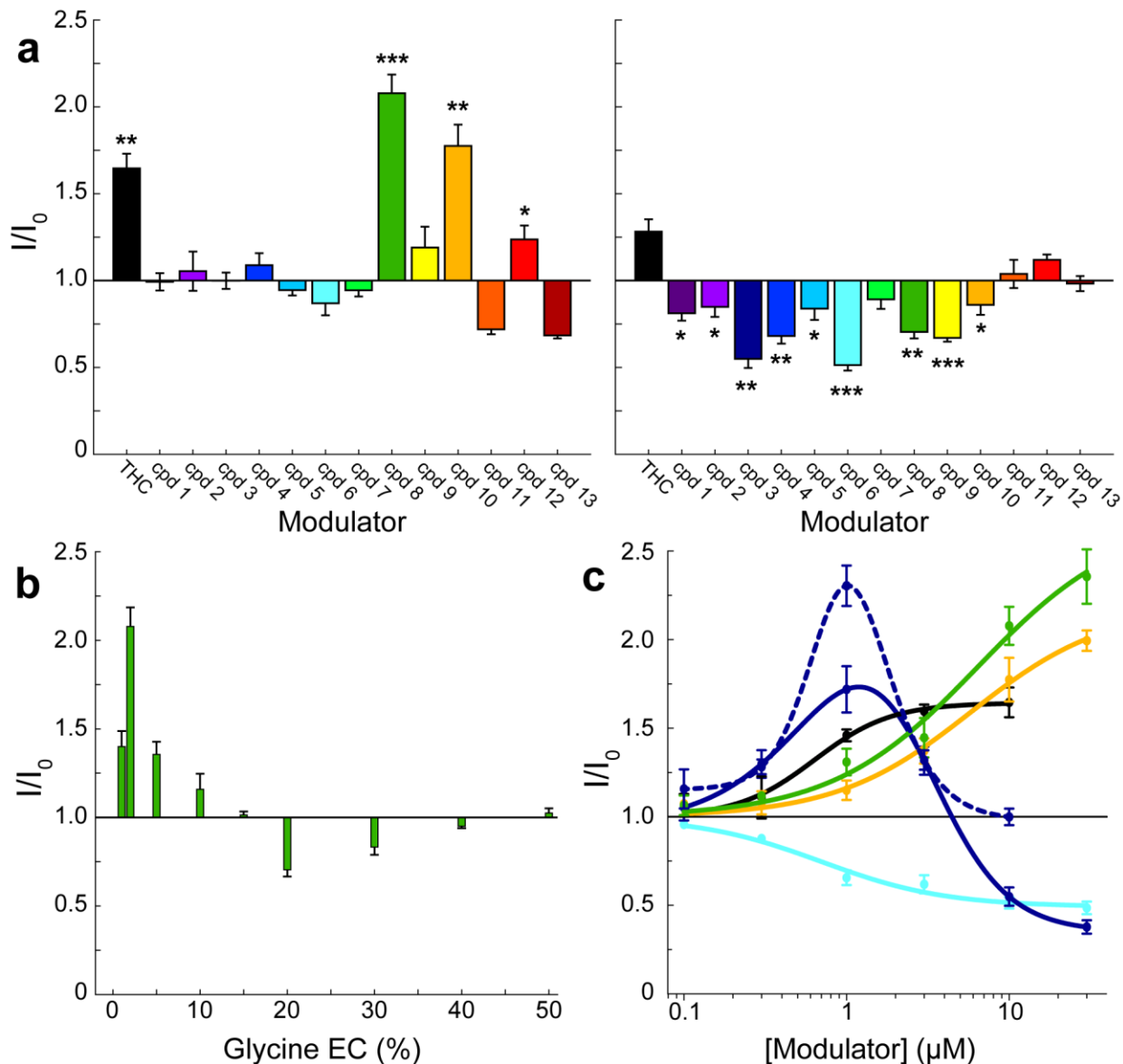
Compound	ZINC ID	Rank*	Structure
8	<i>Confidential Patent Pending</i>	9, 1299	<i>Confidential Patent Pending</i>
9	01713774	10, 30629	
10	05710513	34, 4	
11	01667310	59, 8	
12	01954783	91, 6	
13	01667308	209, 9	
14	05518709	6, 304	
15	05949118	15, 10	

\*Ranks from virtual screening on open-channel, closed-channel conformations of  $\alpha$ 3GlyR.

### 3.3.2 Functional Validation of Screened Compounds

Although THC binding to  $\alpha$ 3GlyR at S296 produces potentiation (52), virtual screening calculations alone cannot predict the functional impact of a novel compound binding to  $\alpha$ 3GlyR. As an initial step, we measured the functional modulation of each of the 13 compounds (10  $\mu$ M) on oocytes expressing  $\alpha$ 3GlyR (**Figure 10a**). All 13 compounds were tested at both EC<sub>2</sub> and EC<sub>20</sub> agonist concentrations in order to observe both potential potentiation and inhibition effects. Indeed, we identified three compounds (8, 10, and 12) that potentiated  $\alpha$ 3GlyR at EC<sub>2</sub> and nine that inhibited  $\alpha$ 3GlyR at EC<sub>20</sub> (1, 2, 3, 4, 5, 6, 8, 9, and 10). Note that compounds 8 and 10 exhibited both positive and negative modulatory effects, depending on the effective concentration of agonist applied during recordings (**Figure 10b**). Although only compounds 8 and 10 were found to potentiate  $\alpha$ 3GlyR to a greater extent than THC at 10  $\mu$ M, 10 of the 13 tested compounds produced significant modulation of  $\alpha$ 3GlyR, positive and/or negative, giving an overall hit rate of 77%.

Of the 10 compounds found to modulate  $\alpha$ 3GlyR, the two strongest potentiators (compounds 8 and 10) and two strongest inhibitors (compounds 3 and 6) were further analyzed by collecting dose response curves at EC<sub>2</sub> and EC<sub>20</sub>, respectively (**Figure 10c**). In agreement with the initial test at 10  $\mu$ M, compounds 8 and 10 exhibited dose-dependent potentiation of  $\alpha$ 3GlyR while compound 6 showed dose-dependent inhibitory effects. Surprisingly, the dose response curve for compound 3 at EC<sub>20</sub> revealed biphasic effects, where low and high concentrations of compound 3 potentiated and inhibited  $\alpha$ 3GlyR, respectively. In light of these new potentiating effects, an additional dose response curve for compound 3 was collected at EC<sub>2</sub> that displayed the same biphasic nature, though positive modulation was enhanced and negative modulation was reduced compared to the effects observed at EC<sub>20</sub> (**Figure 10c**).



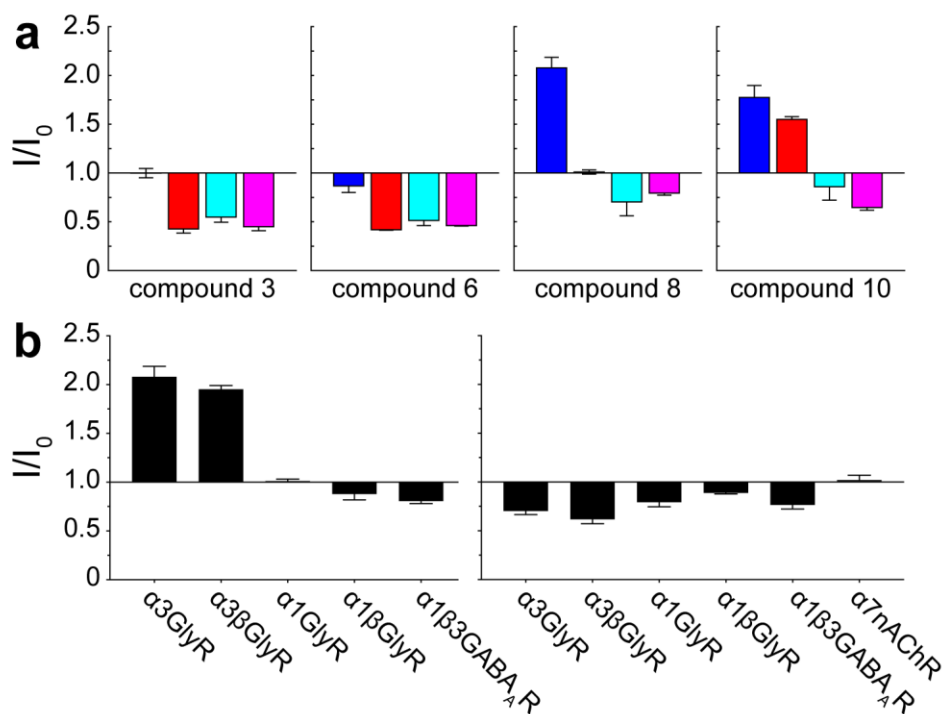
**Figure 10. Functional validation of virtual screening on  $\alpha 3$ GlyR.**

**(a)** *In vitro* electrophysiology responses from oocytes expressing  $\alpha 3$ GlyR activated by 2% (EC<sub>2</sub>, left) and 20% (EC<sub>20</sub>, right) of the maximal effective concentration for glycine. Modulation was measured as the ratio of currents in the presence (I) and absence (I<sub>0</sub>) of 10  $\mu$ M of THC or the indicated top ranked compounds from virtual screening. Stars indicate significant potentiation at EC<sub>2</sub> or inhibition at EC<sub>20</sub> at  $p < 0.05$  (\*),  $p < 0.01$  (\*\*), or  $p < 0.001$  (\*\*\*) by the one-tailed Student's t-test. **(b)** Compound 8 (10  $\mu$ M) modulation of  $\alpha 3$ GlyR at varying effective glycine concentrations, ranging from EC<sub>1</sub> to EC<sub>50</sub>. As the effective glycine concentration increases, compound 8 shifts from positive ( $I/I_0 > 1$ ) to negative ( $I/I_0 < 1$ ) modulation. Maximal potentiation is observed at EC<sub>2</sub> and maximal inhibition

is observed at EC<sub>20</sub>. **(c)** Dose response curves for THC, compound 8, and compound 10 at EC<sub>2</sub> and as well as compound 3 and compound 6 at EC<sub>20</sub> are color-coded as shown in **(a)**. The dose response curve for compound 3 at EC<sub>2</sub> is shown as a dashed line. Error bars in all panels indicate the standard error of the mean (n ≥ 4).

### 3.3.3 pLGIC Selectivity and Off-Target Effects of the Strongest $\alpha$ 3GlyR Modulators

Selectivity of hit compounds for an intended drug target is an important consideration in the drug discovery process: compound activity on unrelated targets drastically increases the likelihood of side effects and clinical toxicity (125-127). Although THC potentiates both  $\alpha$ 3GlyR and  $\alpha$ 1GlyR, only  $\alpha$ 3GlyR has been shown to be essential in cannabinoid-induced analgesia (52).  $\alpha$ 1GlyR subunits are widely expressed throughout the nervous system and mediate the majority of glycinergic neurotransmission; altering the native function of  $\alpha$ 1GlyR generally has severe neurological consequences, such as hyperekplexia and complete muscle rigidity (45). Thus, selectivity for  $\alpha$ 3GlyR over  $\alpha$ 1GlyR is an important consideration in the development of novel analgesics. Selectivity of the four strongest  $\alpha$ 3GlyR modulators (compounds 3, 6, 8 and 10) was measured on oocytes expressing  $\alpha$ 1GlyR (**Figure 11a**). Only compound 8 preferentially potentiated  $\alpha$ 3GlyR over  $\alpha$ 1GlyR. Additional measurements with heteromeric  $\alpha$ 3 $\beta$ GlyR and  $\alpha$ 1 $\beta$ GlyR confirmed that compound 8 is a selective potentiator for  $\alpha$ 3-containing GlyRs (**Figure 11b**). Compound 8 also demonstrated selectivity in potentiating  $\alpha$ 3GlyR over other pLGICs: no potentiation was observed on either  $\alpha$ 1 $\beta$ 3GABA<sub>A</sub>R or  $\alpha$ 7nAChR (**Figure 11b**).

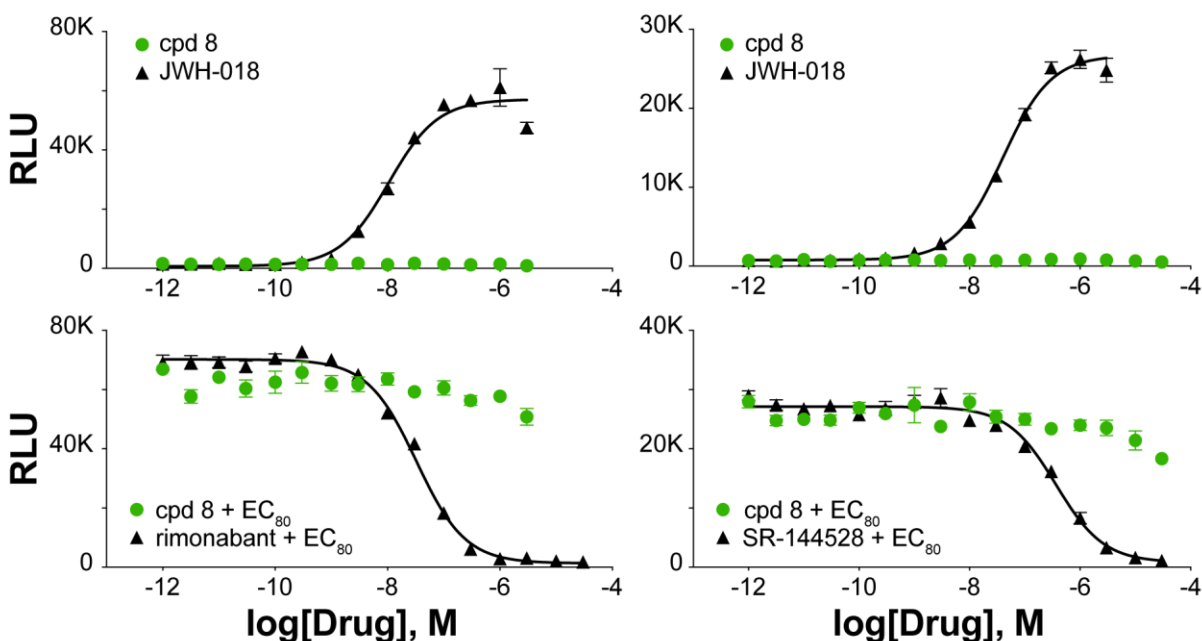


**Figure 11. Modulation of off-target pLGICs.**

**(a)** Comparison of electrophysiology responses from oocytes expressing  $\alpha 3$ GlyR (blue, cyan) or  $\alpha 1$ GlyR (red, pink) activated by 2% (EC<sub>2</sub>; blue, red) or 20% (EC<sub>20</sub>; cyan, pink) of the maximal effective concentration for glycine. Modulation was measured as the ratio of currents in the presence (I) and absence (I<sub>0</sub>) of 10  $\mu$ M of the indicated top ranked compounds from virtual screening. At a low agonist concentration (EC<sub>2</sub>), compound 8 preferentially potentiates  $\alpha 3$ GlyR over  $\alpha 1$ GlyR. **(b)** Modulation by compound 8 (10  $\mu$ M) was measured in oocytes expressing the indicated channel constructs activated by EC<sub>2</sub> (left) or EC<sub>20</sub> (right) agonist concentration. Due to the extremely fast desensitization of  $\alpha 7$ nAChR, EC<sub>50</sub> acetylcholine concentration was used to ensure the accuracy of modulation measurements. Compound 8 selectively potentiates  $\alpha 3$ -containing GlyRs over other pLGICs. Error bars in all panels indicate the standard error of the mean (n  $\geq$  4).

In addition to our own electrophysiology measurements on pLGICs, off-target profiling for compound 8 on other CNS receptors was assessed by the Psychoactive Drug Screening Program (PDSP) at the University of North Carolina at Chapel Hill. Radioligand binding assays (128) showed that compound 8 has no significant cross-reactivity with the pLGICs  $\alpha 4\beta 2$ nAChR or 5-

HT<sub>3A</sub>R, opioid receptors (DOR, KOR, MOR, and NOR), *N*-methyl-D-aspartate receptors (NMDARs), or hERG potassium channels (**Table 5**). More importantly, PRESTO-Tango functional assays (129) revealed that, unlike THC, compound 8 has no agonist or antagonist activity on either CB1 or CB2 receptors (**Figure 12**). Overall, compound 8 was the clear top candidate to move forward into *in vivo* assessments of analgesia.



**Figure 12. Compound 8 does not act on cannabinoid receptors.**

Compound 8 (green) shows no agonist (top) or antagonist (bottom) activity on either CB1 (left) or CB2 (right) receptors in PRESTO-Tango functional assays performed by the PDSP (129). Positive controls for each experiment are shown in black. Activity is reported in relative luminescence units (RLU). Some non-specific antagonist activity is observed at the highest concentrations of compound 8 by the decrease in RLU, but this not considered significant. Error bars in all panels indicate the standard error of the mean (n = 3).

**Table 5. Off-target profiling for compound 8.**

Mean percent inhibition (n = 4) by 10  $\mu$ M compound 8 in radioligand binding assays was measured by the PDSP. Significant inhibition is considered > 50%: compound 8 shows no significant cross-reactivity for any of the tested targets. Negative inhibition represents a non-specific stimulation of binding, likely due to the high concentration of compound 8 used in these preliminary screening assays.

Target	$\alpha$ 4 $\beta$ 2nAChR	5-HT <sub>3A</sub> R	DOR	KOR	MOR	NOR	NMDAR	hERG
% Inhibition	-1.8	-5.7	7.1	5.0	20.6	-16.8	-2.4	-19.5

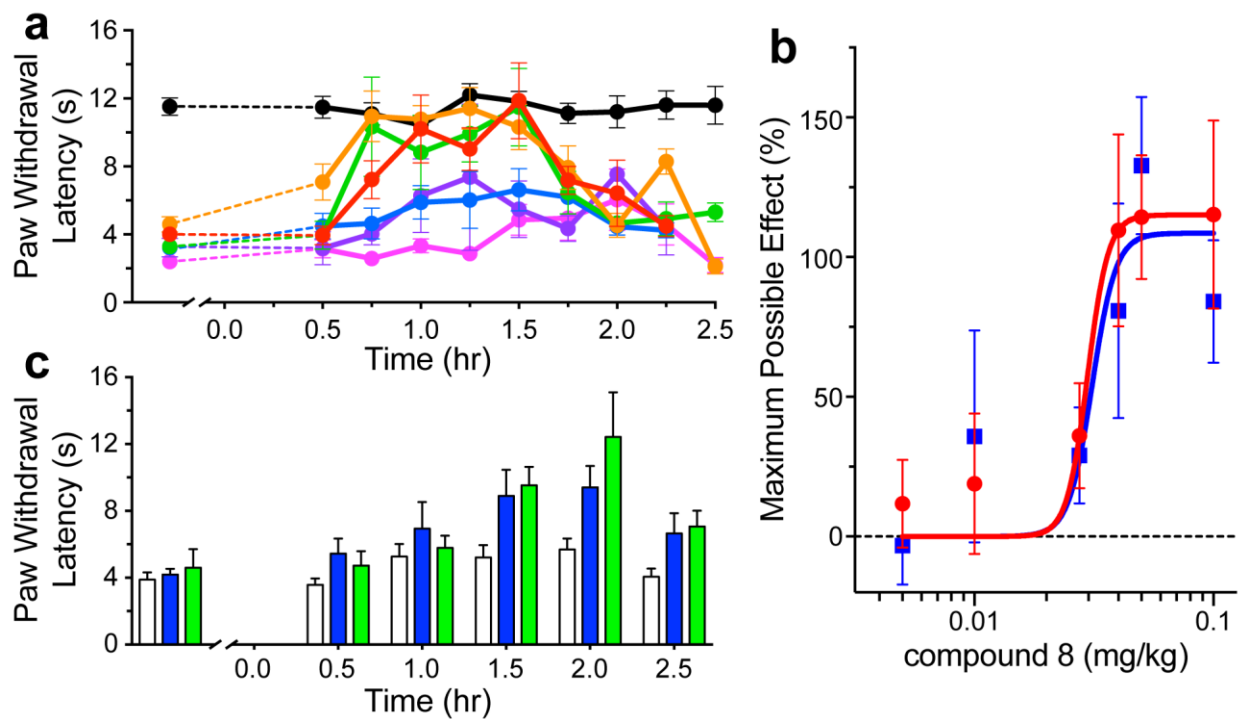
### 3.3.4 Behavioral Evaluation of Lead Compound Analgesia

The *in vivo* results presented here are part of the integrated drug discovery efforts in Professor Pei Tang's and Professor Yan Xu's groups; all animal experiments were performed in collaboration with a postdoctoral scholar, Dr. Joel Caporoso. Since cannabinoid action on  $\alpha$ 3GlyR has previously been shown to alleviate inflammatory pain (100), we first examined the efficacy of compound 8 in a mouse model of CFA-induced inflammatory pain. Injection of CFA into the hind paw generates hypersensitivity to thermal pain, reflected by the significant decrease in paw withdrawal latency (PWL) upon radiant heat stimuli in the Hargreaves Test (124), measured using a radiant heat plantar stimulator meter (IITC Life Science Inc.). Subsequent administration of compound 8 increased PWL, peaking ~1 hour after injection and persisting for ~30 minutes afterward (**Figure 13a**). Maximum efficacy was observed with 0.1 mg/kg of compound 8, where PWL reached the level of non-inflamed mice. The fact that compound 8 did not increase PWL beyond that of naïve mice suggests that it does not interfere with normal nociception and only alleviates pain brought on by inflammation. Dose response curves were calculated for the

maximum possible effect (MPE) at 1 and 1.5 hours after injection and both time points provided an estimated ED<sub>50</sub> of 0.03 mg/kg (**Figure 13b**).

These results suggest that compound 8 is effective in treating acute inflammatory pain, but many inflammatory pain disorders are chronic. Thus, effective therapeutics in this area must be resistant to drug tolerance. We administered 0.1 mg/kg of compound 8 to mice once per day for five consecutive days. CFA was injected on the fourth day and PWL was measured 24 hours later after the last injection of compound 8 using a 980 nm infrared laser (BWT Beijing Ltd.). Repeated doses did not decrease the analgesic effects of compound 8 (**Figure 13c**), where PWL peaked at ~1.5 hours after administration of the fifth dose. The analgesia effects of compound 8 persisted for at least ~1 hour afterward, indicating that no significant drug tolerance was built up from these repeated doses.





**Figure 13. Analgesic effects of compound 8 *in vivo*.**

(a) Responses of mice under maximum CFA-induced hyperalgesic conditions (24 hours post-application of CFA) were assessed by measuring paw withdrawal latency (PWL) to painful radiant heat stimuli before and after application of compound 8 (injection time = 0 hours). Compound 8 was injected at 0.005 (pink), 0.01 (purple), 0.0275 (blue), 0.04 (green), 0.05 (orange), and 0.1 (red) mg/kg. Responses of naïve mice are shown in black for comparison. (b) Dose response curves of mice under maximum CFA-induced hyperalgesic conditions at 1 hour (blue) and 1.5 hours (red) post-application of compound 8. Response was measured as the maximum possible effect (MPE, Equation 1). (c) Responses under maximum CFA-induced hyperalgesic conditions were assessed for mice after a single injection of 0.1 mg/kg compound 8 (blue) or after repeated injections of 0.1 mg/kg compound 8 once per day for five consecutive days (green). PWL to painful laser stimuli was measured before and after the final application of compound 8 (injection time = 0 hours). Responses of CFA-treated mice injected with the vehicle solution are shown in white for comparison. The analgesic effects of the single dose and repeated doses of compound 8 are significantly different from vehicle but not significantly different from each other by repeated measures one-way ANOVA with Fisher's LSD *post hoc* comparisons ( $p < 0.05$ ). Error bars in all panels designate the standard error of the mean ( $n \geq 4$  mice).

### 3.3.5 Compound 8 Binds to the Cannabinoid Site in $\alpha$ 3GlyR

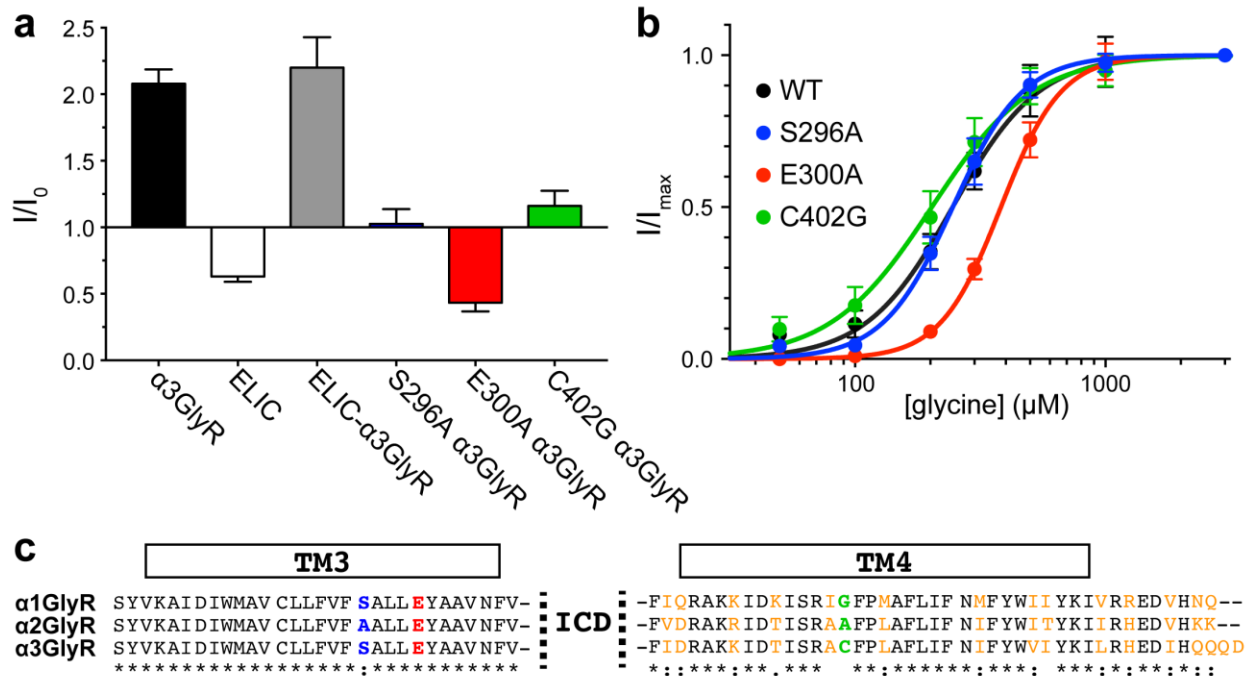
As mentioned above, THC potentiates both  $\alpha$ 3GlyR and  $\alpha$ 1GlyR through hydrogen bond interactions with the shared residue S296 on TM3 (52). Docking calculations from virtual screening predict that compound 8 similarly binds to S296 in  $\alpha$ 3GlyR; however, unlike THC, compound 8 selectively potentiates  $\alpha$ 3GlyR over  $\alpha$ 1GlyR. To experimentally elucidate the binding site and molecular basis for GlyR subtype-selectivity of compound 8, we performed additional electrophysiology experiments with chimeric and mutated  $\alpha$ 3GlyR. First, we constructed a chimeric channel consisting of the  $\alpha$ 3GlyR TMD and the ELIC ECD. Compound 8 inhibits ELIC, but potentiates the ELIC- $\alpha$ 3GlyR chimera to the same extent as observed in the full-length  $\alpha$ 3GlyR (**Figure 14a**), confirming that the potentiation effects of compound 8 are determined by the TMD of  $\alpha$ 3GlyR, and not the ECD or ICD.

Next, we measured the modulation of compound 8 on  $\alpha$ 3GlyR mutants. S296A has been shown to abolish potentiation by THC (52) and the same mutant also destroyed the potentiation by compound 8 (**Figure 14a**). This effect is likely due to the direct binding of compound 8 at S296 and not an allosteric effect resulting from the mutation itself because S296A and WT  $\alpha$ 3GlyR are indistinguishable in their responses to the agonist glycine (**Figure 14b**). To further confirm that THC and compound 8 share a binding site in  $\alpha$ 3GlyR, we tested E300A  $\alpha$ 3GlyR (one helical turn below S296 on TM3) and found that this mutation also abolished potentiation by compound 8. In fact, compound 8 inhibits E300A  $\alpha$ 3GlyR at EC<sub>2</sub> (**Figure 14a**); however, it is important to note that the EC<sub>50</sub> for glycine is significantly increased by this mutation (**Figure 14b**), somewhat complicating the interpretation of these results.

If THC and compound 8 bind to the same site in  $\alpha$ 3GlyR, why doesn't compound 8 also interact with the equivalent binding site in  $\alpha$ 1GlyR? A sequence alignment of  $\alpha$ 3GlyR and  $\alpha$ 1GlyR

residues around the S296 binding site reveals that differences in TM4 are likely the reason why compound 8 exhibits GlyR subtype-selectivity (**Figure 14c**). Specifically, interactions with C402 (equivalent to G402 in  $\alpha$ 1GlyR) likely confer selectivity for  $\alpha$ 3GlyR over  $\alpha$ 1GlyR. Indeed, the C402G mutation in  $\alpha$ 3GlyR abolishes potentiation by compound 8 (**Figure 14a**). Like S296A  $\alpha$ 3GlyR, C402G  $\alpha$ 3GlyR is very similar to WT  $\alpha$ 3GlyR in agonist-elicited channel activation, minimizing the possibility of ambiguous allosteric effects (**Figure 14b**).

To gain further insight into the molecular details of compound 8 binding, we performed three replicate 200-ns molecular dynamics (MD) simulations of compound 8 bound to  $\alpha$ 3GlyR in open- or closed-channel conformations. Although it is unfeasible to run simulations long enough to illuminate the precise mechanism of potentiation, these simulations can clarify the specific interactions involved in compound 8 binding. Since the exact orientation of compound 8 bound to  $\alpha$ 3GlyR is unknown, we unbiasedly docked compound 8 to S296 in the five different subunits of each  $\alpha$ 3GlyR pentamer, resulting in a variety of starting conformations for the drug. Optimized structures were calculated using Gaussian 16 software at the MP2/6-31(G) level of theory. The final parameterization results are confidential due to a pending patent application, but **Tables 6-8** below provide partially censored values as an approximation of the range for the final parameters.



**Figure 14. Compound 8 binds at the cannabinoid site in  $\alpha 3$ GlyR.**

**(a)** Electrophysiology measurements in *Xenopus laevis* oocytes activated by 2% ( $EC_2$ ) of the maximal effective agonist concentration (glycine for  $\alpha 3$ GlyR constructs, propylamine for ELIC and ELIC- $\alpha 3$ GlyR) show that like  $\alpha 3$ GlyR, but unlike ELIC, the chimera ELIC- $\alpha 3$ GlyR is potentiated by 10  $\mu$ M compound 8. The potentiation effects of compound 8 are determined by the TMD of  $\alpha 3$ GlyR, and not the ECD or ICD. Specifically, mutations at the cannabinoid-binding site in full-length  $\alpha 3$ GlyR abolish potentiation by compound 8. Error bars designate the standard error of the mean ( $n \geq 4$  oocytes). **(b)** Glycine activation curves for S296A, E300A, and C402G  $\alpha 3$ GlyR mutants are shown with  $EC_{50} = 246 \pm 9$ ,  $382 \pm 12$ , and  $203 \pm 14$   $\mu$ M, respectively. WT  $\alpha 3$ GlyR is shown for comparison with  $EC_{50} = 245 \pm 15$   $\mu$ M. Only E300A is significantly different from WT by repeated measures one-way ANOVA with Fisher's LSD *post hoc* comparisons ( $p < 0.05$ ). Error bars designate the standard error of the mean ( $n \geq 4$  oocytes). **(c)** Sequence alignment of  $\alpha 1$ ,  $\alpha 2$ , and  $\alpha 3$ GlyR TM3 and TM4 helices reveals a potential mechanism for the subtype-selectivity of compound 8. Sites of critical interactions identified by mutagenesis are highlighted in the corresponding color from **(a)** and **(b)**. All other sites of sequence dissimilarity are highlighted in orange. Hydrogen bonding with S296 likely confers selectivity for  $\alpha 1$  and  $\alpha 3$  subtypes over  $\alpha 2$ GlyR. Hydrogen bonding with C402 likely confers selectivity for  $\alpha 3$ GlyR over both  $\alpha 1$  and  $\alpha 2$ GlyR.

**Table 6. Summary of Lennard-Jones and electrostatic parameters for compound 8.**

Atom	Charge (e)	$\epsilon$ (kcal/mol)	Rmin/2 (Å)
C	-0.094	-0.090	1.90
	-0.817	-0.099	1.86
	-0.230	-0.090	1.90
	-0.220	-0.090	1.90
	-0.173	-0.090	1.90
	-0.171	-0.090	1.90
	-0.153	-0.060	2.02
	-0.120	-0.090	1.90
	-0.118	-0.090	1.90
	-0.117	-0.090	1.90
	-0.106	-0.090	1.90
	-0.071	-0.090	1.90
	-0.004	-0.090	1.90
	0.011	-0.090	1.90
	0.065	-0.056	2.01
	0.342	-0.099	1.86
	0.438	-0.110	2.00
	0.601	-0.110	2.00
N	-0.902	-0.200	1.85
	-0.446	-0.200	1.85
O	-0.460	-0.120	1.70
	-0.427	-0.120	1.70
H	0.090	-0.035	1.34
	0.115	-0.030	1.36
	0.196	-0.030	1.36
	0.202	-0.030	1.36
	0.311	-0.046	0.23
	0.375	-0.046	0.23

**Table 7. Interaction energies and geometries for compound 8-water complexes.**

Deviations between molecular mechanical and quantum mechanical calculations are shown for comparison. Quantum mechanical calculations of compound 8 interactions with water molecules were performed using Gaussian16.

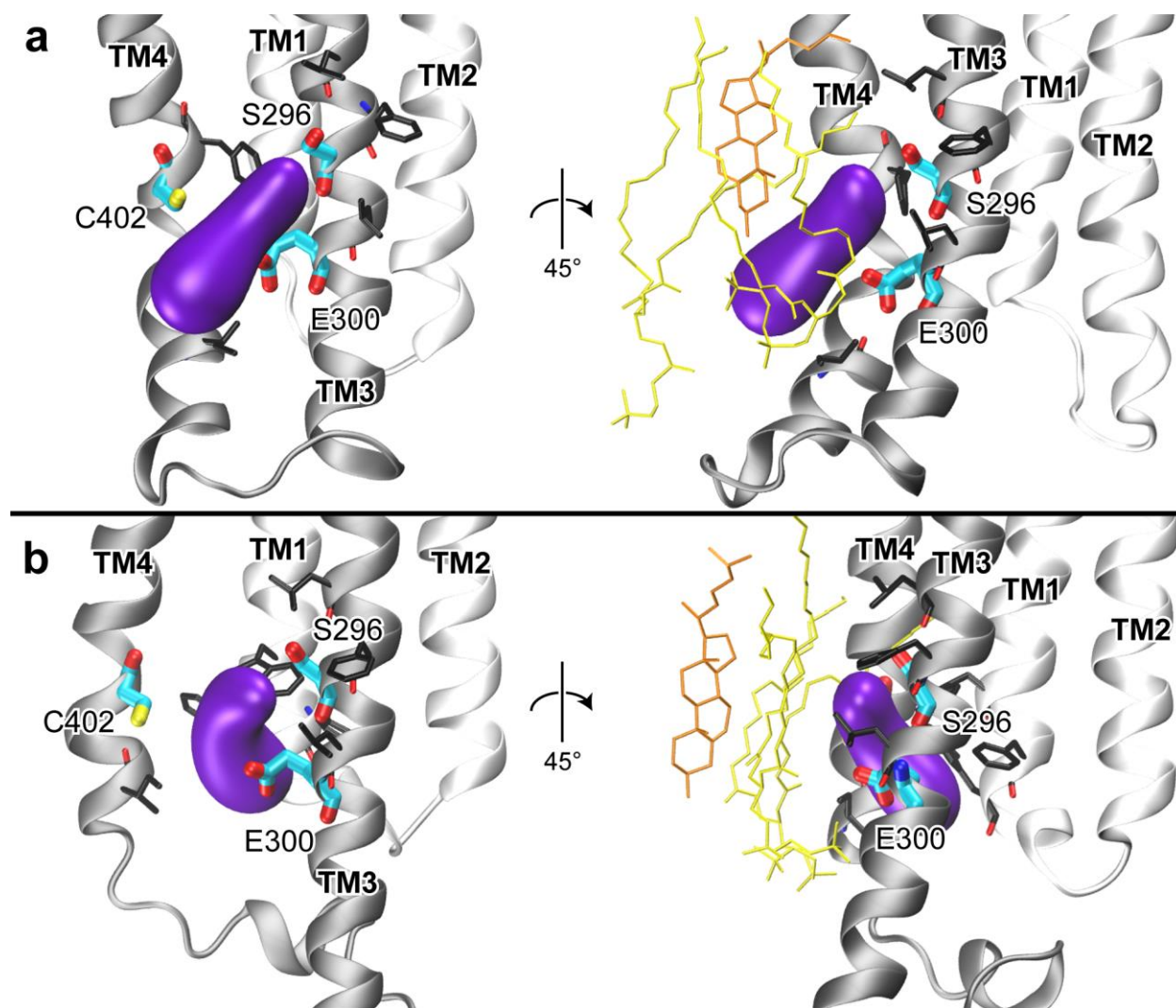
<b>Interaction</b>	$\Delta E_{\text{HF}}$ (kcal/mol)	$\Delta E_{\text{CGenFF}}$ (kcal/mol)	$\Delta \Delta E$ (kcal/mol)	$R_{\text{HF}}$ (Å)	$R_{\text{CGenFF}}$ (Å)	$\Delta R$ (Å)
N...HOH	-3.86	-4.60	-0.74	3.24	3.14	-0.10
	-3.17	-3.26	-0.09	3.29	3.29	0.00
O...HOH	-6.60	-5.67	0.93	2.98	2.93	-0.05
	-6.93	-7.21	-0.28	2.95	2.95	0.00
H...OHH	-6.21	-6.38	-0.17	2.12	2.12	0.00
	-4.97	-4.91	0.06	2.07	2.07	0.00
	-6.38	-6.91	-0.53	2.08	2.08	0.00
Average Deviation			-0.18			-0.02
Root Mean Squared Deviation			0.51			0.04
Absolute Average Deviation			0.40			0.02

**Table 8. Molecular and quantum mechanical dipole moments for compound 8.**

$\mu$ component (Debye)	HF/6-31G(d)	MP2/6-31G(d)	CGenFF
$D_x$	-0.624	0.211	1.079
$D_y$	1.106	0.956	0.529
$D_z$	-1.340	-1.250	-1.562
$D_{\text{Total}}$	1.846	1.588	1.971

MD simulations revealed that compound 8 was only able to form significant interactions when bound to an intra-subunit location between TM3 and TM4 helices. Specifically, compound 8 in this position formed recurring hydrogen bonds with S296 (29.3% occupancy) and occasional hydrogen bonds with E300 (4.7% occupancy) and C402 (3.1% occupancy) in the closed-state

$\alpha$ 3GlyR simulations (**Figure 15a**). Conversely, in the open-channel simulations, compound 8 in this binding site formed consistent hydrogen bonds with E300 (27.1% occupancy) and sporadic hydrogen bonds with S296 (3.0% occupancy) (**Figure 15b**). Although compound 8 binds to a similar intra-subunit site in both channel states, the drug is able to penetrate deeper into the intra-helical space in the open-channel simulations. In particular, the TM4 helix is much further from TM3 in the open-state than in the closed-state, as evidenced by the distance between C $\alpha$  atoms of S296 and C402 in MD simulations ( $15.3 \pm 1.4$  Å and  $11.4 \pm 0.6$  Å, respectively). This separation allows compound 8 to bind deeper within the  $\alpha$ 3GlyR subunit in the open-state; however, in this position compound 8 loses contact with C402. No significant hydrogen bond interactions with residues other than S296, E300, and C402 are observed in either the open- or closed-channel simulations, but compound 8 is stabilized by hydrophobic interactions with other residues in the TM1, TM3, and TM4 helices and with the lipid bilayer (**Figure 15a-b**). Overall, these simulations reveal that compound 8 binds to an intra-subunit site close the protein-lipid interface to establish significant polar interactions with the three critical residues identified in electrophysiology experiments.



**Figure 15. Compound 8 binds an intra-subunit site between TM3 and TM4 helices.**

Representative snapshots of compound 8 (purple) from MD simulations bound to  $\alpha$ 3GlyR in the (a) closed-channel and (b) open-channel conformations. Front views (left) show hydrogen bond interactions with S296, E300, and C402 (residues critical for compound 8 potentiation from mutagenesis experiments), as well as the large separation of TM4 in the open-state simulations. Side views (right) show compound 8 binds much deeper within the  $\alpha$ 3GlyR subunit in open-state than closed-state simulations. Other hydrophobic residues in close contact ( $< 4 \text{ \AA}$ ) with compound 8 are shown in black. POPC and cholesterol molecules that form the outside of the binding pocket are shown in yellow and orange, respectively, in the side views.



### 3.4 DISCUSSION

Through integrated computational virtual screening calculations, *in vitro* electrophysiological measurements, and *in vivo* behavioral tests, we have successfully identified a novel analgesic compound that is effective in treating inflammatory pain and shows no signs of tolerance to repeated exposure (**Figure 13**). While we cannot yet confirm that potentiation of  $\alpha 3$ GlyR is solely responsible for these analgesic effects, all evidence presented in this study supports this conclusion. Our own electrophysiology measurements established that compound 8 is a selective potentiator of  $\alpha 3$ GlyR over other pLGICs and PDSP screening assays found that compound 8 is inactive on several other CNS receptors involved the sensation of pain, including cannabinoid receptors and opioid receptors (**Figure 11-12, Table 5**). These results, combined with previous studies demonstrating the specific role of  $\alpha 3$ GlyR and its inhibition by prostaglandin E<sub>2</sub> (PGE<sub>2</sub>)-induced phosphorylation in inflammatory pain sensation (90-92,100,130), provide strong indications that compound 8 is indeed an  $\alpha 3$ GlyR-selective painkiller. However, additional *in vivo* behavioral experiments with  $\alpha 3$ GlyR knockout mice (90,100) as well as *in vitro* electrophysiology experiments of phosphorylated  $\alpha 3$ GlyR in the presence of PGE<sub>2</sub> (90,92,100) are necessary to confirm this theory.

As mentioned above, off-target profiling indicates that compound 8 does not act on opioid receptors (**Table 5**). More than 10 million adults in the U.S. have been prescribed long-term opioid therapies to treat their chronic pain (131), but opioid medications present serious risks: more than 33,000 overdose deaths in the U.S. in 2015 alone involved opioids of some kind (132). The discovery and development of potent non-opioid analgesics is vital to combat this growing public health epidemic. Although we have not verified that the analgesic effects of compound 8 are

specifically mediated by  $\alpha 3$ GlyR, the fact that this novel drug does not target opioid receptors is an important consideration in the development of effective alternatives for chronic pain.

In addition to *in vitro* selectivity, our *in vivo* results indicate that compound 8 also exhibits high potency for analgesia: 0.1 mg/kg eliminated CFA-induced inflammatory pain in the Hargreaves Test (**Figure 13**). In comparison, 50 mg/kg of the cannabinoid dehydroxyl-cannabidiol were needed to produce similar effects in a comparable behavioral test (100). Outside of the glycinergic pathway for pain treatment, compound 8 still compares favorably to other painkillers. For example, the ED<sub>50</sub>s of morphine, an opioid receptor agonist, and gabapentin, a voltage-gated calcium channel inhibitor, are 5.8 mg/kg and 26 mg/kg, respectively, to treat CFA-induced thermal hyperalgesia in rats (133,134). Both morphine and gabapentin can produce negative side effects, including abuse and addiction (135,136). The ED<sub>50</sub> of compound 8 determined here (0.03 mg/kg), is much smaller than that of either of these FDA-approved analgesics, decreasing the likelihood of adverse effects.

Other recent drug discovery efforts on GlyRs have relied on high-throughput screening (HTS) (137,138), requiring expensive infrastructure and large physical compound libraries. Computational screening is an economical alternative to HTS, with the additional benefit of providing inherent insight into the molecular mechanisms behind drug-target interactions (21,22). A recent HTS for potentiators of  $\alpha 3\beta$ GlyR could not provide any information on the mechanism of action for its hits (137) and a subsequent crystallization study was required to reveal the binding site in  $\alpha 3$ GlyR of the lead compound (97). Here, not only have we identified a novel analgesic compound, but in the process have determined the binding site and critical interactions required for its action on  $\alpha 3$ GlyR (**Figure 14-15**), offering an ideal platform for the rational lead optimization and design of even more potent analgesics.

Although we have detailed information on the binding mode of compound 8, we have not yet fully elucidated the molecular mechanisms of our new analgesic: how exactly does compound 8 binding to the cannabinoid site increase current through  $\alpha 3$ GlyR? As mentioned above, the TM4 helix was significantly separated from the TM3 helix in the open-state simulations of  $\alpha 3$ GlyR bound to compound 8 (**Figure 15**). This aligns with previous results showing that conformational transitions from the closed- to open-channel involve a concerted movement of TM4 away from the TM1/TM3 helices (139,140). Specifically, the aromatic contacts between TM4 and TM1/TM3 are critical for native pLGIC function (141-144), potentially because they facilitate interactions between the C-terminus of TM4 and the Cys-loop in the ECD that are essential for channel gating (59,145). Compound 8 interacts with many aromatic residues in the TM1, TM3, and TM4 helices through its multiple aromatic moieties (**Figure 15**); it is possible that compound 8 facilitates channel opening and/or stabilizes the open conformation by promoting these inter-helical aromatic interactions, thereby enhancing glycinergic currents. Further biochemical studies using  $\alpha 3$ GlyR mutants and non-aromatic analogs of compound 8 are required to verify this hypothesis.

In addition to its potentiation effects, compound 8 can also inhibit  $\alpha 3$ GlyR, depending on the effective glycine concentration applied during the electrophysiology measurements (**Figure 10b**). The ambient concentration of glycine in the extracellular and cerebrospinal fluid of the dorsal horn is usually low, ~5-20  $\mu$ M (146-148), corresponding to an effective concentration below  $EC_5$  (**Figure 14b**). Previous studies have shown that this concentration does not significantly change upon inflammation (148,149) or acute injury (150). Rather, inhibition of  $\alpha 3$ GlyR by intracellular prostaglandins is primarily responsible for increased pain sensitization during inflammation (90). Thus, although compound 8 inhibits  $\alpha 3$ GlyR at higher effective concentrations of glycine, these levels are likely less physiologically relevant for the treatment of pain.

The ultimate goal of this study was to identify potentiators of  $\alpha 3$ GlyR for the treatment of pain; however, our virtual screening calculations could not predict the functional impact of compound binding to  $\alpha 3$ GlyR and many of the top ranked compounds exhibited inhibitory effects on  $\alpha 3$ GlyR in electrophysiology experiments (**Figure 10a**). While inhibition of  $\alpha 3$ GlyR is likely less pharmacologically relevant than potentiation for the treatment of pain (90,92-94), it is important to note that this result was not wholly unexpected. Although many cannabinoids, including THC, cannabidiol, and dehydroxyl-cannabidiol (52,100), exclusively potentiate glycinergic currents, some cannabinoids are inhibitory modulators (151,152). If these structurally similar compounds all act at the same binding site in  $\alpha 3$ GlyR near S296, it is possible that different interactions within this binding site have the ability to positively or negatively modulate channel currents depending on the chemical composition of the compound. This mechanism could be considered analogous to competitive antagonism of the orthosteric binding site in pLGICs, such as strychnine binding to the same site as the agonist glycine in  $\alpha 3$ GlyR to inhibit the channel (95). Alternatively, multiple cannabinoid binding sites may exist in  $\alpha 3$ GlyR with conflicting functional effects and subtle structural changes among the different compounds could affect the relative functional contributions of each binding site. Previous studies have shown that pLGICs and other ion channels can contain multiple binding sites for a single compound and that binding to these different sites can produce varied functional effects (153-157). Additional structural and mechanistic studies are required to distinguish between these two hypotheses and develop a more precise virtual screening protocol to specifically target potentiation of  $\alpha 3$ GlyR.

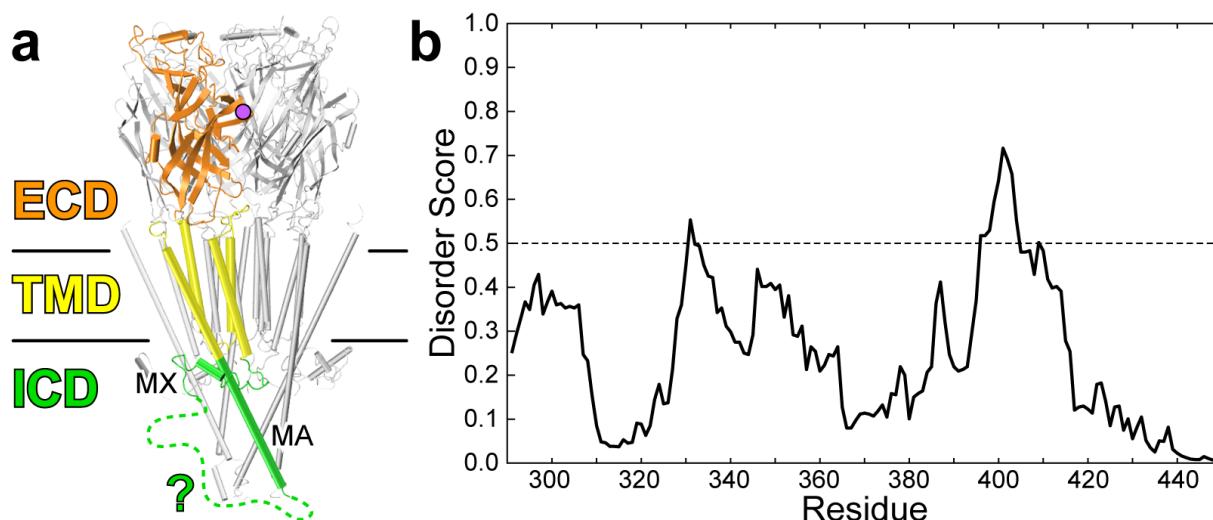
## 4.0 STRUCTURES OF THE A7 NICOTINIC ACETYLCHOLINE RECEPTOR INTRACELLULAR DOMAIN

### 4.1 INTRODUCTION

Nicotinic acetylcholine receptors (nAChRs) are members of the superfamily of pentameric ligand-gated ion channels (pLGICs) that can be activated by the neurotransmitter acetylcholine and the drug nicotine. nAChR subunits are classified as either  $\alpha$  ( $\alpha 2$ - $\alpha 10$ ) or  $\beta$  ( $\beta 2$ - $\beta 4$ ) based on the presence of adjacent cysteines in the extracellular domain (ECD) of only the  $\alpha$  subunits (158). These channels typically assemble as heteromers containing both  $\alpha$  and  $\beta$  subunits, with the exception of  $\alpha 7$ - $\alpha 9$  subunits that can form homomers. nAChRs are non-selective cation channels, permeable to  $\text{Na}^+$ ,  $\text{K}^+$ , and sometimes  $\text{Ca}^{2+}$  ions. In particular, the homomeric  $\alpha 7$ nAChR is 20 $\times$  more permeable for  $\text{Ca}^{2+}$  than other nAChR subtypes (159,160). Along with  $\alpha 4\beta 2$ nAChR,  $\alpha 7$ nAChR is the most abundant nAChR subtype in the mammalian brain (161) and is expressed in many different cortical and subcortical areas, most notably the hippocampus (162).

Like other eukaryotic pLGICs,  $\alpha 7$ nAChR consists of an agonist-binding ECD, a transmembrane domain (TMD) containing four transmembrane helices (TM1-TM4), and a large intracellular domain (ICD) connecting TM3 and TM4 (**Figure 16**). The ICD of pLGICs plays a critical role in trafficking, localization and assembly (32-35) as well as channel conductance and desensitization (163-166).  $\alpha 7$ nAChR is unique among nAChRs in its nearly instantaneous channel

desensitization upon agonist binding; this low opening probability and predisposition for non-conducting desensitized states may be associated with  $\alpha 7$ nAChR's regulation of intracellular signal transduction pathways (167). In addition,  $\alpha 7$ nAChR plays an important role in cognition, learning, and memory (1,168,169) and is an important target for treating Alzheimer's disease (1,170), schizophrenia (171,172), neuroinflammation (173-175), addiction (176), and other neurological disorders (177,178). The involvement of  $\alpha 7$ nAChR in such a wide variety of processes may be mediated by networking with intracellular protein partners associated with diverse downstream signaling pathways (36-41).



**Figure 16. Current structural knowledge of the pLGIC ICD.**

(a) Side view of the pLGIC structure with the highest number of resolved ICD residues up to this point (5-HT<sub>3A</sub>R, PDB: 6BE1) (179). Only the MX and MA helices are visible. Unresolved residues are marked with a dashed line. The orthosteric agonist binding site is marked in purple for reference. (b) Disorder prediction of the  $\alpha 7$ nAChR ICD from the IUPred server (180,181). The ICD is flexible, but only a few residues exceed the 0.5 score threshold for intrinsically disordered proteins.

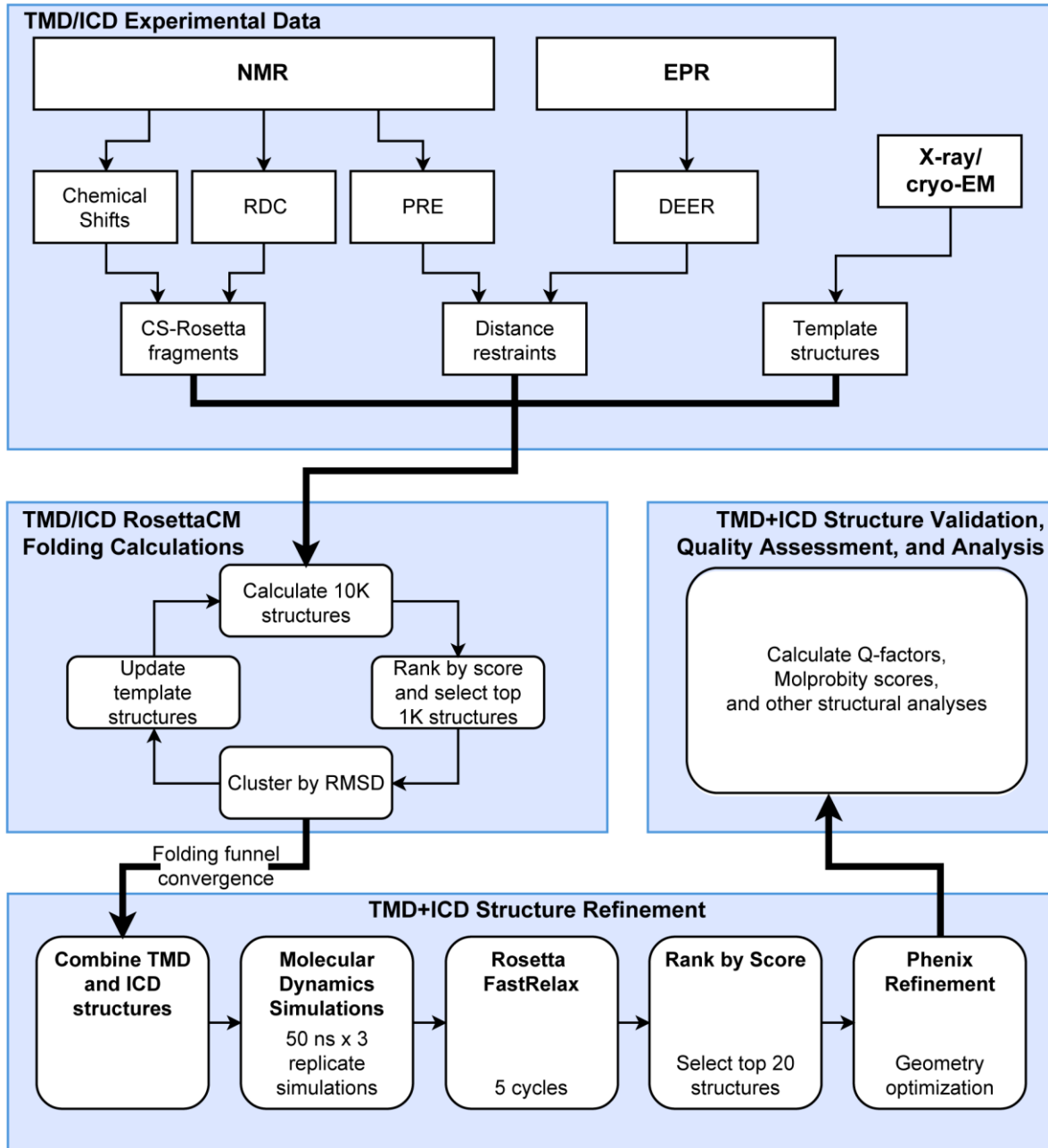
The ICD of  $\alpha 7$ nAChR is a particularly attractive target for therapeutic intervention as this domain is the most diverse in both size and composition among all pLGICs; the high degree of homology among the ECDs and TMDs of different pLGICs makes it difficult to develop drugs with high specificity for a single channel subtype. However, structure-based drug discovery is impossible without accurate structures of this domain. Although the number of high-resolution pLGIC structures has grown rapidly in recent years, no structure obtained so far contains a complete ICD. Large ICDs are naturally absent from the well-studied prokaryotic pLGICs, ELIC (79,182) and GLIC (183,184). X-ray and cryo-EM structures of eukaryotic pLGICs obtained so far are missing either the entire ICD (95-98,185-189) or a majority of ICD residues (179,190-193). In the single case where the full-length ICD was present in the sample, the cryo-EM structure of muscle-type *Torpedo marmorata* nAChR, only a small helical segment directly preceding the TM4 helix could be observed, termed the MA helix (194). It seems that the inherent dynamic nature of the  $\alpha 7$ nAChR ICD excludes the possibility of using traditional X-ray crystallography or cryo-EM to determine its structures. However, both NMR and EPR spectroscopy are well suited to provide structural information on the dynamics of flexible proteins (195-199).

To this end, we have determined structures of the  $\alpha 7$ nAChR TMD+ICD using a variety of restraints from NMR and EPR experiments implemented in Rosetta structure calculations, providing a first look at the complete structure and dynamics of the elusive pLGIC ICD. Many new structural features relevant to ionotropic and metabotropic channel function are revealed, including clear lateral windows for ion exit and insight into the structural basis of interactions between  $\alpha 7$ nAChR and various protein partners. The structures determined in this study lay the foundation for future efforts in a previously untouched area of drug development and the integrated approach described here can also be applied to solve the diverse ICD structures of other pLGICs.

## 4.2 METHODS

An overview of the protocol used to determine structures of the  $\alpha 7$ nAChR TMD+ICD is shown in

**Figure 17.**



**Figure 17.** Overview of the  $\alpha 7$ nAChR TMD+ICD structure determination procedures.



## 4.2.1 Experimental Methods

This work is a collaborative effort and writing for the experimental methods was primarily contributed by other members of Professor Pei Tang's lab. These methods are included here for completeness.

### 4.2.1.1 Sample Preparations

$\alpha 7$ nAChR constructs containing only a single unpaired cysteine were prepared after mutagenesis of the native cysteines as follows: C116S-C219A-C317S-C335S-C342S-C375S-C385S-C390S-C427S-C435S-C443S-C449A-C460A. Two pairs of disulfide-bonded cysteines in the ECD (C128-C142 in the Cys-loop and C190-C191 in loop C) were not mutated in order to retain native function of the channel. Single cysteine constructs used in this study were C219, V311C, C317, C342, C375, C390, L415C, C427, C435, and C449. Site directed mutagenesis was performed using the QuickChange Lightning Kit for single or multi-site mutagenesis (Agilent Technologies). Full-length  $\alpha 7$ nAChR was expressed in *E. coli* and purified by NiNTA chromatography as previously described (200).

For solution NMR experiments, the  $\alpha 7$ nAChR TMD+ICD constructs were expressed in *E. coli* Rosetta 2(DE3) pLysS (Novagen) in  $^{15}\text{N}$  or  $^{15}\text{N}/^{13}\text{C}$  M9 media at 15°C for approximately three days before purification in LDAO using a His-tag affinity column following established protocols (201-203). The final NMR samples contained 0.2 - 0.3 mM protein, 1-2 % (~44 - 87 mM) LDAO detergent or nanodiscs (see below), 5 mM sodium acetate at pH 4.7, and 25 mM NaCl. 5% D<sub>2</sub>O was added for deuterium lock in NMR experiments. To prepare the  $\alpha 7$ nAChR TMD+ICD sample in nanodiscs, DMPC phospholipids were added to the sample in LDAO micelles at a molar ratio of 80 lipids to one pentameric protein. The sample was incubated for 30 minutes before adding the

membrane scaffold protein MSP1D1 to the solution at a molar ratio of 2 (MSP1D1) to 80 (lipids). After incubating for one hour, detergent was removed using Bio-Beads SM-2 Resin (Bio-Rad Laboratories) or Detergent Removal Resin (Thermo Scientific).

For paramagnetic relaxation enhancement (PRE) NMR experiments, single-cysteine  $\alpha 7nAChR$  TMD+ICD constructs in LDAO micelles were treated with the reducing agent TCEP for approximately one hour. After removing TCEP, the sample was labeled with a ~15 to 25-fold molar excess of the nitroxide spin label MTSL (Toronto Research Chemicals) to ensure labeling efficiency above 90%. After a 2-hour incubation at room temperature, the reaction was allowed to proceed overnight at 4°C. To remove free MTSL, the samples were dialyzed with three buffer changes and then subjected to size exclusion chromatography on a Superdex 200 10/300 column (GE Healthcare). For  $^{19}F$  NMR PRE experiments or  $^{19}F$  spin-spin ( $R_2$ ) relaxation measurements, the same single-cysteine TMD+ICD sample was labeled separately with three different reagents: TET ( $^{19}F$  probe), MTSL (paramagnetic probe), or d-MTSL (diamagnetic probe).  $^{19}F$  labeling with TET followed previously published protocols (157,204-206). The site-directed TET-labeled protein (60  $\mu M$ ) was mixed with site-directed MTSL-labeled or d-MTSL-labeled protein (~180  $\mu M$ ) in 10 mM PBS buffer at pH 7.4 with 150 mM NaCl, 1% LDAO, and 5 %  $D_2O$ .

For residual dipolar coupling (RDC) NMR experiments, a lanthanide ion (paramagnetic:  $Tm^{3+}$  and  $Dy^{3+}$ ; diamagnetic:  $Lu^{3+}$ ) was incorporated into the single-cysteine  $\alpha 7nAChR$  TMD+ICD constructs using a thiol-specific disulfide reagent, N-[S-(2-pyridylthio)cysteaminy]-EDTA (Toronto Research Chemicals), following published methods (207,208). Briefly, a two-fold molar excess of  $LnCl_3$  (Sigma-Aldrich) was added to 10 mM N-[S-(2-pyridylthio)cysteaminy]-EDTA in 100 mM Tris buffer at pH 7.3 and incubated for one hour. After incubation, EDTA was added to the solution to adjust the free  $LnCl_3$  amount below a 5% excess and the solution was incubated for

another 30 minutes. The TMD+ICD sample was treated with DTT in a 20-fold molar excess for 30 minutes. After removing DTT, the TMD+ICD sample in 50 mM Tris buffer at pH 8.0, 120 mM NaCl, and 0.5% LDAO was mixed with a 10-fold molar excess of N-[S-(2-pyridylthio)cysteaminy]-EDTA-Ln<sup>3+</sup> and incubated for two hours. All reactions were done at a room temperature in the dark. A desalting column was used to remove free small molecules from the protein sample.

For EPR experiments, the purified  $\alpha 7$ nAChR constructs in LDAO micelles were labeled with MTSL using a protocol similar to that described above for PRE NMR experiments, but with a five-fold molar excess of MTSL. For DEER EPR measurements, the buffer was exchanged to PBS prepared in D<sub>2</sub>O and  $\alpha 7$ nAChR constructs in lipid vesicles were prepared by adding a 100-fold molar excess of asolectin to the detergent solubilized protein. The detergent was then removed by incubation with Bio-Beads SM-2 Resin (Bio-Rad Laboratories) and the vesicles were collected by centrifugation at 200,000  $\times$  g for one hour.

#### **4.2.1.2 Electrophysiology with *Xenopus* Oocytes**

The function of the full-length  $\alpha 7$ nAChR and TMD+ICD constructs was measured using two-electrode voltage clamp electrophysiology experiments in *Xenopus laevis* oocytes (78). For single cysteine  $\alpha 7$ nAChR mutants, 25 ng of  $\alpha 7$ nAChR RNA was co-injected with 25 ng RNA for RIC-3, an  $\alpha 7$ nAChR chaperone that enhances surface expression (209). Responses to acetylcholine activation were compared to that of wild type  $\alpha 7$ nAChR (200). The function of the  $\alpha 7$ nAChR TMD+ICD construct was measured by injecting 100 ng of purified TMD+ICD in asolectin vesicles directly into *Xenopus laevis* oocytes. Ivermectin was used to activate the channel current (203). Channel function was measured 1-3 days post-injection in a 20- $\mu$ L oocyte recording chamber (Automate Scientific) clamped at -60 mV with an OC-725C Amplifier (Warner

Instruments) and a Digidata 1440A digitizer (Axon Instruments). Recording solutions contained 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, and 5 mM HEPES at pH 7.0. Data were collected and processed using Clampex 10 software (Molecular Devices).

#### 4.2.1.3 NMR Spectroscopy

##### *Collection of NMR Spectra*

NMR spectra were recorded on Bruker Avance 700 - 900 MHz spectrometers equipped with a triple-resonance inverse-detection TCI cryoprobe (Bruker Instruments). A relaxation delay of 1 s was used in all NMR experiments unless specified otherwise. <sup>1</sup>H-<sup>15</sup>N TROSY-HSQC spectra were acquired at 25, 30, 35, 40, and 45°C with 1024 and 160 - 200 data points and spectral widths of 13 and 23 - 28 ppm for the <sup>1</sup>H and <sup>15</sup>N dimensions, respectively. <sup>1</sup>H-<sup>13</sup>C HSQC spectra were collected at 45°C with 2048×320 data points and spectral widths of 11 ppm (<sup>1</sup>H) and 64 ppm (<sup>13</sup>C). In all NMR spectra, the <sup>1</sup>H chemical shifts were referenced to the DSS resonance at 0 ppm and the <sup>15</sup>N and <sup>13</sup>C chemical shifts were referenced indirectly (210).

For <sup>1</sup>H, <sup>15</sup>N, and <sup>13</sup>C chemical shift assignment, a suite of three-dimensional (3D) spectra was acquired at 45°C based on TROSY pulse sequences (211-214), including HNCA and HN(CO)CA with data points of 1024×36×80 and corresponding spectral windows of 12×23×28 ppm for the <sup>1</sup>H, <sup>15</sup>N, and <sup>13</sup>C dimensions, respectively; HNCACB and CBCA(CO)NH (1024×36×104 points, 12×23×56 ppm); HNCO (1024×56×64 points, 12×23×12 ppm); and <sup>15</sup>N-edited NOESY (215-217) (1024×40×144 points, 13×23×13 ppm) with a mixing time of 200 ms. The 3D NMR experiments were performed on both the wild-type and single-cysteine mutant TMD+ICD samples at 25°C and/or 45°C.

To determine PRE NMR distance restraints,  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra were acquired on a group of single-cysteine TMD+ICD constructs labeled with MTSL in the absence and presence of a  $\sim 10$ -fold excess of ascorbic acid. To determine inter-subunit distances in the MA helical region,  $^{19}\text{F}$  PRE was quantified through 1D  $^{19}\text{F}$   $R_2$  relaxation NMR spectra of TMD+ICD pentamers containing site-directed  $^{19}\text{F}$  labeled subunits mixed with MTSL or d-MTSL labeled subunits. For each sample, a series of spectra were collected with varied  $\tau$  delays on a Bruker Avance 600 MHz spectrometer ( $^{19}\text{F}$  frequency of 564.6 MHz) at 25°C. Other experimental parameters include 30 ppm spectral width, 8192 complex points, and one second recycle delay.

To determine RDC NMR restraints,  $^1\text{H}$ - $^{15}\text{N}$  IPAP HSQC spectra (218,219) were recorded at 35°C for the single-cysteine TMD+ICD constructs labeled with the paramagnetic N-[S-(2-pyridylthio)cysteaminy]-EDTA- $\text{Tm}^{3+}$ / $-\text{Dy}^{3+}$  or labeled with the diamagnetic N-[S-(2-pyridylthio)cysteaminy]-EDTA- $\text{Lu}^{3+}$ . The spectra were acquired with 2048 $\times$ 512 data points and 16 $\times$ 23 ppm spectral windows in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions, respectively.  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra were similarly acquired.

To investigate the backbone dynamics of the TMD+ICD,  $^{15}\text{N}$  spin-lattice ( $R_1$ ) and spin-spin ( $R_2$ ) relaxation rates as well as  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear Overhauser effects (hetNOEs) were determined. The spectra for  $R_1$  and  $R_2$  measurements were acquired with 1024 $\times$ 176 data points and 13 $\times$ 23 ppm spectral widths in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions, respectively, and a relaxation recycle delay of 3 seconds. Variable delays ( $\tau$ ) ranging from 16 - 640 ms or 20 - 1500 ms were used for the  $R_2$  or  $R_1$  measurements, respectively. Each hetNOE spectrum had 1024 $\times$ 184 data points and 13 $\times$ 23 ppm spectral widths in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions, respectively, and a relaxation recycle delay of 5 seconds.

## *NMR Data Processing and Analysis*

NMR spectra were processed using NMRPipe 4.1 and NMRDraw 1.8 (220) and analyzed using Sparky 3.13 (221). The  $^1\text{H}$ ,  $^{15}\text{N}$ , and  $^{13}\text{C}$  chemical shifts were assigned manually based on the suite of 3D NMR spectra described above and used to obtain secondary structure predictions by TALOS (222). Hydrogen-bonding restraints were generated based on temperature coefficients of the backbone amide proton chemical shifts in the  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra acquired at different temperatures (25, 30, 35, 40, and 45 °C). Residues with coefficients  $< 4.5$  ppb/K were considered to have hydrogen bonding in helical regions (223).  $^{15}\text{N}$   $R_1$  and  $R_2$  relaxation rates and their uncertainties were obtained from the exponential fitting of NMR peak intensities versus the variable delays. The hetNOE values were calculated as the ratios of peak intensities with and without proton saturation and the associated uncertainties were determined from the signal-to-noise ratio (224).

NMR PRE-derived distances between the amide protons and the unpaired electron in the labeled MTSL were obtained from **Equation 2**, where  $K$  is a constant ( $1.23 \times 10^{-32} \text{ cm}^6\text{s}^{-2}$ ) describing the spin properties of the MTSL label,  $\omega_h$  is the Larmor frequency of the proton spin,  $\tau_c$  is the global correlation time of the protein, and  $R_2^{\text{sp}}$  is the paramagnetic enhancement of the transverse relaxation rate as defined in **Equation 3**. NMR peak intensities in the oxidized ( $I_{\text{ox}}$ ) and reduced ( $I_{\text{red}}$ ) samples were used to obtain  $R_2^{\text{sp}}$  in **Equation 3** as described previously (225), where  $t$  is the total INEPT evolution time (10 ms) and  $R_2$  is the amide proton transverse relaxation rate in the reduced state calculated using **Equation 4**. The half-linewidth,  $\Delta\nu_{1/2}$ , is determined from the peak fitting with a Lorentzian line shape.

**Equation 2. PRE NMR distance restraint.**

$$r = \left[ \frac{K}{R_2^{sp}} \left( 4\tau_c + \frac{3\tau_c}{1 + \omega_h^2 \tau_c^2} \right) \right]^{\frac{1}{6}}$$

**Equation 3. PRE NMR intensity ratio.**

$$\frac{I_{ox}}{I_{red}} = \frac{R_2 e^{-R_2^{sp} t}}{R_2 + R_2^{sp}}$$

**Equation 4. NMR transverse relaxation rate.**

$$R_2 = \pi \Delta \nu_{1/2}$$

RDCs for individual resonances were obtained from the difference between the  $^{15}\text{N}$  chemical shifts in paramagnetic and diamagnetic IPAP spectra. The relation of RDC to the orientation of an aligned protein is defined by **Equation 5** (226-228), where  $D_a$  is the magnitude of the axial component of the molecular alignment tensor and  $R$  is the rhombicity. The angles  $\theta$  and  $\phi$  determine the N-H vector orientation relative the alignment tensor. The principal components of the alignment tensor were calculated based on the experimental RDC values using the REDCAT software (229,230).

**Equation 5. RDC NMR angles.**

$$RDC(\theta, \phi) = D_a(3\cos^2\theta - 1) + \frac{3}{2}R * \sin^2\theta\cos 2\phi$$

#### 4.2.1.4 EPR Spectroscopy

CW EPR experiments were performed at room temperature using a Bruker ElexSys 580 CW/FT X-band EPR spectrometer and a Bruker ER4122 SHQE-W1 high-resolution resonator. Aliquots of MTSL-labeled  $\alpha 7nAChR$  in detergent micelles in PBS (approximately 8-10  $\mu L$ ) were drawn into Pyrex capillary sample tubes. CW experiments were performed with a modulation amplitude of 1 G, a modulation frequency of 100 kHz, and a conversion time of 20.48 ms for a total of 1024 points. The doubly-integrated intensity of the CW spectra was compared to a calibration curve to determine the spin concentration and MTSL labeling efficiency.

Four-pulse DEER experiments (195) were performed on either a Bruker ElexSys 580 CW/FT X-band ESR spectrometer or Bruker ElexSys 680 CW/FT X-band ESR spectrometer with a Bruker EN4118X-MD4 or Bruker ER4118-MD5 resonator. Samples contained 100  $\mu L$  of MTSL-labeled  $\alpha 7nAChR$  reconstituted in lipid vesicles in deuterated PBS and 20  $\mu L$  glycerol. Samples were then transferred into a 3 mm quartz tube and flash frozen in liquefied methylacetylene-propadiene and propane (MAPP) gas. The sample temperature was controlled at 80 K using an Oxford ITC503 temperature controller and an Oxford CF935 dynamic continuous-flow cryostat connected to an Oxford LLT 650 low-loss transfer tube. DEER experiments were carried out using the following pulse sequence:  $(\pi/2)_{\nu_1}-\tau_1-(\pi)_{\nu_1}-T-(\pi)_{\nu_2}-\tau_2-(\pi)_{\nu_1}-\tau_2$ -echo (231). The pump frequency ( $\nu_2$ ) was set at the center of the nitroxide triplet with a pulse length  $(\pi)_{\nu_2}$  of 16 ns. The observer frequency ( $\nu_1$ ) was offset by 70 MHz. The observer pulses  $(\pi/2)_{\nu_1}$  and  $(\pi)_{\nu_1}$  were set to 16 ns and 32 ns respectively. DEER data were acquired for 128 points. The time domain DEER signal data were analyzed using DeerAnalysis2016 by model-free Tikhonov regularization (232).



## 4.2.2 Computational Methods

### 4.2.2.1 Rosetta Input Preparation

Three classes of NMR PRE distance restraints were derived and employed in Rosetta structure calculations. In class 1, an upper bound of 15 Å was set for residues undetectable in the oxidized NMR spectrum, i.e.  $I_{\text{ox}}/I_{\text{red}} \leq 0.2$ . In class 2, a lower bound of 25 Å was set for residues with  $I_{\text{ox}}/I_{\text{red}} \geq 1.0$ ; in this case, the spin label was sufficiently far away to not affect the spectrum. In class 3, the distances were calculated using **Equation 2** for detectable residues with  $0.2 < I_{\text{ox}}/I_{\text{red}} < 0.8$ . All three classes of PRE distance restraints were measured between the backbone nitrogen atom and C $\beta$  atom of the MTSL-labeled residue. An error bound of  $\pm 5$  Å was used in the initial calculations and was gradually reduced to  $\pm 3$  Å in the later stages of calculations (round 14+). The error bound partially accounts for the inherent size and motion of the MTSL label itself, which could not be explicitly included in the structure calculations (225,233-235). Class 1 and class 2 restraints were implemented in Rosetta as sigmoid functions, where distances above 15 Å and below 25 Å were penalized, respectively. Class 3 restraints were implemented as harmonic constraints, where distances outside the error bound were penalized. Since PRE data do not distinguish intra- and inter-subunit distances, both class 1 and class 3 restraints were implemented ambiguously, where either intra- or inter-subunit residues could satisfy the restraint in the pentamer. Class 2 restraints were applied unambiguously for both intra- and inter-subunit residues.

EPR DEER-derived quaternary structure restraints for the adjacent and across subunit distances were implemented in Rosetta as harmonic constraints, where distances outside  $\pm 1$  SD from the model-free Tikhonov fitted mean were penalized.

#### 4.2.2.2 TMD/ICD Structure Folding Calculations

The comparative modeling protocol (RosettaCM) (236) in Rosetta 3.7 (237,238) with the talaris2014 energy function (239,240) was used for all structure calculations. Briefly, Rosetta uses a Monte Carlo Metropolis sampling algorithm to efficiently assemble short peptide fragments into a complete structure (241,242). Fragment libraries were generated using CS-Rosetta (243-247) on the Robetta server (248) with input chemical shifts and RDC data from NMR experiments. The RosettaCM protocol combines the *de novo* structure determination approach with known structures of homologous regions: aligned segments are merged with unaligned fragments and the resulting models are optimized by all-atom refinement and evaluated with physics- and knowledge-based potentials as well as their agreement with experimental restraints.

Rosetta has previously been used to generate structures of atomic-level accuracy for proteins of moderate size (249-254), but the quality of models for larger proteins is often compromised by the lack of computational capacity to fully sample all potential conformations (255,256). Experimental restraints (233,243,257-259) and symmetry constraints (260-262) help to restrict the conformational space and significantly improve the accuracy of calculations for larger structures. In addition to applying NMR and EPR experimental restraints and pentameric symmetry constraints in Rosetta structure calculations, we developed an iterative folding protocol (**Figure 17**) to guide conformational sampling towards the global energy minimum while maintaining structural diversity to ensure that calculations do not become trapped in local energy minima (243,263,264). In each iteration of ICD folding (residues I291 to C449), 10,000 structures were calculated using input fragment libraries and restraints as described above, as well as template structures from homologous pLGICs as described below (only used in round 1) or from previous iterations (round 2+). These 10,000 structures were ranked by total score, calculated by **Equation**

6, where  $S_{\text{physics}}$  and  $S_{\text{knowledge}}$  are the standard weighted physics- and knowledge-based potentials from the talaris2014 energy function, respectively (239,240).  $S_{\text{DEER}}$  and  $S_{\text{PRE}}$  are the harmonic and sigmoid restraint potentials described above, with optimized weights ( $w_1 = 10$ ,  $w_2 = 5$ ,  $w_3 = 2$ ) as described in the Results section. The top 1,000 structures were grouped by complete-linkage hierarchical clustering using a 3 Å RMSD cutoff (265). In the first iteration, pairwise RMSD values were calculated only over the backbone atoms of residues within the MX and MA helices (residues I291 to R322 and L411 to C449, respectively). In subsequent iterations, pairwise RMSD values were calculated over an additional ten residues, extending five residues forward from the MX helix and five residues backward from the MA helix, until all ICD residues were covered in the 10<sup>th</sup> iteration. The top ranked structures from each cluster were input as new template structures for the next iteration of folding calculations.

**Equation 6. Total score for Rosetta structures**

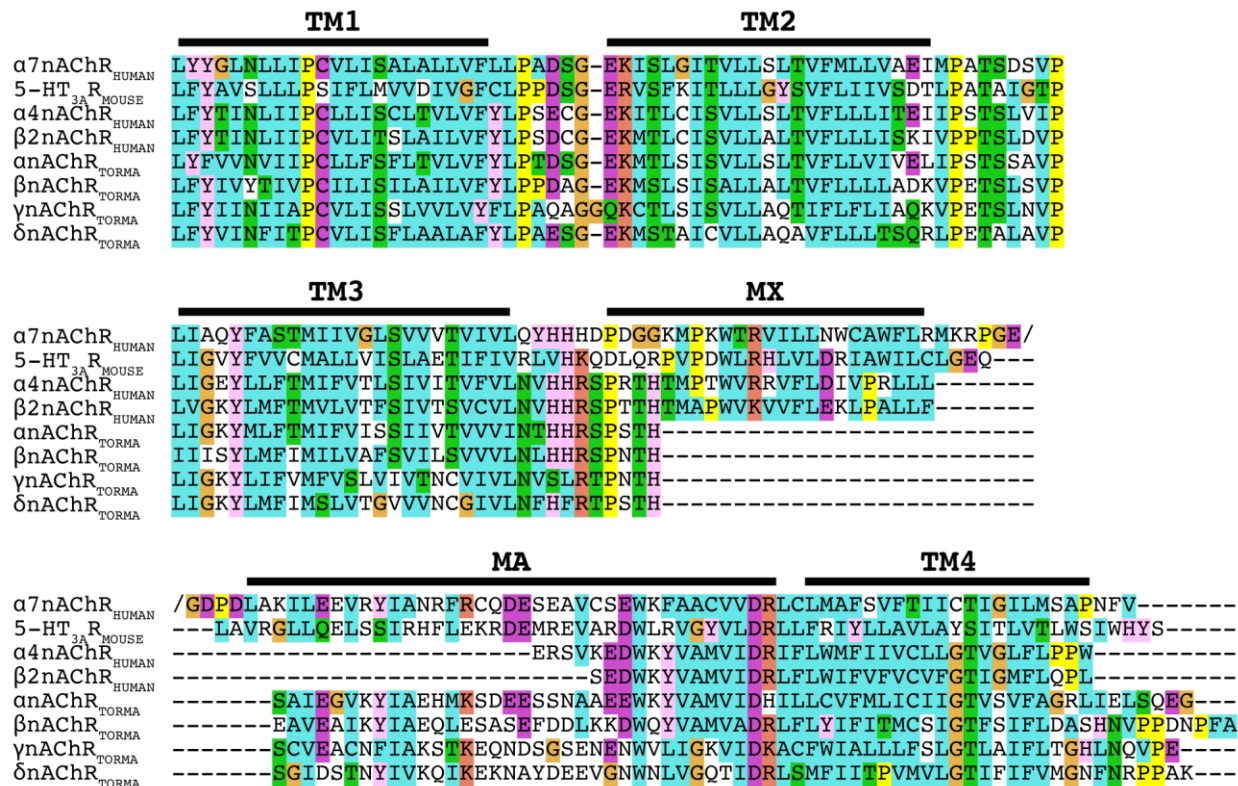
$$S_{\text{total}} = S_{\text{physics}} + S_{\text{knowledge}} + w_1 S_{\text{DEER}} + w_2 (S_{\text{PRE}}^{\text{class1}} + S_{\text{PRE}}^{\text{class3}}) + w_3 S_{\text{PRE}}^{\text{class2}}$$

We have previously determined the NMR structure of the  $\alpha 7\text{nAChR}$  TMD monomer (203). A similar iterative protocol as described above was employed to determine the pentameric structure of the TMD (residues L209 to G303 and D446 to V472, where TM3 and TM4 were connected by the same short linker as used previously (203)). In each iteration, 10,000 structures were calculated using input fragment libraries and experimental restraints as described above. A 2 Å RMSD cutoff was used for complete-linkage hierarchical clustering of the top ranked 1,000 structures in each iteration. Pairwise RMSD values were calculated only over residues within the helical regions of the TMD. The top ranked structures from each cluster were input as new template structures for the next iteration. All folding calculations for both the TMD and ICD were performed

using resources from Open Science Grid (113,114), which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

Previously determined MX and MA helical structures from homologous pLGICs were used as initial input templates in RosettaCM ICD folding calculations. Specifically, the MX and/or MA helices in the x-ray crystal structures of 5-HT<sub>3A</sub>R (PDB: 4PIR) (190) and  $\alpha$ 4 $\beta$ 2nAChR (PDB: 5KXI, MX helices) (193) as well as the cryo-EM structures of 5-HT<sub>3A</sub>R (PDB: 6BE1) (179) and *Torpedo marmorata* nAChR (PDB: 2BG9) (194) were used in round 1 as described above. NMR structures of the  $\alpha$ 7nAChR TMD monomer (PDB: 2MAW) (203) aligned to the quaternary structure of the resting state 5-HT<sub>3A</sub>R and *Torpedo marmorata* nAChR TMDs were used as input templates in the first iteration of Rosetta TMD folding calculations. The quaternary structure of human  $\alpha$ 4 $\beta$ 2nAChR was excluded from these pentameric TMD assembly calculations since it is in a desensitized state. Sequence alignments of these channels with the  $\alpha$ 7nAChR TMD+ICD are shown in **Figure 18**.

The top ranked 1,000 structures from the final round of folding calculations were clustered by pairwise backbone RMSD values. A structure of the  $\alpha$ 7nAChR TMD+ICD was generated by joining the top ranked structures from the major conformation of the TMD and ICD, respectively. The separate domains were aligned along the interfacial residues (I291 to Q294 and D446 to C449) and the ICD replaced the short artificial TM3-TM4 linker in the TMD structure. The resulting TMD+ICD structure was energy minimized and progressed to structure refinement calculations.



**Figure 18. Sequence alignment of  $\alpha 7$ nAChR with input template structures.**

Alignments of the TM1-4, MX, and MA helices for  $\alpha 7$ nAChR with the previously solved x-ray/cryo-EM structures of other pLGICs (179,190,193,194). Only resolved residues for each structure are shown (PDB: 4PIR, 6BEI, 5KXI, 2BG9).  $\alpha 7$ nAChR residues between the MX and MA helices (E328-G407) are not shown for clarity.

#### 4.2.2.3 TMD+ICD Structure Refinement

To refine the  $\alpha 7$ nAChR TMD+ICD structure in an explicit lipid/water environment and optimize the TMD/ICD interface, we performed molecular dynamics (MD)-based refinement of the TMD+ICD structure embedded in a pre-equilibrated POPC lipid bilayer mixed with 20% cholesterol (66,67). The inclusion of cholesterol was based on the knowledge that eukaryotic cell membranes are typically composed of approximately 20% cholesterol (64). Systems were solvated in TIP3P water (266), ionized with 100 mM NaCl, and contained  $\sim 174,000$  atoms. MD simulations were run using GROMACS 5.1 (108) and the CHARMM36 force field (115,119-122). Energy

minimization was performed with 1,000 kJ/mol/nm<sup>2</sup> harmonic position restraints on all non-hydrogen atoms until the maximum force on any single atom was not greater than 100 kJ/mol/nm (~10,000 steps), followed by 3 ns of equilibration, during which harmonic position restraints on all non-C $\alpha$  atoms were gradually reduced from 1,000 to 0 kJ/mol/nm<sup>2</sup>. 1,000 kJ/mol/nm<sup>2</sup> harmonic position restraints were applied to all C $\alpha$  atoms over the entire simulation to maintain the initial protein conformation. Three replicate 50-ns simulations were run at a constant pressure and temperature (NPT) of 1 atm and 310 K. All bonds were constrained using the LINCS algorithm (74). The integration time step used for all simulations was 2 fs. Particle mesh Ewald was used for long-range electrostatic interactions (109). A 12-Å cutoff was used for nonbonded interactions. Full electrostatic and nonbonded interactions were evaluated every step and the neighbor list was updated every five steps. Systems were simulated with periodic boundary conditions in three dimensions.

After MD-based refinement, final energy minimization calculations were performed in Rosetta to score the TMD+ICD structures and reintroduce strict pentameric symmetry. The FastRelax protocol, which consists of cycles of packing and all-atom minimization (267-269), was run using individual monomers extracted every 100 ps from the last 40 ns of each replicate simulation (400 structures  $\times$  3 replicate simulations  $\times$  5 subunits = 6,000 structures per system). Structures were relaxed for five cycles with NMR and EPR distance restraints as described above. The resulting structures were ranked by total score and the top 20 structures were optimized in Phenix 1.13 (270) with Ramachandran restraints. Rosetta FastRelax refinement calculations were performed using resources from Open Science Grid (113,114), which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

#### 4.2.2.4 Structure Validation, Quality Assessment, and Analysis

Agreement between the calculated and observed distances for PRE NMR and DEER EPR restraints was evaluated using Q-factors (271-273) (analogous to the crystallographic R-factor) defined in **Equation 7**, where  $r_{\text{exp}}$  is the experimental restraint distance and  $r_{\text{calc}}$  is the distance observed in the calculated structure. Q-factors were calculated separately for the DEER EPR and PRE NMR restraints ( $Q^{\text{DEER}}$  and  $Q^{\text{PRE}_{\text{work}}}$ , respectively). In addition, ~10% of the PRE restraints (13 and 32 restraints from class 2 and class 3, respectively) were randomly selected to exclude from the structure calculations for validation purposes. Q-factors for these restraints ( $Q^{\text{PRE}_{\text{free}}}$ ) were calculated based on these restraints.

**Equation 7. Q-factors for calculated structures.**

$$Q = \sqrt{\frac{\sum_i (r_{\text{exp}}(i) - r_{\text{calc}}(i))^2}{\sum_i r_{\text{exp}}(i)^2}}$$

The geometry and stereochemistry of the final structures was evaluated by MolProbity (274) implemented in Phenix 1.13 (270). Structure statistics are presented in **Table 10** in **Section 4.3.6**.  $\alpha$ 7nAChR ICD structures were subjected to principal component analysis using the Python package ProDy (275) to calculate the superposition to the mean and the covariance matrix, which was diagonalized to determine the principal modes of structural variations observed in the ensemble. The modes were rank ordered by the size of the corresponding eigenvalues, where Mode 1 is the direction of maximal variance, followed by Mode 2, etc. VMD (76) was used for structure rendering, visualization, and analysis. Pore profiles were calculated using the HOLE program (80).

## 4.3 RESULTS

This work is a collaborative effort and the data from electrophysiology, NMR, and EPR experiments described below were obtained by other members of Professor Pei Tang, Professor Yan Xu, and Professor Sunil Saxena's labs. The results are included here for the completeness.

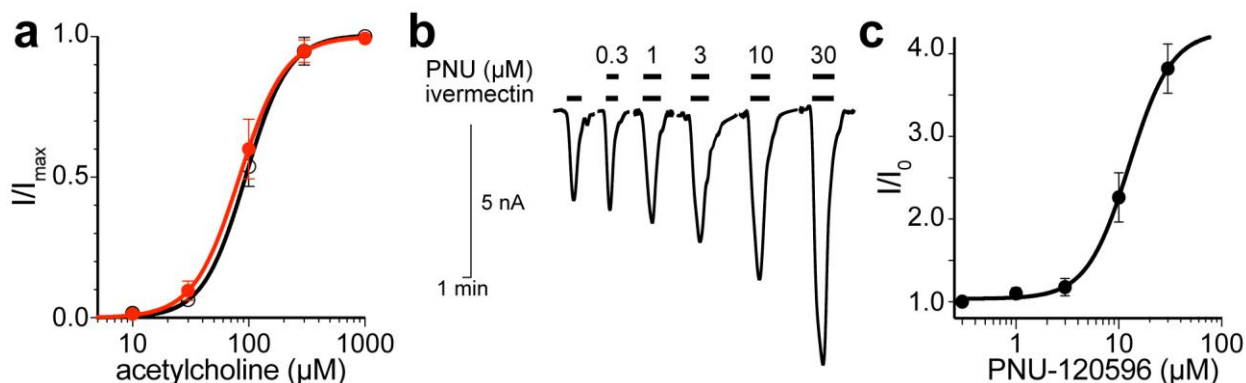
### 4.3.1 Functional Characterization of $\alpha 7$ nAChR Constructs for Structural Studies.

Human  $\alpha 7$ nAChR expressed in *E. coli* and purified by NiNTA chromatography was used for all NMR and EPR experiments as described below. We have previously shown that this full-length WT  $\alpha 7$ nAChR construct can form functional ion channels that retain the pharmacological characteristics of native  $\alpha 7$ nAChR when reconstituted in *Xenopus laevis* oocytes (200). Most  $\alpha 7$ nAChR constructs used for structural studies here contain only a single unpaired cysteine residue for precise site-directed spin labeling. To verify that this cysteine mutagenesis did not affect native channel function, the single-cysteine C317 mutant  $\alpha 7$ nAChR was expressed in *Xenopus laevis* oocytes and responses to the agonist acetylcholine were compared to WT  $\alpha 7$ nAChR (**Figure 19a**). Activation curves for the C317 mutant and WT  $\alpha 7$ nAChR were not significantly different, indicating that the removal of unpaired cysteines in  $\alpha 7$ nAChR does not affect native channel function.

Furthermore, we previously demonstrated that the  $\alpha 7$ nAChR TMD alone can spontaneously assemble into a pentamer and form a functional ion channel in the absence of the ECD and ICD (203). Although the  $\alpha 7$ nAChR TMD in isolation does not contain the native orthosteric agonist-binding site, channel current can be elicited by ivermectin, a positive allosteric modulator that acts through the TMD (276). Similarly, we examined the channel function of the



$\alpha 7$ nAChR TMD+ICD expressed in *Xenopus laevis* oocytes. Like the TMD alone, the TMD+ICD was activated by ivermectin, as well as potentiated by PNU-120596, an  $\alpha 7$ nAChR-specific positive allosteric modulator (**Figure 19b-c**) (86). Thus, the  $\alpha 7$ nAChR constructs used here have pharmacological properties similar to those of  $\alpha 7$ nAChR in the native environment and are suitable for structural studies.



**Figure 19. Functional characterization of  $\alpha 7$ nAChR constructs.**

(a) *In vitro* electrophysiology responses to acetylcholine from *Xenopus laevis* oocytes expressing WT (black) and C317 mutant  $\alpha 7$ nAChR. Fitting to the Hill equation yields  $EC_{50}$  values of  $94 \pm 3 \mu\text{M}$  and  $83 \pm 4 \mu\text{M}$  for WT and C317 mutant  $\alpha 7$ nAChR, respectively ( $n = 5$  oocytes). The C317 mutant  $\alpha 7$ nAChR has all TMD cysteines mutated to alanine and all ICD cysteines mutated to serine except C317. All unpaired ECD cysteines were also mutated to serine. (b) Representative traces from *Xenopus laevis* oocytes injected with vesicles containing the purified  $\alpha 7$ nAChR TMD+ICD construct show activation by  $30 \mu\text{M}$  ivermectin and potentiation by the indicated concentrations of PNU-120596. (c) Dose response curves for PNU-120596 potentiation of *Xenopus laevis* oocytes injected with vesicles containing the purified  $\alpha 7$ nAChR TMD+ICD construct. Modulation was measured as the ratio of currents in the presence ( $I$ ) and absence ( $I_0$ ) of  $30 \mu\text{M}$  ivermectin. Data were fit to the Hill equation and  $EC_{50} = 10 \pm 1 \mu\text{M}$  ( $n = 5$  oocytes).

### 4.3.2 Structural Restraints from NMR Experiments

Preparation of the  $\alpha 7$ nAChR TMD+ICD in LDAO micelles provided the highest-quality NMR spectra and allowed for almost complete chemical shift assignment (**Table 9**). However, LDAO is a relatively harsh detergent that has the potential to destabilize native protein structures (277). To verify that LDAO does not destabilize the structure of  $\alpha 7$ nAChR, NMR spectra were also collected for  $\alpha 7$ nAChR TMD+ICD constructs in DMPC nanodiscs (278), a membrane mimetic similar to the native bilayer environment (279). As expected, fewer  $\alpha 7$ nAChR residues are observed in NMR spectra of DMPC nanodiscs compared to LDAO micelles due to the larger size and slower diffusion of nanodiscs (280) (**Figure 20a**). However, ICD residues that are visible in both micelle and nanodisc spectra have similar chemical shifts ( $\Delta\delta < 0.08\text{ppm}$ ), suggesting that the protein structure is nearly the same in both preparations. Thus, LDAO micelles do not destabilize the native structure of  $\alpha 7$ nAChR and are appropriate for structural studies.

Secondary structures and order parameters of the  $\alpha 7$ nAChR TMD+ICD were analyzed based on chemical shifts (281,282), revealing the four conserved TM helices, as well as MA and MX helices at the N- and C-termini of the ICD, respectively (**Figure 20b**). Additional small regions with  $\alpha$ -helical character were also observed at discrete locations in the ICD between the conserved MA and MX helices, seemingly separated by less-ordered short loops.

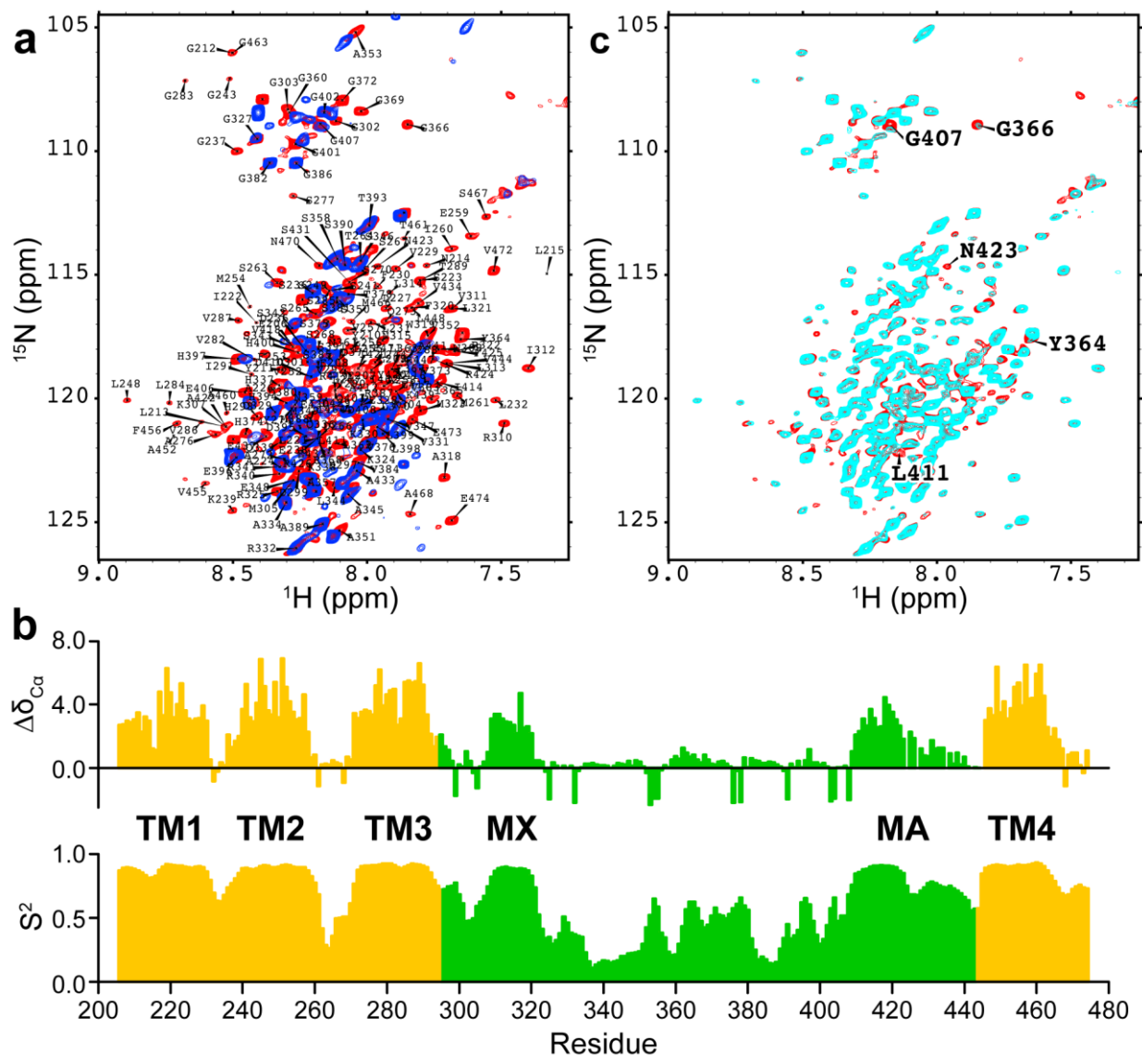
Paramagnetic relaxation enhancement (PRE) from NMR experiments provides long distance restraints ( $\sim 13\text{\AA} \leq r \leq 25\text{\AA}$ ) for structure determination. Briefly, the transverse relaxation rate ( $R_2$ ) is dependent on the inverse sixth power of the distance ( $1/R^6$ ) between the unpaired electron from an MTSL label and an observed proton (225,283-285). A representative  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC NMR spectra of the single-cysteine L415C  $\alpha 7$ nAChR TMD+ICD construct labeled with MTSL in the presence and absence of the reducing agent (ascorbic acid) is shown in

**Figure 20c.** A total of 520 PRE-derived tertiary and quaternary structure distance restraints were obtained from eight different single-cysteine TMD+ICD constructs in LDAO micelles (V311C, C317, C342, C375, C390, L415C, C427, and C435; **Table 9**).

Residual dipolar couplings (RDCs) provide angular information relative to an external magnetic field, which is useful in determining the relative orientation of chemical bonds in a protein, regardless of spatial separation (286). RDCs were measured in  $^1\text{H}$ - $^{15}\text{N}$ -IPAP-HSQC NMR spectra using lanthanide ions chelated to thiol-reactive EDTA (208,287). A total of 462 RDC values for the  $^1\text{H}$ - $^{15}\text{N}$  backbone amide bonds were obtained from three different single-cysteine  $\alpha 7\text{nAChR}$  TMD+ICD constructs (C219, C317, and C427), ranging from -15 to 10 Hz at 900 MHz (**Table 9**).

**Table 9. Structural restraints from NMR and EPR experiments.**

Input		ICD Folding Calculations	TMD+ICD Refinement
NMR Chemical Shifts	N	126	219
	C $\alpha$	147	249
	C	110	187
	C $\beta$	65	145
	HA	84	181
RDC NMR Angle Restraints		336	462
PRE NMR Distance Restraints	Class 1	62	62
	Class 2	127	318
	Class 3	133	140
DEER EPR Distance Restraints		8	10



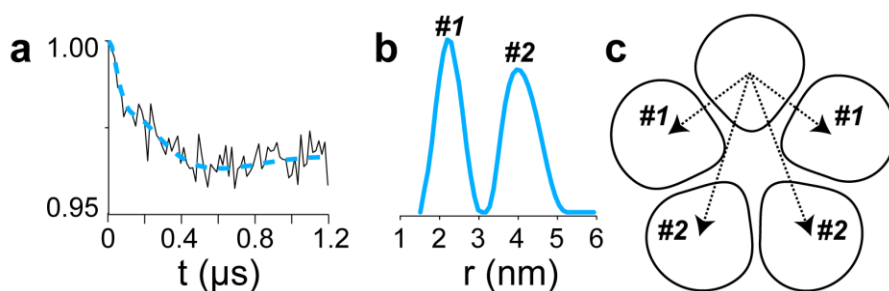
**Figure 20. Structural restraints from NMR experiments.**

**(a)** Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra of single-cysteine L415C  $\alpha 7\text{nAChR}$  TMD+ICD in LDAO micelles (red) and DMPC nanodiscs (blue). Spectra were acquired at 45°C and 900 MHz. Peak assignment is labeled in the LDAO spectrum. **(b)** Differences between the observed  $\text{C}\alpha$  chemical shift and that of the corresponding amino acid in a random coil conformation ( $\Delta\delta_{\text{C}\alpha}$ , top) are shown. Values greater than zero are indicative of  $\alpha$ -helical structure. Order parameters ( $S^2$ , bottom) derived from  $\text{C}\alpha$ ,  $\text{C}\beta$ ,  $\text{CO}$ ,  $\text{N}$ , and  $\text{HN}$  chemical shifts for each residue are shown.  $S^2$  is a general measure of flexibility, where 0 = flexible and 1 = rigid. Large values of  $S^2$  are indicative of  $\alpha$ -helical structure. TMD residues are colored gold and ICD residues are colored green for clarity. **(c)** Overlay of  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra of single-cysteine L415C  $\alpha 7\text{nAChR}$  TMD+ICD labeled with MTSL in the presence (red) and absence (cyan)

of 2 mM ascorbic acid. Spectra were acquired at 45°C and 900 MHz. Labeled peaks exhibit obvious PRE; note that these residues include those both close (G407, L411, N423) and far (Y364, G366) in sequence from the labeling site at L415C.

### 4.3.3 Distance Restraints from EPR Experiments

DEER EPR signal modulation is proportional to the dipolar coupling frequency with  $1/R^3$  dependence, where R is the interspin distance. For pLGICs, DEER spectra can be used to measure distance distributions between adjacent and non-adjacent subunits (204,288,289). **Figure 21** shows representative DEER data of the single-cysteine C449  $\alpha 7nAChR$  TMD+ICD construct labeled with MTSL. DEER-derived quaternary structure distance restraints were obtained for single-cysteine  $\alpha 7nAChR$  TMD+ICD (C219, C375, and C449) and full-length  $\alpha 7nAChR$  (C317 and E437C) constructs (**Table 9**). DEER data from each construct produced two major distributions by model-free Tikhonov regularization (**Figure 21b**), corresponding to the adjacent and across subunit distances (**Figure 21c**).

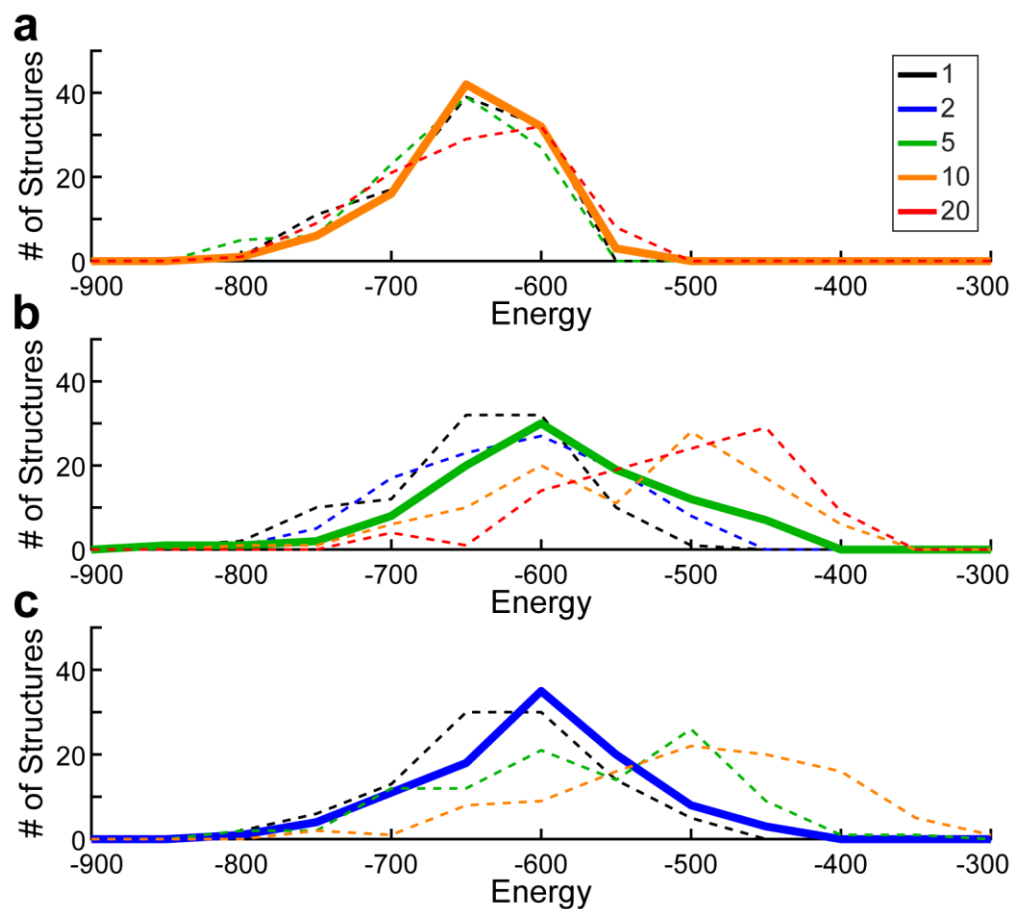


**Figure 21. Distance restraints from EPR experiments.**

(a) DEER EPR data of single-cysteine C449  $\alpha 7nAChR$  TMD+ICD labeled with MTSL obtained at 80K after baseline correction (black) and fit by model-free Tikhonov regularization (blue). (b) Calculated distance distributions reveal two major populations at  $2.2 \pm 0.3$  and  $4.0 \pm 0.4$  nm. (c) Top view of a cartoon  $\alpha 7nAChR$  pentamer showing adjacent (#1) and across (#2) subunit distances corresponding to the two major distance populations shown in (b).

#### 4.3.4 Weight Optimization of Experimental Distance Restraints

The total score (**Equation 6**) was used to rank output structures from Rosetta calculations; obtaining an optimal balance between the standard physics- and knowledge-based potentials and the  $\alpha 7nAChR$ -specific experimental restraint potentials was critical for generating structures that satisfied the restraints while still sampling realistic protein-like topologies (290). Generally, high weights for experimental restraints will increase restraint satisfaction but are likely to produce nonphysical models. Prior to the production Rosetta calculations outlined in **Figure 17**, optimized restraint weights were determined by small-scale empirical calculations of 1,000 ICD structures (**Figure 22**). The talaris2014 energy ( $S_{\text{physics}} + S_{\text{knowledge}}$ ) of the top 100 structures from each calculation was analyzed and final weights were chosen based on the peak position of the energy distribution. The highest weight tested that did not increase the peak (produce higher energy) compared to the distribution for weight = 1 was selected for each type of experimental restraint ( $w_1 = 10$ ,  $w_2 = 5$ , and  $w_3 = 2$  in **Equation 6**).



**Figure 22. Optimization of experimental distance restraint weights in Rosetta.**

Talaris2014 scores ( $S_{\text{physics}} + S_{\text{knowledge}}$ ) of the top 100 structures from small-scale RosettaCM calculations with varied weights for (a) DEER EPR distance restraints, (b) class 1 and class 3 PRE NMR distance restraints, and (c) class 2 PRE NMR distance restraints. The weighting color code shown in (a) applies to all three panels. The highest weight tested that did not shift the peak position towards higher energies compared to the distribution for weight = 1 was selected for each type of experimental restraint. The weight selected for each type of distance restraint is shown as a bold solid line and other weights are shown as dashed lines.

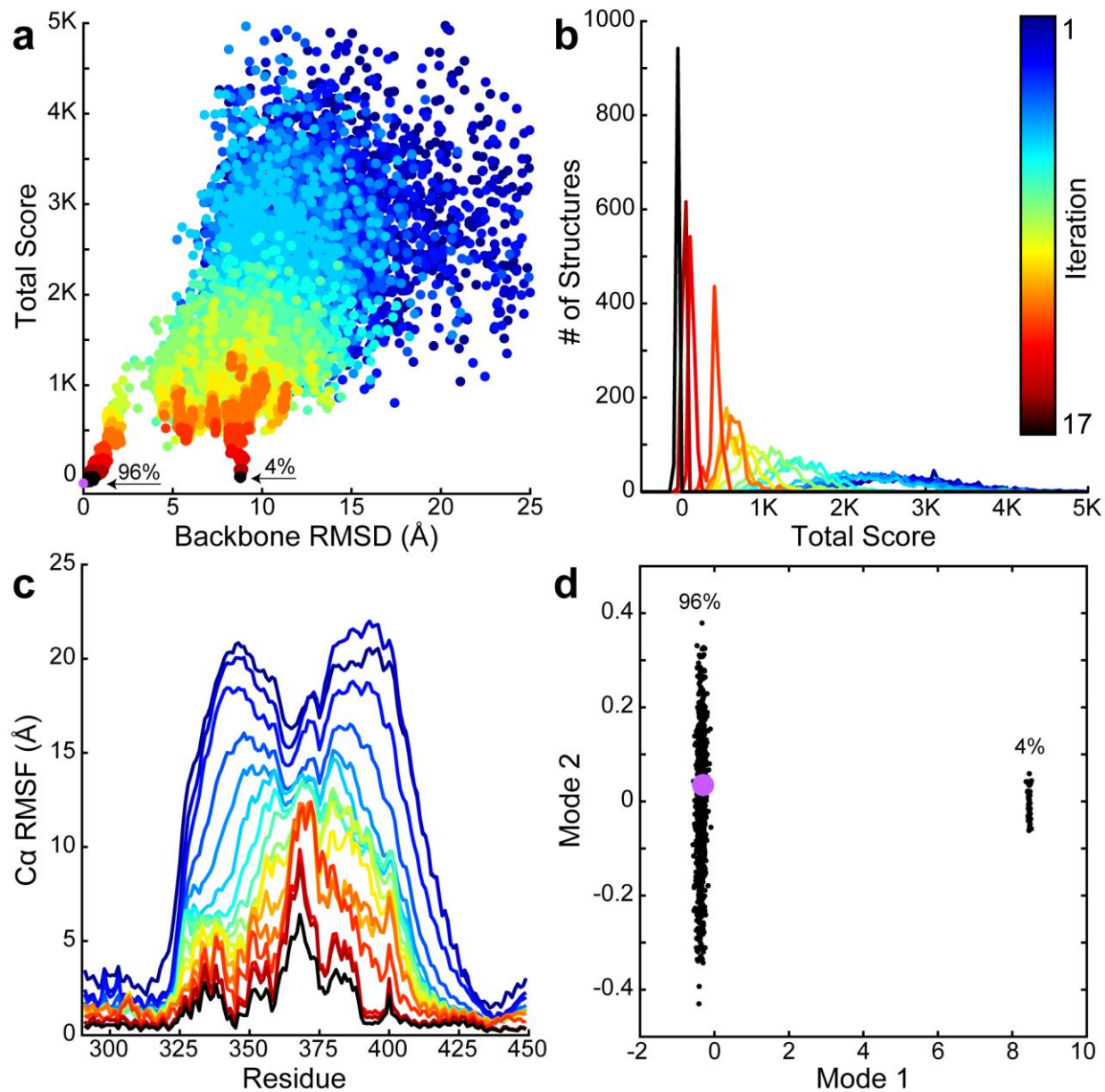
#### 4.3.5 Structure Determination of the $\alpha 7\text{nAChR}$ ICD and TMD

A total of 170,000  $\alpha 7\text{nAChR}$  ICD structures (17 rounds) were generated over the course of iterative RosettaCM folding calculations (Figure 23). In each iteration, the top 1,000 structures

were clustered by RMSD and representative structures were input as new templates to the next iteration of folding calculations. By the 17<sup>th</sup> iteration, 96% of the calculated structures (959/1000) fell into a single major cluster where all structures were  $\leq 3$  Å pairwise backbone RMSD from all other structures in the cluster. As an approximation of the energy landscape, plotting the RMSD of structures from all iterations with respect to the single top ranked structure versus the total score (**Equation 6**) reveals a folding funnel characteristic of native protein structures (291) (**Figure 23a**). Examined individually, both the score distributions and C $\alpha$  RMSF values are gradually reduced with increasing iterations (**Figure 23b-c**), suggesting that the iterative folding protocol was successful in guiding the conformational sampling towards the global energy minimum while maintaining structural diversity to ensure that calculations did not become trapped in local energy minima.

In general, low scoring ICD structures also presented low RMSDs with respect to the single lowest scoring structure; a small deviation from this funnel is observed at  $\sim 9$  Å backbone RMSD, corresponding to the minor cluster from the final iteration that captured the remaining 4% of the calculated structures (41/1000). Principal component analysis of the top 1,000 structures from this last iteration agrees with these RMSD clustering results (**Figure 23d**): plotting the projections of the first two modes (accounting for over 98% of the total variance in the ensemble) reveals two distinct clusters that align exactly with the major and minor clusters described above. Although this minor cluster has signs of developing its own folding funnel, the scores of the top 20 structures from the major cluster are significantly better than the top 20 structures from the minor cluster ( $p < 0.0001$  by the two-sample Kolmogorov-Smirnov test). Thus, only the major cluster was considered for further structural refinement.





**Figure 23. Convergence of the iterative  $\alpha 7$ nAChR ICD folding calculations.**

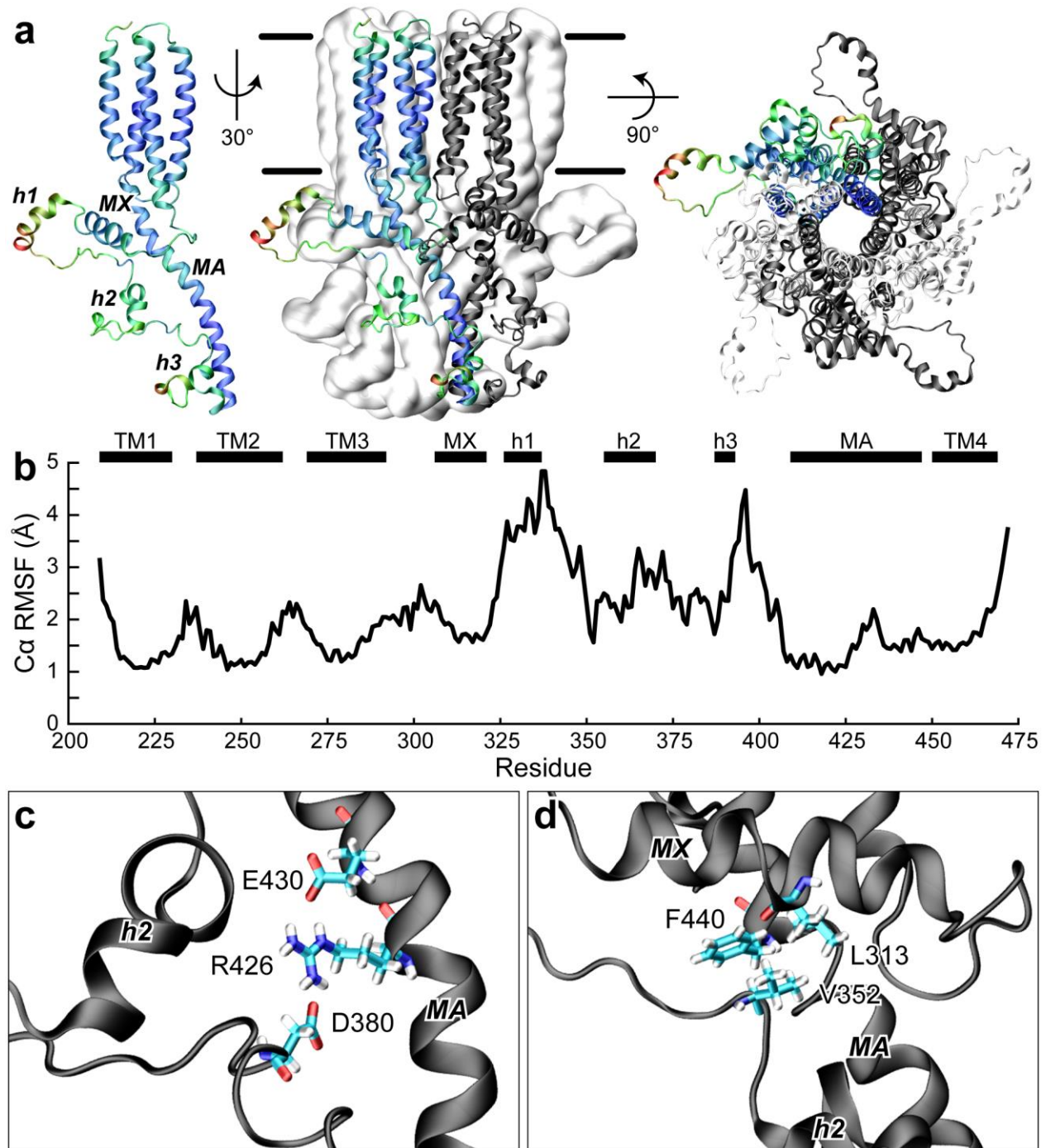
(a) Structural convergence across 17 iterations of RosettaCM calculations for the  $\alpha 7$ nAChR ICD is shown. Only the top 1,000 structures by total score from each round are plotted. RMSD was calculated with respect to the best scoring structure from the final iteration (purple). Labels indicate the proportion of the top 1,000 structures from iteration 17 that fall within each cluster. A folding funnel characteristic of native protein structures is observed. Evolution of (b) scores and (c) C $\alpha$  RMSF values over 17 iterations of RosettaCM calculations for the  $\alpha 7$ nAChR ICD. Distributions of the top 1,000 structures by total score from each round are shown. The color code for different iterations shown in (b)

also applies to **(a)** and **(c)**. **(d)** Projections of the first two modes from principal component analysis of the top 1,000 structures from the 17<sup>th</sup> iteration matches the RMSD clustering results shown in **(a)**, where the lowest scoring structure that progressed to refinement calculations is highlighted in purple.

Similar convergence was obtained for the TMD calculations, where a total of 40,000 structures (4 rounds) were generated and the major cluster in the final iteration corresponded to a folding funnel in the energy landscape. Fewer iterations were required for the TMD than for the ICD because the monomeric structure of the  $\alpha 7$ nAChR TMD was already determined by NMR (203). TMD calculations were only needed to assemble the pentameric complex while taking into account quaternary structure distance restraints from DEER EPR experiments.

#### 4.3.6 Refined $\alpha 7$ nAChR TMD+ICD Structures

The top ranked TMD and ICD structures from the final iterations (rounds 4 and 17, respectively) were combined by aligning along the interfacial residues (I291 to Q294 and D446 to C449). The resulting  $\alpha 7$ nAChR TMD+ICD structure was refined by MD simulations in an explicit lipid bilayer to optimize the TMD/ICD interface. After final energy minimization calculations with Rosetta FastRelax, the top 20 structures were selected based on total score (**Figure 24a**). The average pairwise RMSD in the helical regions is  $2.75 \pm 0.57$  Å for backbone atoms and  $3.13 \pm 0.53$  Å for all heavy atoms. Q-factors for EPR DEER and working PRE NMR distance restraints are low, i.e. calculated structures satisfy the experimental restraints. Q-factors for free PRE NMR distance restraints (excluded from all structure calculations) are only slightly higher than  $Q^{\text{PRE}}_{\text{work}}$ , indicating that the structures are not overfit to the working restraints. Additional detailed structure statistics are provided in **Table 10**.



**Figure 24. Refined structures of the  $\alpha 7$ nAChR TMD+ICD.**

(a) Side (left, middle) and cytoplasmic (right) views of the  $\alpha 7$ nAChR TMD+ICD, averaged over the top 20 structures from refinement calculations, ranked by total score. A single monomer is shown on the left with labeled ICD helices. Residues in (a) are colored based on (b) C $\alpha$  RMSF values calculated across the top 20 structures, where the scale ranges from blue (1  $\text{\AA}$ ) to red (5  $\text{\AA}$ ). Helices are marked at the corresponding residues across the top of the plot. (c)

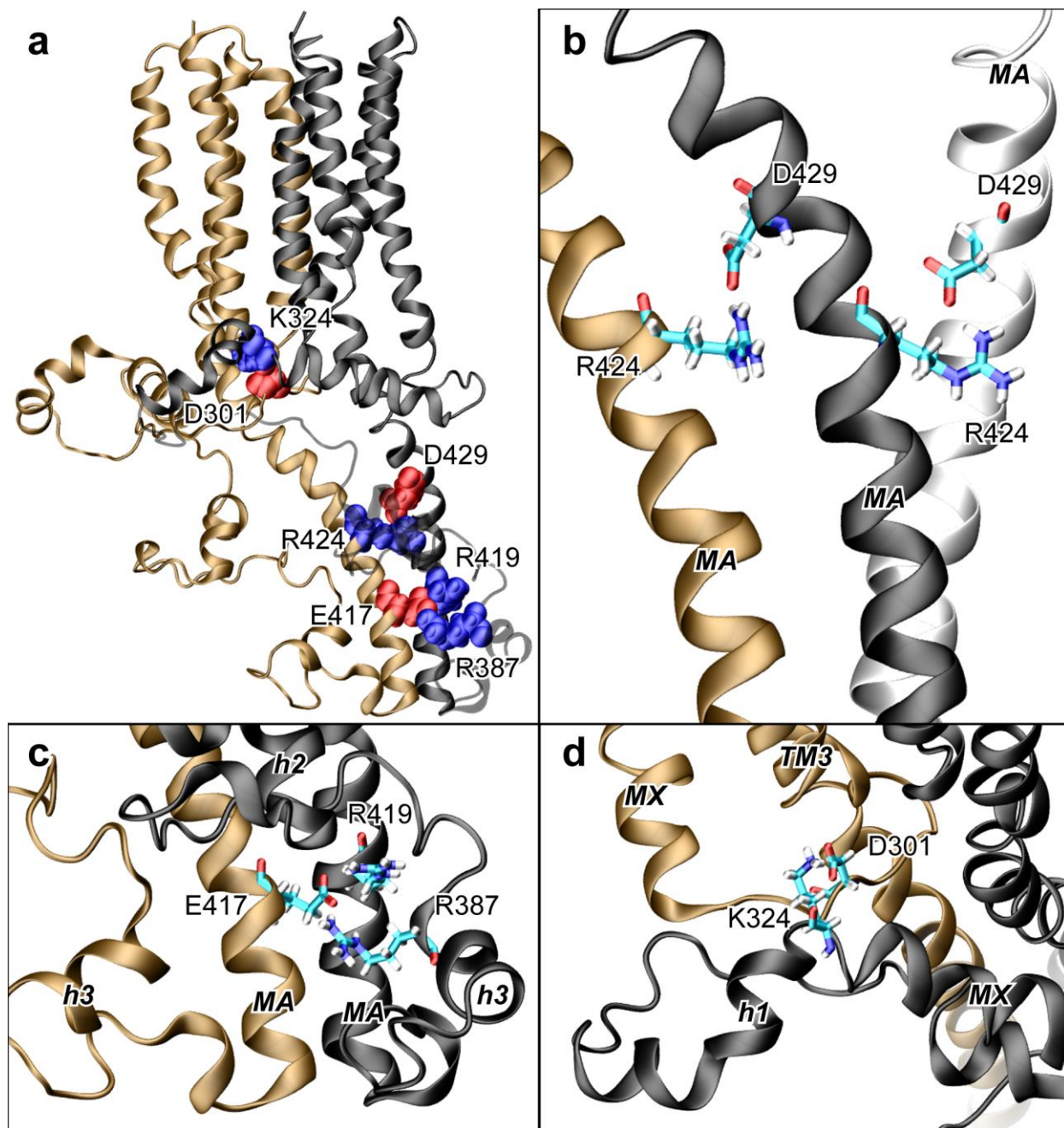
Zoom-in view of the intra-subunit salt bridge between D380 and R426 in a representative  $\alpha 7nAChR$  TMD+ICD structure. A single subunit is shown for clarity. Ten of the top 20 structures show sidechain D380<sub>O</sub>-R426<sub>N</sub> distances less than 4 Å. The average sidechain D380<sub>O</sub>-R426<sub>N</sub> distance across all 20 structures is  $4.32 \pm 1.73$  Å. **(d)** Zoom-in view of intra-subunit hydrophobic packing between L313 in the MX helix, V352, and F440 in the MA helix in a representative  $\alpha 7nAChR$  TMD+ICD structure. A single subunit is shown for clarity.

**Table 10. Structure statistics for the  $\alpha 7nAChR$  TMD+ICD.**

		<b>Backbone</b>	<b>Heavy atom</b>
<b>Helical Pairwise RMSD (Å)</b>	TMD+ICD helices (residues 211 to 230, 237 to 262, 269 to 292, 306 to 321, 326 to 337, 355 to 370, 387 to 393, 404 to 469)	$2.75 \pm 0.57$	$3.13 \pm 0.53$
	TMD helices (residues 211 to 230, 237 to 262, 269 to 292, 448 to 469)	$1.55 \pm 0.41$	$1.88 \pm 0.39$
	ICD helices (residues 306 to 321, 326 to 337, 355 to 370, 387 to 393, 404 to 447)	$3.09 \pm 0.70$	$3.53 \pm 0.66$
<b>Pairwise RMSD (Å)</b>	TMD+ICD (residues 209 to 472)	$3.13 \pm 0.54$	$3.63 \pm 0.52$
	TMD (residues 209 to 292, 448 to 469)	$1.85 \pm 0.43$	$2.29 \pm 0.44$
	ICD (residues 293 to 447)	$3.42 \pm 0.60$	$3.97 \pm 0.59$
<b>Q-factors</b>	$Q^{DEER}$	$0.083 \pm 0.026$	
	$Q^{PRE}_{work}$	$0.034 \pm 0.016$	
	$Q^{PRE}_{free}$	$0.050 \pm 0.021$	
<b>Structure Statistics</b>	MolProbity Score	$2.64 \pm 0.051$	
	Ramachandran Favored (%)	$98.42 \pm 0.89$	
	Ramachandran Outliers (%)	0.00	
	Rotamer Outliers (%)	0.00	

All data are reported as mean  $\pm$  SD of the top 20 structures after refinement.

The secondary and tertiary structure of the  $\alpha 7$ nAChR ICD is characterized by short horizontal MX  $\alpha$ -helices parallel to the plane of the lipid bilayer (P306 to L321) and longer MA  $\alpha$ -helices beginning at L409 and continuing into TM4 in the membrane, analogous to those observed in previously solved partial ICD structures of other pLGICs (179,190,193,194). Three additional helical regions are found in the  $\alpha 7$ nAChR ICD between the MX and MA helices. First, a short  $\alpha$ -helix extending away from the pore axis and angled down into the cytoplasm is observed adjacent to the MX helix spanning residues P326 to H337 (here named the h1 helix). However, despite the stable secondary structure in this region, these residues are highly dynamic across the bundle of 20 TMD+ICD structures, with maximum C $\alpha$  RMSF values of  $\sim 4.8$  Å for H337 (**Figure 24b**). Second, residues P355 to I365 in the middle of the ICD form another small  $\alpha$ -helix (h2), bordered by a short  $3_{10}$  helix spanned by residues R368 to L370. This small  $3_{10}$  helix is positioned at a right angle with respect to the h2 helix so that the solvent-exposure of the hydrophobic sidechains for residues L362, L363, and L370 is minimized. The last small  $\alpha$ -helix in the ICD before the MA helix is found in residues R387 to T393 (h3), positioned perpendicular to the MA helix. Other notable tertiary structural features include a stable intra-subunit salt bridge between D380 and R426 in the MA helix (**Figure 24c**), as well as a distinct C $\alpha$  RMSF minimum at V352 (**Figure 24b**): this residue is consistently involved in tight intra-subunit hydrophobic packing between L313 in the MX helix and F440 in the MA helix (**Figure 24d**).



**Figure 25. Salt bridges stabilize the  $\alpha 7$ nAChR ICD structure.**

**(a)** Overview of inter-subunit salt bridges stabilizing the quaternary structure of the ICD. Only two subunits are shown for clarity. **(b)** Zoom-in view of the R419-E417-R387 salt bridges (cyan). Thirteen of the top 20 structures show sidechain R419<sub>N</sub>-E417<sub>O</sub> distances and/or sidechain E417<sub>O</sub>-R387<sub>N</sub> distances less than 4 Å. **(c)** Zoom-in view of R424-D429 salt bridge (cyan). Fifteen of the top 20 structures show sidechain R424<sub>N</sub>-D429<sub>O</sub> distances less than 4 Å. **(d)**

Zoom-in view of K324-D301 salt bridge (cyan). Eleven of the top 20 structures show sidechain K324<sub>N</sub>-D301<sub>O</sub> distances less than 4 Å.

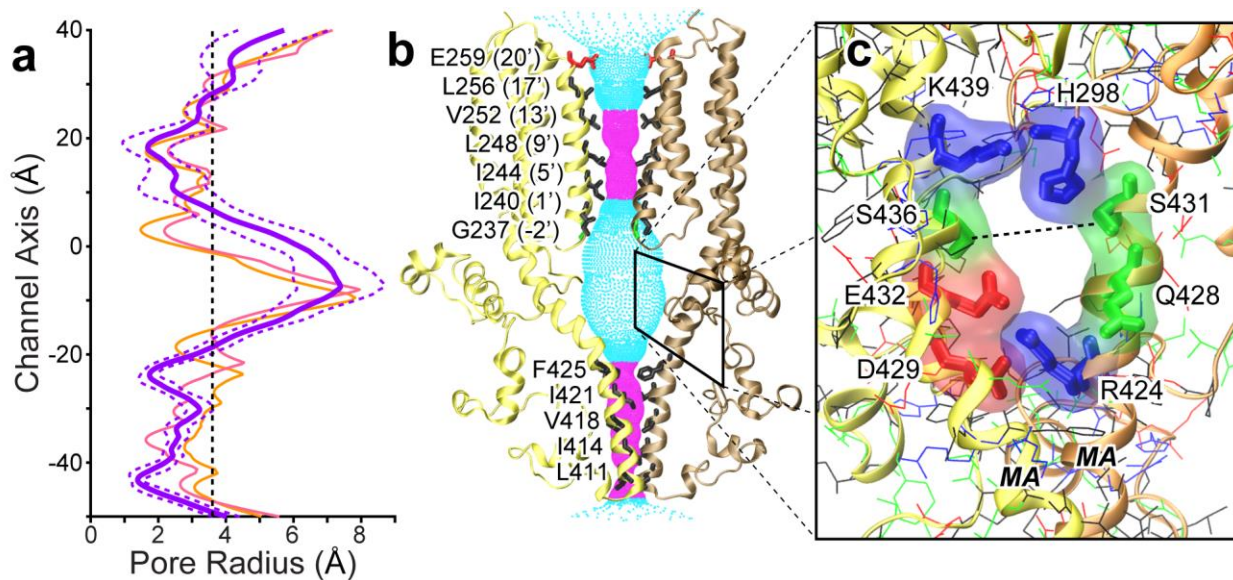
The quaternary structure of the  $\alpha 7$ nAChR ICD is characterized by four distinct inter-subunit salt bridges (**Figure 25a**). Two of these salt bridges are found between adjacent MA helices (E417-R419 and R424-D429), forming the central core structure of the ICD (**Figure 25b-c**). E417 is also involved in another inter-subunit salt bridge with R387 at beginning of the h3 helix (**Figure 25b**). Finally, D301 in the TM3-MX loop forms a salt bridge with K324 in the connecting loop between the MX and h1 helices (**Figure 25d**). This salt bridge acts as an anchor point for the mobile h1 helix, whose residues that are further from the MX helix experience greater fluctuations (**Figure 24b**). A previous study showed that mutations to some of these salt bridge residues can induce significant changes to  $\alpha 7$ nAChR expression (163): D429A mutants showed decreased expression compared to WT while E417A, R419A, and R424A mutants showed increased expression. D301A did not affect  $\alpha 7$ nAChR expression and mutations to K324 were not tested, suggesting that this salt bridge is likely less critical for stabilizing the quaternary structure of the  $\alpha 7$ nAChR ICD.

#### 4.3.7 Pore Conformation and Ion Permeation Pathway

Pore profiles of the  $\alpha 7$ nAChR TMD+ICD structures determined here are in a non-conductive state, as expected since all NMR and EPR structural restraints were obtained in the absence of any channel-activating ligands. The TMD is most constricted at L248 (9') with a radius of  $\sim 1.7$  Å (**Figure 26a**), too small to allow a hydrated Na<sup>+</sup> or Ca<sup>2+</sup> ion to pass through ( $\sim 3.6$  or  $4.1$  Å radius, respectively) (292). This hydrophobic gate at the 9' position has also been observed previously in

other resting state cationic pLGICs (79,179,190,293). However, the second TMD constriction observed in structures of 5-HT<sub>3A</sub>R at the -1' position is not shared by  $\alpha$ 7nAChR, despite their matching pore-lining residues in the lower half of the TMD. In fact, E238 (-1') does not face the pore in  $\alpha$ 7nAChR; rather, G237 (-2') is the final pore-lining residue in the TM2 helix (**Figure 26a**). The  $\alpha$ 7nAChR pore widens at the top of the ICD into a small vestibule before constricting again in the middle of the MA helix at F425 to a radius of  $\sim 1.7$  Å. The rest of ICD below F425 maintains a non-conductive pore radius below 3.2 Å, ending with a final constriction point at L411 where the radius is  $\sim 1.4$  Å. This end of the ICD is characterized by tight inter-subunit hydrophobic packing of residues F425, I421, V418, I414, and L411 along one side of the MA helices, forming a series of hydrophobic girdles that would be impossible for ions to pass through without significant conformational changes (**Figure 26b**). Instead, lateral windows formed by charged and polar residues directly above these hydrophobic girdles are observed, providing a pathway for ions exiting the TMD to reach the cytoplasm (**Figure 26c**). The size of this lateral window was estimated by measuring the distance between the sidechain oxygen atoms of S431 and S436. The average distance over the top 20 structures is  $7.9 \pm 1.5$  Å, sufficiently large for a hydrated Na<sup>+</sup> ion to pass through; the upper range of this distribution would also be large enough for a hydrated Ca<sup>2+</sup> ion to pass through.





**Figure 26. Pore conformation and ion permeation pathway.**

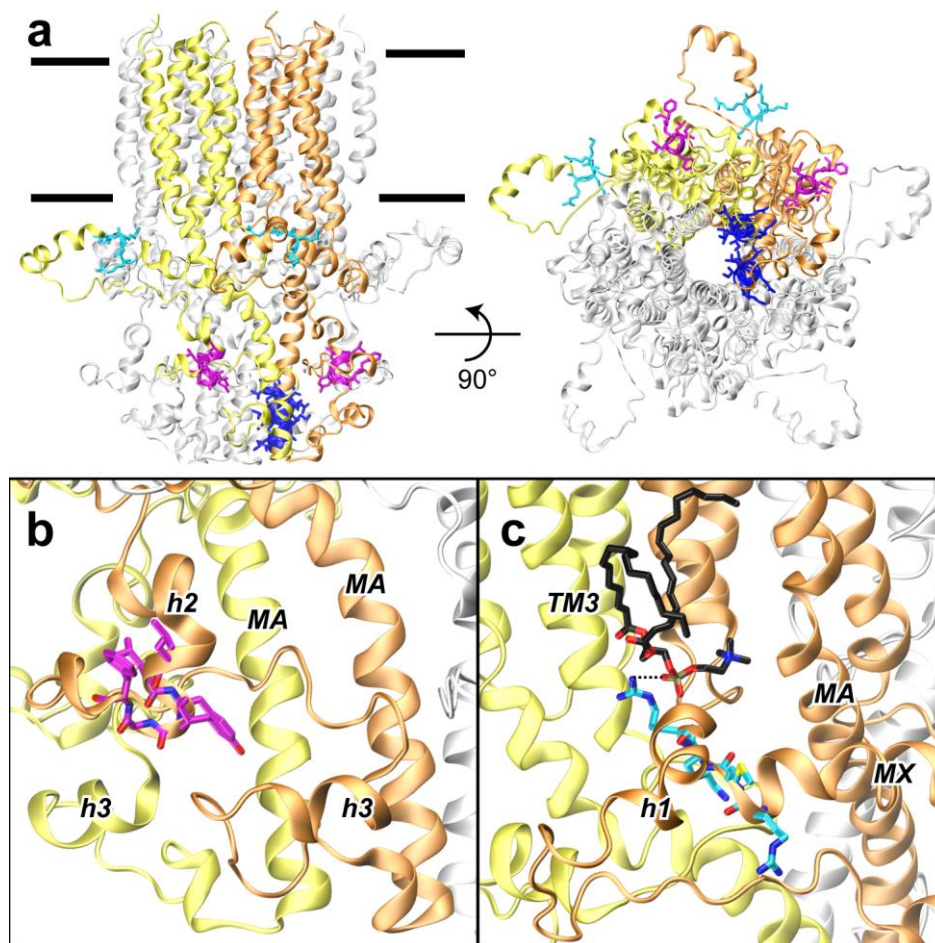
(a) The average pore profile of the top 20  $\alpha 7$ nAChR TMD+ICD structures after refinement is shown in purple with standard deviations as dashed lines. Pore profiles of closed-state 5-HT<sub>3A</sub>R structures are shown for comparison in orange (cryo-EM, PDB: 6BE1 (179)) and pink (x-ray, PDB: 4PIR (190)). The radius of a hydrated Na<sup>+</sup> ion (3.6 Å) is marked for reference. (b) Side view along the pore axis of a representative  $\alpha 7$ nAChR TMD+ICD. Side chains of pore-lining residues in the TMD and hydrophobic girdles in the ICD are highlighted. Cyan and magenta dots define pore radii greater or less than the radius of a hydrated Na<sup>+</sup> ion (3.6 Å), respectively. (c) Zoom-in view of lateral windows formed by charged and polar residues in adjacent MA helices showing a clear pathway for ions exiting the TMD to reach the cytoplasm. The average distance between the sidechain oxygen atoms of S431 and S436 ( $7.9 \pm 1.5$  Å) is marked with a dashed line as an estimate of the window size.

#### 4.3.8 Structural Basis for Metabotropic Signaling

The  $\alpha 7$ nAChR ICD directly interacts with a large number of cytosolic protein partners (36), but the structural basis of these interactions is not well understood. Only a few mutagenesis studies published so far have made progress in this direction by identifying ICD residues critical for interactions with various intracellular proteins. For example, six ICD residues (362LLYIGF367)

were pinpointed to be essential for the selective transport of  $\alpha 7$ nAChR to dendrites over axons in hippocampal neurons (33). Although the exact mechanism by which these residues regulate  $\alpha 7$ nAChR targeting has not been confirmed, this short stretch contains a tyrosine motif (YXX $\phi$ ) known to interact with adaptor protein complexes for transmembrane protein transport (294). Indeed, a later study found that the clathrin adaptor AP-1 complex  $\mu 1$  subunit was responsible for dendritic localization of ACR-16, an  $\alpha 7$ nAChR homolog in *C. elegans* (295). Here, residues 362LLYIGF367 consistently form part of the h2 helix in the  $\alpha 7$ nAChR TMD+ICD structures, surrounded on either side by large stretches of residues in turn/coil conformations (**Figure 27a-b**). Although there are some variations in tertiary folding among these structures, the conserved secondary structure in this region is likely required for efficient interactions with the adaptor protein complexes.

Resistance to inhibitors of cholinesterase 3 (RIC-3) is a chaperone protein found primarily in the endoplasmic reticulum that can dramatically increase the expression of  $\alpha 7$ nAChR on the cell surface (296,297). A mutagenesis study demonstrated that specific residues along the hydrophobic side of the MA helix (L411, V418, and F425) were required for interactions between  $\alpha 7$ nAChR and the chaperone protein RIC-3 (34). Specifically, a later work established that the coiled-coil domain of RIC-3 is required for accurate folding and assembly of  $\alpha 7$ nAChR (298). In the context of the  $\alpha 7$ nAChR TMD+ICD structures determined here (**Figure 26b, 27a**), RIC-3 likely shields the large hydrophobic surface of the MA helix during monomeric folding and/or pentameric assembly in order to prevent degradation or aggregation with other exposed hydrophobic regions (299).



**Figure 27. Protein binding sites in the  $\alpha 7$ nAChR ICD for cytoplasmic signaling.**

(a) Side and cytoplasmic views of a representative  $\alpha 7$ nAChR TMD+ICD structure highlighting ICD residues suggested previously to be critical for protein-protein interactions. The dendritic targeting sequence (362LLYIGF367) (33) is shown in magenta. The RIC-3 binding region in the MA helix (411LAKILEEV418) (34) is shown in blue. The G protein binding site (322RMKR325) (39) is shown in cyan. (b) Zoom-in view of the helical structure in the dendritic targeting residues that can bind clathrin adaptor proteins (magenta). (c) Zoom-in view of the G protein binding residues (cyan) near the lipid bilayer (black) from refinement MD simulations. R325 forms consistent hydrogen bonds with POPC phosphate headgroups throughout the simulations (dashed line).

A recent study revealed that residues 322RMKR325 directly bind the G proteins  $G\alpha_q$  and  $G\beta\gamma$  to activate downstream signaling pathways involving phospholipase C and  $IP_3$ -mediated intracellular calcium release (39). G proteins are membrane-associated and accordingly, residues

322RMKR325 in the  $\alpha 7$ nAChR structures determined in our study are also positioned to be in direct contact with the lipid bilayer. In fact, the R325 sidechain formed hydrogen bonds with POPC phosphate headgroups in  $\sim 73\%$  of the total simulation time during the refinement MD simulations (**Figure 27c**). Such a high propensity for direct lipid interactions in this segment of the ICD, combined with the knowledge that G proteins can modulate their surrounding lipid environment (300), suggest that the lipid bilayer is likely an important factor in the structural basis of  $\alpha 7$ nAChR-G protein interactions.

#### 4.4 DISCUSSION

Here we present the first complete structures of a pLGIC ICD, refuting the notion that the majority of this domain is intrinsically completely unstructured (179,190,194,301). The integrated experimental and computational approach developed here for structure determination of the  $\alpha 7$ nAChR ICD can be applied to solve the structures of other pLGIC ICDs, as well as other flexible proteins that have proven to be difficult targets for traditional structure determination tactics. Thus far, x-ray crystallography and cryo-EM have been unable to capture a complete picture of the elusive pLGIC ICD due to its inherent flexibility (95-98,179,185-194). In fact, the most flexible region of the  $\alpha 7$ nAChR ICD is found immediately following the furthest residues solved in any of the previous structures of partial pLGIC ICDs (**Figure 24**), shedding light on the reason why neither x-ray crystallography or cryo-EM have been able to resolve pLGIC ICD residues past this point. Alternatively, NMR and EPR spectroscopy provided an ideal platform to obtain structural information for the dynamic  $\alpha 7$ nAChR ICD.

While the  $\alpha 7nAChR$  ICD is certainly more flexible than the TMD, as evidenced by the RMSD values from individual domains shown in **Table 10** and RMSF values shown in **Figure 24**, helical secondary structural elements, intra-subunit salt bridges and hydrophobic packing for tertiary structure as well as inter-subunit salt bridges for quaternary structure all provide considerable stability to the ICD (**Figure 24-25**). It is important to note that the dynamics of the ICD calculated here align well with the inherent ICD flexibility measured by EPR and NMR experiments. For example, the magnitude of conformational heterogeneity in a sample is proportional to the width of the distance distribution from DEER EPR (302,303); standard deviations of distance distributions from DEER EPR are  $\sim 4$  Å across all ICD labeling sites (**Figure 21**), matching the  $\sim 4$  Å heavy atom RMSD values for ICD structures determined here (**Table 10**). Order parameters from NMR chemical shifts also show correlation with the resulting ICD structures; the locations of the MX, h1, h2, h3, and MA helices (**Figure 24**) align well with the peak regions for high values of  $S^2$  in the ICD (**Figure 20**).

The  $\alpha 7nAChR$  structures determined here generally resemble the partial ICD structures determined previously for other pLGICs with horizontal MX helices parallel to membrane and longer MA helices connecting to TM4 (179,190,193,194). One important difference is observed between  $\alpha 7nAChR$  and the structures of 5-HT<sub>3A</sub>R at the interface between the MA and TM4 helices. The MA helix is a continuous, unbroken extension of TM4 in 5-HT<sub>3A</sub>R (179,190), but a distinct kink is observed between the MA and TM4 helices in  $\alpha 7nAChR$  at L448-C449. This conformational change leads to a much tighter association between the  $\alpha 7nAChR$  TM4 helix and TM1-TM3 helical bundle compared to 5-HT<sub>3A</sub>R (**Figure 24**). Indeed, previous studies of GLIC chimeras containing the ICD of  $\alpha 7nAChR$  or 5-HT<sub>3A</sub>R demonstrated that GLIC- $\alpha 7nAChR$  constructs were functional without any sequence modifications, whereas GLIC-5-HT<sub>3A</sub>R required

optimization of the TMD-ICD interface in order to produce channel current (304,305). This suggests that significant constraints exist between the MA and TM4 helices in 5-HT<sub>3A</sub>R, but not in  $\alpha$ 7nAChR, supporting the kinked conformation of the MA and TM4 helices in the  $\alpha$ 7nAChR structures determined here.

Another significant difference between the structures of  $\alpha$ 7nAChR and 5-HT<sub>3A</sub>R is the openness of the lateral windows for ion exit between MA helices. These portals are blocked by the TM3-MX loop in both x-ray and cryo-EM structures of 5-HT<sub>3A</sub>R (179,190). Although a few of the  $\alpha$ 7nAChR structures determined here also present lateral windows partially occluded by the TM3-MX loop, most are open directly to the cytoplasm, wide enough for hydrated Na<sup>+</sup> or Ca<sup>2+</sup> ions to pass through (**Figure 26**). The existence of these lateral windows has been previously proposed (194), but the structures presented here provide the first complete look at the pLGIC ion permeation pathway through the ICD and into the cytosol. In addition, a previous mutagenesis study demonstrated that  $\alpha$ 7nAChR conductance could be increased or decreased by mutations to the negatively charged/polar and positively charged residues, respectively, that form the lateral windows for ion exit in the structures determined here. Specifically, Q428A, S431A, and E432A mutations decreased current whereas H298A and R424A increased current compared to WT  $\alpha$ 7nAChR (163). The tight hydrophobic packing seen at the bottom of the MA helix combined with this mutagenesis data strongly support the lateral windows for ion exit observed in the structures of  $\alpha$ 7nAChR.

The  $\alpha$ 7nAChR structures determined here provide the first insight into the structural basis of pLGIC interactions with cytoplasmic protein partners (**Figure 27**). As mentioned above,  $\alpha$ 7nAChR is involved in many different neurological processes that are likely mediated by networking with a variety of intracellular protein partners associated with diverse downstream

signaling pathways (36-41). The structures revealed in this study are the first critical component required to understand the molecular details of these intracellular interactions. This knowledge will be critical for future structure-based drug discovery efforts to target specific  $\alpha 7$ nAChR network interactions.

## BIBLIOGRAPHY

1. Taly, A., Corringer, P. J., Guedin, D., Lestage, P., and Changeux, J. P. (2009) Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nat Rev Drug Discov* **8**, 733-750.
2. Yevenes, G. E., and Zeilhofer, H. U. (2011) Allosteric modulation of glycine receptors. *British journal of pharmacology* **164**, 224-236.
3. Dani, J. A., and Bertrand, D. (2007) Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annu Rev Pharmacol Toxicol* **47**, 699-729.
4. Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A., and Grutter, T. (2012) Ligand-gated ion channels: new insights into neurological disorders and ligand recognition. *Chem Rev* **112**, 6285-6318.
5. Nemezc, A., Prevost, M. S., Menny, A., and Corringer, P. J. (2016) Emerging Molecular Mechanisms of Signal Transduction in Pentameric Ligand-Gated Ion Channels. *Neuron* **90**, 452-470.
6. Woolf, C. J. (2010) Overcoming obstacles to developing new analgesics. *Nat Med* **16**, 1241-1247.
7. Lader, M. (1999) Some adverse effects of antipsychotics: prevention and treatment. *J Clin Psychiatry* **60 Suppl 12**, 18-21.
8. Khawam, E. A., Laurencic, G., and Malone, D. A., Jr. (2006) Side effects of antidepressants: an overview. *Clev Clin J Med* **73**, 351-353, 356-361.
9. Sauguet, L., Shahsavar, A., and Delarue, M. (2015) Crystallographic studies of pharmacological sites in pentameric ligand-gated ion channels. *Biochim Biophys Acta* **1850**, 511-523.
10. Mowrey, D. D., Kinde, M. N., Xu, Y., and Tang, P. (2015) Atomistic insights into human Cys-loop receptors by solution NMR. *Biochim Biophys Acta* **1848**, 307-314.
11. Taly, A., Henin, J., Changeux, J. P., and Cecchini, M. (2014) Allosteric regulation of pentameric ligand-gated ion channels: an emerging mechanistic perspective. *Channels (Austin)* **8**, 350-360.
12. Bajorath, J. (2002) Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* **1**, 882-894.



13. Lacapere, J. J., Pebay-Peyroula, E., Neumann, J. M., and Etchebest, C. (2007) Determining membrane protein structures: still a challenge! *Trends Biochem Sci* **32**, 259-270.
14. Fan, J., Heng, J., Dai, S., Shaw, N., Zhou, B., Huang, B., He, Z., Wang, Y., Jiang, T., Li, X., Liu, Z., Wang, X., and Zhang, X. C. (2011) An efficient strategy for high throughput screening of recombinant integral membrane protein expression and stability. *Protein Expr Purif* **78**, 6-13.
15. Jhoti, H., Rees, S., and Solari, R. (2013) High-throughput screening and structure-based approaches to hit discovery: is there a clear winner? *Expert Opin Drug Discov* **8**, 1449-1453.
16. Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe, E. W., Jr. (2014) Computational methods in drug discovery. *Pharmacol Rev* **66**, 334-395.
17. Bohacek, R. S., McMartin, C., and Guida, W. C. (1996) The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* **16**, 3-50.
18. Lionta, E., Spyrou, G., Vassilatis, D. K., and Cournia, Z. (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* **14**, 1923-1938.
19. Lavecchia, A., and Di Giovanni, C. (2013) Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem* **20**, 2839-2860.
20. Cerqueira, N. M., Sousa, S. F., Fernandes, P. A., and Ramos, M. J. (2009) Virtual screening of compound libraries. *Methods Mol Biol* **572**, 57-70.
21. Macalino, S. J., Gosu, V., Hong, S., and Choi, S. (2015) Role of computer-aided drug design in modern drug discovery. *Arch Pharm Res* **38**, 1686-1701.
22. Cheng, T., Li, Q., Zhou, Z., Wang, Y., and Bryant, S. H. (2012) Structure-based virtual screening for drug discovery: a problem-centric review. *AAPS J* **14**, 133-141.
23. Heusser, S. A., Howard, R. J., Borghese, C. M., Cullins, M. A., Broemstrup, T., Lee, U. S., Lindahl, E., Carlsson, J., and Harris, R. A. (2013) Functional validation of virtual screening for novel agents with general anesthetic action at ligand-gated ion channels. *Mol Pharmacol* **84**, 670-678.
24. Zheng, F., Robertson, A. P., Abongwa, M., Yu, E. W., and Martin, R. J. (2016) The *Ascaris suum* nicotinic receptor, ACR-16, as a drug target: Four novel negative allosteric modulators from virtual screening. *Int J Parasitol Drugs Drug Resist* **6**, 60-73.
25. Akdemir, A., Rucktooa, P., Jongejan, A., Elk, R., Bertrand, S., Sixma, T. K., Bertrand, D., Smit, A. B., Leurs, R., de Graaf, C., and de Esch, I. J. (2011) Acetylcholine binding protein (AChBP) as template for hierarchical in silico screening procedures to identify structurally novel ligands for the nicotinic receptors. *Bioorg Med Chem* **19**, 6107-6119.

26. Ulens, C., Akdemir, A., Jongejan, A., van Elk, R., Bertrand, S., Perrakis, A., Leurs, R., Smit, A. B., Sixma, T. K., Bertrand, D., and de Esch, I. J. (2009) Use of acetylcholine binding protein in the search for novel alpha7 nicotinic receptor ligands. In silico docking, pharmacological screening, and X-ray analysis. *Journal of medicinal chemistry* **52**, 2372-2383.
27. Akdemir, A., Edink, E., Thompson, A. J., Lummis, S. C., Kooistra, A. J., de Graaf, C., and de Esch, I. J. (2012) Identification of novel alpha7 nicotinic receptor ligands by in silico screening against the crystal structure of a chimeric alpha7 receptor ligand binding domain. *Bioorg Med Chem* **20**, 5992-6002.
28. Mahasenan, K. V., Pavlovicz, R. E., Henderson, B. J., Gonzalez-Cestari, T. F., Yi, B., McKay, D. B., and Li, C. (2011) Discovery of Novel alpha4beta2 Neuronal Nicotinic Receptor Modulators through Structure-Based Virtual Screening. *ACS Med Chem Lett* **2**, 855-860.
29. Chakka, N., Andrews, K. L., Berry, L. M., Bregman, H., Gunaydin, H., Huang, L., Guzman-Perez, A., Plant, M. H., Simard, J. R., Gingras, J., and DiMauro, E. F. (2017) Applications of parallel synthetic lead hopping and pharmacophore-based virtual screening in the discovery of efficient glycine receptor potentiators. *Eur J Med Chem* **137**, 63-75.
30. Middendorp, S. J., Puthenkalam, R., Baur, R., Ernst, M., and Sigel, E. (2014) Accelerated discovery of novel benzodiazepine ligands by experiment-guided virtual screening. *ACS Chem Biol* **9**, 1854-1859.
31. Corringer, P. J., Poitevin, F., Prevost, M. S., Sauguet, L., Delarue, M., and Changeux, J. P. (2012) Structure and pharmacology of pentameric receptor channels: from bacteria to brain. *Structure* **20**, 941-956.
32. Kracun, S., Harkness, P. C., Gibb, A. J., and Millar, N. S. (2008) Influence of the M3-M4 intracellular domain upon nicotinic acetylcholine receptor assembly, targeting and function. *British journal of pharmacology* **153**, 1474-1484.
33. Xu, J., Zhu, Y., and Heinemann, S. F. (2006) Identification of sequence motifs that target neuronal nicotinic receptors to dendrites and axons. *J Neurosci* **26**, 9780-9793.
34. Castillo, M., Mulet, J., Gutierrez, L. M., Ortiz, J. A., Castelan, F., Gerber, S., Sala, S., Sala, F., and Criado, M. (2005) Dual role of the RIC-3 protein in trafficking of serotonin and nicotinic acetylcholine receptors. *J Biol Chem* **280**, 27062-27068.
35. Baer, K., Burli, T., Huh, K. H., Wiesner, A., Erb-Vogtli, S., Gockeritz-Dujmovic, D., Moransard, M., Nishimune, A., Rees, M. I., Henley, J. M., Fritschy, J. M., and Fuhrer, C. (2007) PICK1 interacts with alpha7 neuronal nicotinic acetylcholine receptors and controls their clustering. *Mol Cell Neurosci* **35**, 339-355.
36. Paulo, J. A., Brucker, W. J., and Hawrot, E. (2009) Proteomic analysis of an alpha7 nicotinic acetylcholine receptor interactome. *J Proteome Res* **8**, 1849-1858.

37. Jones, A. K., Buckingham, S. D., and Sattelle, D. B. (2010) Proteins interacting with nicotinic acetylcholine receptors: expanding functional and therapeutic horizons. *Trends in pharmacological sciences* **31**, 455-462.
38. Kihara, T., Shimohama, S., Sawada, H., Honda, K., Nakamizo, T., Shibasaki, H., Kume, T., and Akaike, A. (2001) alpha 7 nicotinic receptor transduces signals to phosphatidylinositol 3-kinase to block A beta-amyloid-induced neurotoxicity. *J Biol Chem* **276**, 13541-13546.
39. King, J. R., Nordman, J. C., Bridges, S. P., Lin, M. K., and Kabbani, N. (2015) Identification and Characterization of a G Protein-binding Cluster in alpha7 Nicotinic Acetylcholine Receptors. *J Biol Chem* **290**, 20060-20070.
40. King, J. R., Ullah, A., Bak, E., Jafri, M. S., and Kabbani, N. (2018) Ionotropic and Metabotropic Mechanisms of Allosteric Modulation of alpha7 Nicotinic Receptor Intracellular Calcium. *Mol Pharmacol* **93**, 601-611.
41. King, J. R., and Kabbani, N. (2018) Alpha 7 nicotinic receptors attenuate neurite development through calcium activation of calpain at the growth cone. *PLoS One* **13**, e0197247.
42. Olsen, R. W. (2018) GABAA receptor: Positive and negative allosteric modulators. *Neuropharmacology* **136**, 10-22.
43. Bertrand, D., and Gopalakrishnan, M. (2007) Allosteric modulation of nicotinic acetylcholine receptors. *Biochemical pharmacology* **74**, 1155-1163.
44. Changeux, J. P., and Christopoulos, A. (2016) Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell* **166**, 1084-1102.
45. Lynch, J. W. (2009) Native glycine receptor subtypes and their physiological roles. *Neuropharmacology* **56**, 303-309.
46. Laube, B., Maksay, G., Schemm, R., and Betz, H. (2002) Modulation of glycine receptor function: a novel approach for therapeutic intervention at inhibitory synapses? *Trends in pharmacological sciences* **23**, 519-527.
47. McCool, B. A., and Chappell, A. (2007) Strychnine and taurine modulation of amygdala-associated anxiety-like behavior is 'state' dependent. *Behavioural brain research* **178**, 70-81.
48. Costa, B. (2007) On the pharmacological properties of Delta9-tetrahydrocannabinol (THC). *Chem Biodivers* **4**, 1664-1677.
49. Elphick, M. R., and Egertova, M. (2001) The neurobiology and evolution of cannabinoid signalling. *Philos Trans R Soc Lond B Biol Sci* **356**, 381-408.

50. Pacher, P., Batkai, S., and Kunos, G. (2006) The endocannabinoid system as an emerging target of pharmacotherapy. *Pharmacol Rev* **58**, 389-462.
51. Zimmer, A., Zimmer, A. M., Hohmann, A. G., Herkenham, M., and Bonner, T. I. (1999) Increased mortality, hypoactivity, and hypoalgesia in cannabinoid CB1 receptor knockout mice. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 5780-5785.
52. Xiong, W., Cheng, K., Cui, T., Godlewski, G., Rice, K. C., Xu, Y., and Zhang, L. (2011) Cannabinoid potentiation of glycine receptors contributes to cannabis-induced analgesia. *Nature chemical biology* **7**, 296-303.
53. Hejazi, N., Zhou, C., Oz, M., Sun, H., Ye, J. H., and Zhang, L. (2006) Delta9-tetrahydrocannabinol and endogenous cannabinoid anandamide directly potentiate the function of glycine receptors. *Mol Pharmacol* **69**, 991-997.
54. Betz, H., and Laube, B. (2006) Glycine receptors: recent insights into their structural organization and functional diversity. *J Neurochem* **97**, 1600-1610.
55. Pless, S. A., and Lynch, J. W. (2008) Illuminating the structure and function of Cys-loop receptors. *Clinical and experimental pharmacology & physiology* **35**, 1137-1142.
56. Mowrey, D. D., Cui, T., Jia, Y., Ma, D., Makhov, A. M., Zhang, P., Tang, P., and Xu, Y. (2013) Open-channel structures of the human glycine receptor alpha1 full-length transmembrane domain. *Structure* **21**, 1897-1904.
57. Barrantes, F. J. (2002) Lipid matters: nicotinic acetylcholine receptor-lipid interactions (Review). *Molecular membrane biology* **19**, 277-284.
58. daCosta, C. J., Ogrel, A. A., McCardy, E. A., Blanton, M. P., and Baenziger, J. E. (2002) Lipid-protein interactions at the nicotinic acetylcholine receptor. A functional coupling between nicotinic receptors and phosphatidic acid-containing lipid bilayers. *J Biol Chem* **277**, 201-208.
59. daCosta, C. J., Dey, L., Therien, J. P., and Baenziger, J. E. (2013) A distinct mechanism for activating uncoupled nicotinic acetylcholine receptors. *Nature chemical biology* **9**, 701-707.
60. Brannigan, G., Henin, J., Law, R., Eckenhoff, R., and Klein, M. L. (2008) Embedded cholesterol in the nicotinic acetylcholine receptor. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14418-14423.
61. Tillman, T. S., and Cascio, M. (2003) Effects of membrane lipids on ion channel structure and function. *Cell biochemistry and biophysics* **38**, 161-190.
62. Cheng, L. S., Amaro, R. E., Xu, D., Li, W. W., Arzberger, P. W., and McCammon, J. A. (2008) Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *Journal of medicinal chemistry* **51**, 3878-3894.

63. Durrant, J. D., and McCammon, J. A. (2011) Molecular dynamics simulations and drug discovery. *Bmc Biol* **9**, 71.
64. Weinstein, D. B., Marsh, J. B., Glick, M. C., and Warren, L. (1969) Membranes of animal cells. IV. Lipids of the L cell and its surface membrane. *J Biol Chem* **244**, 4103-4111.
65. Wolf, M. G., Hoefling, M., Aponte-Santamaria, C., Grubmuller, H., and Groenhof, G. (2010) *g\_membed*: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *Journal of computational chemistry* **31**, 2169-2174.
66. Hub, J. S., Winkler, F. K., Merrick, M., and de Groot, B. L. (2010) Potentials of mean force and permeabilities for carbon dioxide, ammonia, and water flux across a Rhesus protein channel and lipid membranes. *J Am Chem Soc* **132**, 13251-13263.
67. Wennberg, C. L., van der Spoel, D., and Hub, J. S. (2012) Large influence of cholesterol on solute partitioning into lipid membranes. *J Am Chem Soc* **134**, 5351-5361.
68. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **4**, 435-447.
69. Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845-854.
70. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry* **24**, 1999-2012.
71. Jambeck, J. P., and Lyubartsev, A. P. (2012) Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J Phys Chem B* **116**, 3164-3179.
72. Jambeck, J. P., and Lyubartsev, A. P. (2012) An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *J Chem Theory Comput* **8**, 2938-2948.
73. Jambeck, J. P., and Lyubartsev, A. P. (2013) Another Piece of the Membrane Puzzle: Extending Slipids Further. *J Chem Theory Comput* **9**, 774-784.
74. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of computational chemistry* **18**, 1463-1472.
75. Heyer, L. J., Kruglyak, S., and Yooseph, S. (1999) Exploring expression data: identification and analysis of coexpressed genes. *Genome research* **9**, 1106-1115.

76. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38.
77. Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31**, 455-461.
78. Dascal, N. (2001) Voltage clamp recordings from *Xenopus* oocytes. *Current protocols in neuroscience / editorial board, Jacqueline N. Crawley ... [et al.]* **Chapter 6**, Unit 6 12.
79. Pan, J., Chen, Q., Willenbring, D., Yoshida, K., Tillman, T., Kashlan, O. B., Cohen, A., Kong, X. P., Xu, Y., and Tang, P. (2012) Structure of the pentameric ligand-gated ion channel ELIC cocrystallized with its competitive antagonist acetylcholine. *Nat Commun* **3**, 714.
80. Smart, O. S., Neduvilil, J. G., Wang, X., Wallace, B. A., and Sansom, M. S. (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* **14**, 354-360, 376.
81. Seeman, P. (2002) Atypical antipsychotics: mechanism of action. *Can J Psychiatry* **47**, 27-38.
82. Baell, J. B., and Holloway, G. A. (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of medicinal chemistry* **53**, 2719-2740.
83. Backman, T. W., Cao, Y., and Girke, T. (2011) ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic Acids Res* **39**, W486-491.
84. Norton, W. L., and Meisinger, M. A. (1977) An overview of nonsteroidal antiinflammatory agents (NSAIA). *Inflammation* **2**, 37-46.
85. Gillespie, R. J., Adams, D. R., Bebbington, D., Benwell, K., Cliffe, I. A., Dawson, C. E., Dourish, C. T., Fletcher, A., Gaur, S., Giles, P. R., Jordan, A. M., Knight, A. R., Knutsen, L. J., Lawrence, A., Lerpiniere, J., Misra, A., Porter, R. H., Pratt, R. M., Shepherd, R., Upton, R., Ward, S. E., Weiss, S. M., and Williamson, D. S. (2008) Antagonists of the human adenosine A2A receptor. Part 1: Discovery and synthesis of thieno[3,2-d]pyrimidine-4-methanone derivatives. *Bioorganic & medicinal chemistry letters* **18**, 2916-2919.
86. Hurst, R. S., Hajos, M., Raggenbass, M., Wall, T. M., Higdon, N. R., Lawson, J. A., Rutherford-Root, K. L., Berkenpas, M. B., Hoffmann, W. E., Piotrowski, D. W., Groppi, V. E., Allaman, G., Ogier, R., Bertrand, S., Bertrand, D., and Arneric, S. P. (2005) A novel positive allosteric modulator of the alpha7 neuronal nicotinic acetylcholine receptor: in vitro and in vivo characterization. *J Neurosci* **25**, 4396-4405.

87. Oz, M., Zhang, L., Ravindran, A., Morales, M., and Lupica, C. R. (2004) Differential effects of endogenous and synthetic cannabinoids on alpha7-nicotinic acetylcholine receptor-mediated responses in *Xenopus Oocytes*. *The Journal of pharmacology and experimental therapeutics* **310**, 1152-1160.
88. Wells, M. M., Tillman, T. S., Mowrey, D. D., Sun, T., Xu, Y., and Tang, P. (2015) Ensemble-based virtual screening for cannabinoid-like potentiators of the human glycine receptor alpha1 for the treatment of pain. *Journal of medicinal chemistry* **58**, 2958-2966.
89. Lu, J., Fan, S., Zou, G., Hou, Y., Pan, T., Guo, W., Yao, L., Du, F., Homanics, G. E., Liu, D., Zhang, L., and Xiong, W. (2018) Involvement of glycine receptor alpha1 subunits in cannabinoid-induced analgesia. *Neuropharmacology* **133**, 224-232.
90. Harvey, R. J., Depner, U. B., Wassle, H., Ahmadi, S., Heindl, C., Reinold, H., Smart, T. G., Harvey, K., Schutz, B., Abo-Salem, O. M., Zimmer, A., Poisbeau, P., Welzl, H., Wolfer, D. P., Betz, H., Zeilhofer, H. U., and Muller, U. (2004) GlyR alpha3: an essential target for spinal PGE2-mediated inflammatory pain sensitization. *Science* **304**, 884-887.
91. Harvey, V. L., Caley, A., Muller, U. C., Harvey, R. J., and Dickenson, A. H. (2009) A Selective Role for alpha3 Subunit Glycine Receptors in Inflammatory Pain. *Front Mol Neurosci* **2**, 14.
92. Acuna, M. A., Yevenes, G. E., Ralvenius, W. T., Benke, D., Di Lio, A., Lara, C. O., Munoz, B., Burgos, C. F., Moraga-Cid, G., Corringer, P. J., and Zeilhofer, H. U. (2016) Phosphorylation state-dependent modulation of spinal glycine receptors alleviates inflammatory pain. *J Clin Invest* **126**, 2547-2560.
93. Breitingner, U., and Breitingner, H. G. (2016) Augmentation of glycine receptor alpha3 currents suggests a mechanism for glucose-mediated analgesia. *Neurosci Lett* **612**, 110-115.
94. Zeilhofer, H. U., Acuna, M. A., Gingras, J., and Yevenes, G. E. (2018) Glycine receptors and glycine transporters: targets for novel analgesics? *Cell Mol Life Sci* **75**, 447-465.
95. Huang, X., Chen, H., Michelsen, K., Schneider, S., and Shaffer, P. L. (2015) Crystal structure of human glycine receptor-alpha3 bound to antagonist strychnine. *Nature* **526**, 277-280.
96. Huang, X., Chen, H., and Shaffer, P. L. (2017) Crystal Structures of Human GlyRalpha3 Bound to Ivermectin. *Structure* **25**, 945-950 e942.
97. Huang, X., Shaffer, P. L., Ayube, S., Bregman, H., Chen, H., Lehto, S. G., Luther, J. A., Matson, D. J., McDonough, S. I., Michelsen, K., Plant, M. H., Schneider, S., Simard, J. R., Teffera, Y., Yi, S., Zhang, M., DiMauro, E. F., and Gingras, J. (2017) Crystal structures of human glycine receptor alpha3 bound to a novel class of analgesic potentiators. *Nat Struct Mol Biol* **24**, 108-113.

98. Du, J., Lu, W., Wu, S., Cheng, Y., and Gouaux, E. (2015) Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526**, 224-229.
99. Gonzalez-Gutierrez, G., Wang, Y., Cymes, G. D., Tajkhorshid, E., and Grosman, C. (2017) Chasing the open-state structure of pentameric ligand-gated ion channels. *J Gen Physiol* **149**, 1119-1138.
100. Xiong, W., Cui, T., Cheng, K., Yang, F., Chen, S. R., Willenbring, D., Guan, Y., Pan, H. L., Ren, K., Xu, Y., and Zhang, L. (2012) Cannabinoids suppress inflammatory and neuropathic pain by targeting alpha3 glycine receptors. *J Exp Med* **209**, 1121-1134.
101. Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nature chemical biology* **5**, 789-796.
102. Changeux, J. P., and Edelstein, S. (2011) Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep* **3**, 19.
103. Korb, O., Olsson, T. S., Bowden, S. J., Hall, R. J., Verdonk, M. L., Liebeschuetz, J. W., and Cole, J. C. (2012) Potential and limitations of ensemble docking. *J Chem Inf Model* **52**, 1262-1274.
104. Ganesan, A., Coote, M. L., and Barakat, K. (2017) Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discov Today* **22**, 249-269.
105. De Vivo, M., Masetti, M., Bottegoni, G., and Cavalli, A. (2016) Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of medicinal chemistry* **59**, 4035-4061.
106. Webb, B., and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **86**, 2 9 1-2 9 37.
107. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815.
108. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25.
109. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092.
110. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R. G. (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* **52**, 1757-1768.
111. Lipinski, C. A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* **44**, 235-249.



112. Pajouhesh, H., and Lenz, G. R. (2005) Medicinal chemical properties of successful central nervous system drugs. *NeuroRx* **2**, 541-553.
113. Pordes, R., Petravick, D., Kramer, B., Olson, D., Livny, M., Roy, A., Avery, P., Blackburn, K., Wenaus, T., and Würthwein, F. (2007) The open science grid. in *Journal of Physics: Conference Series*, IOP Publishing
114. Sfiligoi, I., Bradley, D. C., Holzman, B., Mhashilkar, P., Padhi, S., and Wurthwein, F. (2009) The pilot way to grid resources using glideinWMS. in *Computer Science and Information Engineering, 2009 WRI World Congress on*, IEEE
115. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell, A. D., Jr. (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of computational chemistry* **31**, 671-690.
116. Vanommeslaeghe, K., and MacKerell, A. D., Jr. (2012) Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* **52**, 3144-3154.
117. Vanommeslaeghe, K., Raman, E. P., and MacKerell, A. D., Jr. (2012) Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* **52**, 3155-3168.
118. Mayne, C. G., Saam, J., Schulten, K., Tajkhorshid, E., and Gumbart, J. C. (2013) Rapid parameterization of small molecules using the Force Field Toolkit. *Journal of computational chemistry* **34**, 2757-2770.
119. Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and Mackerell, A. D., Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput* **8**, 3257-3273.
120. Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., Jr., and Pastor, R. W. (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* **114**, 7830-7843.
121. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmuller, H., and MacKerell, A. D., Jr. (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **14**, 71-73.
122. Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., and Lindahl, E. (2010) Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J Chem Theory Comput* **6**, 459-466.

123. Pribilla, I., Takagi, T., Langosch, D., Bormann, J., and Betz, H. (1992) The atypical M2 segment of the beta subunit confers picrotoxinin resistance to inhibitory glycine receptor channels. *EMBO J* **11**, 4305-4311.
124. Hargreaves, K., Dubner, R., Brown, F., Flores, C., and Joris, J. (1988) A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. *Pain* **32**, 77-88.
125. Lahlou, M. (2013) The Success of Natural Products in Drug Discovery. *Pharmacology & Pharmacy* **04**, 17-31.
126. Kirchmair, J., Goller, A. H., Lang, D., Kunze, J., Testa, B., Wilson, I. D., Glen, R. C., and Schneider, G. (2015) Predicting drug metabolism: experiment and/or computation? *Nat Rev Drug Discov* **14**, 387-404.
127. Hughes, J. P., Rees, S., Kalindjian, S. B., and Philpott, K. L. (2011) Principles of early drug discovery. *British journal of pharmacology* **162**, 1239-1249.
128. Besnard, J., Ruda, G. F., Setola, V., Abecassis, K., Rodriguiz, R. M., Huang, X. P., Norval, S., Sassano, M. F., Shin, A. I., Webster, L. A., Simeons, F. R., Stojanovski, L., Prat, A., Seidah, N. G., Constam, D. B., Bickerton, G. R., Read, K. D., Wetsel, W. C., Gilbert, I. H., Roth, B. L., and Hopkins, A. L. (2012) Automated design of ligands to polypharmacological profiles. *Nature* **492**, 215-220.
129. Kroeze, W. K., Sassano, M. F., Huang, X. P., Lansu, K., McCorvy, J. D., Giguere, P. M., Sciaky, N., and Roth, B. L. (2015) PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat Struct Mol Biol* **22**, 362-369.
130. Ahmadi, S., Lippross, S., Neuhuber, W. L., and Zeilhofer, H. U. (2002) PGE(2) selectively blocks inhibitory glycinergic neurotransmission onto rat superficial dorsal horn neurons. *Nat Neurosci* **5**, 34-40.
131. Boudreau, D., Von Korff, M., Rutter, C. M., Saunders, K., Ray, G. T., Sullivan, M. D., Campbell, C. I., Merrill, J. O., Silverberg, M. J., Banta-Green, C., and Weisner, C. (2009) Trends in long-term opioid therapy for chronic non-cancer pain. *Pharmacoepidemiol Drug Saf* **18**, 1166-1175.
132. Seth, P., Scholl, L., Rudd, R. A., and Bacon, S. (2018) Overdose Deaths Involving Opioids, Cocaine, and Psychostimulants - United States, 2015-2016. *MMWR Morb Mortal Wkly Rep* **67**, 349-358.
133. Zhang, W. S., Xu, H., Xu, B., Sun, S., Deng, X. M., and Zhang, Y. Q. (2009) Antihyperalgesic effect of systemic dexmedetomidine and gabapentin in a rat model of monoarthritis. *Brain Res* **1264**, 57-66.
134. Armendariz, A., and Nazarian, A. (2018) Morphine antinociception on thermal sensitivity and place conditioning in male and female rats treated with intraplantar complete freund's adjuvant. *Behavioural brain research* **343**, 21-27.

135. Mersfelder, T. L., and Nichols, W. H. (2016) Gabapentin: Abuse, Dependence, and Withdrawal. *Ann Pharmacother* **50**, 229-233.
136. Weeks, J. R. (1962) Experimental morphine addiction: method for automatic intravenous injections in unrestrained rats. *Science* **138**, 143-144.
137. Stead, C., Brown, A., Adams, C., Nickolls, S. J., Young, G., Kammonen, J., Pryde, D., and Cawkill, D. (2016) Identification of Positive Allosteric Modulators of Glycine Receptors from a High-Throughput Screen Using a Fluorescent Membrane Potential Assay. *J Biomol Screen* **21**, 1042-1053.
138. Bregman, H., Simard, J. R., Andrews, K. L., Ayube, S., Chen, H., Gunaydin, H., Guzman-Perez, A., Hu, J., Huang, L., Huang, X., Krolikowski, P. H., Lehto, S. G., Lewis, R. T., Michelsen, K., Pegman, P., Plant, M. H., Shaffer, P. L., Teffera, Y., Yi, S., Zhang, M., Gingras, J., and DiMauro, E. F. (2017) The Discovery and Hit-to-Lead Optimization of Tricyclic Sulfonamides as Potent and Efficacious Potentiators of Glycine Receptors. *Journal of medicinal chemistry* **60**, 1105-1125.
139. Chen, X., Webb, T. I., and Lynch, J. W. (2009) The M4 transmembrane segment contributes to agonist efficacy differences between alpha1 and alpha3 glycine receptors. *Molecular membrane biology* **26**, 321-332.
140. Mitra, A., Bailey, T. D., and Auerbach, A. L. (2004) Structural dynamics of the M4 transmembrane segment during acetylcholine receptor gating. *Structure* **12**, 1909-1918.
141. Haeger, S., Kuzmin, D., Detro-Dassen, S., Lang, N., Kilb, M., Tsetlin, V., Betz, H., Laube, B., and Schmalzing, G. (2010) An intramembrane aromatic network determines pentameric assembly of Cys-loop receptors. *Nat Struct Mol Biol* **17**, 90-98.
142. Carswell, C. L., Henault, C. M., Murlidaran, S., Therien, J. P. D., Juranka, P. F., Surujballi, J. A., Brannigan, G., and Baenziger, J. E. (2015) Role of the Fourth Transmembrane alpha Helix in the Allosteric Modulation of Pentameric Ligand-Gated Ion Channels. *Structure* **23**, 1655-1664.
143. Therien, J. P., and Baenziger, J. E. (2017) Pentameric ligand-gated ion channels exhibit distinct transmembrane domain archetypes for folding/expression and function. *Sci Rep* **7**, 450.
144. Henault, C. M., Juranka, P. F., and Baenziger, J. E. (2015) The M4 Transmembrane alpha-Helix Contributes Differently to Both the Maturation and Function of Two Prokaryotic Pentameric Ligand-gated Ion Channels. *J Biol Chem* **290**, 25118-25128.
145. daCosta, C. J., and Baenziger, J. E. (2009) A lipid-dependent uncoupled conformation of the acetylcholine receptor. *J Biol Chem* **284**, 17819-17825.
146. Skilling, S. R., Smullin, D. H., Beitz, A. J., and Larson, A. A. (1988) Extracellular amino acid concentrations in the dorsal spinal cord of freely moving rats following veratridine and nociceptive stimulation. *J Neurochem* **51**, 127-132.

147. Whitehead, K. J., Manning, J. P., Smith, C. G., and Bowery, N. G. (2001) Determination of the extracellular concentration of glycine in the rat spinal cord dorsal horn by quantitative microdialysis. *Brain Res* **910**, 192-194.
148. Malmberg, A. B., and Yaksh, T. L. (1995) Cyclooxygenase inhibition and the spinal release of prostaglandin E2 and amino acids evoked by paw formalin injection: a microdialysis study in unanesthetized rats. *J Neurosci* **15**, 2768-2776.
149. Maeda, Y., Lisi, T. L., Vance, C. G., and Sluka, K. A. (2007) Release of GABA and activation of GABA(A) in the spinal cord mediates the effects of TENS in rats. *Brain Res* **1136**, 43-50.
150. Zahn, P. K., Sluka, K. A., and Brennan, T. J. (2002) Excitatory amino acid release in the spinal cord caused by plantar incision in the rat. *Pain* **100**, 65-76.
151. 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. *Biochemical pharmacology* **76**, 1014-1023.
152. Lozovaya, N., Yatsenko, N., Beketov, A., Tsintsadze, T., and Burnashev, N. (2005) Glycine receptors in CNS neurons as a target for nonretrograde action of cannabinoids. *J Neurosci* **25**, 7499-7506.
153. Wang, Y., Yang, E., Wells, M. M., Bondarenko, V., Woll, K., Carnevale, V., Granata, D., Klein, M. L., Eckenhoff, R. G., Dailey, W. P., Covarrubias, M., Tang, P., and Xu, Y. (2018) Propofol inhibits the voltage-gated sodium channel NaChBac at multiple sites. *J Gen Physiol*
154. Bromstrup, T., Howard, R. J., Trudell, J. R., Harris, R. A., and Lindahl, E. (2013) Inhibition versus potentiation of ligand-gated ion channels can be altered by a single mutation that moves ligands between intra- and intersubunit sites. *Structure* **21**, 1307-1316.
155. Bondarenko, V., Mowrey, D., Liu, L. T., Xu, Y., and Tang, P. (2013) NMR resolved multiple anesthetic binding sites in the TM domains of the alpha4beta2 nAChR. *Biochim Biophys Acta* **1828**, 398-404.
156. Chen, Q., Wells, M. M., Tillman, T. S., Kinde, M. N., Cohen, A., Xu, Y., and Tang, P. (2017) Structural Basis of Alcohol Inhibition of the Pentameric Ligand-Gated Ion Channel ELIC. *Structure* **25**, 180-187.
157. Kinde, M. N., Bondarenko, V., Granata, D., Bu, W., Grasty, K. C., Loll, P. J., Carnevale, V., Klein, M. L., Eckenhoff, R. G., Tang, P., and Xu, Y. (2016) Fluorine-19 NMR and computational quantification of isoflurane binding to the voltage-gated sodium channel NaChBac. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 13762-13767.
158. Zoli, M., Pistillo, F., and Gotti, C. (2015) Diversity of native nicotinic receptor subtypes in mammalian brain. *Neuropharmacology* **96**, 302-311.

159. Seguela, P., Wadiche, J., Dineley-Miller, K., Dani, J. A., and Patrick, J. W. (1993) Molecular cloning, functional properties, and distribution of rat brain alpha 7: a nicotinic cation channel highly permeable to calcium. *J Neurosci* **13**, 596-604.
160. Ragozzino, D., Barabino, B., Fucile, S., and Eusebi, F. (1998) Ca<sup>2+</sup> permeability of mouse and chick nicotinic acetylcholine receptors expressed in transiently transfected human cells. *J Physiol* **507 ( Pt 3)**, 749-757.
161. Dani, J. A. (2015) Neuronal Nicotinic Acetylcholine Receptor Structure and Function and Response to Nicotine. *Int Rev Neurobiol* **124**, 3-19.
162. Dominguez del Toro, E., Juiz, J. M., Peng, X., Lindstrom, J., and Criado, M. (1994) Immunocytochemical localization of the alpha 7 subunit of the nicotinic acetylcholine receptor in the rat central nervous system. *J Comp Neurol* **349**, 325-342.
163. Castelan, F., Mulet, J., Aldea, M., Sala, S., Sala, F., and Criado, M. (2007) Cytoplasmic regions adjacent to the M3 and M4 transmembrane segments influence expression and function of alpha7 nicotinic acetylcholine receptors. A study with single amino acid mutants. *J Neurochem* **100**, 406-415.
164. Hales, T. G., Dunlop, J. I., Deeb, T. Z., Carland, J. E., Kelley, S. P., Lambert, J. J., and Peters, J. A. (2006) Common determinants of single channel conductance within the large cytoplasmic loop of 5-hydroxytryptamine type 3 and alpha4beta2 nicotinic acetylcholine receptors. *J Biol Chem* **281**, 8062-8071.
165. O'Toole, K. K., and Jenkins, A. (2011) Discrete M3-M4 intracellular loop subdomains control specific aspects of gamma-aminobutyric acid type A receptor function. *J Biol Chem* **286**, 37990-37999.
166. Papke, D., and Grosman, C. (2014) The role of intracellular linkers in gating and desensitization of human pentameric ligand-gated ion channels. *J Neurosci* **34**, 7238-7252.
167. Williams, D. K., Peng, C., Kimbrell, M. R., and Papke, R. L. (2012) Intrinsically low open probability of alpha7 nicotinic acetylcholine receptors can be overcome by positive allosteric modulation and serum factors leading to the generation of excitotoxic currents at physiological temperatures. *Mol Pharmacol* **82**, 746-759.
168. Leiser, S. C., Bowlby, M. R., Comery, T. A., and Dunlop, J. (2009) A cog in cognition: how the alpha 7 nicotinic acetylcholine receptor is geared towards improving cognitive deficits. *Pharmacology & therapeutics* **122**, 302-311.
169. Lendvai, B., Kassai, F., Szajli, A., and Nemethy, Z. (2013) alpha7 nicotinic acetylcholine receptors and their role in cognition. *Brain research bulletin* **93**, 86-96.
170. Parri, H. R., Hernandez, C. M., and Dineley, K. T. (2011) Research update: Alpha7 nicotinic acetylcholine receptor mechanisms in Alzheimer's disease. *Biochemical pharmacology* **82**, 931-942.

171. Ishikawa, M., and Hashimoto, K. (2011) alpha7 nicotinic acetylcholine receptor as a potential therapeutic target for schizophrenia. *Current pharmaceutical design* **17**, 121-129.
172. Wallace, T. L., and Bertrand, D. (2013) Alpha7 neuronal nicotinic receptors as a drug target in schizophrenia. *Expert opinion on therapeutic targets* **17**, 139-155.
173. de Jonge, W. J., and Ulloa, L. (2007) The alpha7 nicotinic acetylcholine receptor as a pharmacological target for inflammation. *British journal of pharmacology* **151**, 915-929.
174. Wang, H., Yu, M., Ochani, M., Amella, C. A., Tanovic, M., Susarla, S., Li, J. H., Wang, H., Yang, H., Ulloa, L., Al-Abed, Y., Czura, C. J., and Tracey, K. J. (2003) Nicotinic acetylcholine receptor alpha7 subunit is an essential regulator of inflammation. *Nature* **421**, 384-388.
175. Alsharari, S. D., Freitas, K., and Damaj, M. I. (2013) Functional role of alpha7 nicotinic receptor in chronic neuropathic and inflammatory pain: studies in transgenic mice. *Biochemical pharmacology* **86**, 1201-1207.
176. Feduccia, A. A., Chatterjee, S., and Bartlett, S. E. (2012) Neuronal nicotinic acetylcholine receptors: neuroplastic changes underlying alcohol and nicotine addictions. *Front Mol Neurosci* **5**, 83.
177. Dineley, K. T., Pandya, A. A., and Yakel, J. L. (2015) Nicotinic ACh receptors as therapeutic targets in CNS disorders. *Trends in pharmacological sciences* **36**, 96-108.
178. Wallace, T. L., and Porter, R. H. (2011) Targeting the nicotinic alpha7 acetylcholine receptor to enhance cognition in disease. *Biochemical pharmacology* **82**, 891-903.
179. Basak, S., Gicheru, Y., Samanta, A., Molugu, S. K., Huang, W., Fuente, M., Hughes, T., Taylor, D. J., Nieman, M. T., Moiseenkova-Bell, V., and Chakrapani, S. (2018) Cryo-EM structure of 5-HT3A receptor in its resting conformation. *Nat Commun* **9**, 514.
180. Dosztanyi, Z. (2018) Prediction of protein disorder based on IUPred. *Protein Sci* **27**, 331-340.
181. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433-3434.
182. Hilf, R. J., and Dutzler, R. (2008) X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452**, 375-379.
183. Hilf, R. J., and Dutzler, R. (2009) Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457**, 115-118.
184. Bocquet, N., Nury, H., Baaden, M., Le Poupon, C., Changeux, J. P., Delarue, M., and Corringer, P. J. (2009) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457**, 111-114.

185. Miller, P. S., and Aricescu, A. R. (2014) Crystal structure of a human GABAA receptor. *Nature* **512**, 270-275.
186. Althoff, T., Hibbs, R. E., Banerjee, S., and Gouaux, E. (2014) X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature* **512**, 333-337.
187. Hibbs, R. E., and Gouaux, E. (2011) Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474**, 54-60.
188. Miller, P. S., Scott, S., Masiulis, S., De Colibus, L., Pardon, E., Steyaert, J., and Aricescu, A. R. (2017) Structural basis for GABAA receptor potentiation by neurosteroids. *Nat Struct Mol Biol* **24**, 986-992.
189. Laverty, D., Thomas, P., Field, M., Andersen, O. J., Gold, M. G., Biggin, P. C., Gielen, M., and Smart, T. G. (2017) Crystal structures of a GABAA-receptor chimera reveal new endogenous neurosteroid-binding sites. *Nat Struct Mol Biol* **24**, 977-985.
190. Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M. B., Hovius, R., Graff, A., Stahlberg, H., Tomizaki, T., Desmyter, A., Moreau, C., Li, X. D., Poitevin, F., Vogel, H., and Nury, H. (2014) X-ray structure of the mouse serotonin 5-HT<sub>3</sub> receptor. *Nature* **512**, 276-281.
191. Kudryashev, M., Castano-Diez, D., Deluz, C., Hassaine, G., Grasso, L., Graf-Meyer, A., Vogel, H., and Stahlberg, H. (2016) The Structure of the Mouse Serotonin 5-HT<sub>3</sub> Receptor in Lipid Vesicles. *Structure* **24**, 165-170.
192. 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 alpha4beta2 nicotinic receptor. *Nature* **557**, 261-265.
193. Morales-Perez, C. L., Noviello, C. M., and Hibbs, R. E. (2016) X-ray structure of the human alpha4beta2 nicotinic receptor. *Nature* **538**, 411-415.
194. Unwin, N. (2005) Refined structure of the nicotinic acetylcholine receptor at 4Å resolution. *J Mol Biol* **346**, 967-989.
195. Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H. W. (2000) Dead-time free measurement of dipole-dipole interactions between electron spins. *J Magn Reson* **142**, 331-340.
196. Jeschke, G., and Polyhach, Y. (2007) Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical chemistry chemical physics : PCCP* **9**, 1895-1910.
197. Eliezer, D. (2009) Biophysical characterization of intrinsically disordered proteins. *Curr Opin Struct Biol* **19**, 23-30.
198. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev* **104**, 3607-3622.

199. Drescher, M. (2012) EPR in protein science : intrinsically disordered proteins. *Topics in current chemistry* **321**, 91-119.
200. Tillman, T. S., Alvarez, F. J., Reinert, N. J., Liu, C., Wang, D., Xu, Y., Xiao, K., Zhang, P., and Tang, P. (2016) Functional Human alpha7 Nicotinic Acetylcholine Receptor (nAChR) Generated from Escherichia coli. *J Biol Chem* **291**, 18276-18282.
201. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**, 71-75.
202. Bondarenko, V., Mowrey, D., Tillman, T., Cui, T., Liu, L. T., Xu, Y., and Tang, P. (2012) NMR structures of the transmembrane domains of the alpha4beta2 nAChR. *Biochim Biophys Acta* **1818**, 1261-1268.
203. Bondarenko, V., Mowrey, D. D., Tillman, T. S., Seyoum, E., Xu, Y., and Tang, P. (2014) NMR structures of the human alpha7 nAChR transmembrane domain and associated anesthetic binding sites. *Biochim Biophys Acta* **1838**, 1389-1395.
204. Kinde, M. N., Chen, Q., Lawless, M. J., Mowrey, D. D., Xu, J., Saxena, S., Xu, Y., and Tang, P. (2015) Conformational Changes Underlying Desensitization of the Pentameric Ligand-Gated Ion Channel ELIC. *Structure* **23**, 995-1004.
205. Mehta, V. D., Kulkarni, P. V., Mason, R. P., Constantinescu, A., and Antich, P. P. (1994) Fluorinated proteins as potential <sup>19</sup>F magnetic resonance imaging and spectroscopy agents. *Bioconjugate chemistry* **5**, 257-261.
206. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C., and Wuthrich, K. (2012) Biased signaling pathways in beta2-adrenergic receptor characterized by <sup>19</sup>F-NMR. *Science* **335**, 1106-1110.
207. Ebright, Y. W., Chen, Y., Pendergrast, P. S., and Ebright, R. H. (1992) Incorporation of an EDTA-metal complex at a rationally selected site within a protein: application to EDTA-iron DNA affinity cleaving with catabolite gene activator protein (CAP) and Cro. *Biochemistry* **31**, 10664-10670.
208. Kamen, D. E., Cahill, S. M., and Girvin, M. E. (2007) Multiple alignment of membrane proteins for measuring residual dipolar couplings using lanthanide ions bound to a small metal chelator. *J Am Chem Soc* **129**, 1846-1847.
209. Halevi, S., Yassin, L., Eshel, M., Sala, F., Sala, S., Criado, M., and Treinin, M. (2003) Conservation within the RIC-3 gene family. Effectors of mammalian nicotinic acetylcholine receptor expression. *J Biol Chem* **278**, 34411-34417.
210. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135-140.



211. Schulte-Herbruggen, T., and Sorensen, O. W. (2000) Clean TROSY: compensation for relaxation-induced artifacts. *J Magn Reson* **144**, 123-128.
212. Eletsky, A., Kienhofer, A., and Pervushin, K. (2001) TROSY NMR with partially deuterated proteins. *J Biomol NMR* **20**, 177-180.
213. Salzmänn, M., Wider, G., Pervushin, K., Senn, H., and Wüthrich, K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *Journal of the American Chemical Society* **121**, 844-848.
214. Salzmänn, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 13585-13590.
215. Zhu, G., Kong, X. M., and Sze, K. H. (1999) Gradient and sensitivity enhancement of 2D TROSY with water flip-back, 3D NOESY-TROSY and TOCSY-TROSY experiments. *J Biomol NMR* **13**, 77-81.
216. Czisch, M., and Boelens, R. (1998) Sensitivity enhancement in the TROSY experiment. *J Magn Reson* **134**, 158-160.
217. Pervushin, K. V., Wider, G., and Wüthrich, K. (1998) Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [<sup>15</sup>N,<sup>1</sup>H]-TROSY. *J Biomol NMR* **12**, 345-348.
218. Nietlispach, D. (2005) Suppression of anti-TROSY lines in a sensitivity enhanced gradient selection TROSY scheme. *J Biomol NMR* **31**, 161-166.
219. Yang, D., and Kay, L. E. (1999) Improved <sup>1</sup>HN-detected triple resonance TROSY-based experiments. *J Biomol NMR* **13**, 3-10.
220. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293.
221. Goddard, T., and Kneller, D. (2001) Sparky 3 NMR software. *University of California, San Francisco*
222. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**, 289-302.
223. Baxter, N. J., and Williamson, M. P. (1997) Temperature dependence of <sup>1</sup>H-1 chemical shifts in proteins. *Journal of biomolecular NMR* **9**, 359-369.

224. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry* **33**, 5984-6003.
225. Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355-5365.
226. Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111-1114.
227. Losonczi, J. A., Andrec, M., Fischer, M. W., and Prestegard, J. H. (1999) Order matrix analysis of residual dipolar couplings using singular value decomposition. *J Magn Reson* **138**, 334-342.
228. Bax, A., Kontaxis, G., and Tjandra, N. (2001) Dipolar couplings in macromolecular structure determination. *Methods Enzymol* **339**, 127-174.
229. Valafar, H., and Prestegard, J. H. (2004) REDCAT: a residual dipolar coupling analysis tool. *J Magn Reson* **167**, 228-241.
230. Schmidt, C., Irausquin, S. J., and Valafar, H. (2013) Advances in the REDCAT software package. *BMC Bioinformatics* **14**, 302.
231. Martin, R. E., Pannier, M., Diederich, F., Gramlich, V., Hubrich, M., and Spiess, H. W. (1998) Determination of End-to-End Distances in a Series of TEMPO Diradicals of up to 2.8 nm Length with a New Four-Pulse Double Electron Electron Resonance Experiment. *Angew Chem Int Ed Engl* **37**, 2833-2837.
232. Jeschke, G., Chechik, V., Ionita, P., Godt, A., Zimmermann, H., Banham, J., Timmel, C. R., Hilger, D., and Jung, H. (2006) DeerAnalysis2006 - a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* **30**, 473-498.
233. Alexander, N., Bortolus, M., Al-Mestarihi, A., McHaourab, H., and Meiler, J. (2008) De novo high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure* **16**, 181-195.
234. Gottstein, D., Reckel, S., Dotsch, V., and Guntert, P. (2012) Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR. *Structure* **20**, 1019-1027.
235. Sjodt, M., and Clubb, R. T. (2017) Nitroxide Labeling of Proteins and the Determination of Paramagnetic Relaxation Derived Distance Restraints for NMR Studies. *Bio Protoc* **7**
236. Song, Y., DiMaio, F., Wang, R. Y., Kim, D., Miles, C., Brunette, T., Thompson, J., and Baker, D. (2013) High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-1742.

237. Fleishman, S. J., Leaver-Fay, A., Corn, J. E., Strauch, E. M., Khare, S. D., Koga, N., Ashworth, J., Murphy, P., Richter, F., Lemmon, G., Meiler, J., and Baker, D. (2011) RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**, e20161.
238. Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P. D., Smith, C. A., Sheffler, W., Davis, I. W., Cooper, S., Treuille, A., Mandell, D. J., Richter, F., Ban, Y. E., Fleishman, S. J., Corn, J. E., Kim, D. E., Lyskov, S., Berrondo, M., Mentzer, S., Popovic, Z., Havranek, J. J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J. J., Kuhlman, B., Baker, D., and Bradley, P. (2011) ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* **487**, 545-574.
239. Leaver-Fay, A., O'Meara, M. J., Tyka, M., Jacak, R., Song, Y., Kellogg, E. H., Thompson, J., Davis, I. W., Pache, R. A., Lyskov, S., Gray, J. J., Kortemme, T., Richardson, J. S., Havranek, J. J., Snoeyink, J., Baker, D., and Kuhlman, B. (2013) Scientific benchmarks for guiding macromolecular energy function improvement. *Methods Enzymol* **523**, 109-143.
240. O'Meara, M. J., Leaver-Fay, A., Tyka, M. D., Stein, A., Houlihan, K., DiMaio, F., Bradley, P., Kortemme, T., Baker, D., Snoeyink, J., and Kuhlman, B. (2015) Combined covalent-electrostatic model of hydrogen bonding improves structure prediction with Rosetta. *J Chem Theory Comput* **11**, 609-622.
241. Rohl, C. A., Strauss, C. E., Misura, K. M., and Baker, D. (2004) Protein structure prediction using Rosetta. *Methods Enzymol* **383**, 66-93.
242. Das, R., and Baker, D. (2008) Macromolecular modeling with rosetta. *Annu Rev Biochem* **77**, 363-382.
243. Raman, S., Lange, O. F., Rossi, P., Tyka, M., Wang, X., Aramini, J., Liu, G., Ramelot, T. A., Eletsky, A., Szyperski, T., Kennedy, M. A., Prestegard, J., Montelione, G. T., and Baker, D. (2010) NMR structure determination for larger proteins using backbone-only data. *Science* **327**, 1014-1018.
244. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4685-4690.
245. Rohl, C. A., and Baker, D. (2002) De novo determination of protein backbone structure from residual dipolar couplings using Rosetta. *J Am Chem Soc* **124**, 2723-2729.
246. Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* **43**, 63-78.

247. Kontaxis, G., Delaglio, F., and Bax, A. (2005) Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol* **394**, 42-78.
248. Kim, D. E., Chivian, D., and Baker, D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32**, W526-531.
249. Bonneau, R., Tsai, J., Ruczinski, I., Chivian, D., Rohl, C., Strauss, C. E., and Baker, D. (2001) Rosetta in CASP4: progress in ab initio protein structure prediction. *Proteins Suppl* **5**, 119-126.
250. Bradley, P., Malmstrom, L., Qian, B., Schonbrun, J., Chivian, D., Kim, D. E., Meiler, J., Misura, K. M., and Baker, D. (2005) Free modeling with Rosetta in CASP6. *Proteins* **61 Suppl 7**, 128-134.
251. Das, R., Qian, B., Raman, S., Vernon, R., Thompson, J., Bradley, P., Khare, S., Tyka, M. D., Bhat, D., Chivian, D., Kim, D. E., Sheffler, W. H., Malmstrom, L., Wollacott, A. M., Wang, C., Andre, I., and Baker, D. (2007) Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta@home. *Proteins* **69 Suppl 8**, 118-128.
252. Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., Kim, D., Kellogg, E., DiMaio, F., Lange, O., Kinch, L., Sheffler, W., Kim, B. H., Das, R., Grishin, N. V., and Baker, D. (2009) Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins* **77 Suppl 9**, 89-99.
253. Ovchinnikov, S., Kim, D. E., Wang, R. Y., Liu, Y., DiMaio, F., and Baker, D. (2016) Improved de novo structure prediction in CASP11 by incorporating coevolution information into Rosetta. *Proteins* **84 Suppl 1**, 67-75.
254. Ovchinnikov, S., Park, H., Kim, D. E., DiMaio, F., and Baker, D. (2018) Protein structure prediction using Rosetta in CASP12. *Proteins* **86 Suppl 1**, 113-121.
255. Kim, D. E., Blum, B., Bradley, P., and Baker, D. (2009) Sampling bottlenecks in de novo protein structure prediction. *J Mol Biol* **393**, 249-260.
256. Bradley, P., Misura, K. M., and Baker, D. (2005) Toward high-resolution de novo structure prediction for small proteins. *Science* **309**, 1868-1871.
257. Rohl, C. A. (2005) Protein structure estimation from minimal restraints using Rosetta. *Methods Enzymol* **394**, 244-260.
258. Hirst, S. J., Alexander, N., McHaourab, H. S., and Meiler, J. (2011) RosettaEPR: an integrated tool for protein structure determination from sparse EPR data. *J Struct Biol* **173**, 506-514.

259. Rossi, P., Shi, L., Liu, G., Barbieri, C. M., Lee, H. W., Grant, T. D., Luft, J. R., Xiao, R., Acton, T. B., Snell, E. H., Montelione, G. T., Baker, D., Lange, O. F., and Sgourakis, N. G. (2015) A hybrid NMR/SAXS-based approach for discriminating oligomeric protein interfaces using Rosetta. *Proteins* **83**, 309-317.
260. DiMaio, F., Leaver-Fay, A., Bradley, P., Baker, D., and Andre, I. (2011) Modeling symmetric macromolecular structures in Rosetta3. *PLoS One* **6**, e20450.
261. Das, R., Andre, I., Shen, Y., Wu, Y., Lemak, A., Bansal, S., Arrowsmith, C. H., Szyperski, T., and Baker, D. (2009) Simultaneous prediction of protein folding and docking at high resolution. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18978-18983.
262. Sgourakis, N. G., Lange, O. F., DiMaio, F., Andre, I., Fitzkee, N. C., Rossi, P., Montelione, G. T., Bax, A., and Baker, D. (2011) Determination of the structures of symmetric protein oligomers from NMR chemical shifts and residual dipolar couplings. *J Am Chem Soc* **133**, 6288-6298.
263. Brunette, T. J., and Brock, O. (2008) Guiding conformation space search with an all-atom energy potential. *Proteins* **73**, 958-972.
264. Qian, B., Raman, S., Das, R., Bradley, P., McCoy, A. J., Read, R. J., and Baker, D. (2007) High-resolution structure prediction and the crystallographic phase problem. *Nature* **450**, 259-264.
265. Bender, B. J., Cisneros, A., 3rd, Duran, A. M., Finn, J. A., Fu, D., Lokits, A. D., Mueller, B. K., Sangha, A. K., Sauer, M. F., Sevy, A. M., Sliwoski, G., Sheehan, J. H., DiMaio, F., Meiler, J., and Moretti, R. (2016) Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. *Biochemistry* **55**, 4748-4763.
266. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **79**, 926-935.
267. Tyka, M. D., Keedy, D. A., Andre, I., DiMaio, F., Song, Y., Richardson, D. C., Richardson, J. S., and Baker, D. (2011) Alternate states of proteins revealed by detailed energy landscape mapping. *J Mol Biol* **405**, 607-618.
268. Nivon, L. G., Moretti, R., and Baker, D. (2013) A Pareto-optimal refinement method for protein design scaffolds. *PLoS One* **8**, e59004.
269. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E., and Baker, D. (2014) Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Sci* **23**, 47-55.

270. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.
271. Iwahara, J., Anderson, D. E., Murphy, E. C., and Clore, G. M. (2003) EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. *J Am Chem Soc* **125**, 6634-6635.
272. Iwahara, J., Schwieters, C. D., and Clore, G. M. (2004) Ensemble approach for NMR structure refinement against <sup>1</sup>H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *J Am Chem Soc* **126**, 5879-5896.
273. Clore, G. M., Tang, C., and Iwahara, J. (2007) Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol* **17**, 603-616.
274. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.
275. Bakan, A., Meireles, L. M., and Bahar, I. (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575-1577.
276. Krause, R. M., Buisson, B., Bertrand, S., Corringer, P. J., Galzi, J. L., Changeux, J. P., and Bertrand, D. (1998) Ivermectin: a positive allosteric effector of the alpha7 neuronal nicotinic acetylcholine receptor. *Mol Pharmacol* **53**, 283-294.
277. Prive, G. G. (2007) Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397.
278. Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., and Sligar, S. G. (2004) Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J Am Chem Soc* **126**, 3477-3487.
279. Denisov, I. G., and Sligar, S. G. (2016) Nanodiscs for structural and functional studies of membrane proteins. *Nat Struct Mol Biol* **23**, 481-486.
280. Susac, L., Horst, R., and Wuthrich, K. (2014) Solution-NMR characterization of outer-membrane protein A from *E. coli* in lipid bilayer nanodiscs and detergent micelles. *Chembiochem* **15**, 995-1000.
281. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* **222**, 311-333.

282. Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .1. Theory and Range of Validity. *Journal of the American Chemical Society* **104**, 4546-4559.
283. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J Mol Biol* **268**, 170-184.
284. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol* **268**, 158-169.
285. Liang, B., Bushweller, J. H., and Tamm, L. K. (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *J Am Chem Soc* **128**, 4389-4397.
286. Prestegard, J. H., al-Hashimi, H. M., and Tolman, J. R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q Rev Biophys* **33**, 371-424.
287. Ikegami, T., Verdier, L., Sakhaii, P., Grimme, S., Pescatore, B., Saxena, K., Fiebig, K. M., and Griesinger, C. (2004) Novel techniques for weak alignment of proteins in solution using chemical tags coordinating lanthanide ions. *J Biomol NMR* **29**, 339-349.
288. Basak, S., Schmandt, N., Gicheru, Y., and Chakrapani, S. (2017) Crystal structure and dynamics of a lipid-induced potential desensitized-state of a pentameric ligand-gated channel. *Elife* **6**
289. Dellisanti, C. D., Ghosh, B., Hanson, S. M., Raspanti, J. M., Grant, V. A., Diarra, G. M., Schuh, A. M., Satyshur, K., Klug, C. S., and Czajkowski, C. (2013) Site-directed spin labeling reveals pentameric ligand-gated ion channel gating motions. *PLoS Biol* **11**, e1001714.
290. Kim, D. E., Dimaio, F., Yu-Ruei Wang, R., Song, Y., and Baker, D. (2014) One contact for every twelve residues allows robust and accurate topology-level protein structure modeling. *Proteins* **82 Suppl 2**, 208-218.
291. Koga, N., Tatsumi-Koga, R., Liu, G., Xiao, R., Acton, T. B., Montelione, G. T., and Baker, D. (2012) Principles for designing ideal protein structures. *Nature* **491**, 222-227.
292. Volkov, A. G., Paula, S., and Deamer, D. W. (1997) Two mechanisms of permeation of small neutral molecules and hydrated ions across phospholipid bilayers. *Bioelectrochemistry and Bioenergetics* **42**, 153-160.
293. Sauguet, L., Shahsavar, A., Poitevin, F., Huon, C., Menny, A., Nemezc, A., Haouz, A., Changeux, J. P., Corringer, P. J., and Delarue, M. (2014) Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 966-971.

294. Bonifacino, J. S., and Traub, L. M. (2003) Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu Rev Biochem* **72**, 395-447.
295. Margeta, M. A., Wang, G. J., and Shen, K. (2009) Clathrin adaptor AP-1 complex excludes multiple postsynaptic receptors from axons in *C. elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1632-1637.
296. Lansdell, S. J., Gee, V. J., Harkness, P. C., Doward, A. I., Baker, E. R., Gibb, A. J., and Millar, N. S. (2005) RIC-3 enhances functional expression of multiple nicotinic acetylcholine receptor subtypes in mammalian cells. *Mol Pharmacol* **68**, 1431-1438.
297. Halevi, S., McKay, J., Palfreyman, M., Yassin, L., Eshel, M., Jorgensen, E., and Treinin, M. (2002) The *C. elegans* ric-3 gene is required for maturation of nicotinic acetylcholine receptors. *EMBO J* **21**, 1012-1020.
298. Wang, Y., Yao, Y., Tang, X. Q., and Wang, Z. Z. (2009) Mouse RIC-3, an endoplasmic reticulum chaperone, promotes assembly of the alpha7 acetylcholine receptor through a cytoplasmic coiled-coil domain. *J Neurosci* **29**, 12625-12635.
299. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins. *Science* **286**, 1888-1893.
300. Vogler, O., Barcelo, J. M., Ribas, C., and Escriba, P. V. (2008) Membrane interactions of G proteins and other related proteins. *Biochim Biophys Acta* **1778**, 1640-1652.
301. Kukhtina, V., Kottwitz, D., Strauss, H., Heise, B., Chebotareva, N., Tsetlin, V., and Hucho, F. (2006) Intracellular domain of nicotinic acetylcholine receptor: the importance of being unfolded. *J Neurochem* **97 Suppl 1**, 63-67.
302. Sale, K., Song, L., Liu, Y. S., Perozo, E., and Fajer, P. (2005) Explicit treatment of spin labels in modeling of distance constraints from dipolar EPR and DEER. *J Am Chem Soc* **127**, 9334-9335.
303. Jeschke, G. (2018) The contribution of modern EPR to structural biology. *Emerging Topics in Life Sciences* **2**, 9-18.
304. Mnatsakanyan, N., Nishtala, S. N., Pandhare, A., Fiori, M. C., Goyal, R., Pauwels, J. E., Navetta, A. F., Ahrorov, A., and Jansen, M. (2015) Functional Chimeras of GLIC Obtained by Adding the Intracellular Domain of Anion- and Cation-Conducting Cys-Loop Receptors. *Biochemistry* **54**, 2670-2682.
305. Goyal, R., Salahudeen, A. A., and Jansen, M. (2011) Engineering a prokaryotic Cys-loop receptor with a third functional domain. *J Biol Chem* **286**, 34635-34642.