Non-Canonical Amino Acids As Biochemical Probes of Ligand-Gated Ion Channel Structure and Function

Thesis by

Matthew Rienzo

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2017

(Defended August 17, 2016)

For Mom, Dad, and Mia

Acknowledgments

I first need to thank my adviser, Prof. Dennis Dougherty. I joined Dennis's group knowing that I would be challenged by the independence he gives his students. Graduate school at Caltech can be a grueling experience, but the group dynamic Dennis fosters has helped me build confidence in both research and communication beyond what I thought myself capable of. Dennis genuinely leads by example, and I have deep respect for his wisdom, compassion, and professionalism.

I would also like to recognize my thesis committee, Prof. Shu-ou Shan, Prof. Robert Grubbs, and Prof. David Tirrell. Despite their busy schedules, these three have always been generous with their time and attention. I have enjoyed talking with them about my research and ideas throughout my time here, and I appreciate all of their input. Prof. Henry Lester serves as de facto co-advisor for all of Dennis's students. Henry's enthusiasm for science is infectious, and his extraordinary knowledge of molecular neurobiology has been indispensible for our group.

We are also fortunate to have an ongoing collaboration with the incomparable Prof. Sarah Lummis of Cambridge University. I thank Sarah for giving me the chance to work with her on the GLIC projects, leading to both of my publications during graduate school.

I owe a great deal to all of the members of the Dougherty group, who have largely shaped my day-to-day experience as a graduate student. It seems that our group attracts mainly brilliant, collaborative, and fun people, and much of what I know about chemical biology has come out of our interactions. I briefly overlapped with Dr. Angela Blum, Dr. Kay Limapichat, Dr. Nyssa Puskar, and Erin Lamb at the start of my time at Caltech. Angela attained legendary status in the group, and I have to thank her for inspiring my ASICs projects - she designed several of the experiments described in Chapter 4 for her proposals

exam. Although diminutive in stature, Kay Limapichat is a force to be reckoned with. I thank her for trying to warn me of the perils of making new amino acids with fluorinated carboxylic acid side chains, even though I didn't listen. Nyssa was a friendly and encouraging presence in the group, and I thank her for initiating my first and last rolling chair race tournament. Though I worked only briefly with Erin Lamb, it was long enough to appreciate the strange and colorful sense of humor she brought to the lab.

Dr. Ethan Van Arnam, Dr. Kristina Daeffler, Dr. Noah Duffy, and Dr. Tim Miles were, in retrospect, incredibly patient with me as an eager and needy first-year. I feel extremely lucky to have had their friendship over the last five years. Ethan and I shared neighboring desks for several years, during which I learned things like how not to flood the Opus, all the kinds of canned fish you can eat on toast, and how long toes can be. Ethan is a very talented scientist and an all-around great person, and his good-natured wit always made it fun to be in lab. Kristina has a true talent for making cynicism an endearing trait. Her laid-back attitude belies her remarkable productivity and sharply analytical mind. We have shared many delightful yoga sessions and iPhone game addictions, and I have missed her company tremendously. Fortunately, Noah has stayed in town and in touch since his graduation. Noah is a hilarious and big-hearted fellow who never ceases to surprise me with his interests and talents. He is the first among us to have entered the world of industry, and his advice and support helped make my job search and interview process quick and (nearly) painless. I am sure we will stay in touch, and I hope he decides to move to the Bay Area. Tim's charms include his dazzling vocabulary, minimalistic but joyous lifestyle, and ridiculous humility. I definitely look up to Tim as an independent and creative thinker, and I am certain great things lie ahead for him.

Dr. Ximena da Silva and I co-taught several courses in my early years. Ximena is a kind soul and conscientious lab-mate, and has an intimidating reputation in the online poker community.

Dr. Chris Marotta has dedicated countless hours to keeping our group running smoothly, and I have rarely encountered anyone exercising the same level of experimental consistency. Chris played an important role in introducing me to science in the group, and has been a big help with all matters regarding electrophysiology and nicotinic receptors.

As the only two students grappling with proton-gated ion channels, Dr. Oliver Shafaat and I have had a number of occasions to commiserate. I have often turned to Oliver for help with concerns about photochemistry, and his expertise on the subject has been invaluable for the biologists and chemists alike.

Dr. Michael Post, Betty Wong, and I entered Caltech in the same year. Mike has been a constant, caring friend to me since we first met. Mike's optimism toward science and the world helped lift me from some of my darkest moods, and he has never failed to celebrate my successes, even when I myself have minimized them. I know that his positive energy and charisma will lead him to success, and I wish him the best in all his endeavors. Betty's shrewd entrepreneurship has brought a unique perspective to the lab. I have appreciated our meaningful conversations about building careers in the real world, and am grateful to her for sharing her connections and insights.

Dr. Clint Regan and Paul Walton, the denizens of the South Bay, have energized the photochemistry branch of the lab with their new ideas, feverish discussions, and intense motivation. Clint was the only permanent inhabitant of the chemistry lab when I arrived in the group, and I am grateful for his company there and in the South Bay during my early days here. I already miss our philosophical discussions and occasional NMR puzzles. Paul is,

simply put, a fascinating creature. I am constantly impressed by his originality and talent in all aspects of life. The hours we've spent together (along with his awesome fiancée, Jenna Bush) getting delirious on absurd hiking and camping adventures have been some of my favorite graduate school memories. I will miss both of them very much, and I hope that we will remain in touch.

It is hard to believe that Matt Davis and Kayla Busby joined our lab four whole years ago. Matt has been keeping the computational chemistry in the group going strong, and his positive attitude and good sense of humor (despite the occasional lousy pun) have always kept things cheery in the North Bay. Matt is a talented musician, and it has been fun to geek out with him over reed instruments. Although Kayla is now at UCSD, I am glad that I have had the opportunity to see her at a number of conferences and group trips. Kayla is a ray of sunshine and a very talented researcher, and I look forward to seeing where her career takes her.

Since Bryce Jarman has joined the lab, I have admired his innovation, resourcefulness, and determination in research. Bryce has been bold in his exploration of new techniques and systems in our group, and his fearlessness has been a great inspiration for me as well. It is unlikely that I would have begun my work on the estrogen receptor were it not for his experience and helpful discussion, and he has been a terrific friend as well.

Catie Blunt shares my love of skiing and of Boston, and I have enjoyed our collective grumbling about the general state of things. Catie has taken on some very difficult problems in the lab, and it has been great watching her grow as a chemist.

Richard Mosesso almost always makes me laugh, even when I feel guilty about it. I am always impressed with his confidence as an experimentalist, his productivity, and of course, his height. He has been a welcome addition to the group dynamic. Richard's first-

year protégée, Malayney Young, has also been a lot of fun to have around. I look forward to seeing what the future holds in store for her research.

I have had the pleasure of working with two talented young graduate students, Stephanie Threatt and Stephen Grant. Stephanie was kind enough to simultaneously assist me with the grunt-work of the GLIC proline project *and* laugh at my bad jokes, which is a testament to her patience and congeniality. Stephen Grant is a fellow Long Islander, and possesses unbelievable drive and natural athletic talent; I am almost certain that he can run up the Sam Merrill trail faster than I can run down it. I have had a lot of fun training him in the art of crosslinking, and I know that with him, the project will be in good hands.

Annet Blom has served double as my awesome labmate and awesome roommate. Annet has been one of my closest friends at Caltech, and I am slowly becoming convinced that she is superhuman. Annet is efficient and organized in every way imaginable, and I have the impression that the student organizations on campus are mostly held together by her doing. This has not stopped her from achieving record numbers of Opus runs per week or getting to bed early enough that she can make it to CrossFit at 6 AM. It has also been great fun getting to know our roommate Matt Smarte. I admire Matt for his strong moral compass and willingness to speak up against injustice without hesitation.

Nick O'Connor, whom I also lived with for several years, is really one-of-a-kind. Nick works the typically insane hours of a synthetic organic chemist and operates on a different sleep cycle than I do. Although we only bumped into each other occasionally while living together, our late-night impromptu living room chats were always entertaining.

The Lester lab has a history of helping our group with biology, and this still holds. Many thanks to Dr. Brandon Henderson for his expertise on nicotinic receptors and assistance with mammalian cell electrophysiology experiments, to Dr. Rell Parker for training

me to use a vibratome, to Dr. Matt Mulcahy for his insight on mass spectrometry of membrane receptors, to Jonathan Wang for being a tireless frog surgeon and an excellent oocyte sorter, and to Purnima Deshpande for keeping the whole operation running.

Thanks also to Matt Griffin in the Hsieh-Wilson lab for his kind assistance with my protein biochemistry experiments. Matt and I have become quite good friends in the past year, and his special brand of comic relief has made the grind of writing a bit more tolerable.

I have made countless friends at Caltech who have contributed to my happiness here in many ways – too many to list here. However, I would like to generally thank all of my cohort for their warmth and for the richness of the community in the chemistry department here.

Instrumentation is crucial at Caltech, and the often thankless job of maintaining all our gadgets falls to an exceptional group of scientists. Dr. Dave VanderVelde is CCE's resident NMR-whisperer, and even after finishing my synthetic work, I am endlessly amused by the dry humor of his emails. Dr. Mona Shahgholi, our mass spectrometry guru, has collaborated with our group on a number of projects over the years and has made our tRNA characterization possible. The Proteome Exploration Laboratory staff (Dr. Sonja Hess, Dr. Annie Moradian, Dr. Mike Sweredoski) have been very accommodating, and keep an impressive arsenal of spectrometers running at peak performance.

A big thank you to Linda Syme, who juggles two busy advisers and all of their students, and with style. Joe Drew in the chemistry stockroom takes care of all CCE's behind-the-scenes facilities needs, and always puts the students first. Grad students are not always the neatest people, and so I am very grateful to Camilo Toribio, who cleans up after all of us and has always been a friendly face around Crellin.

I would like to give special thanks to two of my scientist friends from outside Caltech. I have known Madeleine Jensen for about 20(!) years, and I find it amazing that the two of us have both ended up doing biophysics-related Ph.D. programs. It is a joy to see that we still have so much in common, and I thank her for her insights on single-molecule fluorescence imaging. I have shared so many wonderful times with Hannah Payne since we met in middle school, and I have the highest respect for her as a scientist and as a person. I owe her many thanks for her fantastic public speaking advice, and for her perspective as a neuroscientist on my research.

I cannot imagine having gone through graduate school without the love and companionship of my incredible boyfriend, Marcus Low. Marcus is fun-loving, generous, and brilliant, and living with him for the past couple of years has made Pasadena feel like a true home. In addition to teaching me how to cook world-class cuisine and the undisputed most efficient way to evolve my Pokémon, Marcus has thoughtfully critiqued much of my work and kept me company on lots of late nights in lab. I look forward to many more adventures together in the years to come.

Finally, I want to thank my parents and sister, who have been the ultimate support system. I credit them for teaching me to seek the wonder and beauty in the world, which is, I think, what has brought me to pursue this degree in the first place.

Abstract

This dissertation describes several different chemical-scale studies of proteins involved in cellular signaling. The primary focus of this work is on ligand-gated ion channels, an important family of membrane receptors. In each study, the incorporation of structurally diverse non-canonical amino acids have been used to attain a high level of precision in probing the mechanisms of molecular processes. The first chapter provides an introduction to the nonsense suppression methodology used to genetically encode these probes, and surveys the classes of proteins studied herein.

The second and third chapters are concerned with the mechanism of activation of a prokaryotic receptor, *Gloeobacter violaceus* ligand-gated ion channel. In these experiments, novel histidine derivatives were designed, synthesized, and incorporated to test the functional importance of acid-base titration at several positions in the receptor. Then, a battery of proline analogs were used to identify necessary structural features of several critical proline residues, providing clues to conformational changes that occur during receptor activation.

The fourth chapter discusses studies of a different class of receptors, the Acid-Sensing Ion Channels. Several fluorinated aspartic acid and glutamic acid derivatives were targeted to modulate the acidity of putative proton binding sites. Attempts at preparation of these compounds for incorporation into proteins are detailed. Additional sections describe, efforts to elucidate factors in the binding selectivity of the tarantula venom psalmotoxin and selectivity of cation permeability.

In the final chapter, early efforts at developing crosslinking assays for protein-protein interactions in mammalian cells are outlined. These assays involve the introduction of orthogonal tRNA/synthetase pairs for genetically encoding photoreactive phenylalanine

analogs. The most successful studies involve the ligand-dependent dimerization of the soluble nuclear receptor, estrogen receptor α. However, the ultimate goal of this work is to to probe protein-protein interactions among membrane receptors and other proteins. Progress toward extension of the photocrosslinking assay into membrane receptors is described.

Published Content and Contributions

1. Rienzo, M., Lummis, S.C.R., and Dougherty, D.A. Structural Requirements in the Transmembrane Domain of GLIC Revealed by Incorporation of Noncanonical Histidine Analogs. *Chem Biol* **21**, 1700-1706 (2014). doi: 10.1016/j.chembiol.2014.10.019.

M.R. participated in the conception of the project, performed the experiments, prepared the data, and participated in the writing of the manuscript.

2. Rienzo, M., Rocchi, A.R., Threatt, S.D., Dougherty, D.A., and Lummis, S.C.R. Perturbation of Critical Prolines in *Gloeobacter violaceus* Ligand-Gated Ion Channel (GLIC) Supports Conserved Gating Motions Among Cys-Loop Receptors. *J. Biol. Chem* (2016). **291**, 6272-6280. doi: 10.1074/jbc.M115.694372.

M.R. participated in the conception of the project, performed most of the experiments, prepared the data, and participated in the writing of the manuscript.

Table of Contents

Acknowledgments ..... **v**

Abstract ..... **xiii**

Published Content and Contributions ..... **xv**

Table of Contents ..... **xvi**

Chapter 1: Introduction **1**

1.1 Signal transduction in neuronal cells: the ligand-gated ion channels 1

1.1.1 Pentameric ligand-gated ion channels 2

1.1.2 Acid-sensing ion channels 5

1.2 Intracellular signaling via transcriptional control: the nuclear receptors 6

1.3 Tools for mechanistic analysis of signaling proteins 8

1.3.1 Non-canonical amino acid mutagenesis 8

1.3.2 Electrophysiology 10

1.3.3 Probing protein-protein interaction interfaces 14

1.4 Summary of dissertation work 16

1.5 References 17

Chapter 2: Probing Structural Requirements in the Transmembrane Domain of GLIC **23**

2.1 Abstract 23

2.2 Introduction 23

2.3 Results and discussion 26

2.3.1 His234 is the only His residue critical for GLIC function 26

2.3.2 Probing the His234-Ile258 hydrogen bond 27

2.3.3 Synthesis and incorporation of non-canonical histidine analogs 30

2.4 Conclusions and recent developments 35

2.5 Experimental procedures 37

2.5.1 Molecular biology 37

2.5.2 Oocyte preparation and RNA injection 38

2.5.3 Electrophysiology 39

2.5.4 Immunofluorescence Experiments 40

2.5.5 Chemical synthesis and characterization 40

2.6 References 47

Chapter 3: Interrogation of Gating Motions of GLIC via Non-Canonical Amino Acid Mutagenesis at Sensitive Prolines **49**

3.1 Abstract 49

3.2 Introduction 49

3.3 Results 51

3.3.1 Conventional mutagenesis identifies four sensitive proline sites 51

3.3.2 The Cys-loop Tyr/Phe-Pro motif shows sensitivity to Pro cis bias and hydrophobicity 53

3.3.3 A helical distortion in M1 is involved in activation 57

3.3.4 The M4 Pro residue....60
3.4 Discussion....62
3.5 Conclusions....65
3.6 Experimental procedures....66
3.6.1 Molecular biology....66
3.6.2 Oocyte preparation and RNA injection....66
3.6.3 Electrophysiology....67
3.7 References....68

Chapter 4: Tools for Exploring Acid-Sensing Ion Channel Function....71
4.1 Abstract....71
4.2 Introduction....71
4.3 Results and discussion....75
4.3.1 Exploring the origins of sodium selectivity*....75
4.3.2 A noncovalent interaction with Phe351 influences psalmotoxin selectivity for ASIC1*....79
4.3.3 Efforts toward probing titratable residues with fluorinated glutamate and aspartate analogs....82
4.4 Conclusions....86
4.5 Experimental procedures....87
4.5.1 Molecular biology....87
4.5.2 Oocyte preparation and RNA injection....88
4.5.3 Electrophysiology....89
4.5.4 Chemical synthesis and characterization....89
4.6 References....95

Chapter 5: Progress Toward Mapping Protein-Protein Interfaces With Photocrosslinking Non-Canonical Amino Acids....99
5.1 Abstract....99
5.2 Introduction....99
5.2.1 Covalent crosslinking with non-canonical amino acids....99
5.2.2 Estrogen receptor α....102
5.2.3 Pentameric ligand-gated ion channels (pLGICs)....104
5.3 Results and discussion....106
5.3.1 Exploring the ERα ligand-binding domain dimer interface....106
5.3.2 Progress toward intersubunit crosslinking in the human 5-HT_3A receptor....112
5.3.3 Nonsense suppression in nAChRs....117
5.4 Conclusions....120
5.5 Experimental procedures....121
5.5.1 Molecular biology....121
5.5.2 Mammalian cell culture and protein expression....121
5.5.3 Photocrosslinking....123
5.5.4 Lysis, immunoprecipitation, SDS-PAGE, and immunoblotting....123
5.5.6 Membrane potential dye assay....125
5.5.7 Proteolysis and HPLC-MS analysis....126
5.6 References....126

Chapter 1: Introduction

1.1 Signal transduction in neuronal cells: the ligand-gated ion channels

Communication among cells in the nervous systems of higher organisms is conducted by electrochemical impulses called action potentials (**Figure 1.1**).[1] These signals propagate down the length of a neuron, and then stimulate or inhibit other action potentials in thousands of neighboring neurons at regions of cellular contact, called synapses. This process, which gives rise to sensory perception, muscle control, and conscious thought, is coordinated by the release of small-molecule neurotransmitters at the synapse. Each neurotransmitter activates particular membrane-bound proteins on the post-synaptic cell, triggering a corresponding response. A major aim of our work as a molecular neuroscience lab has been to illuminate the chemical-scale factors that influence neurotransmitter binding and receptor activation in the context of neuronal signaling. This dissertation details the development and implementation of several novel tools for this purpose.

Among the proteins involved in synaptic transmission are ion channels – membrane-spanning proteins that mediate the spontaneous flow of ions into or out of the cell. Ion channels may be constitutively active,[2] or may be opened by a small-molecule ligand,[3,4] a mechanical stimulus,[5] or a change in the local membrane electrochemical potential.[6]

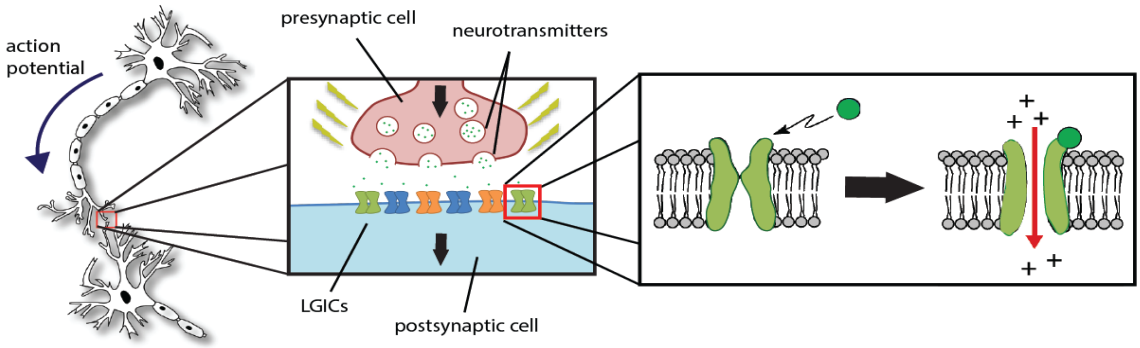


Figure 1.1. Synaptic transmission and ligand-gated ion channels. pLGICs mediate the relay of an action potential from one neuron to the next by binding to small-molecule neurotransmitters released during an action potential.

The studies in this work are mainly concerned with ligand-gated ion channels (LGICs). In most cases, a neurotransmitter binds to an LGIC at a pocket in the extracellular domain of the protein. The noncovalent interactions made by the ligand at this site then enable the channel to undergo a transition into an ion-conducting state. The resulting current produces a depolarization or hyperpolarization of the membrane, which respectively favors or disfavors initiation of an action potential.

The LGICs are phylogenetically diverse and can be categorized by their structures, ligand specificity, and other functional characteristics. Here we focus on two classes of LGICs – the pentameric ligand-gated ion channels and the acid-sensing ion channels.

1.1.1 Pentameric ligand-gated ion channels

The pentameric ligand-gated ion channels (pLGICs) are present in the nervous systems of bilaterally symmetric animals, as well as some prokaryotes.[3,4] In animals, pLGICs include the excitatory, cation-selective nicotinic acetylcholine receptors (nAChRs)[7] and type 3 serotonin receptors (5-HT_3Rs),[8] and the inhibitory, anion-selective glycine receptors (GlyRs),[9] type A γ-aminobutyric acid receptors ($GABA_A$Rs),[10] and glutamate-gated chloride channel (GluCl).[11] nAChRs are well known for mediating signaling at the neuromuscular junction, as well as for their key roles in nicotine addiction, Parkinson's disease, and other neurodegenerative and psychological disorders.[12] Thus, they have been important targets for medicinal chemists and a dominant subject of research in our lab.

Efforts to control the activity of pLGICs have been heavily guided in recent years by emerging structural data (**Figure 1.2**). Until 2001, knowledge of the ligand-binding sites was based upon mutagenesis and biochemistry experiments.[13] At this point, the first X-ray diffraction data were obtained for acetylcholine binding proteins, soluble proteins with high

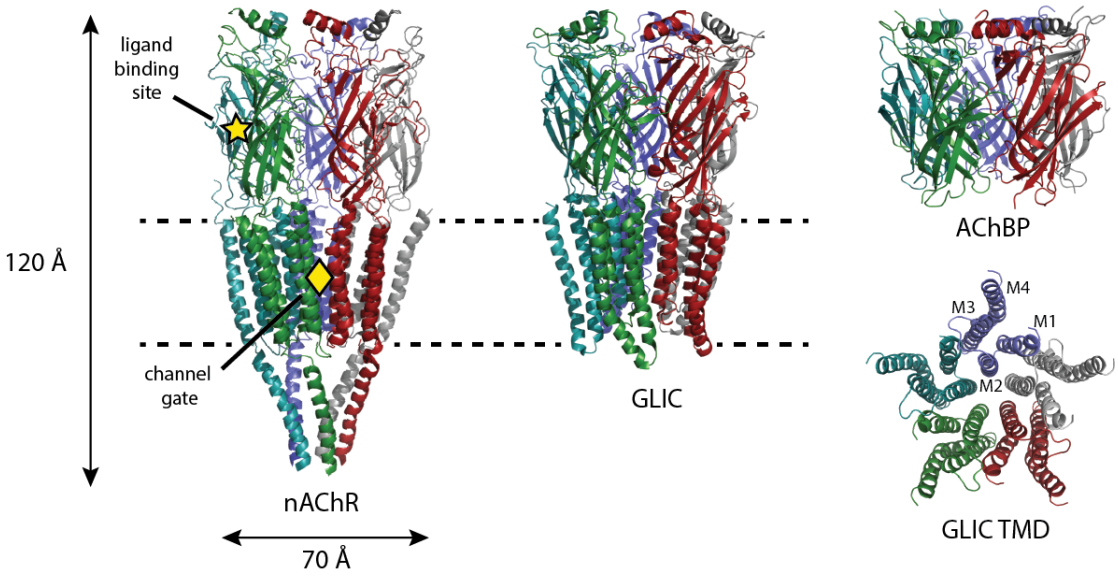


Figure 1.2. Structures of the pentameric ligand-gated ion channels. (Left) Cryo-EM structure of the nicotinic acetylcholine receptor from the *Torpedo marmorata* electric organ (PDB 2BG9). (Center) The prokaryotic proton-activated ion channel from *Gloeobacter violaceus* (PDB 3EHZ) has been most extensively structurally characterized. (Top right) The soluble acetylcholine binding protein from *Lymnaea stagnalis* (PDB 1UX2) has provided a convenient model for the ligand-binding domain. (Bottom right) Cross-section of the transmembrane domain from GLIC (PDB 3EHZ).

structural homology to the ligand-binding domain of the nAChRs.[14,15] Advances in techniques for the purification of oligomeric membrane proteins and in cryo-electron microscopy have since increased the feasibility of obtaining high-resolution structures of full-length eukaryotic receptors, yielding data for the muscle-type torpedo nAChR,[16] *Caenorhabditis elegans* GluCl,[11] mouse 5-HT_{3A}R,[17] homomeric human β3 $GABA_A$R,[18] and human α3 GlyR.[19]

The pLGICs share a common global architecture, comprising five identical or homologous subunits arranged in a barrel configuration around a C_5-axis of pseudosymmetry. Each subunit bears a large N-terminal extracellular domain primarily consisting of β-strands, a membrane-spanning domain made up of four α-helices (M1-M4), and a highly variable intracellular M3-M4 loop. In nearly all pLGICs, the ligand-binding site is at a conserved location at the interface of subunits in the extracellular domain.[20] Extensive

mutagenesis experiments in many pLGICs have allowed us to rationalize trends in ligand potency and selectivity based on noncovalent binding interactions.[21]

A more challenging goal has been to obtain a chemical-scale model of the protein movements that occur during pLGIC activation. The kinetic and thermodynamic parameters of these transitions are critical to the tuning of receptor activity by allosteric modulators and interactions with other proteins. Understanding the energetic landscape of receptor activation will be key in the design of modulating ligands and treatment of disease states. However, global changes in protein shape are often more difficult to probe using site-directed mutagenesis than local noncovalent interactions.

In 2004, the discovery of prokaryotic pLGICs that are more conducive to crystallization enabled the resolution of receptors under a number of different conditions, yielding structures that are thought to represent several mechanistically relevant states.[22-26] These channels, *Gloeobacter* ligand-gated ion channel (GLIC) and *Erwinia* ligand-gated ion channel (ELIC) share the same global architecture as the eukaryotic receptors, but lack the large intracellular M3-M4 loop. GLIC has garnered an especially large amount of attention, with over 50 PDB structures available. This wealth of information provides a convenient testing ground to reconcile functional models obtained using structural and electrophysiological approaches. Studies of GLIC therefore constitute a significant portion of the work described herein.

Although most pharmacological characterization of pLGICs is performed on individual receptors expressed in heterologous model systems, interactions among pLGICs and other endogenous proteins are well documented. These proteins include other LGICs, such as the purinergic P2X receptors,[27-41] cytoplasmic G proteins,[42-44] and extracellular proteins, such as the three-finger snake neurotoxins and the endogenous nAChR modulator

Lynx1.[45] We are interested in developing temporally and spatially controlled assays for the investigation of such complexes, and our progress toward this aim will be discussed.

1.1.2 Acid-sensing ion channels

Reductions in extracellular pH in the nervous system can occur under a number of conditions including neurotransmission by acidic vesicles, inflammation, or restriction of blood supply (ischemia).[46] Application of low-pH stimulus to some neurons is known to produce action potentials, and in many of these cases, pH sensitivity has been linked to a class of receptors known as the acid-sensing ion channels (ASICs).[47] Studies of ASICs have revealed connections with mechanisms of nociception and mechanosensation in sensory neurons.[48,49] Mouse behavioral models for fear, anxiety, and learning have pointed to possible CNS-related roles for ASICs in synaptic plasticity and in susceptibility to psychiatric disorders.[46,50,51]

Several X-ray crystal structures of members of the ASIC family are available, depicting the receptor in what have been interpreted as three different functional states (**Figure 1.3**).[52-56] These structures show a pseudosymmetric trimer, with each subunit having cytoplasmic N- and C-termini, two membrane-spanning α-helices, and a large extracellular domain. It is challenging to identify the site(s) of ligand binding from structural data alone, as ASICs are activated by protons rather than small-molecule ligands. However, along with conventional mutagenesis experiments, the structures have highlighted a region of the extracellular domain rich in carboxylic acid residues, referred to as the "acid pocket", which has been proposed to be titrated by decreasing extracellular pH.[52] These side chains also appear to interact with cationic groups on peptide toxin inhibitors, as demonstrated by cocrystallization.[55,56]

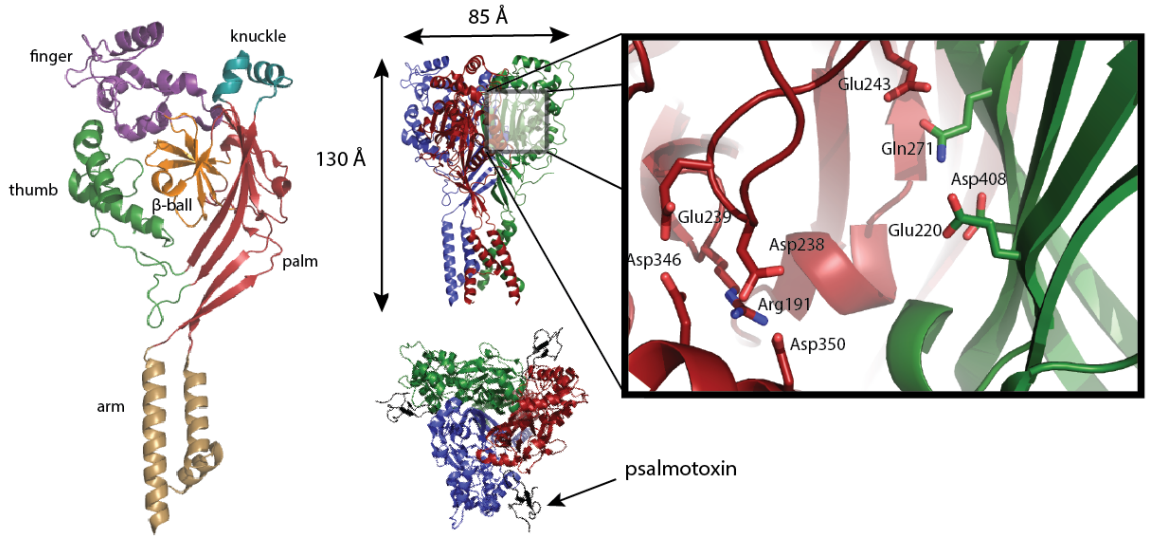


Figure 1.3. The acid-sensing ion channels. (Left) A cartoon representation of the subunit architecture of ASIC1a. (Center) Side and top views of the quaternary structure of ASIC1a homotrimers. The bottom pane shows the binding site for psalmotoxin, a peptide tarantula venom. (Right) Zoomed in view of the "acid pocket", with residues potentially involved in pH sensitivity highlighted.

Our work on ASICs has used unnatural amino acid mutagenesis to approach several different questions about the function of these receptors. These include an investigation of the mechanism of proton detection, identification of determinants for binding and modulation by a peptide toxin, and structural features influencing the relative permeabilities of monovalent cations.

1.2 Intracellular signaling via transcriptional control: the nuclear receptors

Signal transduction via small molecules can also be mediated by soluble protein receptors. Many hormones operate by directly binding to transcription factors called nuclear receptors, which in turn modulate gene expression by associating with DNA and recruiting other transcriptional machinery.[57]

Our group has sought to gain insight into the function of estrogen receptor α (ERα), which is classified as a Type I nuclear receptor (**Figure 1.4A**).[58] ERα resides in the

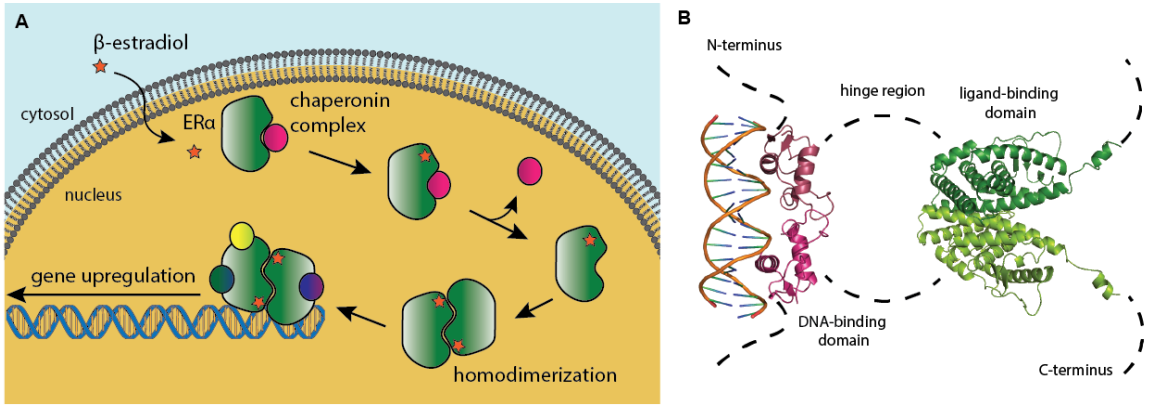


Figure 1.4. Structure and function of estrogen receptor α. (A) Schematic of the canonical mechanism of action for ERα. (B) Cartoon representation of the DNA-binding (PDB 4AA6) and ligand-binding domains of ERα (PDB 1A52). Other domains are disordered and represented by dashed curves.

nucleus in its resting state as a complex with heat shock protein 90 and several other co-chaperones, maintaining a conformation that allows binding of steroid hormones.[59] Upon ligand binding, ERα dissociates from the chaperone complex and forms a homodimer, which can interact with various coactivators or inhibitors. These cofactors influence the ability of ERα to modulate transcription of genes under the control of an estrogen response element (ERE).

Each Type I nuclear receptor is composed of architecturally conserved DNA-binding and ligand-binding domains connected by a disordered linker, along with unstructured N- and C-terminal regions. Crystal structures of the DNA-binding and ligand-binding domains have revealed the regions of the protein that are involved in dimerization interfaces (**Figure 1.4B**).[60,61] They have also provided evidence for conformational changes induced by agonist or antagonist binding, which are thought to confer the differential effects of the ligands on gene expression.[62-64] Validation of these mechanisms has mostly been driven by assays confirming the general colocalization of proteins (e.g. FRET experiments),[65,66] or by indirect assays that depend on downstream activity (e.g. production of reporter proteins).[67] We are interested in developing higher-resolution methods for probing the specific structural transitions that are relevant in the cellular environment during

activation. The approach we have selected will be discussed in detail in section 1.3.3 and Chapter 5.

1.3 Tools for mechanistic analysis of signaling proteins

1.3.1 Non-canonical amino acid mutagenesis

Site-directed mutagenesis has been an invaluable technique for biochemists, allowing for the identification of specific chemical functionalities that confer function to a protein. However, the level of precision afforded by this tool is limited by the selection of amino acids encoded by the standard genetic code. Expansion of the genetic code to enable incorporation of non-canonical amino acids into proteins in living cells has enabled our group and others to approach mechanistic quandaries about proteins from the perspective of an organic chemist.[68,69] Introduction of subtle structural modifications has unambiguously identified critical noncovalent interactions and conformational requirements;[70,71] biochemical affinity handles or fluorescent dyes have been installed for isolation or imaging purposes;[72,73] and side chains with desired chemical reactivity, such as strain-promoted cycloadditions or photochemical radical crosslinking, have frequently been incorporated.[74]

Site-specific non-canonical amino acid mutagenesis is accomplished via nonsense suppression – the interruption of a ribosomal response to a nonsense mutation (i.e. a stop codon or frameshift mutation intervening in a gene).[68,75,76] Nonsense suppression depends upon the presence of an engineered "suppressor tRNA" that recognizes the chosen nonsense mutation during translation and incorporates a non-canonical amino acid at the site of the mutation (**Figure 1.5A**). It is critical that the suppressor tRNA be orthogonal to the host system to prevent charging by endogenous aminoacyl tRNA synthetases

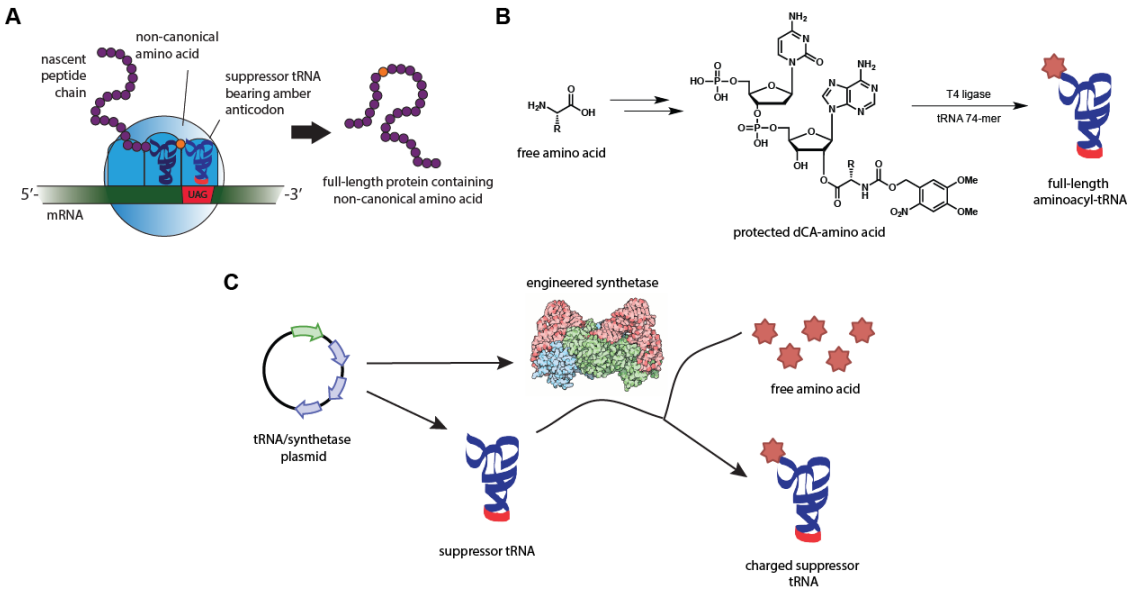


Figure 1.5. Non-canonical amino acid mutagenesis via nonsense suppression. (A) An engineered suppressor tRNA can insert nonproteinogenic amino acid analogues at the site of a nonsense mutation, allowing for expansion of the genetic code. (B) Suppressor tRNAs can be prepared in advance by chemically appending the amino acid of interest to a dCA dinucleotide, which is then enzymatically ligated to the remainder of the tRNA. (C) Aminoacyl suppressor tRNA can be prepared in situ by providing the expression system with an engineered orthogonal tRNA synthetase.

and misincorporation of canonical amino acids at the desired position.

Two methods are generally used to generate charged suppressor tRNA. In the first, the non-canonical amino acid is chemically appended to the suppressor tRNA (**Figure 1.5B**). To ensure that the amino acid is attached at the proper site, it is reacted with a dCA dinucleotide mimicking the final two bases in every tRNA. This aminoacyl dCA is then enzymatically ligated with the remainder of the tRNA and introduced into the system. The chemical acylation method provides a great deal of flexibility in the functionalities that can be incorporated, with the drawback that the yield of protein depends stoichiometrically upon the quantity of charged tRNA supplied.

In the second method, the suppressor tRNA is introduced together with an engineered aminoacyl tRNA synthetase that selectively recognizes the suppressor tRNA and charges it preferentially with the desired non-canonical amino acid (**Figure 1.5C**). Again, it is important that both the tRNA and the synthetase are orthogonal to the endogenous tRNAs

and synthetases, to prevent the frequent misincorporation of amino acids. Such a tRNA-synthetase pair often requires a great deal of optimization through rational mutagenesis and/or directed evolution. This system is therefore convenient for encoding a limited number of amino acids previously optimized by other research groups. However, the regeneration of charged tRNA enables the production of much larger quantities of desired protein product, constituting a distinct advantage over the chemical acylation method. In this dissertation, both methods have been used at various points for different applications.

1.3.2 Electrophysiology

Since the late 1970s, molecular neuroscience has relied heavily upon electrophysiology as a tool for the study of ion channels.[77] This family of techniques evaluates the properties of proteins, neurons, and neural networks by observing electrical currents and potentials across lipid bilayers. Electrophysiology can be used to investigate endogenous activity in living animals, channels expressed in heterologous systems, or channels reconstituted into artificial membrane bilayers. As our research interests are centered on the relationship between chemical structure and protein function, the work described here examines the function of LGICs through heterologous expression in two systems: oocytes from *Xenopus laevis* (the African clawed frog) and cultured human cells.

X. laevis oocytes are a classic model system for pharmacological characterization of ion channels, due to their economy, convenient size, and relatively robust expression of eukaryotic proteins.[78] Protein expression is carried out by microinjecting the appropriate mRNA or plasmid cDNA into the cells, followed by an incubation period during which the cells perform translation, assembly, and surface trafficking of receptors (**Figure 1.6**). This system is amenable to the nonsense suppression methodology, as chemically acylated tRNA

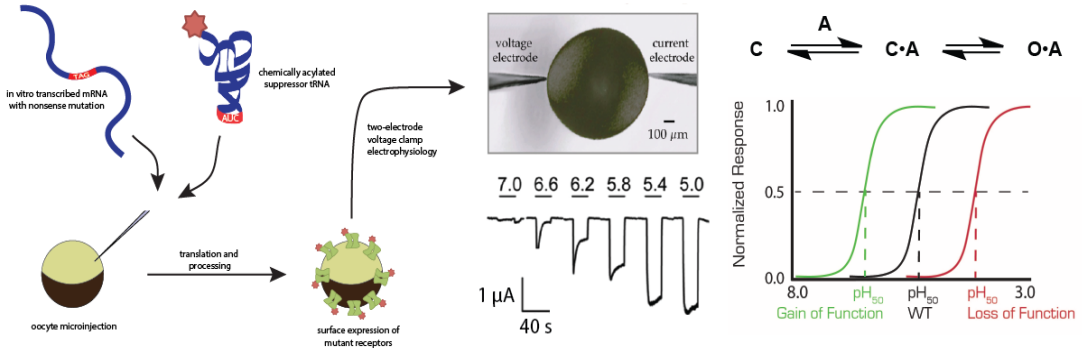


Figure 1.6. Nonsense suppression in *Xenopus laevis* oocytes. Ligand-gated ion channels containing non-canonical amino acids are expressed by coinjecting oocytes with charged suppressor tRNA and mRNA with a nonsense mutation at the site of interest. After expression, mutants are assayed for function using a dose-response electrophysiology assay in two-electrode voltage clamp mode. C = closed, O = open, A = agonist.

can be easily coinjected along with the gene containing the nonsense mutation.

In mammalian cells, LGIC expression is accomplished using standard transfection vectors and reagents. An additional plasmid can be cotransfected to allow transcription of a suppressor tRNA and expression of an engineered synthetase for the purpose of nonsense suppression. Following transfection, the cells are incubated in a solution containing the desired non-canonical amino acid, which then permeates the plasma membrane and is recognized by the engineered synthetase.

Different aspects of LGIC function may be illuminated using various electrophysiological assays. In the projects discussed here, the primary goals of our experiments in *X. laevis* oocytes were to verify functionality of mutants and reveal any effects that the mutations induce upon the receptors' conformational equilibria. A simplified model of the activation process is shown in **Figure 1.6**: the cooperative binding of one or more agonist molecules allows the receptor to access an ion-permeable state, a process referred to as gating. Mutations that influence the energetics of agonist binding or gating are expected to shift these equilibria, resulting in an increase or decrease in the sensitivity of the LGIC to its agonist.[79]

This effect can be quantified by measuring the current passing through the population of receptors occupying the open state. Such a whole-cell current measurement can be quite easily obtained using the two-electrode voltage clamp (TEVC) configuration. In this setup, two glass electrodes impale the cell. The first measures the internal potential of the cell relative to an external reference electrode in the solution bathing the oocyte, while the second electrode responds to fluctuations in this voltage by injecting an appropriate current to maintain a "clamped" membrane potential difference. Monitoring this injected current provides an indication of the charge passing through channels in the open state, a quantity proportional to the number of activated LGICs. Our group is fortunate to have an eight-channel recording rig with simultaneous voltage-clamping, buffer perfusion, and liquid handling abilities, dramatically increasing the throughput of our TEVC experiments in oocytes.

Typically, a dose-response experiment is performed in which an oocyte expressing the LGIC is treated with various concentrations of agonist. The resulting whole-cell currents are plotted, normalizing to the maximal current observed (**Figure 1.6**). These data are fit to the Hill equation, which provides a simplified model for cooperative ligand binding:

$$I([A]) = \frac{I_{\max}}{1+\left(\frac{[A]}{\mathrm{EC_{50}}}\right)^{n_{\mathrm{H}}}}, \qquad \text{(Eq. 1.1)}$$

where I is the whole cell current, [A] is the agonist concentration, I_{max} is the current observed at saturating concentrations of agonist, EC_{50} is the agonist concentration eliciting a half-maximal response, and n_H is the Hill coefficient. The EC_{50} is then used as a readout for the magnitude of the energetic perturbation caused by the mutation, where log(fold-change in EC_{50}) varies roughly linearly with the $\Delta\Delta G$ value.

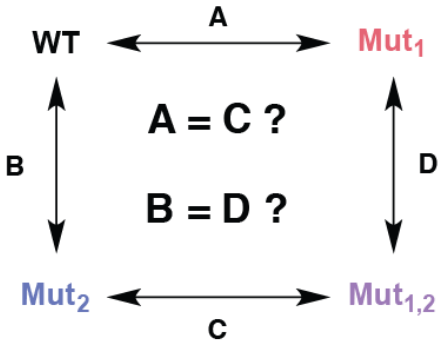


Figure 1.7. Double-mutant cycle analysis. If the functional perturbation caused by mutation 1 is identical in the absence or presence of mutation 2, the effects are additive and therefore independent. Nonadditivity indicates functional coupling of the mutations.

In some cases, we wish to determine whether a particular perturbation originates from disruption of a specific interaction. We often proceed by performing a double mutant cycle analysis, in which the effects of two mutations expected to disrupt the interaction are observed alone and in combination (**Figure 1.7**). If the values for log(fold-change in EC_{50}) from the two mutations are additive, it appears that the effects of the modifications are independent of one another. If the values are significantly nonadditive, then the mutated positions are interpreted to interact in either a cooperative or a redundant fashion, depending on the magnitude of the coupling parameter:

$$\Omega = \frac{EC_{50}(mut_{1,2}) \bullet EC_{50}(WT)}{EC_{50}(mut_1) \bullet EC_{50}(mut_2)}. \qquad \text{(Eq. 1.2)}$$

Although similar electrophysiology techniques are commonly used to assay LGICs in mammalian systems, these experiments are much more difficult to automate. We have made use of an alternative technique that indirectly records the membrane potential from a monolayer of adherent cells. In this, the fluorescent imaging plate reader (FLIPR) assay, a commercial membrane potential dye is loaded into the cells following expression of the LGICs of interest.[80] Quenching of the dye is reduced by membrane depolarization, due to reorientation of the dye molecules in the lipid bilayer. Thus, the activation of ion channels

can be visualized by changes in fluorescence of the dye in response to the application of various concentrations of ligand. Compared with the TEVC, the lack of a voltage clamp in the FLIPR assay makes it somewhat challenging to quantitatively interpret the influence of mutations on channel activation equilibria. Nevertheless, it is a simple and reliable tool for validating the functionality of mutants, monitoring relative expression levels, or comparing agonist sensitivities.

1.3.3 Probing protein-protein interaction interfaces

Proper operation of most proteins is contingent upon interactions with other biomacromolecules. Our understanding of protein-protein interactions among soluble proteins is dramatically advanced by the ability to cocrystallize them and obtain X-ray diffraction data. However, the acquisition of reliable structures of membrane protein complexes is hampered by the dependence of their conformation on their lipid environment, especially in cases where associations are weak and transient. Because LGICs are poised to transition among several conformations upon association of a small molecule, it is to be expected that crystal structures may not be capable of revealing all physiologically important details.

Chemical crosslinking offers a unique solution to this problem. An unobtrusive crosslinking group allows proteins to associate normally, and then irreversibly links them at a site of contact, preserving information about the interaction of interest even after extraction out of the cellular environment (**Figure 1.8**). An ideal experimental setup should feature both temporal and spatial resolution. Temporal resolution requires a trigger for the crosslinking reaction, most conveniently supplied by an external UV light source to initiate decaging of a reactive group or generation of a reactive radical intermediate. Different extents of spatial resolution can be accomplished in several ways. Chemoselectivity is

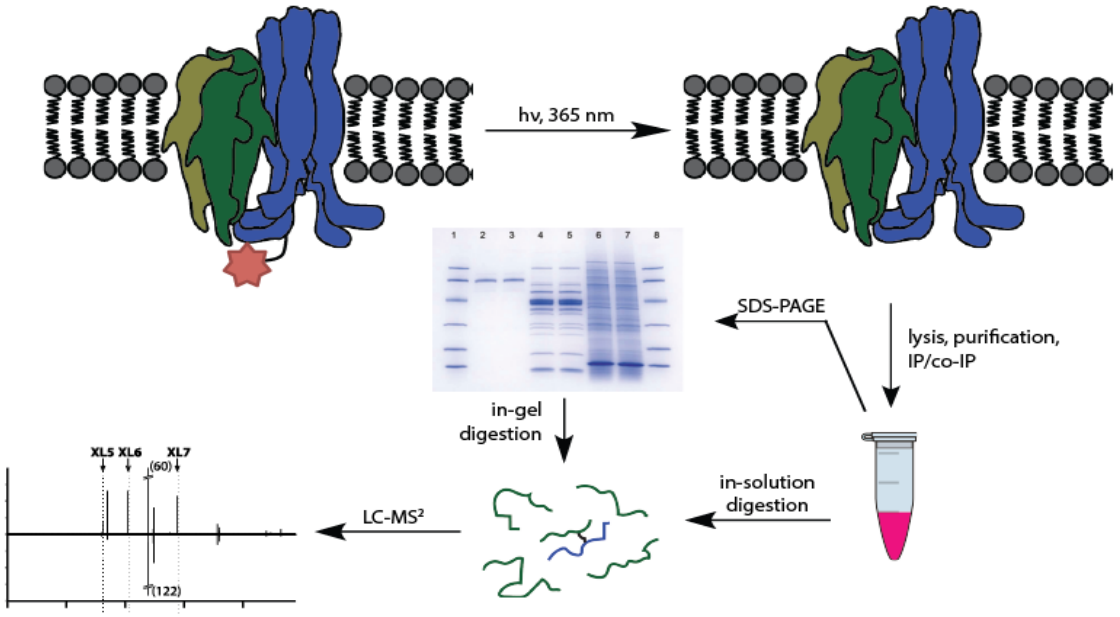


Figure 1.8. Probing protein-protein interfaces with site-specific crosslinking. Photoreactive crosslinking amino acids are incorporated at positions of interest using nonsense suppression. After crosslinking proteins by UV irradiation, cells are homogenized, and lysates are processed. Photoadducts are analyzed by SDS-PAGE and/or proteolysis followed by tandem mass spectrometry.

exploited in many cases to reduce the complexity of crosslinked product mixtures – for example, an accessible reactive cysteine or lysine residue can be introduced at a position of interest and used in conjunction with an amine- or thiol-reactive bifunctional crosslinker. However, the most precise method of spatial resolution uses the site-specific installation of a crosslinker.

Other groups have previously established a convenient procedure for doing this using nonsense suppression with engineered tRNA/synthetase pairs.[81,82] In our crosslinking experiments, we have made use of two such pairs optimized for 4-benzoylphenylalanine and 4-azidophenylalanine. These are UV-activated zero-length crosslinking groups with somewhat differing modes of reaction and chemoselectivity, which will be described in detail in a later chapter. Due to the minor steric perturbation these probes impose relative to phenylalanine or tyrosine, we have selected them as the optimal choice for our investigation of protein-protein interaction interfaces.

Following incorporation of the crosslinking amino acids and irradiation to trigger the photoreaction, a number of possible approaches can be used to extract information from the resulting adducts (**Figure 1.8**). These include various enrichment and fractionation steps, electrophoresis and western blotting for visualization, and mass spectrometric analysis. While mass spectrometry has in several instances been used in conjunction with crosslinking by genetically encoded reagents,[83-85] two out of the three studies were carried out in vitro with purified proteins, and none involved membrane proteins. This is reflective of the challenging nature of the problem, despite the wealth of information that could likely be uncovered through improvement of current methods.

1.4 Summary of dissertation work

In this dissertation, non-canonical amino acid mutagenesis is used in conjunction with other biochemistry techniques to probe the structure and function of proteins with a high level of precision. These studies include work on ligand-gated ion channels and nuclear receptors, using heterologous model systems from *Xenopus laevis* and mammalian cell lines.

Chapters 2 and 3 describe work focused on the prokaryotic proton-gated pentameric ligand-gated ion channel, GLIC. These experiments were conducted by expressing mutant receptors in *Xenopus laevis* oocytes and assaying their responses to changes in pH through electrophysiology. In Chapter 2, the first incorporation of two unnatural histidine analogs via nonsense suppression is presented, and it is demonstrated that this method can be effectively applied to determine the functional importance of histidine side chain basicity in a protein. Chapter 3 surveys the roles of critical proline sites in GLIC through incorporation of a wide array of unnatural amino acid analogs. The observed effects are then discussed in the context of other similar proline experiments from our group and the accumulated mechanistic knowledge for the pLGIC family.

Chapter 4 applies the nonsense suppression methodology to the study of ASIC1a, also in the *Xenopus laevis* oocyte expression system. The attempted syntheses of two fluorinated analogs of aspartic and glutamic acid are discussed, and the intended experimental design for their use in probing the acid pocket is outlined. Backbone ester mutagenesis was used to investigate the importance of carbonyl hydrogen bond-acceptors in the ion selectivity filter region of the receptor, and preliminary results of these experiments are presented. Attempts were also made to modulate an interaction with the tarantula venom psalmotoxin, using a fluorinated phenylalanine residue to probe a possible cation-π interaction at the protein-peptide interface.

Chapter 5 details our group's first efforts at using site-specific photocrosslinking for the characterization of protein-protein binding interfaces. An assay for direct detection of ligand-induced dimerization of estrogen receptor α in living mammalian cells is established with the aid of an engineered orthogonal tRNA/synthetase pair. Finally, progress towards the use of the crosslinking strategy for studying interactions among membrane receptors is described. These experiments include the incorporation of photocrosslinking groups into nAChRs containing α6 and α4 subunits, as well as into the 5-HT_{3A} homopentamer.

1.5 References

1. Kandel, E. R., Schwartz, J. H., Jessell, T. M. & Siegelbaum, S. A. *Principles of Neural Science* Edn. 4 (McGraw-Hill, New York, New York, USA, 2000).
2. Enyedi, P. & Czirják, G. Molecular Background of Leak K^+ Currents: Two-Pore Domain Potassium Channels. *Physiol. Rev.* **90,** 559–605 (2010).
3. Dacosta, C. J. B. & Baenziger, J. E. Gating of pentameric ligand-gated ion channels: structural insights and ambiguities. *Structure* **21,** 1271–1283 (2013).
4. Nys, M., Kesters, D. & Ulens, C. Structural insights into Cys-loop receptor function and ligand recognition. *Biochem. Pharmacol.* **86,** 1042–1053 (2013).
5. Ranade, S. S., Syeda, R. & Patapoutian, A. Mechanically Activated Ion Channels. *Neuron* **87,** 1162–1179 (2015).
6. Catterall, W. A. Voltage-gated sodium channels at 60: structure, function and pathophysiology. *J. Physiol. (Lond.)* **590,** 2577–2589 (2012).
7. Albuquerque, E. X., Pereira, E. F. R., Alkondon, M. & Rogers, S. W. Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiol. Rev.* **89,** 73–120 (2009).

8. Thompson, A. J. & R Lummis, S. C. 5-HT_3 Receptors. *Curr. Pharm. Design* **12,** 3615-3630 (2006).
9. Danysz, W. & Parsons, C. G. Glycine and N-Methyl-D-Aspartate Receptors: Physiological Significance and Possible Therapeutic Applications. *Pharmacol. Rev.* **50,** 597–664 (1998).
10. Olsen, R. W. & Sieghart, W. $GABA_A$ receptors: Subtypes provide diversity of function and pharmacology. *Neuropharmacology* **56,** 141–148 (2009).
11. Althoff, T., Hibbs, R. E., Banerjee, S. & Gouaux, E. X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature* **512,** 333–337 (2014).
12. Kalamida, D. *et al.* Muscle and neuronal nicotinic acetylcholine receptors. *The FEBS Journal* **274,** 3799–3845 (2007).
13. Dougherty, D. A. Cation-π interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **271,** 163–168 (1996).
14. Sixma, T. K. & Smit, A. B. Acetylcholine binding protein (AChBP): a secreted glial protein that provides a high-resolution model for the extracellular domain of pentameric ligand-gated ion channels. *Annu. Rev. Biophys. Biomol. Struct.* **32,** 311–334 (2003).
15. Brejc, K. *et al.* Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411,** 269–276 (2001).
16. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4Å resolution. *J. Mol. Biol.* **346,** 967–989 (2005).
17. Hassaine, G. *et al.* X-ray structure of the mouse serotonin 5-HT_3 receptor. *Nature* **512,** 276–281 (2014).
18. Miller, P. S. & Aricescu, A. R. Crystal structure of a human $GABA_A$ receptor. *Nature* **512,** 270–275 (2014).
19. Huang, X., Chen, H., Michelsen, K., Schneider, S. & Shaffer, P. L. Crystal structure of human glycine receptor-α3 bound to antagonist strychnine. *Nature* **526,** 277–280 (2015).
20. Sauguet, L., Shahsavar, A. & Delarue, M. Crystallographic studies of pharmacological sites in pentameric ligand-gated ion channels. *Biochim. Biophys. Acta* **1850,** 511–523 (2015).
21. Van Arnam, E. B. & Dougherty, D. A. Functional probes of drug-receptor interactions implicated by structural studies: Cys-loop receptors provide a fertile testing ground. *J. Med. Chem.* **57,** 6289–6300 (2014).
22. Tasneem, A., Iyer, L. M., Jakobsson, E. & Aravind, L. Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels. *Genome Biology* **6,** R4 (2004).
23. Hilf, R. J. C. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457,** 115–118 (2009).
24. Hilf, R. J. C. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452,** 375–379 (2008).
25. Sauguet, L. *et al.* Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 966–971 (2014).
26. Prevost, M. S. *et al.* A locally closed conformation of a bacterial pentameric proton-gated ion channel. *Nat. Struct. Mol. Biol.* **19,** 642–649 (2012).
27. Khakh, B. S. An Angstrom Scale Interaction between Plasma Membrane ATP-Gated P2X2 and α4β2 Nicotinic Channels Measured with Fluorescence Resonance Energy Transfer and Total Internal Reflection Fluorescence Microscopy. *J. Neurosci.* **25,** 6911–6920 (2005).
28. Shrivastava, A. N., Triller, A., Sieghart, W. & Sarto-Jackson, I. Regulation of $GABA_A$ Receptor Dynamics by Interaction with Purinergic P2X2 Receptors. *J. Biol. Chem.* **286,** 14455–14468 (2011).
29. Wieskopf, J. S. *et al.* The nicotinic α6 subunit gene determines variability in chronic pain sensitivity via cross-inhibition of P2X2/3 receptors. *Sci. Transl. Med.* **7,** 287ra72 (2015).
30. Toulmé, E. *et al.* An intracellular motif of P2X3 receptors is required for functional cross-talk with $GABA_A$ receptors in nociceptive DRG neurons. **102,** 1357–1368 (2007).
31. Boué-Grabot, E., Toulmé, E., Emerit, M. B. & Garret, M. Subunit-specific coupling between gamma-aminobutyric acid type A and P2X2 receptor channels. *J. Biol. Chem.* **279,** 52517–52525 (2004).
32. Boué-Grabot, E., Emerit, M. B., Toulmé, E., Séguéla, P. & Garret, M. Cross-talk and co-trafficking between ρ1/GABA receptors and ATP-gated channels. *J. Biol. Chem.* **279,** 6967–6975 (2004).
33. Boué-Grabot, E. *et al.* Intracellular cross talk and physical interaction between two classes of neurotransmitter-gated channels. *J. Neurosci.* **23,** 1246-1253 (2003).
34. Nakazawa, K. ATP-activated current and its interaction with acetylcholine-activated current in rat sympathetic neurons. *J. Neurosci.* **14,** 740-750 (1994).

35. Sokolova, E., Nistri, A. & Giniatullin, R. Negative cross talk between anionic $GABA_A$ and cationic P2X ionotropic receptors of rat dorsal root ganglion neurons. *J. Neurosci.* **21,** 4958-4968 (2001).
36. Decker, D. A. & Galligan, J. J. Cross-inhibition between nicotinic acetylcholine receptors and P2X receptors in myenteric neurons and HEK-293 cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **296,** G1267–76 (2009).
37. Decker, D. A. & Galligan, J. J. Molecular mechanisms of cross-inhibition between nicotinic acetylcholine receptors and P2X receptors in myenteric neurons and HEK-293 cells. *Neurogastroenterol. Motil.* **22,** 901–8– e235 (2010).
38. Karanjia, R. *et al.* Cross-inhibitory interactions between $GABA_A$ and P2X channels in myenteric neurones. *Eur. J. Neurosci.* **23,** 3259–3268 (2006).
39. Barajas-López, C., Espinosa-Luna, R. & Zhu, Y. Functional interactions between nicotinic and P2X channels in short-term cultures of guinea-pig submucosal neurons. *J. Physiol. (Lond.)* **513,** 671–683 (1998).
40. Barajas-López, C., Montaño, L. M. & Espinosa-Luna, R. Inhibitory interactions between $5\text{-}HT_3$ and P2X channels in submucosal neurons. *Am. J. Physiol. Gastrointest. Liver Physiol.* **283,** G1238–48 (2002).
41. Zhou, X. & Galligan, J. J. Non-additive interaction between nicotinic cholinergic and P2X purine receptors in guinea-pig enteric neurons in culture. *J. Physiol. (Lond.)* **513,** 685–697 (1998).
42. Yevenes, G. E., Moraga-Cid, G., Peoples, R. W., Schmalzing, G. & Aguayo, L. G. A selective G betagamma-linked intracellular mechanism for modulation of a ligand-gated ion channel by ethanol. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 20523–20528 (2008).
43. Yevenes, G. E. & Zeilhofer, H. U. Allosteric modulation of glycine receptors. *Brit. J. Pharmacol.* **164,** 224-236 (2011).
44. Fischer, H., Liu, D.-M., Lee, A., Harries, J. C. & Adams, D. J. Selective Modulation of Neuronal Nicotinic Acetylcholine Receptor Channel Subunits by Go-Protein Subunits. *J. Neurosci.* **25,** 3571–3577 (2005).
45. Tsetlin, V. I. Three-finger snake neurotoxins and Ly6 proteins targeting nicotinic acetylcholine receptors: pharmacological tools and endogenous modulators. *Trends Pharmacol. Sci.* **36,** 109–123 (2015).
46. Wemmie, J. A., Taugher, R. J. & Kreple, C. J. Acid-sensing ion channels in pain and disease. *Nat. Rev. Neurosci.* **14,** 461–471 (2013).
47. Waldmann, R., Champigny, G., Bassilana, F., Heurteaux, C. & Lazdunski, M. A proton-gated cation channel involved in acid-sensing. *Nature* **386,** 173–177 (1997).
48. Omerbašić, D., Schuhmacher, L.-N., Bernal Sierra, Y.-A., Smith, E. S. J. & Lewin, G. R. ASICs and mammalian mechanoreceptor function. *Neuropharmacology* **94,** 80–86 (2015).
49. Deval, E. & Lingueglia, E. Acid-Sensing Ion Channels and nociception in the peripheral and central nervous systems. *Neuropharmacology* **94,** 49–57 (2015).
50. Sluka, K. A., Winter, O. C. & Wemmie, J. A. Acid-sensing ion channels: A new target for pain and CNS diseases. *Current opinion in drug discovery & development* **12,** 693–1394 (2009).
51. Huang, Y. *et al.* Two aspects of ASIC function: Synaptic plasticity and neuronal injury. *Neuropharmacology* **94,** 42–48 (2015).
52. Jasti, J., Furukawa, H., Gonzales, E. B. & Gouaux, E. Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature* **449,** 316–323 (2007).
53. Gonzales, E. B., Kawate, T. & Gouaux, E. Pore architecture and ion sites in acid-sensing ion channels and P2X receptors. *Nature* **460,** 599–604 (2009).
54. Baconguis, I. & Gouaux, E. Structural plasticity and dynamic selectivity of acid-sensing ion channel–spider toxin complexes. *Nature* **489,** 400–405 (2012).
55. Baconguis, I., Bohlen, C. J., Goehring, A., Julius, D. & Gouaux, E. X-ray structure of acid-sensing ion channel 1-snake toxin complex reveals open state of a Na(+)-selective channel. *Cell* **156,** 717–729 (2014).
56. Dawson, R. J. P. *et al.* Structure of the Acid-sensing ion channel 1 in complex with the gating modifier Psalmotoxin 1. *Nat Comms* **3,** 936 (2012).
57. McKenna, N. J. & O'Malley, B. W. Combinatorial Control of Gene Expression by Nuclear Receptors and Coregulators. *Cell* **108,** 465–474 (2002).
58. Heldring, N. *et al.* Estrogen Receptors: How Do They Signal and What Are Their Targets. *Physiol. Rev.* **87,** 905–931 (2007).
59. Fliss, A. E., Benzeno, S., Rao, J. & Caplan, A. J. Control of estrogen receptor ligand binding by Hsp90. *J. Steroid Biochem.* **72,** 223-230 (2000).
60. Schwabe, J. W., Chapman, L. & Rhodes, D. The oestrogen receptor recognizes an imperfectly

palindromic response element through an alternative side-chain conformation. *Structure* **3,** 201–213 (1995).

61. Tanenbaum, D. M., Wang, Y., Williams, S. P. & Sigler, P. B. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 5998–6003 (1998).
62. Allegretto, E. A. in *Selective Estrogen Receptor Modulators* 19–28 (Humana Press, Totowa, New Jersey, 2002).
63. Gronemeyer, H., Gustafsson, J.-Å. & Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discov.* **3,** 950–964 (2004).
64. Tata, J. R. Signalling through nuclear receptors. *Nat. Rev. Mol. Cell Biol.* **3,** 702–710 (2002).
65. Tamrazi, A., Carlson, K. E., Daniels, J. R., Hurth, K. M. & Katzenellenbogen, J. A. Estrogen Receptor Dimerization: Ligand Binding Regulates Dimer Affinity and Dimer Dissociation Rate. *Mol. Endocrinol.* **16,** 2706–2719 (2013).
66. Gunther, J. R. *et al.* A Set of Time-Resolved Fluorescence Resonance Energy Transfer Assays for the Discovery of Inhibitors of Estrogen Receptor-Coactivator Binding. *J. Biomol. Screen.* **14,** 181–193 (2009).
67. Paech, K. *et al.* Differential Ligand Activation of Estrogen Receptors ERα and ERβ at AP1 Sites. *Science* **277,** 1508–1510 (1997).
68. Dougherty, D. A. & Van Arnam, E. B. In Vivo Incorporation of Non-canonical Amino Acids by Using the Chemical Aminoacylation Strategy: A Broadly Applicable Mechanistic Tool. *Chembiochem* **15,** 1710–1720 (2014).
69. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Ann. Rev. Biochem.* **79,** 413-444 (2010).
70. Lummis, S. C. R. *et al.* Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438,** 248–252 (2005).
71. Xiu, X., Puskar, N. L., Shanata, J. A. P., Lester, H. A. & Dougherty, D. A. Nicotine binding to brain receptors requires a strong cation-pi interaction. *Nature* **458,** 534–537 (2009).
72. Gallivan, J. P., Lester, H. A. & Dougherty, D. A. Site-specific incorporation of biotinylated amino acids to identify surface-exposed residues in integral membrane proteins. *Chem. Biol.* **4,** 739–749 (1997).
73. Pantoja, R., Rodriguez, E. A., Dibas, M. I., Dougherty, D. A. & Lester, H. A. Single-Molecule Imaging of a Fluorescent Unnatural Amino Acid Incorporated Into Nicotinic Receptors. *Biophys. J.* **96,** 226–237 (2009).
74. Kim, C. H., Axup, J. Y. & Schultz, P. G. Protein conjugation with genetically encoded unnatural amino acids. *Curr. Opin. Chem. Biol.* **17,** 412–419 (2013).
75. Nowak, M. W. *et al.* [28] In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Method. Enzymol.* **293,** 504–529 (1998).
76. Ryu, Y. & Schultz, P. G. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Meth.* **3,** 263–265 (2006).
77. Zhao, Y. *et al.* Patch clamp technique: Review of the current state of the art and potential contributions from nanoengineering. *P. I. Mech. Eng. N-J. Nano.* **222,** 1–11 (2008).
78. Papke, R. L. & Smith-Maxwell, C. High throughput electrophysiology with Xenopus oocytes. *Comb. Chem. High Throughput Screen.* **12,** 38–50 (2009).
79. Gleitsman, K. R., Shanata, J. A. P., Frazier, S. J., Lester, H. A. & Dougherty, D. A. Long-Range Coupling in an Allosteric Receptor Revealed by Mutant Cycle Analysis. *Biophys. J.* **96,** 3168–3178 (2009).
80. Schroeder, K. S. & Neagle, B. D. FLIPR: A New Instrument for Accurate, High Throughput Optical Screening. *J. Biomol. Screen.* **1,** 75–80 (1996).
81. Hino, N. *et al.* Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nat. Meth.* **2,** 201–206 (2005).
82. Farrell, I. S., Toroney, R., Hazen, J. L., Mehl, R. A. & Chin, J. W. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat. Meth.* **2,** 377–384 (2005).
83. Schwarz, R. *et al.* Monitoring Conformational Changes in Peroxisome Proliferator-Activated Receptor α by a Genetically Encoded Photoamino Acid, Cross-Linking, and Mass Spectrometry. *J. Med. Chem.* **56,** 4252–4263 (2013).
84. Forné, I., Ludwigsen, J., Imhof, A., Becker, P. B. & Mueller-Planitz, F. Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. *Mol. Cell Proteomics* **11,** M111.012088 (2012).

85. Berg, M., Michalowski, A., Palzer, S., Rupp, S. & Sohn, K. An In Vivo Photo-Cross-Linking Approach Reveals a Homodimerization Domain of Aha1 in S. cerevisiae. *PLoS ONE* **9,** e89436 (2014).

Chapter 2: Probing Structural Requirements in the Transmembrane Domain of GLIC*

2.1 Abstract

The cyanobacterial pentameric ligand-gated ion channel GLIC, a homolog of the Cys-loop receptor superfamily, has provided useful structural and functional information about its eukaryotic counterparts. X-ray diffraction data and site-directed mutagenesis have previously implicated a transmembrane histidine residue (His234) as essential for channel function. Here, we investigate the role of His234 via synthesis and incorporation of histidine analogues and α-hydroxy acids using in vivo nonsense suppression. Receptors were expressed heterologously in *Xenopus laevis* oocytes, and whole-cell voltage-clamp electrophysiology was used to monitor channel activity. We show that an interhelix hydrogen bond involving His234 is important for stabilization of the open state, and that the shape and basicity of its side chain are highly sensitive to perturbations. In contrast, our data show that two other His residues are not involved in the acid-sensing mechanism.

2.2 Introduction

Cys-loop receptors are a class of pentameric ligand-gated ion channels (pLGICs) that mediate synaptic transmission among neurons in the central and peripheral nervous systems of eukaryotes.[1,2] The family has been extensively characterized using electrophysiological and biochemical techniques, and includes the nicotinic acetylcholine (nACh), glycine (Gly), serotonin (5-HT_3), and γ-aminobutyric acid ($GABA_A$) receptors. These receptors are essential for normal neuronal function, and an appreciation of the mechanisms by which

---

Parts of this chapter are adapted with permission from: Rienzo, M., Lummis, S.C.R., and Dougherty, D.A. Structural Requirements in the Transmembrane Domain of GLIC Revealed by Incorporation of Noncanonical Histidine Analogs. *Chem Biol* (2014) **21**, 1700-1706.

they recognize ligands and undergo structural changes associated with channel gating is crucial for a full understanding of intercellular signaling in the nervous system. Moreover, many Cys-loop receptors have been identified as playing a role in various neurodegenerative diseases, establishing them as targets for drug development.[3,4] Structure-function studies of the Cys-loop family have been guided by electron microscopy images of nAChRs,[5] and by crystal structures of a family of soluble proteins, the acetylcholine binding proteins, which display high sequence homology to the extracellular domain of nAChR subunits.[6,7]

Crystallization of full-length Cys-loop receptors has proven to be challenging, and the only members of the family for which a high-resolution X-ray crystallography structure has been reported are the invertebrate, glutamate-gated chloride channel (GluCl),[8,9] the human $GABA_A$ receptor (β3 subunit homopentamer),[10] and recently the mouse 5-HT_3 receptor (A subunit homopentamer).[11] There are also two prokaryotic pentameric channels whose structures have been solved; they were identified in the cyanobacterium *Gloeobacter violaceus* (GLIC) and the bacterium *Erwinia chrysanthemi* (ELIC).[12-15] GLIC is of particular interest for the study of gating transitions, in that its structure has been elucidated in multiple conformational states under a number of conditions.[15-20] Although molecular details for these homologs likely differ from those of eukaryotic Cys-loop receptors (including the fact that the Cys-loop is absent), the global conformational changes that occur upon channel activation should be informative for the entire family. The prokaryotic proteins thus provide an opportunity to study these energetic landscapes in a context where structural data are already available.

Unlike most Cys-loop receptors, activation of GLIC is mediated by titration of ionizable residues at low pH, rather than by binding of a small molecule. Direct observation of the ligand-binding site(s) via crystallography is therefore not possible, leaving mutagenesis

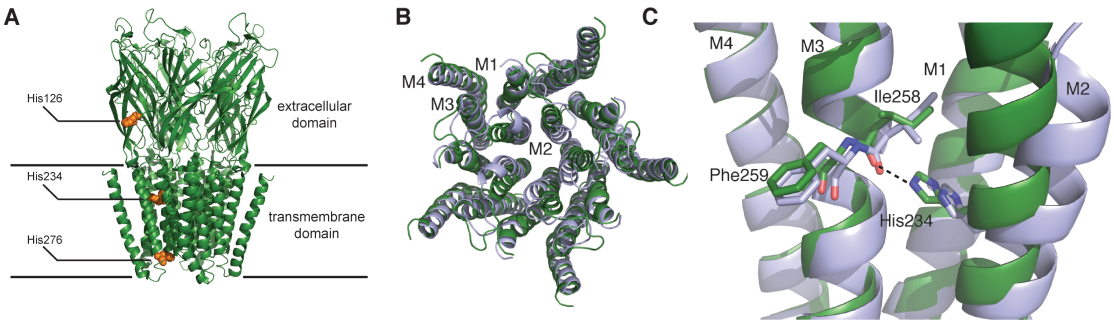


Figure 2.1. Histidine residues in GLIC. (A) Crystal structure of GLIC in a presumed open state (PDB 3EHZ). Histidines studied here are highlighted. (B) Top-down overlay of a locally closed GLIC mutant structure (blue, PDB 3TLU) with the open channel structure (green), omitting the extracellular domain. Note the movement of the M2 helix associated with channel opening. (C) Side view overlay of locally closed (blue) and open (green) GLIC structures. The orientation of the depicted residues suggests the formation of a hydrogen bond (shown as a dashed line) between His234 and the backbone carbonyl of Ile258 upon channel activation.

and biochemical probes of the receptor as the best tools to discern which residues are responsible for detection of a pH change and how resulting conformational changes are propagated to give an ion-conducting state. Recent studies have implicated a transmembrane histidine residue (His234 or His11' using the M2 prime labeling system) as playing an essential role in GLIC function.[17,21] Specifically, the histidine side chain was proposed to form an interhelix hydrogen bond with the backbone carbonyl of Ile258, stabilizing the closer association of the M2 and M3 helices that is observed in the open state of the channel, compared with locally closed mutants and structures obtained at neutral pH (**Figure 2.1**). Conventional mutagenesis at this histidine has revealed high sensitivity to perturbation, both for channel function and surface expression of the receptor. No receptors with point mutations at His234 were shown to access an ion-conducting state.[21] Additionally, molecular dynamics simulations have suggested that protonation of His234 occurs concomitantly with a transition to the open state.[21] In spite of these insights, a functional, proton-gated fusion consisting of the extracellular domain of GLIC and the transmembrane domain from the human α1 glycine receptor (which does not contain an analogous His residue) has also been reported.[22] While functional, this chimera produced a shifted pH-response curve relative to full-length GLIC, a more dramatic shift than is typically seen for comparable chimeras of

mammalian Cys-loop receptors where extracellular domain activation reasonably mimics that of the parent receptor.[23,24] Thus, while some aspects of the pH sensitivity of GLIC may be conferred by an as-yet-unidentified proton-binding region in the extracellular domain, such a region may play a major role only in the chimera. This is reminiscent of proton activation of KcsA, where are there are 2 proton sensing regions: His25 in M1 (likely the major sensor) and a network of ionizable residues at the transmembrane/cytoplasmic interface.[25-28]

Since conventional mutagenesis at His234 is disruptive to GLIC function, we have implemented the subtler probe of non-canonical amino acid mutagenesis in the transmembrane domain. None of the canonical amino acids closely resemble the properties of histidine, and conventional mutagenesis does not allow for modification of the protein backbone. The small systematic perturbations to steric and electrostatic interactions enabled by non-canonical amino acid mutagenesis could allow for a deeper understanding of the origins of its sensitivity to substitution. In this work, we describe the synthesis and incorporation of two non-canonical histidine analogues into GLIC; residues designed specifically to probe the protonation of His234. In addition, we use backbone mutagenesis to probe formation of the M2-M3 interhelix hydrogen bond. Our data suggest that both are important for channel function.

2.3 Results and discussion

2.3.1 His234 is the only His residue critical for GLIC function

GLIC is activated when extracellular pH is lowered, and the half-maximal pH for activation (pH_{50}) is 5.5, implicating a histidine ($pK_a \sim 6$) as a critical ionizable residue. There are three His residues in GLIC (**Figure 2.1A**), and previous data have shown that mutation of either His126 or His234 ablates function, while alteration of His276 does not.[21] Thus, His126 and His234 were initially proposed as candidates for a proton binding site. However,

Table 2.1. Conventional mutagenesis of His126. (Experiments performed by Prof. Sarah Lummis).

Mutant	pH_{50}	Hill Coefficient (n_H)	Number of Oocytes Tested
wt GLIC	5.50 ± 0.02	1.4 ± 0.1	16
H126A	5.24 ± 0.10	2.3 ± 0.3	3
H126E	5.38 ± 0.07	2.2 ± 0.2	3
H126K	5.34 ± 0.06	2.0 ± 0.2	3
H126L	NR		3
H126R	5.40 ± 0.14	2.0 ± 0.4	3
H126S	5.52 ± 0.05	2.0 ± 0.2	3
H126Y	5.35 ± 0.04	2.6 ± 0.3	3

subsequent expression studies in HEK cells showed that lack of function with His126 mutants was due to the role of this His in protein folding, subunit oligomerization, and/or transport to the cell surface.[21] To further probe the importance of this site, we created a range of GLIC His126 mutants and evaluated them in *Xenopus* oocytes, which are often able to express proteins that are not amenable to expression in mammalian cells. Our data revealed that these mutant receptors are functional, with pH_{50} and n_H values similar to those of wild type for nearly all the His126 mutants we tested (**Table 2.1**). The data therefore indicate that His126 is not important for proton activation. Our results also support the earlier study that suggested this residue has a role in expression, as we observed very different maximal currents for the mutant receptors (**Figure 2.2**).

2.3.2 Probing the His234-Ile258 hydrogen bond

Crystallographic studies suggest a functionally important hydrogen bond between the side chain of His234 on helix M2 and the backbone carbonyl of Ile258 on M3 (**Figure 2.1C**). We have frequently employed a general strategy for probing hydrogen bonds to the protein backbone, using nonsense suppression methodology to incorporate α-hydroxy acid analogues of α-amino acids.[29-31] Briefly, a stop codon is introduced at the position of interest, and mRNA is prepared by in vitro runoff transcription. The desired α-hydroxy acid is ligated

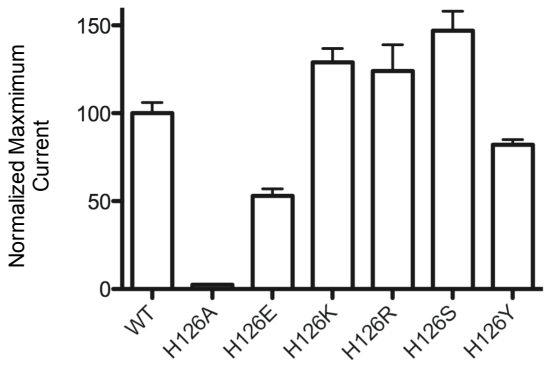


Figure 2.2. Maximal currents observed for mutants at His126. These data were obtained from the same batch of oocytes to minimize oocyte variability. Values are relative to wild type (100%); data = mean ± SEM, n_H = 3-5. (Experiments performed by Prof. Sarah Lummis).

to a suppressor tRNA bearing the anticodon corresponding to the stop codon used. The tRNA and mRNA are then coinjected into *Xenopus* oocytes, and the cells are incubated to allow for expression of the protein containing the α-hydroxy acid.[32,33]

The consequences of substitution with an α-hydroxy acid are twofold (**Figure 2.3**). Replacement of a residue with its α-hydroxy analog results in the loss of the backbone N-H bond, eliminating hydrogen bond donation. In addition, replacement of the backbone amide with an ester diminishes the hydrogen bond-accepting ability of the carbonyl at the *i*-1 residue relative to the mutation. Previous studies have shown that both effects can produce substantial perturbations of protein function.[33] In the present case, to probe the Ile258 carbonyl we replaced Phe259 with phenyllactic acid (Fah), the α-hydroxy acid analogue of Phe.

The backbone mutation to modulate the proposed M2-M3 hydrogen bond produced functional channels that expressed well (**Table 2.2**). The substitution did result in a shift in pH_{50} by about 1 pH unit toward a more acidic value. This suggests that weakening the M2-M3 hydrogen bond results in a relative destabilization of the open state of the channel, consistent with the model of **Figure 2.1**. Of course, by introducing a backbone mutation

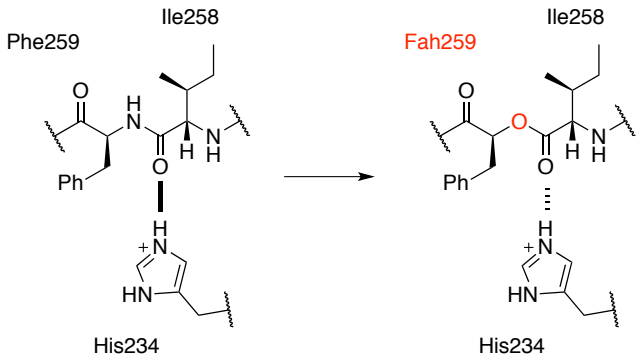


Figure 2.3. Backbone mutagenesis. Incorporation of phenyllactic acid at Phe259 introduces a backbone ester mutation, converting a strong hydrogen bond (solid line) to a weaker hydrogen bond (dashed line).

near the middle of the M3 α-helix, we are impacting several potential hydrogen bonds. Interestingly, in the open channel structure (PDB 3EHZ) there is a distortion of the M3 helix in the region of the Ile258 carbonyl, such that this carbonyl appears to make a stronger hydrogen bond to the M2 His side chain (O•••N distance 3.0Å) than to the backbone NH of the *i*+3 residue in M3 (Tyr262; O•••N distance 3.5Å). At the same time, the NH of Phe259 (which is deleted in the α-hydroxy experiment) makes a strong hydrogen bond to the *i*-4 residue (Gly255; O•••N 3.1Å). When we perturbed this hydrogen bond in the same manner – conversion of Ala256 to its α-hydroxy analogue – we were unable to observe functional channels. Nonsense suppression can be effective at the Ala256 site, however, as we were able to rescue wild type behavior by incorporation of Ala. Further experiments are required to tease out the role of any hydrogen bond networks in this region of the protein, although currently there is no evidence to indicate that hydrogen bonds other than those formed by His234 are actively involved in channel gating. At present then, we can say that our results are consistent with the presence of an important hydrogen bond between the backbone carbonyl of Ile258 on helix M3 and the side chain of His234 on helix M2, but we cannot rule out other interactions.

Table 2.2. Backbone mutagenesis experiments. Effects of perturbation of the interhelix hydrogen bond between His234 and Ile258.

Mutant	pH_{50}	Hill Coefficient (n_H)	Number of Oocytes Tested
F259TAG + Phe	5.54 ± 0.07	1.1 ± 0.1	14
F259TAG + Fah	4.47 ± 0.07	2.4 ± 0.3	10
A256TAG + Ala	5.43 ± 0.03	1.6 ± 0.6	8
A256TAG + Aah	NR		8

2.3.3 Synthesis and incorporation of non-canonical histidine analogs

We next hoped to determine the specific molecular requirements at His234 necessary for the function of GLIC. In particular, we were curious whether GLIC could tolerate subtle changes in the size or basicity of the imidazole side chain of His234. Recall that all conventional mutations of His234 produce nonfunctional channels. We first considered 2-methylhistidine (2-CH_3His), an analogue that would be expected to have increased steric bulk but only a minor (~0.7 pK_a unit) increase in basicity relative to histidine, based upon the effect of adding a methyl group to the corresponding position on imidazole.[34] For the synthesis of the desired non-canonical amino acid, we adapted a literature procedure for radical alkylation of histidine (**Figure 2.4**).[35] While the original reaction conditions resulted in only trace quantities of 2-CH_3His, we found that using acetic acid as a cosolvent and increasing the concentrations of silver nitrate and ammonium persulfate provided a low, but acceptable yield of the desired product, with high regioselectivity for the 2-position. This provided us with relatively rapid access to 2-CH_3His, which could then be appropriately protected and appended to suppressor tRNA as previously described.[30]

Interestingly, coinjection of H234TAG mRNA with 2-CH_3His-charged tRNA gave no evidence of functional channels containing 2-CH_3His, even after long incubation times and multiple injections of large quantities of material. Control experiments in which 2-CH_3His was incorporated at the two other histidine sites (His126 and His276) produced

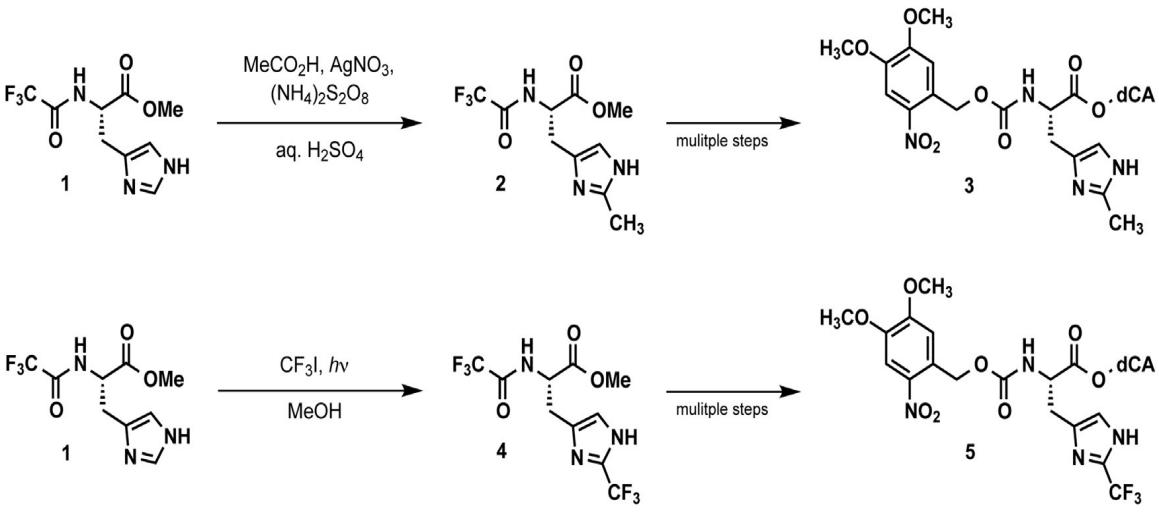


Figure 2.4. Preparation of 2-substituted histidine analogs. Preparation of (top) 2-CH_3His and (bottom) 2-CF_3His for ligation to suppressor tRNA.

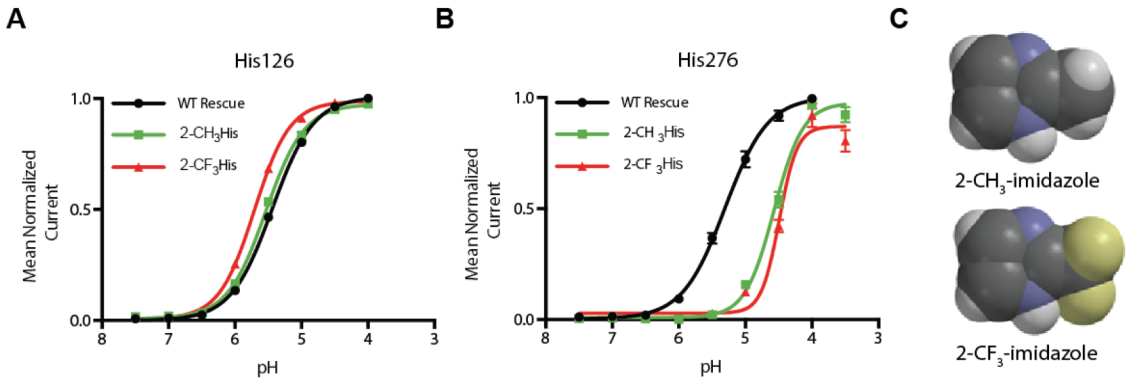


Figure 2.5. Histidine substitutions at control sites. (A and B) Control experiments demonstrate robust nonsense suppression and membrane trafficking with histidine (dotted) and synthetic histidine anlaogs 2-CH_3His (solid) and 2-CF_3His (dashed) at His126 (A) and His276 (B). Mean normalized current responses are plotted from 12-20 oocytes per condition, with standard error indicated. (C) Space filling models of and 2-CH_3-imidazole (top) and 2-CF_3-imidazole (bottom), showing they are comparable in size.

functional channels, establishing that 2-CH_3His can be readily incorporated via nonsense suppression (**Figure 2.5**, **Table 2.3**). Given the modest effect on His pK_a caused by CH_3-substitution, we considered these results to suggest that the introduction of even a minimal steric clash in the vicinity of His234 may disrupt either protein folding, membrane trafficking, or the gating process.

Examination of structural data from the "open" state of GLIC (PDB 3EHZ) suggests possible unfavorable interactions between a group installed at the 2-position of the

Table 2.3. Effects of nonsense suppression with histidine analogs at non-essential histidine sites.

Mutant	pH_{50}	Hill Coefficient (n_H)	Number of Oocytes Tested
H126TAG + His	5.56 ± 0.03	1.5 ± 0.1	13
H126TAG + 2-CH_3His	5.61 ± 0.02	1.6 ± 0.1	21
H126TAG + 2-CF_3His	5.72 ± 0.01	1.6 ± 0.1	18
H276TAG + His	5.41 ± 0.03	1.7 ± 0.1	12
H276TAG + 2-CH_3His	4.52 ± 0.03	2.2 ± 0.1	15
H276TAG + 2-CF_3His	4.52 ± 0.08	2.7 ± 0.4	16

His234 side chain and the side chains of Ile261 and Ile258 (**Figure 2.6A**). We therefore introduced additional mutations in combination with H234TAG in an effort to alleviate crowding around the substituted histidine (**Figure 2.6B**). The I261G mutation alone produced functional channels. Concentration-response curves showed some cell-to-cell variation (**Figure 2.7**), but currents are first observed upon acidification at pH 6, and we estimate pH_{50} to be approximately 5.0. Combining the I216G mutation with 2-CH_3His incorporation at His234 did produce functional channels. The steric sensitivity of this region is emphasized by the fact that 2-CH_3His did not produce functional channels in the presence of the I261A mutation (**Figure 2.6B**). For the 2-CH_3His/I261G double mutant, currents were first observed upon acidification at pH 5.5, and increased roughly linearly as pH was lowered to 3.5. While the pH profiles of these currents were consistent, the oocytes did not tolerate proton concentrations high enough to reach a maximal current. As such, we were unable to obtain a reliable pH_{50} value for this mutant.

In order to demonstrate that the mechanism of channel activation is not in some way modified by the I261G mutation, perhaps via a disruption of the M3 helix, we combined I261G with the conventional mutations H234L and H234F, and observed negligible currents (data not shown). Thus, His234 still plays an essential role in gating, even in the presence of the I261G mutation, and the observed function of the His234(2-CH_3His) mutant appears to

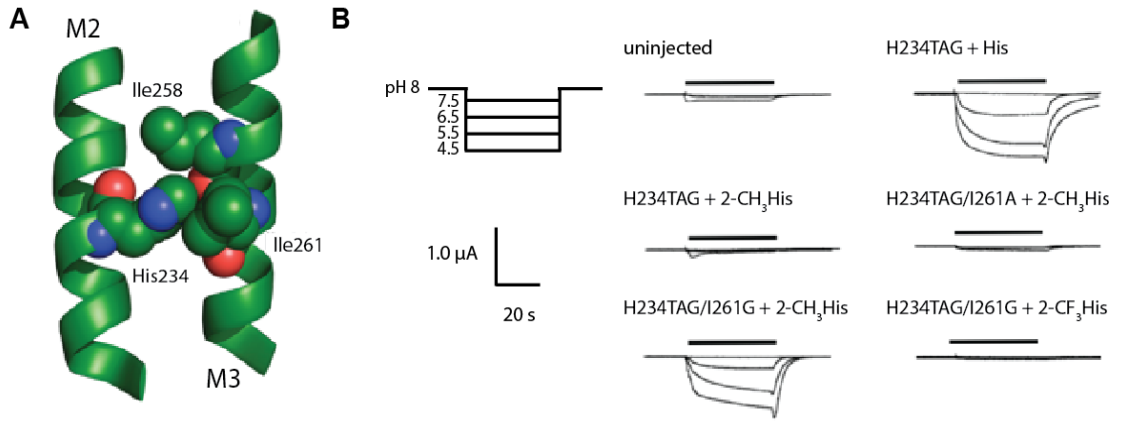


Figure 2.6. Histidine substitution at His234. (A) Space-filling model of His234 and surrounding residues in the presumed open state structure (PDB 3EHZ). (B) Whole-cell current responses of oocytes expressing GLIC mutants to buffer applications at pH 7.5, 6.5, 5.5, and 4.5.

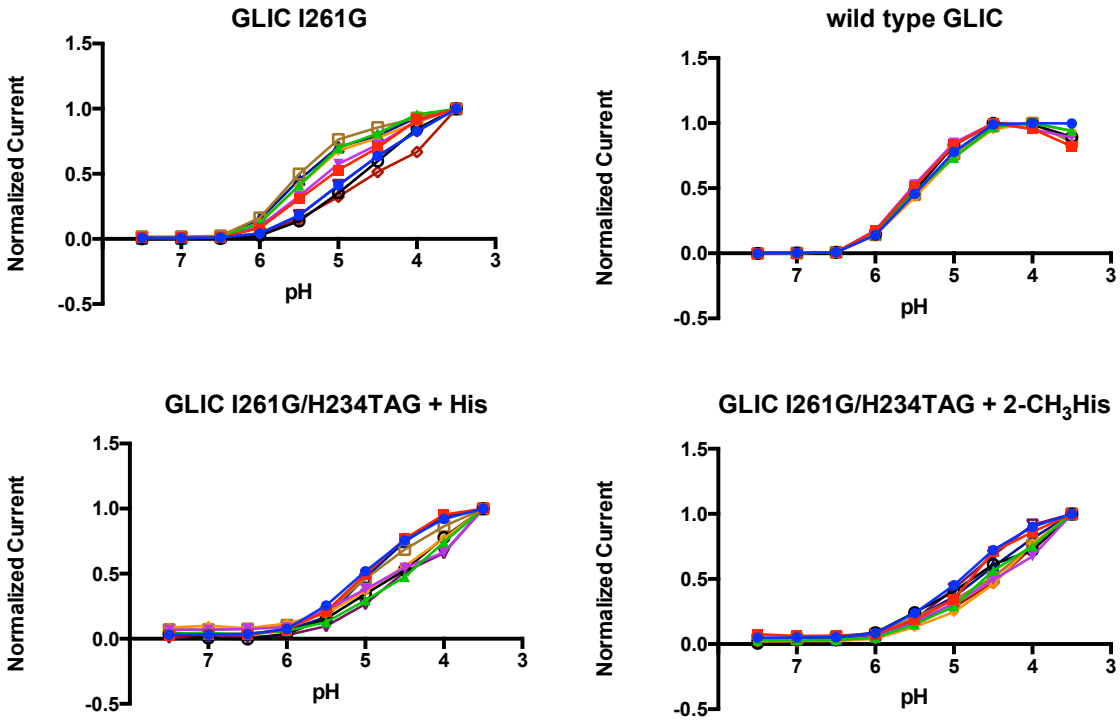


Figure 2.7. Representative dose-response data for selected mutants. Each color represents peak current data recorded from a single oocyte. Cells expressing receptors containing the I261G mutation gave inconsistent pH-response curves with low Hill coefficients. Saturating H_3O^+ concentrations were not generally tolerated by the oocytes.

be permitted specifically by reduction of steric bulk in the proximity of the 2-position. Given the sensitivity to structural perturbation in the region of His234, we could not expect the subtle pK_a shift associated with the addition of the methyl group to be discernable. As such,

Table 2.4. Effects of mutation at His234. pH_{50} and n_H data are shown as mean ± SEM. NR, nonresponsive.

Mutant	pH_{50}	Hill Coefficient (n_H)	Number of Oocytes Tested
H234F	NR		17
H234L	NR		12
H234TAG + His	5.37 ± 0.04	1.5 ± 0.1	23
H234TAG + 2-CH_3His	NR		20
I261A	5.05 ± 0.03	1.6 ± 0.1	11
I261A/H234TAG + His	4.96 ± 0.02	1.3 ± 0.1	15
I261A/H234TAG + 2-CH_3His	NR		19
I261G[a]	~ 5.0	~ 0.9	25
I261G/H234F	NR		17
I261G/H234L	NR		10
I261G/H234TAG + His[a]	≤ 4.5	~ 0.8	9
I261G/H234TAG + 2-CH_3His[a]	≤ 4.5	~ 0.8	14
I261G/H234TAG + 2-CF_3His	NR		23

[a]Oocytes did not tolerate saturating proton concentrations; pH_{50} and n_H values given are estimated.

we sought a more dramatic electronic perturbation, and we chose 2-CF_3His. As shown in **Figure 2.5C**, 2-CH_3His and 2-CF_3His are similar in size. However, the two have very different basicities – while the conjugate acid of 2-methylimidazole has a pK_a of 7.75, the corresponding value for 2-trifluoromethylimidazole is 2.06.[34] Thus 2-CF_3His is a very weak base that would almost certainly remain in the neutral form at any pH accessible in the electrophysiology assay. We therefore prepared 2-trifluoromethylhistidine (2-CF_3His) using a literature procedure (**Figure 2.4**),[36] and appended it to a suppressor tRNA as described for 2-CH_3His. Again, positive control experiments at His126 and His276 indicated that nonsense suppression with 2-CF_3His occurs robustly (**Figure 2.5B**; **Figure 2.8**). The concentration-response curves for the 2-CF_3His and the 2-CH_3His mutants overlay at both sites indicating that: (1) any steric perturbation due to trifluoromethylation is similar to that introduced by methyl substitution at these sites and (2) as the pK_as of these compounds are very different, the protonation state of neither His126 nor His276 affects GLIC function.

Attempts at incorporation of 2-CF_3His at His234 in combination with the I261G

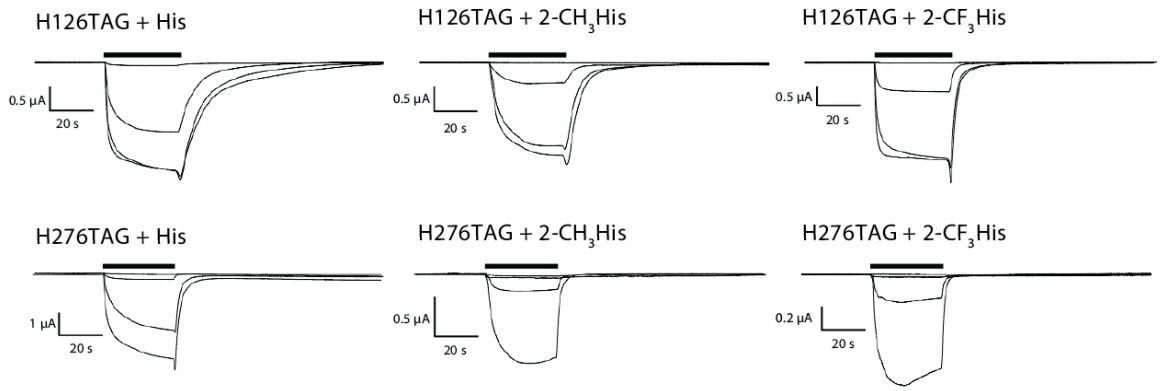


Figure 2.8. Representative current traces for nonsense suppression experiments at His126 and His276. Running buffer was at pH 8; responses are shown for acidification to pH 7, 6, 5, and 4, indicated by the black bar.

mutation, however, gave no significant currents. Due to the inherently low expression levels associated with nonsense suppression experiments in *Xenopus* oocytes, it was not possible to establish whether receptors were successfully synthesized and delivered to the cell surface by immunofluorescence (**Figure 2.9**). However, given that we have established that 2-CF_3His is compatible with our nonsense suppression methodology, and that 2-CH_3His at this site can produce functional receptors, we consider the most reasonable interpretation of the 2-CF_3His results to be that the non-canonical amino acid was incorporated, but that the resulting receptors are nonfunctional because protonation cannot occur.

2.4 Conclusions and recent developments

Structural and functional data for GLIC have shown the importance of His234 in the transmembrane region of this protein. Here we have shown that even subtle steric perturbation of His234 will ablate function, and this can be partially rescued by compensatory steric alterations in the adjacent M3 helix. In addition, we have produced two lines of evidence to support a key role for protonation of His234 in receptor function. First, the subtle mutation by backbone mutagenesis of the putative hydrogen bond partner to the protonated His234 alters function in the expected direction. Second, introducing a His

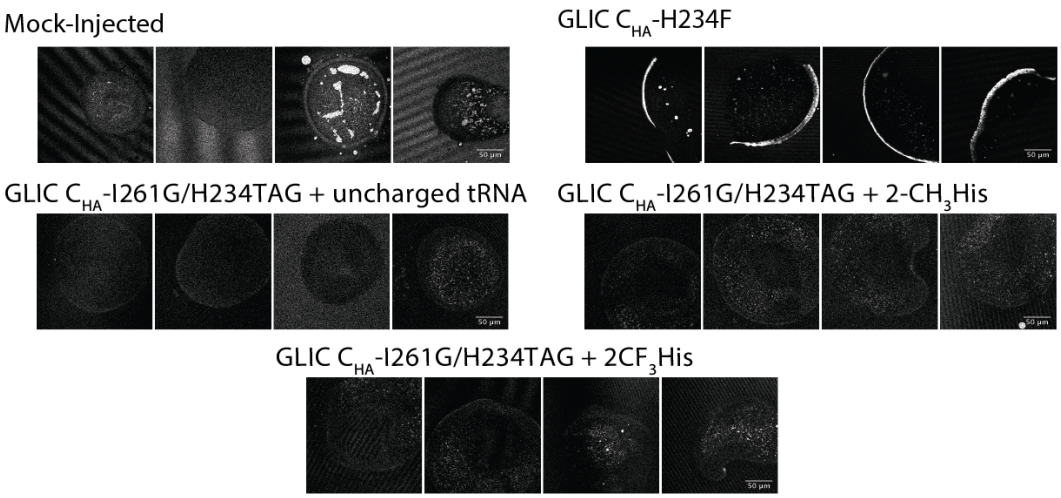


Figure 2.9. Immunofluorescence experiments. Oocytes expressing each mutant were immunolabeled and sliced, then imaged by confocal microscopy (see Supplemental Experimental Procedures). While it was possible to verify surface expression of the nonfunctional H234F mutant, expression from nonsense suppression experiments was too low to distinguish levels of fluorescence from those observed in the negative controls.

analogue that would not be expected to be protonated at pH > 2 (2-CF_3His) produces nonfunctional channels. Overall these experiments show that formation of the M2-M3 hydrogen bond is important for stabilizing the ion-conducting state, and support the *in silico* experiments which show that His234 must exist in the protonated form for GLIC to attain its activated conformation.[21]

Since the publication of the results described here, additional mechanistic work on GLIC by other research groups has further illuminated the role played by His234 and other residues. Interestingly, it has been shown that GLIC mutants containing both H234F and the gain-of-function mutation I9'A produce large acid-induced currents.[37] This indicates that His234 is not an indispensible proton binding site for detection of reduced pH. Rather, it is likely that destruction of the hydrogen bond examined here causes a relative destabilization of the open state too great to be observed within the pH range used in our experiments, but one that can be reversed by a strong gain-of-function mutation at another position.

In the same study, a chimeric receptor consisting of a fusion between the ECD from ELIC and the TMD from GLIC was also shown to produce smaller currents induced either

by reduced pH or by primary amines (the activating ligands for ELIC).[37] The currents persisted even in the presence of the combined H234F and I9'A mutations. Considered together with the previous results involving the GLIC-GlyR chimeras,[22] this is further evidence that no single proton-binding site in either the ECD or TMD is fully responsible for the pH sensitivity of GLIC. In the most likely case, the concerted protonation of a number of residues results in global conformational changes during channel activation. Taken alone, it is doubtful that either crystal structures or simple mutagenesis experiments will be capable of identifying the critical titrated sites, if such a well-defined group of sites indeed exists. Future studies of GLIC will require careful coordination of structural, biochemical, and computational efforts to develop models for the energetic landscape of the receptor in its dynamic membrane environment.

2.5 Experimental procedures

2.5.1 Molecular biology

The cDNA for GLIC was inserted into the pGEMhe plasmid using standard subcloning techniques. Site-directed mutagenesis was performed using the Stratagene QuikChange protocol to generate the appropriate codon. For non-canonical amino acid mutants and conventional mutants generated by nonsense suppression, the site of interest was mutated to the TAG stop codon. Plasmids were linearized with the SbfI restriction enzyme, and receptor mRNA was then prepared by in vitro runoff transcription using the Ambion T7 mMessage mMachine kit.

Hydroxy or amino acid-dCA couples were enzymatically ligated to truncated 74mer THG73 tRNA as previously described.[29,32] The 74mer tRNA was prepared using the Ambion T7MEGAshortscript kit by transcription from a modified DNA oligonucleotide

template as described in the literature to enhance RNA transcript homogeneity.[38] Crude tRNA-amino acid or tRNA-hydroxy acid product was used without desalting, and the product was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry on a 3-hydroxypicolinic acid matrix. Deprotection of the NVOC group on the tRNA-amino acids was carried out by photolysis for 5 minutes on a 300 W high-pressure Hg arc lamp with WG-335 and UG-11 filters immediately prior to injection.

Note that we are using residue numbering consistent with PDB 3EHZ (**Figure 2.1**).[15] In some other published structures of GLIC, residues are offset by 1.

2.5.2 Oocyte preparation and RNA injection

Stage V-VI oocytes of *Xenopus laevis* were harvested and injected with RNAs as described previously.[32] For nonsense suppression experiments, each cell was injected with 50-100 ng each of receptor mRNA and appropriate tRNA approximately 48 h before recording. Mutants yielding small responses required 72 h of incubation, with a second injection of mRNA and tRNA 48 h before recording.

For wild type experiments and conventional mutants, each cell received a single injection of 1-25 ng of receptor mRNA approximately 24 h before recording. Injection volumes for each injection session were 50-100 nL per cell.

As a negative control for suppression experiments at each site, unacylated full-length tRNA was co-injected with mRNA in the same manner as charged tRNA. These experiments yielded negligible responses for all sites. Wild-type recovery conditions (injecting tRNA charged with the appropriate amino acid to regenerate a wild type channel via nonsense suppression at a TAG stop codon) were injected alongside mutant nonsense suppression conditions as a positive control. THG73-phenyllactic acid was unprotected and was injected directly without irradiation.

2.5.3 Electrophysiology

Oocyte recordings were made in two-electrode voltage clamp mode using the OpusXpress 6000A (Axon Instruments). Oocyte equilibration and washes were performed with Ca-free ND96 (96 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 5 mM HEPES) adjusted to pH 8 with 1 N NaOH. The pH of buffers for concentration-response curves were adjusted accordingly with NaOH or HCl, and buffers with pH below 6.8 contained 5 mM MES in place of HEPES. Initial holding potential was -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Oocytes were equilibrated for 30 s at 1 mL/min before each pH application. The pH buffer applications lasted for 15 s at 4 mL/min, followed by a 15 s wait period to allow the cells to attain a peak current. Cells were then washed for 40 s at 3 mL/min before the following equilibration. Concentration-response data were obtained for 9 buffer pHs, for a minimum of two cell batches, and for a minimum of 8 cells total. The concentration-response relations for each cell were fitted to the Hill equation:

$$I([A]) = \frac{I_{\max}}{1+\left(\frac{[A]}{\mathrm{EC}_{50}}\right)^{n_{\mathrm{H}}}}, \tag{2.1}$$

where I_{norm} is the normalized current peak at $[H_3O^+] = A$, EC_{50} is the value of $[H_3O^+]$ that elicits a half-maximum response, and n_H is the Hill coefficient. The values for each cell that displayed a strong fit to the Hill equation were then averaged to give the reported values, where $pH_{50} = -\log(EC_{50})$. At pH < 4, uninjected cells gave current responses that were occasionally up to ~ 200 nA. As such, data were only reported from cells that gave responses above pH 4, or that gave larger, more robust currents below pH 4, and other cells were considered non-responsive (NR).

2.5.4 Immunofluorescence Experiments

Following injection and incubation, oocytes were fixed in a 3.7% solution of paraformaldehyde in phosphate buffered saline (PBS) buffer (0.01 *M*, pH 7.4) for 3 h. The cells were then washed in a 3% solution of bovine serum albumin (BSA) in the same buffer (3 x 10 min) at room temperature, and then labeled overnight at 4 °C with monoclonal mouse-anti-HA (1:1000 in 3% BSA/PBS). Cells were washed in 3% BSA/PBS at room temperature (3 x 10 min), then incubated 3 h at room temperature with anti-mouse IgG secondary antibody conjugated with Alexa Fluor 555 (1:500 in 3% BSA/PBS). Cells were then washed in PBS buffer (3 x 5 min) and embedded in blocks of 3% low melting agarose in PBS buffer before sectioning into 50 μm slices with a vibratome. The slices were then mounted in 70% glycerol in PBS on glass slides. Slices were imaged at room temperature using an Eclipse C1si laser-scanning confocal microscope (Nikon Instruments) equipped with a 10x, 0.45 NA, Plan Apo objective and 32 photomultiplier tubes. Alexa Fluor 555 was excited at 561 nm. Fluorescence data were collected from 570-730 nm, and images were averaged over three scans. Images were unmixed using standard spectra acquired from slices of mock-injected cells and Alexa Fluor 555 alone.

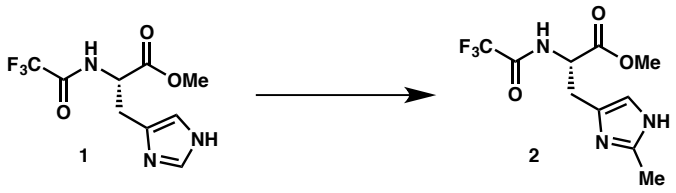
2.5.5 Chemical synthesis and characterization

2.5.5.1 Materials. Commercial reagents and solvents, purchased from Sigma Aldrich and VWR, were used as received. Triethylamine was freshly distilled from CaH_2 prior to use. Deuterated solvents for NMR experiments were purchased from Cambridge Isotope Laboratories and used without further purification. Compounds **1**,[35] and **4**,[36] and pdCpA[39] were synthesized as previously reported. Reactions involving air- or moisture-sensitive reagents were performed in anhydrous solvents under an atmosphere of Ar. Flash chromatography was performed using silica gel from Merck (230-400 mesh). Analytical TLC

was performed using plates from Sigma-Aldrich (60 Å pore size, F_{254}, 0.25 mm) and was visualized by UV irradiation (254 nm) or staining with potassium permanganate.

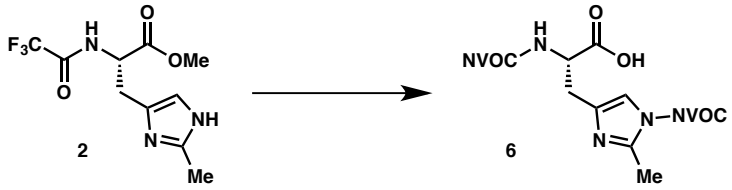
2.5.5.2 Instrumentation. ^{1}H and ^{13}C NMR spectra were obtained using Varian NMR instruments. Chemical shifts are reported in parts per million (ppm, δ) referenced to the residual ^{1}H resonance of the solvent. ^{13}C NMR spectra were referenced to the residual ^{13}C resonance of the solvent. Splitting patterns are designated as follows: s, singlet; br, broad; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet. Mass spectroscopic data were acquired using an Agilent 1100 Series LC/MSD equipped with an ESI-APCI source. HPLC preparative purification was performed using a Waters 1525 binary pump, a Waters 2996 photodiode array detector, and an Atlantis Prep T3 OBD 5μm 19x150mm column.

2.5.5.3 Reaction conditions.



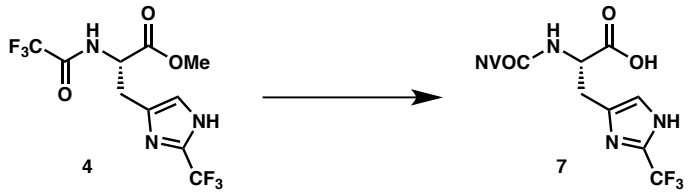
(*S*)-methyl 3-(2-methyl-1*H*-imidazol-4-yl)-2-(2,2,2-trifluoroacetamido)propanoate (**2**). To a solution of **1** (3.0 g, 11.3 mmol) in a mixture of 0.38 *M* aqueous sulfuric acid (30 mL, 11.3 mmol) and glacial acetic acid (30 mL) was added silver nitrate (960 mg, 5.65 mmol) at room temperature. The mixture was heated to 80 °C with stirring, and ammonium persulfate (12.9 g, 56.5 mmol) was slowly added as a solution in 60 mL of water over 10 minutes via cannula. After another 20 minutes, the reaction was poured over 100 g ice and allowed to come to room temperature. The mixture was basified with concentrated aqueous ammonium hydroxide to pH 10, then extracted with ethyl acetate (3 x 100 mL). The combined organic layers were dried over sodium sulfate, then filtered and concentrated. The desired product

was partially purified by column chromatography (SiO_2, 10% MeOH/DCM) to give **2** as a white solid (105 mg, 3%). ^{1}H-NMR (300 MHz; CD_3OD) δ: 6.70 (s, 1H), 4.72 (dd, J = 9.5, 5.0 Hz, 1H), 3.73 (s, 3H), 3.14 (dd, J = 14.9, 5.0 Hz, 1H), 2.97 (dd, J = 14.8, 9.6 Hz, 1H), 2.32 (s, 3H); ^{13}C NMR (126 MHz; CD_3OD) δ: 172.0, 158.8 (q, J_F = 37.60), 146.0, 133.9, 117.2 (q, J_F = 286.59), 117.1, 54.2, 53.0, 29.5, 13.3; ^{19}F NMR (282 MHz; CD_3OD) δ: -78.9; LRMS (ESI) 278.0 $[M-H]^-$.

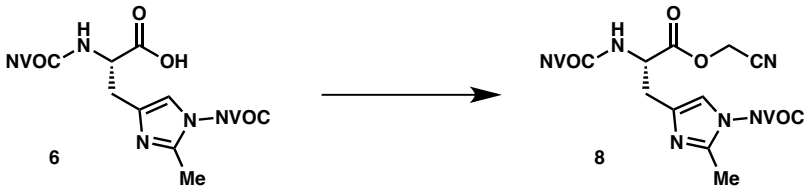


(*S*)-3-(1-(((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)-2-methyl-1*H*-imidazol-4-yl)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)propanoic acid (**6**). A solution of 6 *N* hydrochloric acid (10 mL) was added to a scintillation vial charged with **2** (190 mg, 0.68 mmol). A stir bar was added, and the solution was heated to 80 °C and stirred overnight. The reaction was then cooled to room temperature and concentrated to a brown solid, which was dissolved in methanol and concentrated several times to remove residual hydrochloric acid. The crude product was then directly dissolved in a mixture of water (6 mL) and dioxane (6 mL). Sodium carbonate (505 mg, 4.76 mmol) and 6-nitroveratryl chloroformate (375 mg, 1.36 mmol) were added in one portion, and the reaction was allowed to return to room temperature with stirring for 5 h. The mixture was then concentrated to a solid in vacuo, and the residue was suspended in water (50 mL) and acidified to pH ~2 with 1 N aqueous hydrochloric acid, then extracted with ethyl acetate (5 x 30 mL). The combined organic extracts were dried over sodium sulfate, filtered, and concentrated. The residue was purified on a plug of silica (2% AcOH/EtOAc) to give (**6**) as a tan solid (138 mg, 0.21

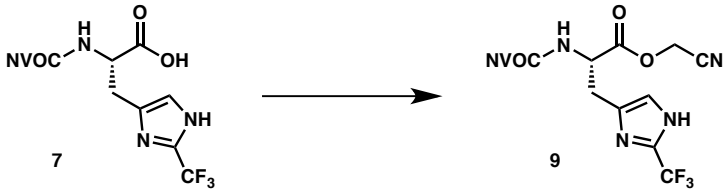
mmol, 34% yield). ^{1}H NMR (400 MHz; d_6-DMSO) δ: 7.72 (s, 1H), 7.61 (s, 1H), 7.22 (s, 1H), 7.17 (s, 1H), 7.10 (s, 1H), 7.02 (d, *J* = 6.3 Hz, 1H), 5.62 (d, *J* = 13.2 Hz, 1H), 5.57 (d, *J* = 14.0 Hz, 1H), 5.29 (d, *J* = 15.4 Hz, 1H), 5.23 (d, *J* = 15.7 Hz, 1H), 3.92 (s, 3H), 3.85-3.95 (m, 1H), 3.88 (s, 3H), 3.82 (s, 3H), 3.78 (s, 3H), 2.94 (dd, *J* = 14.8, 2.8 Hz, 1H), 2.71 (dd, *J* = 15.2, 8.2 Hz, 1H), 2.45 (s, 3H); ^{13}C NMR (126 MHz; d_6-DMSO) δ: 173.3, 155.1, 153.5, 153.3, 148.6, 148.3, 147.4, 145.9, 139.7, 138.9, 138.6, 128.9, 124.7, 114.0, 111.6, 109.7, 108.2, 107.8, 65.7, 61.9, 56.4, 56.10, 56.06, 56.0, 54.9, 31.2, 16.2; LRMS (ESI) 646.2 $[M-H]^-$.



(*S*)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-(2-(trifluoromethyl)-1*H*-imidazol-4-yl)propanoic acid (**7**). Prepared according to the procedure for **6**, starting with **4** (230 mg, 0.87 mg). The crude mixture was purified by precipitation from a hexanes/ethyl acetate mixture, giving **7** as a tan solid (202 mg, 0.29 mmol, 33% yield over 2 steps). ^{1}H NMR (500 MHz; d_6-DMSO) δ: 7.89 (d, *J* = 8.1 Hz, 1H), 7.69 (s, 1H), 7.13 (s, 1H), 7.08 (s, 1H), 5.38 (d, *J* = 15.1 Hz, 1H), 5.31 (d, *J* = 15.1 Hz, 1H), 4.29 (td, *J* = 8.8, 4.5 Hz, 1H), 3.87 (s, 3H), 3.86 (s, 4H), 3.05 (dd, *J* = 14.7, 3.9 Hz, 1H), 2.93 (dd, *J* = 14.6, 9.9 Hz, 1H); ^{13}C NMR (126 MHz; d_6-DMSO) δ: 172.8, 155.5, 153.5, 147.6, 139.0, 133.7 (q, J_F = 39.5 Hz), 128.0, 118.9 (q, J_F = 268.3), 110.0, 108.2, 62.3, 56.1, 53.7; ^{19}F NMR (282 MHz; CD_3CN) δ: -64.0; LRMS (ESI) 461.1 $[M - H]^-$.

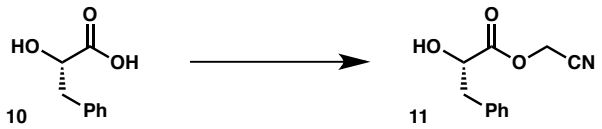


(*S*)-4,5-dimethoxy-2-nitrobenzyl 4-(3-(cyanomethoxy)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-oxopropyl)-2-methyl-1*H*-imidazole-1-carboxylate (**8**). To a solution of **6** (80 mg, 0.12 mmol) in a mixture of DMF (1.2 mL) and chloroacetonitrile (1.2 mL) was added triethylamine (54 μL, 0.39 mmol). The reaction was stirred for 24 h at room temperature, then diluted with ether (10 mL) and washed with water (3 x 10 mL). The organic layer was then concentrated to an oil, which was dissolved in ethyl acetate. The product was precipitated by addition of hexanes, giving **8** as a tan solid (25 mg, 29%). ^{1}H NMR (500 MHz; d_6-DMSO) δ: 8.09 (d, *J* = 7.8, 1H), 7.72 (s, 1H), 7.66 (s, 1H), 7.28 (s, 2H), 7.11 (s, 1H), 5.64 (s, 2H), 5.35 (d, *J* = 15.0, 1H), 5.31 (d, *J* = 14.7, 1H), 5.01 (s, 2H), 4.46-4.42 (m, 1H), 3.90 (s, 3H), 3.89 (s, 3H), 3.85 (s, 6H), 2.89-2.87 (m, 2H), 2.48 (s, 3H); ^{13}C NMR (126 MHz; d_6-DMSO) δ: 170.8, 155.5, 153.4, 153.2, 148.5, 148.4, 147.7, 146.8, 140.1, 139.1, 135.8, 127.6, 124.3, 115.6, 115.5, 112.4, 110.3, 108.3, 108.1, 66.0, 62.7, 56.4, 56.2, 56.11, 56.06, 53.2, 49.5, 29.1, 16.3; LRMS (ESI) 685.1 $[M-H]^-$.

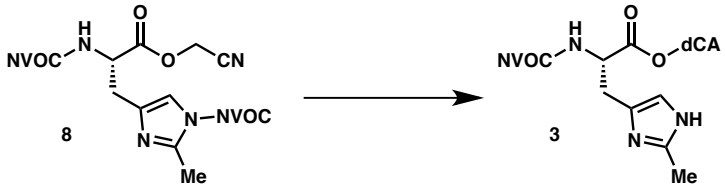


(*S*)-cyanomethyl 2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-(2-(trifluoromethyl)-1*H*-imidazol-4-yl)propanoate (**9**). Prepared according to the procedure for **8**, starting with **7** (150 mg, 0.32 mmol). Following precipitation, **9** was obtained as a tan solid (80 mg, 0.16 mmol, 49% yield). ^{1}H-NMR (300 MHz; $CDCl_3$) δ: 9.71 (s, 1H), 7.73 (s, 1H),

7.07 (s, 1H), 7.03 (d, *J* = 2.0 Hz, 1H), 6.50 (d, *J* = 8.8 Hz, 1H), 5.63 (d, *J* = 15.5 Hz, 1H), 5.49 (d, *J* = 15.8 Hz, 1H), 4.80 (d, *J* = 15.7 Hz, 1H), 4.77-4.72 (m, 1H), 4.67 (d, *J* = 15.7 Hz, 1H), 3.97 (d, *J* = 10.3 Hz, 6H), 3.28 (dd, *J* = 15.1, 4.8 Hz, 1H), 3.16 (dd, *J* = 14.5, 4.8 Hz, 1H); ^{13}C-NMR (126 MHz; d_6-DMSO) δ: 170.6, 155.5, 153.4, 147.8, 139.2, 137.53, 133.9 (q, J_F = 38 Hz), 127.5, 118.8 (q, J_F = 265 Hz), 117.3, 115.4, 110.3, 108.2, 62.7, 56.09, 56.04, 53.76, 49.5, 29.3; ^{19}F NMR (282 MHz; $CDCl_3$) δ: -63.7; LRMS (ESI) 500.1 $[M - H]^-$.

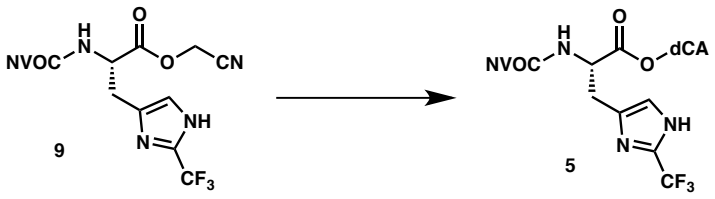


(*S*)-cyanomethyl 2-hydroxy-3-phenylpropanoate (**11**). Prepared according to the procedure for **8**, starting with **10** (100 mg, 0.60 mmol). Following precipitation, **11** was obtained as a colorless oil (85 mg, 0.41 mmol, 69% yield). ^{1}H NMR (500 MHz; $CDCl_3$) δ: 7.33-7.30 (m, 2H), 7.28-7.25 (m, 1H), 7.22-7.20 (m, 2H), 4.75 (d, J = 15.7, 1H), 4.72 (d, J = 15.7, 1H), 4.53 (q, J = 5.6, 1H), 3.14 (dd, J = 14.0, 4.8, 1H), 3.02 (dd, J = 14.0, 6.7, 1H), 2.70 (d, J = 6.5, 1H); ^{13}C NMR (126 MHz; $CDCl_3$) δ: 172.5, 135.4, 129.5, 128.7, 127.3, 113.7, 71.3, 48.9, 40.4; LRMS (ESI) 204.0 $[M-H]^-$.

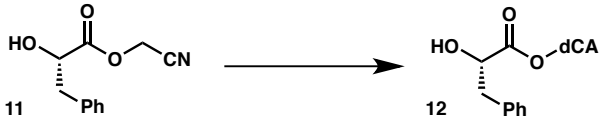


4,5-dimethoxy-2-nitrobenzyl 4-((2*S*)-3-((5-(((((5-(4-amino-2-oxopyrimidin-1(2*H*)-yl)-2-((phosphonooxy)methyl)tetrahydrofuran-3-yl)oxy)(hydroxy)phosphoryl)oxy)-methyl)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxytetrahydrofuran-3-yl)oxy)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-oxopropyl)-2-methyl-1*H*-imidazole-1-carboxylate (**3**). A 0.5-dram vial was charged with 10 mg pdCpA•2.4 TBAH and a stir bar,

then evacuated and backfilled several times with argon. A 0.02 *M* solution of **8** in DMF (0.5 mL, 0.01 mmol) was then added, and the mixture was stirred under argon overnight. The product was then purified using reverse-phase HPLC to give **3** as a white solid. MALDI-MS (6-aza-2-thiothymine matrix) 1026.9 $[M\text{-NVOC}+H]^+$.



(2*S*)-5-(((((5-(4-amino-2-oxopyrimidin-1(2*H*)-yl)-2-((phosphonooxy)methyl) tetrahydrofuran-3-yl)oxy)(hydroxy)phosphoryl)oxy)methyl)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxytetrahydrofuran-3-yl 2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl) amino)-3-(2-(trifluoromethyl)-1*H*-imidazol-4-yl)propanoate (**5**). Prepared according to the procedure for **3**, starting with **9**. MALDI-MS (6-aza-2-thiothymine matrix) 1081.2 $[M+H]^+$.



(2*S*)-5-(((((5-(4-amino-2-oxopyrimidin-1(2*H*)-yl)-2-((phosphonooxy)methyl) tetrahydrofuran-3-yl)oxy)(hydroxy)phosphoryl)oxy)methyl)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxytetrahydrofuran-3-yl 2-hydroxy-3-phenylpropanoate (**12**). Prepared according to the procedure for **3**, starting with **11**. MALDI-MS (6-aza-2-thiothymine matrix) 785.4 $[M+H]^+$.

2.6 References

1. Lester, H. Cys-loop receptors: new twists and turns. *Trends Neurosci.* **27,** 329–336 (2004).
2. Sine, S. M. & Engel, A. G. Recent advances in Cys-loop receptor structure and function. *Nature* **440,** 448–455 (2006).
3. Jensen, A. A., Frølund, B., Liljefors, T. & Krogsgaard-Larsen, P. Neuronal nicotinic acetylcholine receptors: structural revelations, target identifications, and therapeutic inspirations. *J. Med. Chem.* **48,** 4705–4745 (2005).
4. Lemoine, D. *et al.* Ligand-gated ion channels: new insights into neurological disorders and ligand recognition. *Chem. Rev.* **112,** 6285–6318 (2012).
5. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4Å resolution. *J. Mol. Biol.* **346,** 967–989 (2005).
6. Brejc, K. *et al.* Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411,** 269–276 (2001).
7. Sixma, T. K. & Smit, A. B. Acetylcholine binding protein (AChBP): a secreted glial protein that provides a high-resolution model for the extracellular domain of pentameric ligand-gated ion channels. *Annu. Rev. Biophys. Biomol. Struct.* **32,** 311–334 (2003).
8. Hibbs, R. E. & Gouaux, E. Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* **474,** 54–60 (2011).
9. Althoff, T., Hibbs, R. E., Banerjee, S. & Gouaux, E. X-ray structures of GluCl in apo states reveal a gating mechanism of Cys-loop receptors. *Nature* **512,** 333–337 (2014).
10. Miller, P. S. & Aricescu, A. R. Crystal structure of a human $GABA_A$ receptor. *Nature* **512,** 270–275 (2014).
11. Hassaine, G. *et al.* X-ray structure of the mouse serotonin 5-HT3 receptor. *Nature* **512,** 276–281 (2014).
12. Tasneem, A., Iyer, L. M., Jakobsson, E. & Aravind, L. Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels. *Genome Biol.* **6,** R4 (2004).
13. Bocquet, N. *et al.* A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature* **445,** 116–119 (2007).
14. Hilf, R. J. C. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452,** 375–379 (2008).
15. Hilf, R. J. C. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457,** 115–118 (2009).
16. Sauguet, L. *et al.* Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proc. Natl. Acad. Sci.* **111,** 966–971 (2014).
17. Prevost, M. S. *et al.* A locally closed conformation of a bacterial pentameric proton-gated ion channel. *Nat. Struct. Mol. Biol.* **19,** 642–649 (2012).
18. Bocquet, N. *et al.* X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457,** 111–114 (2008).
19. Nury, H. *et al.* X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature* **469,** 428–431 (2011).
20. Gonzalez-Gutierrez, G., Cuello, L. G., Nair, S. K. & Grosman, C. Gating of the proton-gated ion channel from Gloeobacter violaceus at pH 4 as revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 18716–18721 (2013).
21. Wang, H.-L., Cheng, X. & Sine, S. M. Intramembrane proton binding site linked to activation of bacterial pentameric ion channel. *J. Biol. Chem.* **287,** 6482–6489 (2012).
22. Duret, G. *et al.* Functional prokaryotic–eukaryotic chimera from the pentameric ligand-gated ion channel family. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 12143-12148 (2011).
23. Eisele, J. L. *et al.* Chimaeric nicotinic-serotonergic receptor combines distinct ligand binding and channel specificities. *Nature* **366,** 479–483 (1993).
24. Grutter, T. *et al.* Molecular tuning of fast gating in pentameric ligand-gated ion channels. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 18207–18212 (2005).
25. Posson, D. J., Thompson, A. N., McCoy, J. G. & Nimigean, C. M. Molecular interactions involved in proton-dependent gating in KcsA potassium channels. *J. Gen. Physiol.* **142,** 613–624 (2013).
26. Cuello, L. G. *et al.* Structural basis for the coupling between activation and inactivation gates in K(+) channels. *Nature* **466,** 272–275 (2010).
27. Thompson, A. N., Posson, D. J., Parsa, P. V. & Nimigean, C. M. Molecular mechanism of pH sensing

in KcsA potassium channels. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 6900–6905 (2008).
28. Takeuchi, K., Takahashi, H., Kawano, S. & Shimada, I. Identification and characterization of the slowly exchanging pH-dependent conformational rearrangement in KcsA. *J. Biol. Chem.* **282,** 15179–15186 (2007).
29. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Lett.* **40,** 6189–6192 (1999).
30. Noren, C., Anthony-Cahill, S., Griffith, M. & Schultz, P. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244,** 182–188 (1989).
31. England, P. M., Zhang, Y., Dougherty, D. A. & Lester, H. A. Backbone Mutations in Transmembrane Domains of a Ligand-Gated Ion Channel. *Cell* **96,** 89–98 (1999).
32. Nowak, M. W. *et al.* [28] In vivo incorporation of unnatural amino acids into ion channels in Xenopus oocyte expression system. *Method. Enzymol.* **293,** 504–529 (Elsevier, 1998).
33. Dougherty, D. A. & Van Arnam, E. B. In Vivo Incorporation of Non-canonical Amino Acids by Using the Chemical Aminoacylation Strategy: A Broadly Applicable Mechanistic Tool. *Chembiochem* **15,** 1710–1720 (2014).
34. Bruice, T. C. & Schmir, G. L. Imidazole Catalysis. II. The Reaction of Substituted Imidazoles with Phenyl Acetates in Aqueous Solution. *J. Am. Chem. Soc.* **80,** 148–156 (1958).
35. Jain, R., Cohen, L. A., El-Kadi, N. A. & King, M. M. Regiospecific alkylation of histidine and histamine at C-2. *Tetrahedron* **53,** 2365–2370 (1997).
36. Kimoto, H., Fujii, S. & Cohen, L. A. Photochemical trifluoromethylation of some biologically significant imidazoles. *J. Org. Chem.* **49,** 1060-1064 (1984).
37. Schmandt, N. *et al.* A chimeric prokaryotic pentameric ligand-gated channel reveals distinct pathways of activation. *J. Gen. Physiol.* **146,** 323–340 (2015).
38. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5,** 1268–1272 (1999).
39. Robertson, S. A., Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. The use of 5'-phospho-2 deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA. *Nucleic Acids Res.* **17,** 9649–9660 (1989).

Chapter 3: Interrogation of Gating Motions of GLIC via Non-Canonical Amino Acid Mutagenesis at Sensitive Prolines*

3.1 Abstract

Gloeobacter violaceus ligand-gated ion channel (GLIC) has served as a valuable structural and functional model for the eukaryotic Cys-loop receptor superfamily. In Cys-loop and other receptors, we have previously demonstrated the crucial roles played by several conserved prolines. Here we explore the role of prolines in the gating transitions of GLIC. As conventional substitutions at some positions resulted in nonfunctional proteins, we used in vivo non-canonical amino acid mutagenesis to determine the specific structural requirements at these sites. Receptors were expressed heterologously in *Xenopus laevis* oocytes, and whole-cell electrophysiology was used to monitor channel activity. Pro119 in the Cys-loop, Pro198 and Pro203 in the M1 helix, and Pro299 in the M4 helix were sensitive to substitution, and distinct roles in receptor activity were revealed for each. In the context of the available structural data for GLIC, the behaviors of Pro119, Pro203, and Pro299 mutants are consistent with earlier proline mutagenesis work. However, the Pro198 site displays a unique phenotype that gives evidence of the importance of the region surrounding this residue for the correct functioning of GLIC.

3.2 Introduction

The Cys-loop family of ligand-gated ion channels (LGICs) includes nicotinic acetylcholine receptors (nAChRs), serotonin receptors (5-HT_3Rs), $GABA_A$ receptors ($GABA_A$Rs), and glycine receptors (GlyRs). These receptors are pentameric membrane-spanning proteins that mediate fast synaptic transmission throughout the nervous system.[1,2]

This chapter is adapted with permission from:* Rienzo, M., Rocchi, A.R., Threatt, S.D., Dougherty, D.A., and Lummis, S.C.R. Perturbation of Critical Prolines in *Gloeobacter violaceus* Ligand-Gated Ion Channel (GLIC) Supports Conserved Gating Motions Among Cys-Loop Receptors. *J. Biol. Chem.* **291, 6272-6280 (2016).

They are members of a larger family, the pentameric ligand-gated ion channels (pLGICs), all of which share a pentameric structure and a similar global layout. Individual subunits contain a large extracellular N-terminal domain (ECD) and four transmembrane helices (M1–M4). Neurotransmitters generally bind in the ECD, at the interface between two adjacent subunits, triggering a conformational change that opens the pore, which is lined by the five M2 helices. The transduction of binding information to the pore region is not yet fully understood, partly because structural data for this family have lagged behind those for many other ion channels. However, relevant data are now becoming available, and recent publications describing the high-resolution structures of the first mammalian Cys-loop receptors ($GABA_A$, $5\text{-}HT_3$, and glycine receptors) have contributed to our growing knowledge.[3-5] Nevertheless, obtaining structural data on these proteins is challenging, and many in the field have turned to prokaryotic pLGICs[6,7] as model systems due to their facile expression in heterologous systems and relative ease of crystallization. The most widely used pLGICs are those from *Erwinia chrysanthemi*, now known as *Dickeya dadantii* (*Erwinia* ligand-gated ion channel; ELIC)[8] and *Gloeobacter violaceus* (*Gloeobacter* ligand-gated ion channel; GLIC).[9–14] These prokaryotic receptors share the general structure of the Cys-loop receptors, but lack the large, variable intracellular domain as well as the cysteines in the eponymous Cys-loop. Electrophysiological characterization has revealed that these prokaryotic channels also share many functional attributes with the eukaryotic pLGICs, including conductance, cation selectivity, sensitivity to similar channel blockers, and comparable effects of pore mutations.[7,15] Over 40 structures of GLIC have been obtained under various crystallization conditions, and they are presumed to represent several different thermodynamic states of the receptor. This has allowed for evidence-based predictions of global gating transitions

conserved among the pLGICs. GLIC thus offers a testing ground for corroboration of functional data from mutant receptors with hypothetical conformational changes.

Unusual among the pLGICs, GLIC is activated by a decrease in extracellular pH, rather than by a small-molecule ligand. The receptor is also unique in having proline residues that are not found in eukaryotic pLGICs in a number of locations. Proline is set apart from the other amino acids by its limited hydrogen bonding capability, increased steric bulk at the backbone nitrogen, and greater propensity to exist in a cis conformation at the backbone peptide bond. These features are essential for function in many proteins, including some eukaryotic Cys-loop receptors.[16–19] Understanding the roles of the proline residues in GLIC could allow for elucidation of structural transitions during activation and could help clarify whether GLIC is an appropriate model system to understand structure-function relationships in vertebrate Cys-loop receptors. Here we have substituted each of the proline residues in GLIC with alanine to identify any sensitive sites. Because conventional mutagenesis at some positions appears to ablate GLIC expression, trafficking, and/or function, we have also implemented the subtler probe of non-canonical amino acid mutagenesis, which permits small systematic perturbations to steric and electrostatic interactions. We identify several trends that agree with those previously observed in corresponding sites of the eukaryotic receptors,[17–19] as well as several unusual effects that reveal regions important for the activation process.

3.3 Results

3.3.1 Conventional mutagenesis identifies four sensitive proline sites

GLIC contains 18 proline residues, 12 of which are also present in related receptors (**Figure 3.1**). Upon mutation to alanine, 14 sites gave functional receptors with only minor decreases in pH_{50} (**Figure 3.2**), suggesting that these residues are not critical for channel

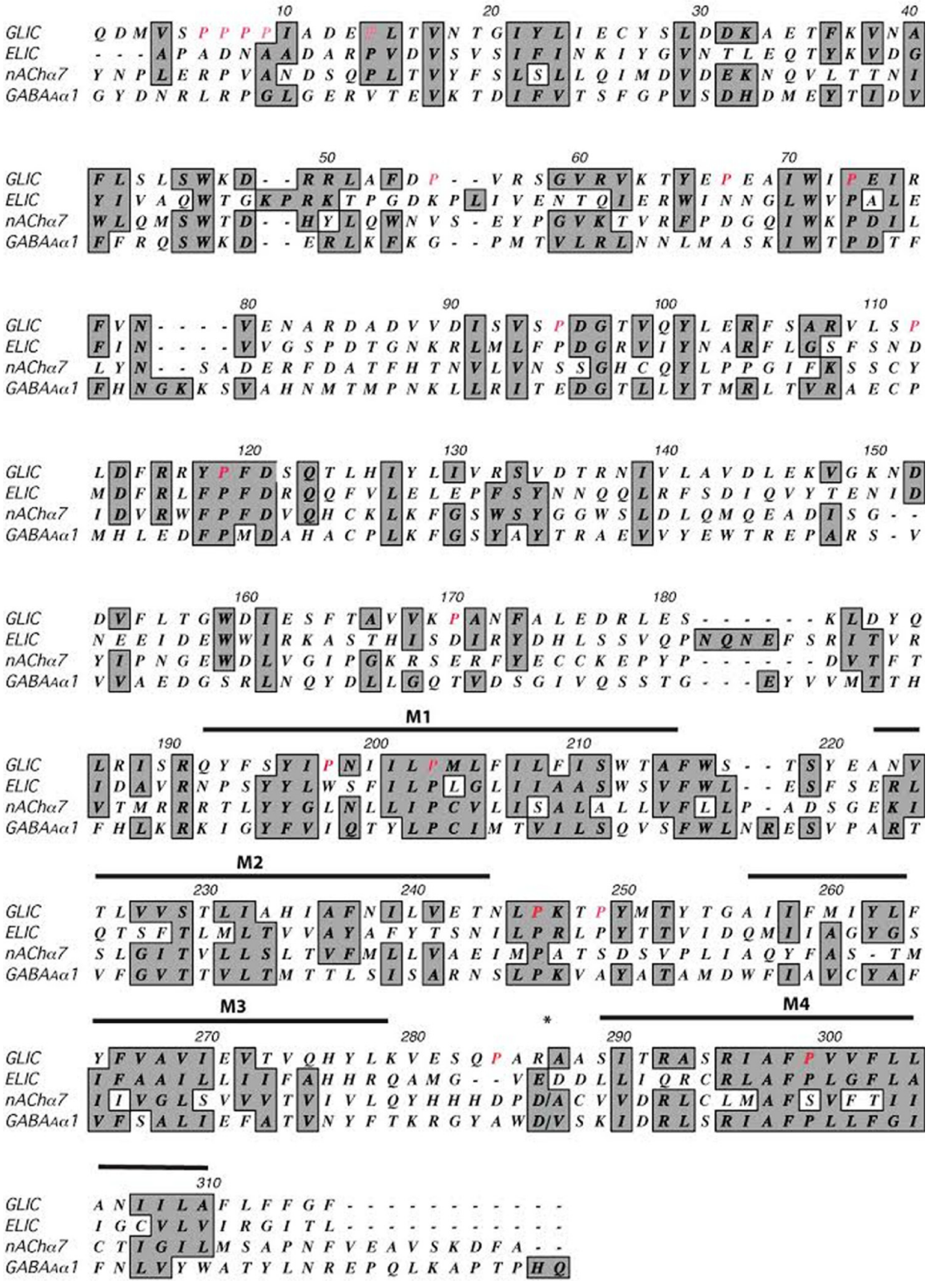


Figure 3.1. Sequence alignment of GLIC and eukaryotic pLGICs. Prolines examined in this study are shown in red. Shaded residues are conserved or similar in side-chain character among sequences. The M3/M4 loop of the eukaryotic receptors has been removed at *.

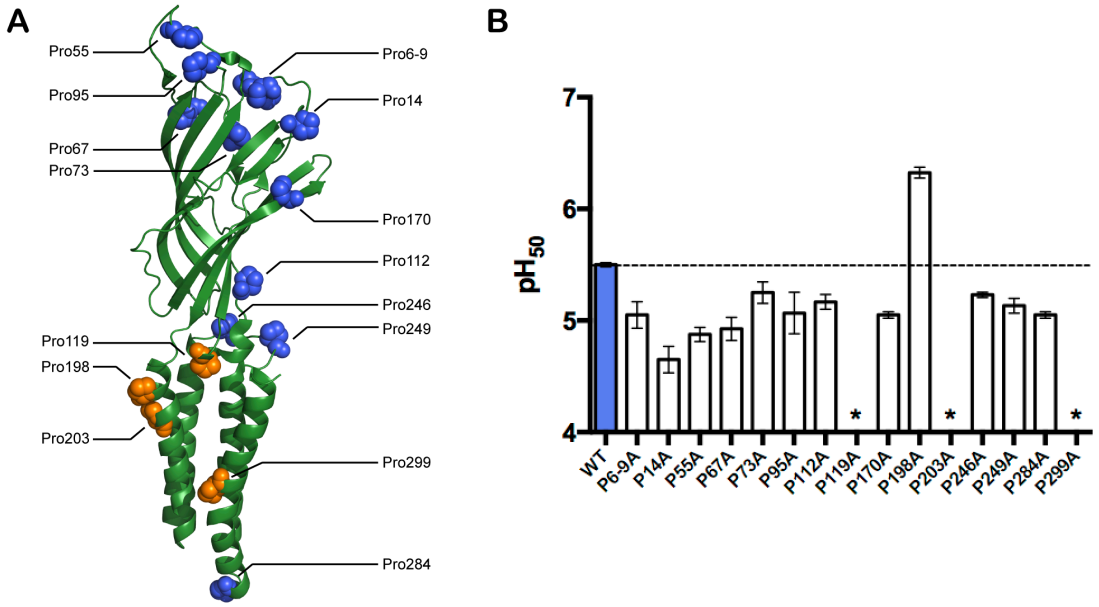


Figure 3.2. Prolines sites examined in this study. (A) GLIC subunit structure, indicating proline residues (PDB 3EHZ). (B) Conventional mutagenesis of GLIC Pro residues. Substitution with Ala mostly had little or no effect on pH_{50} values, although P119A, P203A, and P299 were nonfunctional, and P198A had increased sensitivity. P6-9A indicates simultaneous substitution of prolines at sites 6, 7, 8 and 9 with Ala; individual Ala substitutions at these sites yielded similar pH_{50} values (not shown). Data shown are mean ± SEM, n = 3-4. *pH-induced responses were comparable to uninjected cells. Typical maximal responses for WT and all Ala-containing mutants were 20-40 μA, except for P14A, where maximal responses were 2-5 μA.

gating. However, conventional mutations at 3 conserved sites (Pro119, Pro203, and Pro299) ablated acid-induced currents, and mutation at Pro198, unique to GLIC, resulted in an increased sensitivity to protons (i.e. higher pH_{50}). We therefore proceeded to examine these four sites more closely with the aid of non-canonical amino acid analogs.

3.3.2 The Cys-loop Tyr/Phe-Pro motif shows sensitivity to Pro cis bias and hydrophobicity

The Cys-loop, a disulfide-containing region at the ECD/TM interface, is a unifying trait of the eukaryotic pLGICs and has long been thought to contribute to transduction of ligand-binding information to the transmembrane domain.[16,23,24] Although GLIC lacks the two cysteine residues usually found in this loop, the region is structurally similar to the vertebrate receptors. The first proline we identified as essential for function, Pro119, is located in this loop, and is a highly conserved feature of the Cys-loop of the eukaryotic pLGICs (**Figure 3.3A**). Notably, this proline is generally flanked by 2 aromatic residues. It is

well established that aromatic residues immediately preceding proline sites preferentially stabilize the cis backbone conformation of the proline.[17,25] Indeed, the highest-resolution X-ray crystal structure of GLIC shows this proline in the cis configuration.[26] In other pLGIC structures, including one of GLIC, this proline is in a trans conformation,[4, 8–10] but this is a subtle distinction that may be beyond the resolution of some structures. At present, we cannot rule out the possibility that both conformers can exist in functional pLGICs, bringing up the intriguing possibility that they may be linked to different receptor states.

Previous studies from our group have investigated the potential of various proline residues in membrane receptors to undergo cis-trans isomerization as a step in the process of channel gating, including a site homologous to Pro119 in a nAChR.[16,17] To determine whether such an isomerization at Pro119 might be involved in GLIC gating, we performed mutagenesis with a number of canonical and non-canonical amino acids. Our data revealed this proline to be highly sensitive, yielding no functional receptors when substituted with Ala, Leu, or Gly (**Table 3.1**). Many prolines serve to disrupt secondary structure, because their backbone nitrogen lacks the hydrogen bond-donating site present in the other amino acids. To see whether this condition would be sufficient to allow the channel to function, we introduced Aah and Lah, the α-hydroxy acid analogs of Ala and Leu (**Figure 3.4**), which also lack a backbone hydrogen bond-donor. Such mutations at Pro sites often produce fully functional receptors. However, at Pro119 of GLIC these analogs resulted in no functional responses. This is a phenotype that would be expected if interconversion between cis and trans forms were required for channel activation because all of these substitutions would disfavor isomerization, but it is also consistent with a cis orientation being essential for function. Interestingly, two analogs with increased intrinsic cis-biases, the six-membered rings Pip and Mor, engendered an increase in sensitivity to acid. This effect was similar to

Table 3.1. Unnatural proline mutagenesis in GLIC. pH_{50} values and Hill coefficients of unnatural proline mutations examined in this study.[a]

Site	Mutation	pH_{50}	Hill Slope	Site	Mutation	pH_{50}	Hill Slope
Pro119	Ala	NR	NR	Pro203	Ala	NR	NR
	Leu	NR	NR		Leu	NR	NR
	Gly	NR	NR		Gly	NR	NR
	Pro	5.40 ± 0.06	1.6 ± 0.1		Pro	5.61 ± 0.01	1.9 ± 0.1
	Aah	NR	NR		Aah	NR	NR
	Lah	NR	NR		Lah	NR	NR
	N-MeAla	NR	NR		N-MeAla	5.24 ± 0.03	1.7 ± 0.1
	Aze	NR	NR		Aze	NR	NR
	Pip	6.34 ± 0.02	2.1 ± 0.1		Pip	NR	NR
	Mor	6.35 ± 0.02	3.1 ± 0.1		Mor	4.73 ± 0.02	2.0 ± 0.1
	P3m[b]	6.47 ± 0.01	1.5 ± 0.1		P3m[c]	< 4.5	
	Dhp	NR	NR		Dhp	5.12 ± 0.04	1.4 ± 0.1
	Flp	5.58 ± 0.07	1.7 ± 0.2		Flp	NR	
	flp	NR	NR		flp	5.63 ± 0.02	2.0 ± 0.1
Pro198	Ala	6.70 ± 0.03	0.7 ± 0.1	Pro299	Ala	NR	
	Leu	6.37 ± 0.03	2.4 ± 0.2		Leu	NR	
	Gly	6.48 ± 0.02	2.4 ± 0.2		Gly	NR	
	Pro	5.64 ± 0.02	1.6 ± 0.1		Pro	5.52 ± 0.02	1.6 ± 0.1
	Aah	5.29 ± 0.02	1.4 ± 0.1		Aah	NR	
	Lah	5.33 ± 0.05	1.4 ± 0.1		Lah	5.44 ± 0.02	1.3 ± 0.1
	N-MeAla	5.33 ± 0.05	1.2 ± 0.1		N-MeAla	5.53 ± 0.05	1.5 ± 0.1
	Aze	5.88 ± 0.04	1.7 ± 0.1		Aze	NR	
	Pip	5.53 ± 0.03	1.2 ± 0.1		Pip	5.77 ± 0.03	1.6 ± 0.1
	Mor	5.93 ± 0.02	1.7 ± 0.1		Mor	5.75 ± 0.02	1.6 ± 0.1
	P3m	5.76 ± 0.06	1.2 ± 0.1		P3m	5.60 ± 0.03	1.4 ± 0.1
	Dhp	5.42 ± 0.03	1.3 ± 0.1		Dhp	5.41 ± 0.04	1.4 ± 0.1
	Flp	6.10 ± 0.05	1.5 ± 0.1		Flp	6.01 ± 0.02	2.2 ± 0.1
	flp[b]	6.52 ± 0.02	1.9 ± 0.2		flp	5.91 ± 0.01	1.6 ± 0.1

[a]Data shown are mean ± SEM. NR indicates that acid-induced responses were comparable to those recorded from uninjected oocytes. Ala, Leu, and Gly mutants were generated by conventional mutagenesis. All other mutants were generated via nonsense suppression.
[b]A biphasic curve was observed with a low-sensitivity component having $pH_{50} < 4.5$.
[c]Currents were observed, but pH_{50} was too low to measure (< 4.5), due to nonspecific acid-induced currents.

that which was observed in the nAChR (**Table 3.1**; **Figure 3.3B**). We considered an additional analog, Aze, which is also expected to have an increased cis-bias, but we were unable to observe receptor function when we attempted to incorporate Aze.

In the case of the nAChR, receptor function markedly favored hydrophobicity at both the proline and the preceding aromatic residue.[17] Introduction of methyl groups on either side chain resulted in lower EC_{50} values than were expected based on the proline cis preference, whereas the more polar Mor and flp gave the opposite effect. Although a similar

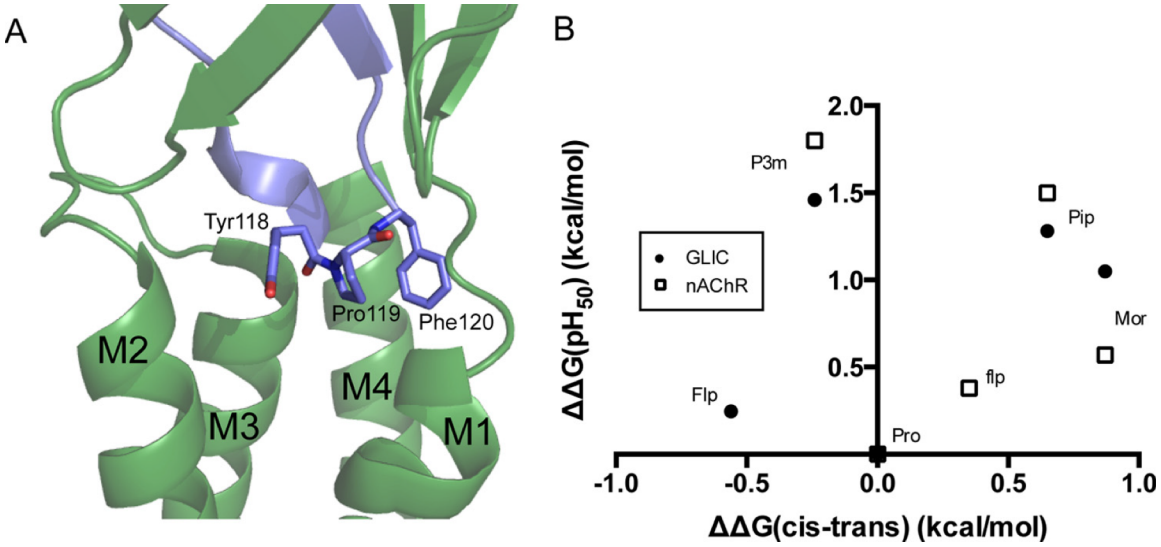


Figure 3.3. The GLIC YPF motif. (A) The TM/ECD interface in GLIC, highlighting the YPF motif in the cis (PDB 3EAM). (B) Relationship between mean pH_{50} values and cis-trans preferences for Pro analogues at position 119 in GLIC (filled circles), compared with the analogous position (Pro136) in the muscle-type nAChR (open squares).[17] Typical maximal currents for functional mutants generated by nonsense suppression were 0.5-2 μA, with P119Pip giving responses as high as 10 μA.

trend was observed with P3m and Flp at the GLIC site, Mor showed nearly the same pH_{50} as Pip. The aromatic residue preceding Pro119 in GLIC is tyrosine, but it is phenylalanine in the nAChR and other Cys-loop receptors. We therefore repeated the experiment in combination with the Y118F mutation to determine if this difference in polarity could account for the altered hydrophobicity effect. In the Y118F background, the Mor- and Pip-containing mutants still yielded very similar pH_{50} values (**Table 3.2**), indicating that other factors intrinsic to GLIC must be responsible for the observed phenotype.

Flp and flp maintain roughly equivalent steric and polarity contributions, mainly varying in the direction of pucker in the cyclic side chain.[27] This effect can in turn influence the cis-trans equilibrium of the proline, as has been previously shown in a model peptide containing a Tyr-Pro motif; Flp, the trans stereoisomer, favors an exo pucker and has a strong trans bias (relative to Pro, $\Delta\Delta G$ = –0.56 kcal/mol), while the epimer flp favors an endo pucker and a relative cis bias ($\Delta\Delta G$ = 0.35 kcal/mol).[28] In contrast to the nAChR, we observed currents only from receptors containing Flp, the trans stereoisomer, and not from

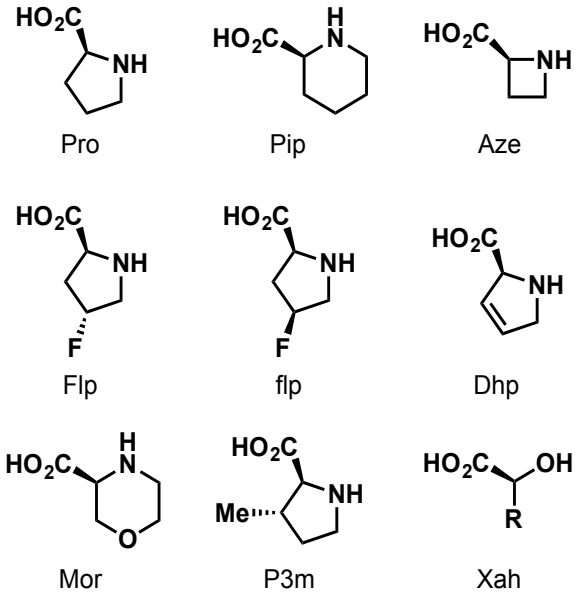


Figure 3.4. Structures of amino acid analogs used in this study.

flp. Based on the structures showing either cis or trans conformations at this site, the fluoro group on Flp is expected to point into the Tyr side chain, whereas the fluoro group on flp should point in the opposite direction, toward the subsequent Phe. This suggests a high degree of sensitivity not only to the polarity of the side chain, but also to the particular orientation and localization of the dipole.

A number of the mutations at Pro119 resulted in negligible currents, and so disruption of surface expression must also be considered as a possibility, especially for the more dramatic conventional mutations. However, as several proline analogs that introduce moderate steric perturbations are readily tolerated, it seems most likely that the lack of current seen in other cases is due to an effect on gating transitions.

3.3.3 A helical distortion in M1 is involved in activation

The second conserved proline that was intolerant to alanine substitution, Pro203, is located in the M1 helix. A similar proline is present in all pLGICs, and has been shown to be essential for receptor activity in nAChRs and 5-HT_3 receptors.[18,19] In the Cys-loop receptors

Table 3.2 Mutant cycle analyses for interactions at the extracellular terminus of M1.[a]

Mutant	pH_{50}	Hill Slope	Double Mutant	pH_{50}	Hill Slope	$\Delta\Delta\Delta G_{int}$ (kcal/mol)
wild type	5.50 ± 0.02	1.4 ± 0.1				
S195Sah	5.15 ±0.03	1.7 ± 0.1				
P198L	6.37 ± 0.03	2.4 ± 0.2	S195Sah/P198L	5.76 ± 0.05	1.8 ± 0.1	-0.35
N199A	4.64 ± 0.02	1.5 ± 0.1	P198L/N199A	6.40 ± 0.02	2.1 ± 0.1	1.21
N199L	5.21 ± 0.02	1.6 ± 0.1	P198L/N199L	5.98 ± 0.03	2.5 ± 0.1	-0.14
P119Pip	6.34 ± 0.02	2.1 ± 0.1				
P119Mor	6.35 ± 0.02	3.1 ± 0.1				
Y118F	4.88 ± 0.03	1.9 ± 0.1				
Y118F/P119Pip	5.34 ± 0.04	1.2 ± 0.1	Y118F/P119Mor[b]	5.31 ± 0.04	1.3 ± 0.1	0.19

[a]$\Delta\Delta\Delta G_{interaction}$ was calculated from the formula:

$$\Delta\Delta\Delta G_{\text{interaction}} = -RT\ln\left(\frac{EC_{50}(WT) \bullet EC_{50}(Mut1,2)}{EC_{50}(Mut1) \bullet EC_{50}(Mut2)}\right)$$

where $EC_{50} = 10^{-pH_{50}}$.

[b]Interaction energy was calculated using P119Pip, rather than wild type, as a reference point.

this proline displays a distinctive phenotype: conventional mutants give a nonfunctional channel, but incorporation of α-hydroxy residues, regardless of side chain, yields robustly expressing, ligand-activated receptors. This establishes that of the several unique features of a proline residue, lack of a backbone hydrogen bond-donor is the essential requirement here.

In GLIC, conventional mutants at Pro203 similarly gave negligible current responses. Surface expression of the P203L mutant was verified by western blotting of stripped oocyte membranes (data not shown), indicating that mutations at this site are capable of rendering the receptor nonfunctional, rather than interrupting assembly or trafficking. However, we were unable to rescue function by ablation of the backbone NH; α-hydroxy analogs gave nonfunctional channels. In fact, Pro203 proved to be extremely sensitive to substitution (**Table 3.1**; **Figure 3.5D**). Even the very close proline analogs Pip and Aze did not produce functional receptors. In contrast, cells injected with N-methylalanine (N-MeAla), which ablates the backbone NH but also lacks proline's cyclic side chain, did yield large currents. The substituted proline analogs P3m and flp, as well as the unsaturated Dhp, also gave functional channels.

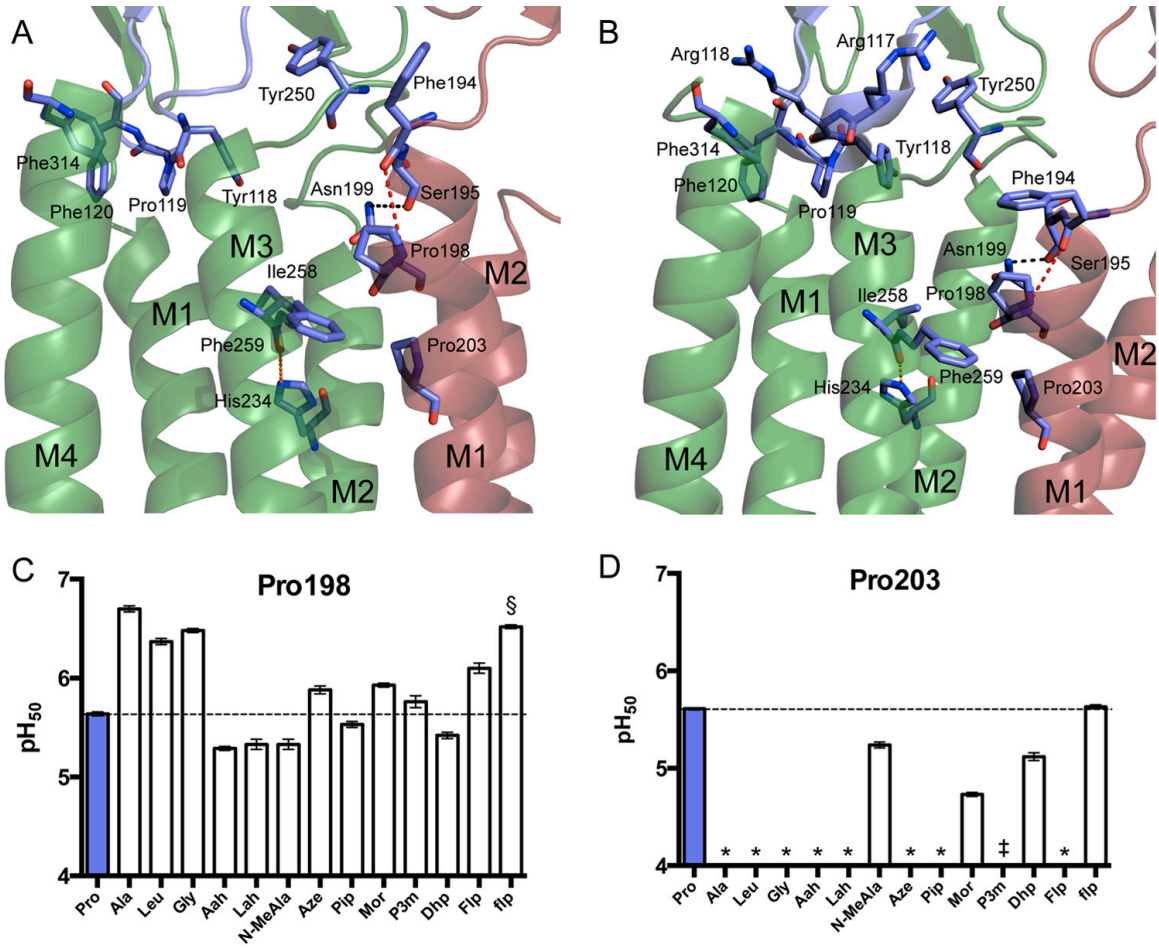


Figure 3.5. The GLIC M1 prolines. (A and B) Pro198 and Pro203 in the top of the M1 helix at pH 7 (PDB 4NPQ) (A) and pH 4 (PDB 3EHZ) (B). Shown are a hydrogen bond between Asn199 and the main-chain carbonyl of Ser195 (black), the disrupted hydrogen bond hypothesized to be restored by conventional mutagenesis at Pro198 (red), and the interhelix hydrogen bond between His234 and the Ile258 carbonyl (green). (C and D) pH_{50} values for GLIC Pro198 (C) and Pro203 (D) mutants. Data shown are mean ± SEM, with n = 8-20. Conventional mutants at Pro198 gave currents comparable with wild type GLIC. Typical maximal currents from functional mutants generated by nonsense suppression were 0.5-3 μA, with P198Lah giving responses as high as 20 μA. *pH-induced responses comparable to uninjected cells; §a biphasic curve was observed with a low-sensitivity component having pH_{50} < 4.5; ‡currents were observed, but pH_{50} was too low to measure (< 4.5), due to nonspecific acid-induced currents.

This difference in phenotype between GLIC and the eukaryotic Cys-loop receptors could be related to a unique feature of the GLIC M1 helix: a second proline, Pro198, located a single helix turn away from Pro203. The presence of these two juxtaposed prolines results in a more dramatic helix deformation than is normally observed in structures of pLGIC M1 helices (**Figure 3.5A**,**B**). Pro198 mutants generally retained function (**Figure 3.5C**). Most perturbations at Pro198 that preserved proline's lack of a backbone hydrogen bond-donor displayed a pH_{50} within 0.3 units of the wild type receptor. However, mutations to the

canonical amino acids Gly, Ala, or Leu resulted in a gain of function (activation at lower acid concentrations). These mutants could potentially allow formation of a backbone hydrogen bond between the introduced NH and the carbonyl of Phe194, perhaps stabilizing the M1 helix and favorably influencing the open conformation. To test this hypothesis, we weakened the hydrogen bonding ability of the Phe194 carbonyl via incorporation of an α-hydroxy acid at the adjacent residue, Ser195. The resulting mutant P198L/S195Sah does indeed show a more basic pH_{50} value than that of the single mutant P198L (**Table 3.2**). However, examination of the S195Sah mutation alone reveals that the effects of the two mutations are nearly additive, suggesting that this hydrogen-bonding interaction may not be responsible for the observed gain of function in the conventional Pro198 mutants.

In another effort to examine the importance of hydrogen bonding in this region, we mutated Asn199, the side chain of which forms a strong hydrogen bond with the Ser195 main-chain carbonyl in the crystal structure (**Figure 3.5A**,**B**). Although the isosteric N199L mutant produced only a minor effect, removal of the bulky side chain in N199A dramatically reduced pH_{50}, indicating important steric requirements in the intersubunit region immediately surrounding Pro198. Furthermore, introduction of the N199A mutation in the presence of P198L results in no change in pH_{50} compared with the proline mutation alone. This suggests that the effects of a conventional mutation at Pro198 obviate the need for steric bulk at the side chain of Asn199, pointing to a modified intersubunit interaction as the origin of the gain of function.

3.3.4 The M4 Pro residue

Recent data have indicated that the M4 helix is a highly important region of pLGICs that is involved in allosteric modulation of receptor function via interaction with lipids, the

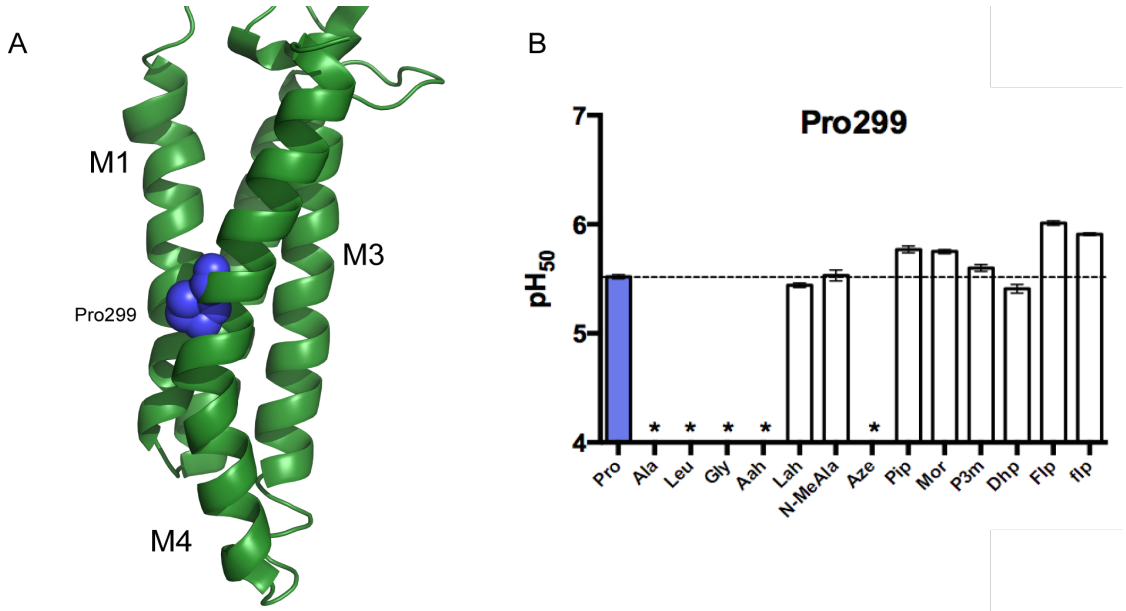


Figure 3.6. GLIC Pro299. (A) The GLIC TMD, highlighting Pro299 (PDB 3EHZ). (B) pH_{50} values for GLIC Pro299 mutants. Data shown are mean ± SEM, with n = 7-20. Typical maximal currents from functional mutants generated by nonsense suppression were 0.5.-5 μA. *pH-induced responses comparable with uninjected cells.

adjacent M1 and M3 helices, and the extracellular domain.[29,30] The final sensitive residue identified in GLIC, Pro299, is located in the middle of the M4 helix (**Figure 3.6A**), and has recently been identified as functionally important in an alanine screen of the M4 helix.[31] Although a homologous proline does not appear in the nAChRs or 5-HT_3Rs, it is conserved in the $GABA_A$Rs, GlyR, and ELIC. In each of the available structures containing a proline in this position, a clear and dramatic helix kink is observed. Conventional mutants at this site resulted in nonfunctional receptors. In contrast, with the exception of Aah and Aze, most analogs lacking a backbone NH rescued function, giving close to wild type characteristics (**Figure 3.6B**). Similar to the behavior typically observed for the M1 proline of Cys-loop receptors (but not in the case of GLIC), this pattern indicates that disruption of the helix by deletion of a hydrogen bond is sufficient for maintaining the observed kink, and therefore, channel function.

3.4 Discussion

Previous studies concerning the functions of conserved prolines in transmembrane receptors[16-19,32] have given rise to several commonly observed phenotypes: (1) those for which cis bias influences activity, (2) those for which ablation of a backbone hydrogen bond suffices to produce near wild type function, and (3) those for which steric bulk at the amine is necessary. Of the prolines examined above, Pro119 conforms to the first group, Pro299 conforms to the second, and P203 conforms to the last. Examination of the available GLIC structural data also reveals that the conformational effects conferred by these residues are consistent with those of previously studied prolines with the same phenotypes: Pro119 is located in a potentially flexible loop; Pro299 forms a kink in M4 due to simple ablation of a hydrogen bond; P203 exists below a prominent bulge and displays steric requirements we have previously reported for a conserved bulge-inducing proline in the D2 dopamine receptor.[32]

Although structural information from GLIC in a number of states has been obtained, it remains to be determined whether the gating motions observed in GLIC mirror the global conformational changes that govern activation of the eukaryotic Cys-loop receptors. Despite the lack of the disulfide in the loop containing Pro119, both the structure of the loop and the effects of cis bias and hydrophobicity at Pro119 closely resemble what we have observed at the analogous site in the nAChR. This supports the notion that during activation, any reorganization involving the Tyr/Phe-Pro motif is likely conserved, at least between nAChRs and GLIC. In the open GLIC structure (PDB 3EHZ), this loop directly contacts the M2-M3 loop via hydrophobic interactions and a cation-π interaction between Arg116 and Tyr250 (**Figure 3.5B**), which suggests that the conformation of the "Cys-loop" could be coupled to the gating motions of M2. As pH_{50} data for Aze could not be obtained,

it is difficult to conclude whether cis conformation at this site is required for channel function, or whether a cis-trans isomerization is linked to gating transitions. However, considering the failure of this site to tolerate residues without an appreciable cis contributor and the gain of function observed upon incorporation of several cis-biased analogs, one of these possibilities is likely.

The M1 proline corresponding to Pro203 in GLIC is conserved throughout the pLGIC superfamily, and in two Cys-loop receptors we find a common phenotype: any residue that lacks a backbone NH produces an essentially wild type receptor.[18,19] In GLIC, however, this proline faces somewhat more stringent requirements: the necessity for substitution at the backbone nitrogen of Pro203 suggests that the pronounced bulge in M1 observed in GLIC crystal structures must be enforced for activation to occur. Examination of the closed (pH 7) and open (pH 4) crystal structures[13,14] has identified a difference in the interface between the extracellular end of M1 and the M2-M3 loop of the adjacent subunit. In the open structures, the pre-M1 region is tucked under the adjacent M2-M3 loop, with the nonconserved Pro249 sandwiched between Phe194 and Gln192. In contrast, the closed structure shows Phe194 pointed away from M1 into lipid, while the M2-M3 loop has shifted so that Pro249 has moved past Gln192. It is likely that this shifting occurs concurrently with or prior to the bend observed in M2 in the open structures. Substitution of Pro203 with analogs bearing a hydrogen bond donor may result in a more ordered helical structure at the top of M1, eliminating any flexibility that is enabled by the presence of the bulge in the wild type receptor. This could prevent the motion of the M2-M3 loop, which may be required to initiate the bending of M2 observed in the open state.

Alternatively, or in addition, perturbation of Pro203 could interfere with the correct alignment of a hydrogen bond that we have previously shown to be crucial for stabilization

of the open state.[33] In that work, we showed that a backbone amide-to-ester substitution at Phe259 attenuates the hydrogen bond-accepting ability of the Ile258 carbonyl, and that tight steric complementarity makes this entire region highly intolerant to modification. The side chain of Phe259 is in van der Waals contact with both Pro198 and Pro203; thus the bulge created by Pro203 allows a direct interaction with the M3 helix of the adjacent subunit, which may be essential for function.

Interestingly, the nearby Pro198 does not seem to be required to maintain this disruption in M1, and does not suffice to rescue function when Pro203 is substituted. The distinct and robust increase in proton sensitivity conferred by conventional mutation of Pro198 lies in stark contrast with the extreme sensitivity of Pro203. Because mutations that introduce a backbone NH at Pro198 cause a gain of function, whereas those that maintain proline's lack of a hydrogen bond-donor produce receptors with wild type-like properties, it is tempting to argue that the effect we observe is simply due to an increased ordering of the N-terminus of M1. However, the nearly additive effects we observed in the mutant cycle analysis with S195Sah make it difficult to conclude whether this is the primary cause of the shift in activity. The P198L and N199A mutations, on the other hand, are clearly non-independent. The importance of stabilizing interactions in the intersubunit region near the top of M1 has previously been established by structural studies on ethanol potentiation of GLIC mutants.[34] Restructuring of this region caused by the conventional Pro198 mutations could potentially allow for formation of comparable interactions in the absence of ethanol, leading to the observed pH sensitivity increase.

If the available structures accurately represent the open and resting conformations of GLIC, it does not appear that any major changes occur in the conformation or position of the M4 helix between these states. However, recent mutagenesis experiments examining the

GLIC M4 helix have indicated that interactions of M4 with M1, M3, and the ECD are critical for receptor function.[31] In that study, mutations of two or more aromatic residues in this region at the interface with M1 or M3 resulted in nonfunctional receptors, and so it is most likely that elimination of the M4 kink simply prevents the helix from adopting a conformation where these crucial interactions can be maintained. It is also possible that the correct structure of the GLIC M4 helix is essential for folding and trafficking of receptors, as proper folding of M4 is required for cell-surface trafficking in some other pLGICs. Various nAChR subunits, for example, have endoplasmic reticulum retention motifs in the pre-M1 region, which are masked when M4 binds to M1/M3, permitting cell-surface expression.[35] However, these residues are not well conserved in GLIC, and it is unclear whether a similar mechanism is operative.

3.5 Conclusions

We have identified several highly sensitive proline residues in GLIC. We show that the gating energetics for GLIC appear to depend significantly on the structure of the YPF motif of the "Cys-loop", similar to what has previously been seen in eukaryotic nAChRs. We find that the highly conserved M1 proline maintains its critical role, albeit with a higher level of stringency than in the 5-HT_3Rs or nAChRs. Both these results help to validate GLIC as a relevant homolog in the study of the gating process of the Cys-loop receptors. The M1 and M4 sites obey previously observed phenotypes for bulge- and kink-inducing prolines, respectively, reinforcing our devised classification strategy. Finally, we highlight the importance of stabilizing interactions in the intersubunit space adjoining the extracellular terminus of M1 based on unusual effects of substitutions to the Pro198 site. Further studies

will be required to delineate the structural consequences of mutations at this position, and may illuminate poorly understood aspects of the activation process.

3.6 Experimental procedures

3.6.1 Molecular biology

The cDNA for GLIC was in the pGEMhe plasmid. Site-directed mutagenesis was performed using the Stratagene QuikChange protocol to generate the appropriate codon. For non-canonical amino acid mutants and conventional mutants generated by nonsense suppression, the site of interest was mutated to the TAG stop codon. Plasmids were linearized with the SbfI restriction enzyme, and receptor mRNA was then prepared by in vitro runoff transcription using the Ambion T7 mMessage mMachine kit.

Hydroxy or amino acid-dCA conjugates were enzymatically ligated to truncated 74mer THG73 tRNA as previously described.[20,21] The 74mer tRNA was prepared using the Ambion T7MEGAshortscript kit by transcription from a DNA oligonucleotide template modified to enhance RNA transcript homogeneity, as described in the literature.[22] Crude tRNA-amino acid or tRNA-hydroxy acid product was used without desalting, and the product was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry on a 3-hydroxypicolinic acid matrix. Deprotection of the NVOC group on the tRNA-amino acids was carried out by photolysis for 5 minutes on a 300-watt high-pressure mercury arc lamp with WG-335 and UG-11 filters immediately prior to injection.

3.6.2 Oocyte preparation and RNA injection

Stage V-VI oocytes of *Xenopus laevis* were harvested and injected with RNAs as described previously.[21] For nonsense suppression experiments, each cell was injected with 50-100 ng each of receptor mRNA and appropriate tRNA approximately 48 h before

recording. Mutants yielding small responses required 72 h of incubation, with a second injection of mRNA and tRNA 48 h before recording.

For wild type experiments and conventional mutants, each cell received a single injection of 1-25 ng of receptor mRNA approximately 24 h before recording. Injection volumes for each injection session were 50-100 nL per cell.

As a negative control for suppression experiments at each site, unacylated full-length tRNA was co-injected with mRNA in the same manner as charged tRNA. These experiments yielded negligible responses for all sites. Wild-type recovery conditions (injecting tRNA charged with the appropriate amino acid to regenerate a wild type channel via nonsense suppression at a TAG stop codon) were injected alongside mutant nonsense suppression conditions as a positive control.

3.6.3 Electrophysiology

Oocyte recordings were made in two-electrode voltage clamp mode using the OpusXpress 6000A (Axon Instruments). Oocyte equilibration and washes were performed with Ca-free ND96 (96 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 5 mM HEPES) adjusted to pH 8 with 1 N NaOH. The pH of buffers for concentration-response curves were adjusted accordingly with NaOH or HCl, and buffers with pH below 6.8 contained 5 mM MES in place of HEPES. Initial holding potential was -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Oocytes were equilibrated for 30 s at 1 mL/min before each pH application. The pH buffer applications lasted for 15 s at 4 mL/min, followed by a 15-s wait period to allow the cells to attain a peak current. Cells were then washed for 40 s at 3 mL/min before the following equilibration. Concentration-response data were obtained for 9 buffer pHs, for a minimum of two cell batches, and for a minimum of 7 cells total (or 3-4

cells from 1-3 batches for alanine scanning). The concentration-response relations for each cell were fitted to the Hill equation:

$$I([A]) = \frac{I_{\max}}{1+\left(\frac{[A]}{\mathrm{EC}_{50}}\right)^{n_{\mathrm{H}}}}, \quad \text{(Eq. 3.1)}$$

where I_{norm} is the normalized current peak at $[H_3O^+] = A$, EC_{50} is the value of $[H_3O^+]$ that elicits a half-maximum response, and n_H is the Hill coefficient. The values for each cell were then averaged to give the reported values, where $pH_{50} = -\log(EC_{50})$. At pH < 4, uninjected cells gave current responses that were occasionally up to ~ 200 nA. Thus, data were only reported from cells that gave responses above pH 4, or that gave larger, more robust currents below pH 4, and other cells were considered non-responsive (NR).

3.7 References

1. Dacosta, C. J. B. & Baenziger, J. E. Gating of pentameric ligand-gated ion channels: structural insights and ambiguities. *Structure* **21**, 1271–1283 (2013).
2. Nys, M., Kesters, D. & Ulens, C. Structural insights into Cys-loop receptor function and ligand recognition. *Biochem. Pharmacol.* **86**, 1042–1053 (2013).
3. Miller, P. S. & Aricescu, A. R. Crystal structure of a human $GABA_A$ receptor. *Nature* **512**, 270–275 (2014).
4. Hassaine, G. *et al.* X-ray structure of the mouse serotonin 5-HT3 receptor. *Nature* **512**, 276–281 (2014).
5. Du, J., Lü, W., Wu, S., Cheng, Y. & Gouaux, E. Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526**, 224-229 (2015).
6. Tasneem, A., Iyer, L. M., Jakobsson, E. & Aravind, L. Identification of the prokaryotic ligand-gated ion channels and their implications for the mechanisms and origins of animal Cys-loop ion channels. *Genome Biol.* **6**, R4 (2004).
7. Bocquet, N. *et al.* A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature* **445**, 116–119 (2007).
8. Hilf, R. J. C. & Dutzler, R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* **452**, 375–379 (2008).
9. Hilf, R. J. C. & Dutzler, R. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* **457**, 115–118 (2009).
10. Bocquet, N. *et al.* X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* **457**, 111–114 (2008).
11. Gonzalez-Gutierrez, G., Cuello, L.G., Nair, S.K. & Grosman, C. Gating of the proton-gated ion channel from Gloeobacter violaceus at pH 4 as revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18716–18721 (2013).
12. Nury, H. *et al.* X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature* **469**, 428–431 (2011).
13. Prevost, M. S. *et al.* A locally closed conformation of a bacterial pentameric proton-gated ion channel. *Nat. Struct. Mol. Biol.* **19**, 642–649 (2012).
14. Sauguet, L. *et al.* Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 966–971 (2014).
15. AlQazzaz, M. *et al.* Cys-Loop Receptor Channel Blockers Also Block GLIC. *Biophys. J.* **101**, 2912–2918

(2011).
16. Lummis, S. C. R. *et al.* Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248–252 (2005).
17. Limapichat, W., Lester, H. A. & Dougherty, D. A. Chemical scale studies of the Phe-Pro conserved motif in the cys loop of Cys loop receptors. *J. Biol. Chem.* **285**, 8976–8984 (2010).
18. England, P. M., Zhang, Y., Dougherty, D. A. & Lester, H. A. Backbone Mutations in Transmembrane Domains of a Ligand-Gated Ion Channel. *Cell* **96**, 89–98 (1999).
19. Dang, H. *et al.* Probing the role of a conserved M1 proline residue in 5-hydroxytryptamine(3) receptor gating. *Mol. Pharmacol.* **57**, 1114–1122 (2000).
20. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Lett.* **40**, 6189–6192 (1999).
21. Nowak, M. W. *et al.* [28] In vivo incorporation of unnatural amino acids into ion channels in *Xenopus* oocyte expression system. *Method. Enzymol.* **293**, 504–529 (1998).
22. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268–1272 (1999).
23. Bouzat, C. *et al.* Coupling of agonist binding to channel gating in an ACh-binding protein linked to an ion channel. *Nature* **430**, 896–900 (2004).
24. Jha, A., Cadugan, D. J., Purohit, P. & Auerbach, A. Acetylcholine receptor gating at extracellular transmembrane domain interface: the cys-loop and M2-M3 linker. *J. Gen. Physiol.* **130**, 547–558 (2007).
25. Reimer, U. *et al.* Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* **279**, 449–460 (1998).
26. Sauguet, L. *et al.* Structural basis for ion permeation mechanism in pentameric ligand-gated ion channels. *EMBO J.* 32, 728–741 (2013).
27. Bretscher, L. E., Jenkins, C. L. & Taylor, K. M. Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc.* **123**, 777-778 (2001).
28. Pandey, A. K., Naduthambi, D., Thomas, K. M. & Zondlo, N. J. Proline editing: a general and practical approach to the synthesis of functionally and structurally diverse peptides. Analysis of steric versus stereoelectronic effects of 4-substituted prolines on conformation within peptides. *J. Am. Chem. Soc.* **135**, 4333–4363 (2013).
29. Hénault, C. M. *et al.* The role of the M4 lipid-sensor in the folding, trafficking, and allosteric modulation of nicotinic acetylcholine receptors. *Neuropharmacology* **96**, 157–168 (2015).
30. Carswell, C. L. *et al.* Role of the Fourth Transmembrane α Helix in the Allosteric Modulation of Pentameric Ligand-Gated Ion Channels. *Structure* **23**, 1655–1664 (2015).
31. Hénault, C. M., Juranka, P. F. & Baenziger, J. E. The M4 transmembrane α-helix contributes differently to both the maturation and function of two prokaryotic pentameric ligand-gated ion channels. *J. Biol. Chem.* 10.1074/jbc.M115.676833 (2015).
32. Van Arnam, E. B., Lester, H. A. & Dougherty, D. A. Dissecting the Functions of Conserved Prolines within Transmembrane Helices of the D2 Dopamine Receptor. *ACS Chem. Biol.* **6**, 1063–1068 (2011).
33. Rienzo, M., Lummis, S. C. R. & Dougherty, D. A. Structural Requirements in the Transmembrane Domain of GLIC Revealed by Incorporation of Noncanonical Histidine Analogs. *Chem. Biol.* **21**, 1700–1706 (2014).
34. Sauguet, L. *et al.* Structural basis for potentiation by alcohols and anaesthetics in a ligand-gated ion channel. *Nat. Comms.* **4**, 1697 (2013).
35. Wang, J.-M. *et al.* A transmembrane motif governs the surface trafficking of nicotinic acetylcholine receptors. *Nat. Neurosci.* **5**, 963–970 (2002).

Chapter 4: Tools for Exploring Acid-Sensing Ion Channel Function*

4.1 Abstract

We aim to understand the molecular basis of pH sensing and channel gating in the acid-sensing ion channels (ASICs). ASICs are trimeric ion channels found in the central nervous systems of chordates, with implications in pain sensation, ischemic stroke, and cognitive function. Crystal structures have provided clues to the origins of ion selectivity in ASIC1a, and we have tested these models using α-hydroxy analogs of pore-lining amino acids. We also use fluorination of an aromatic amino acid to determine molecular factors involved in the selectivity of the active component of a tarantula venom, psalmotoxin, for ASIC1a over other ASIC subtypes. Based on structural data, ASIC1a is thought to detect pH changes by binding protons among several pairs of carboxylic acid residues. We planned to investigate these sites using fluorinated variants of glutamate and aspartate, and discuss progress toward the syntheses of these probes.

4.2 Introduction

ASICs are a class of trimeric, membrane-spanning proteins[1] and have been implicated in ischemic stroke, pain perception, mechanosensation, learning, and memory.[2-5] These channels are members of the amiloride-sensitive epithelial sodium channel/degenerin (ENaC/DEG) family, and are expressed in neurons throughout the central and peripheral nervous systems in multicellular eukaryotes.[6] ASICs are classified by sequence homology, and despite their name, are not all acid-sensitive.[7-9] There are four known genes coding for ASICs in humans (ASIC1-4), three of which exist in different splice variants.[6] The subunits can form functional homomeric and heteromeric channels (with the exception of ASIC2b

Several experiments in this chapter were conceived by a group alum, Dr. Angela Blum, as part of her final propositions examination. The related sections are indicated by asterisks.

and ASIC4, which are only functional in heteromers with other subunits), which exhibit different acid sensitivities, ion selectivities, and desensitization kinetics.[10] Homomers and heteromers are both formed under native conditions.

Several nonspecific inhibitory ligands for ASICs exist, including amiloride, aspirin, and ibuprofen. Specific peptide binders of ASICs include psalmotoxin 1 (PcTx1), a component of the venom from the Trinidad chevron tarantula that targets ASIC1a homomers;[11] mambalgins, a class of three-finger peptide toxins from the black mamba that target ASIC1a homomers and heteromeric ASIC1a/2a or ASIC1a/2b channels;[12] and MitTx, the active species in Texas coral snake venom, which shows some selectivity for ASIC1a and 2a homomers.[13] It has been shown that activation of ASIC1a during ischemic stroke plays a role in neuronal damage, and treatment of cells with PcTx1 to inhibit the channel has demonstrated a neuroprotective effect.[14] Thus, understanding the mechanisms by which ASIC1a detects extracellular acidosis and is inhibited by peptides like these may lead to improved therapies for ischemic stroke and chronic pain.

This project focuses first on understanding the function of the homomeric ASIC1a channel, due to its physiological interest and the availability of several X-ray crystal structures.[15-19] ASIC1 is the most abundant gene of its family in the central nervous system.[20] It gives rise to two protein subunits, ASIC1a and ASIC1b, although the latter protein is only produced in the peripheral nervous system.[21] Although we are primarily interested in homomers of ASIC1a, the ASIC1a subunits are also known to form heteromers with subunits of other types.[22] ASIC1a homomers are the most acid-sensitive of the ASIC channels, displaying current responses at pHs as high as 6.9.[23] Currents are sharply dependent upon pH, with maximum whole-cell responses observed at around pH 6 (**Figure 4.1**). The high Hill coefficient of ASIC1a (~5.5) is characteristic of channels that bind a large

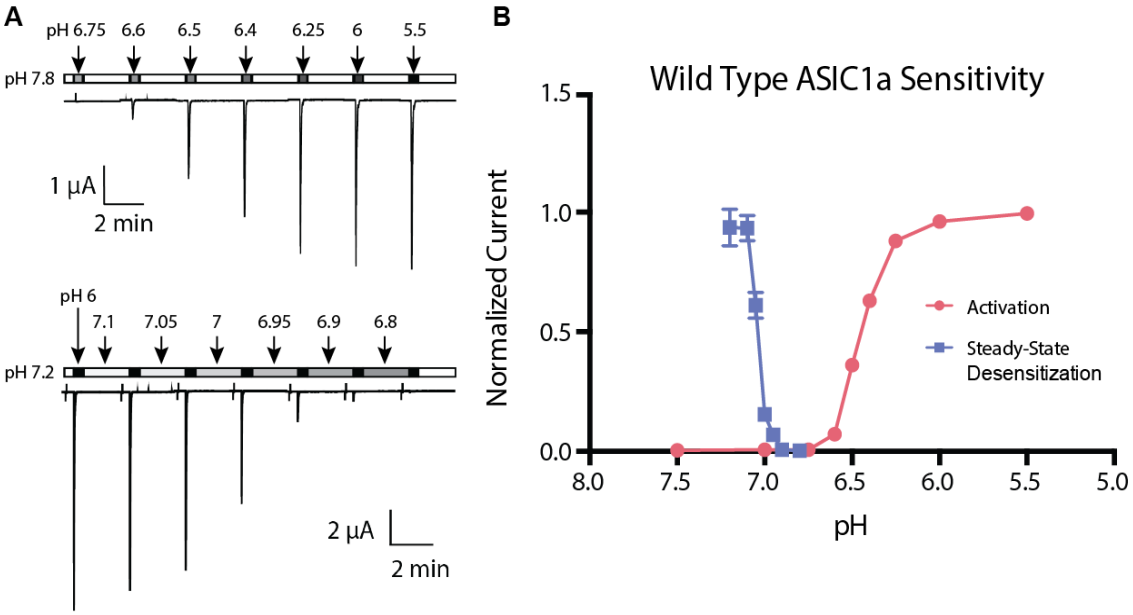


Figure 4.7. Acid-sensitivity of wild type ASIC1a. (A) Representative whole-cell current traces from two-electrode voltage clamp of *Xenopus laevis* oocytes expressing ASIC1a. The top trace shows the observed current upon application of decreasing pH buffers, equilibrating with pH 7.8 buffer between test applications. The bottom trace shows the steady-state desensitization achieved after equilibrating in decreasing pH buffers, assaying the proportion of activatable channels with applications of pH 6 test buffer. (B) Line graph showing the peak current data collected from traces such as those shown in (A). Mean normalized currents are plotted from 8-12 cells, with standard errors indicated.

number of ligands in a cooperative fashion, and has been interpreted as an indication of a large number of proton-binding sites in the extracellular domain. Upon prolonged exposure to acid, these channels undergo a second conformational change into a non-conducting state, called the desensitized state. Channels in this state are not visibly activated by changes in pH, and must be incubated at high pH before a macroscopic response can be observed again. ASIC1a desensitizes with a time constant of 1-2 s, recovering from desensitization after several seconds at elevated pH.

This desensitization complicates the thermodynamic model generally used for LGICs, since it is unclear whether the channel population reaches an equilibrium between the closed and open states before channels begin to desensitize. If no equilibrium is reached, the maximum current observed will not necessarily provide accurate thermodynamic information about the relative energies of the closed and open states. Instead, it provides a kinetic parameter – a quantity proportional to the maximum number of receptors open

during the course of the treatment. In spite of these complications, most reports of ASIC1a biophysical properties fit the maximum observed current responses at each pH to the Hill equation, assuming a traditional binding model.

In addition to desensitization following current responses, ASIC1a also exhibits desensitization without any observable macroscopic current. In other words, after equilibration at pHs below 7.4, treatment with pH 6 buffer triggers a diminished current response.[20,21] This has been termed steady-state desensitization (SSD), and can also be fit to a Hill function. There is evidence supporting a model in which channels proceed through the open state to the desensitized state, wherein the accumulation of open channels is small enough that currents are invisible to a whole-cell voltage clamp. Namely, mutations that shift the EC_{50} of ASIC1a also often have comparable effects on the SSD curve.[23] However, treatment of ASIC1a with mambalgin-1 appears to cause inhibition of acid-induced currents without influencing the SSD curve, suggesting that SSD may indeed occur by an independent pathway, perhaps without involving any open state.[12]

Members of the ENaC/DEG family share a common general structure (**Figure 1.3**). Each subunit comprises two transmembrane helical domains linked by a large extracellular region, with short intracellular N- and C-termini.[24] Overall, each ASIC subunit is approximately 450 amino acids in length. There are several crystal structures of chicken ASIC1a truncations available, both at high and low pH, including several with peptide toxins bound.[15-19] The three subunits are arranged about the pore with 3-fold symmetry. Each subunit resembles an upright forearm and fist, with the extracellular domain connected to the transmembrane helices at the base of a large β-sheet domain. Intersubunit contacts are formed primarily between the palm and thumb domains. At this interface are located several acidic residues expected to be titrated by diminished pH, providing a pH-sensing mechanism

for the receptor. A number of highly conserved disulfide linkages give structural stability to the thumb domain, which is thought to play a role in conveying conformational changes from these proton binding sites to the transmembrane helices.[15] A constriction in the transmembrane domain in the low-pH structures has been suggested to form the occlusion that prevents ion permeation in the desensitized state. Three lateral openings provide access to the central pore just above a cavity in the extracellular region along the central axis.[16]

Numerous mutagenesis experiments have been used to identify important structural features of ASICs and uncover their roles in the mechanism of action of these channels. In this chapter, we discuss our efforts to approach the same problem, but using the more subtle probes made available by non-canonical amino acid mutagenesis.

4.3 Results and discussion

*4.3.1 Exploring the origins of sodium selectivity**

ASICs and ENaCs are both sodium-selective, with G_{Na}/G_K ~ 13 and 100-500, respectively.[25] Prior to the first structural elucidation of ASICs, mutagenesis studies revealed that a transmembrane GAS motif highly conserved in Na-selective channels is crucial for maintaining this discrimination.[6] At the time we performed the studies described here, the available crystal structures of ASIC1a (PDB 3HGC, 3IJ4) explained little about the function of these residues, as they appear to be positioned at the side of the helix facing away from the channel pore.[15,16] However, later structures of ASIC1a bound to Texas coral snake toxin, MitTx, revealed an unusual structural motif in the transmembrane domain that was not previously noticed by the authors: the C-terminal helix of each subunit is discontinuous, with the intracellular third positioned at the base of the helix of the adjacent subunit (**Figure 4.2A**).[19] The GAS motif connects these two portions of the helix, creating a ring around the channel pore. This unique feature is now believed to allow ENaC/DEG channels to attain

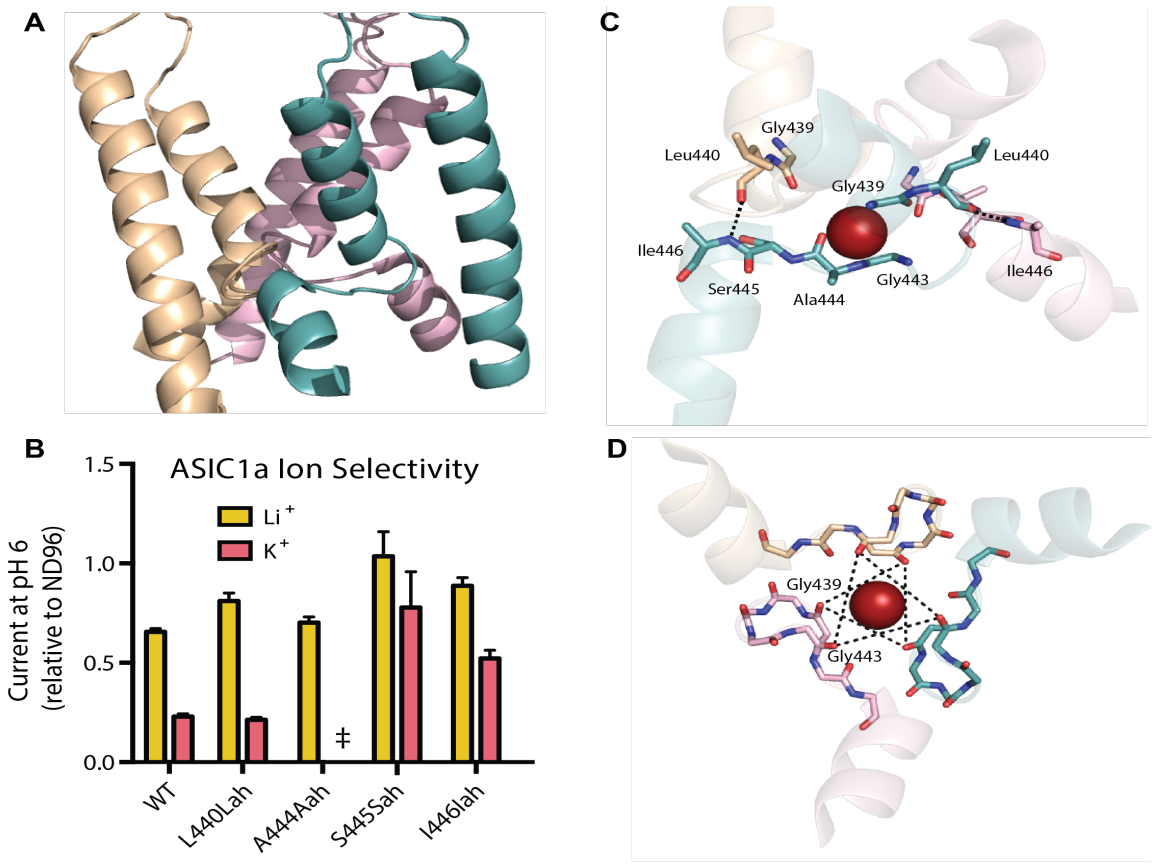


Figure 4.8. The ASIC1a selectivity filter. (A) Cartoon representation of the ASIC1a transmembrane domain, showing the discontinuous M2 helix from each subunit (PDB 4NTY). (B) Effects on ion selectivity from incorporation of α-hydroxy acids in the pore-lining region. Currents are normalized to the peak current at pH 6 in the presence of 96 mM Na^+ (ND96). Mean currents from 3-7 oocytes are plotted with standard errors indicated. ‡Currents were too small to reliably determine the relative permeability of potassium. (C) Side view of the transmembrane domain, highlighting residues predicted to be involved in ion coordination and sodium selectivity. A structurally important hydrogen bond perturbed by the I446Iah mutation is highlighted. (D) View of the selectivity filter from the cytoplasmic end of the pore. The two sets of glycine carbonyls thought to coordinate ions at this site are indicated.

high sodium selectivity, despite having only three transmembrane helices lining the pore. Though the studies in this section were designed before its discovery, our interpretation will assume this ring to be included in the physiologically relevant protein conformation.

A strategy often employed by channels to conduct ions across the membrane involves positioning negative charge density in an ordered fashion along the pore. The selectivity filter generally functions by providing an ion coordination site that selectively stabilizes desired cations using the appropriate number, strength, and arrangement of coordinating ligands.[25] Lewis basic backbone carbonyl groups often satisfy the role of coordinating ligand. To investigate the importance of Lewis basicity in the selectivity filter,

we used non-canonical amino acid mutagenesis to target backbone carbonyls at four positions: the three residues composing the GAS motif (Gly443, Ala444, and Ser445) and Gly439, which coordinates cesium in some structures. Our group has commonly incorporated α-hydroxy analogs of amino acids as a subtle way of modulating the hydrogen bonding ability of backbone carbonyls in receptors.[26-30] Because an ester carbonyl is a weaker Lewis base than an amide carbonyl, we anticipated that such a substitution at residue *i* would impact ion selectivity if the carbonyl at position *i*-1 were involved in coordinating ions at the selectivity filter. Therefore, the four sites we selected for incorporation were Leu440, Ala444, Ser445, and Ile446.

In each ion selectivity experiment, receptors were expressed in *Xenopus laevis* oocytes, incorporating the α-hydroxy acid at the site of interest. The channels in each oocyte were then equilibrated in standard recording solution (96 mM Na^+), and subsequently activated in a pH 6 buffer. This procedure was repeated, except in buffers replacing all other cations with the cation of interest (Li^+ or K^+), as described in earlier studies.[31,32] Relative conductivities of these cations were then compared with sodium.

In the context of the structures available at the time, these results were somewhat unexpected (**Figure 4.2B**). Two of the substitutions in the GAS motif reduced selectivity, increasing the relative permeabilities of K^+ and Li^+. This observation reinforced the importance of these residues as the selectivity filter, but the affected carbonyl groups did not appear to point toward the channel pore.[15,16] For the A444Aah substitution, which should have affected the Gly443 carbonyl, currents were too low to obtain reliable values for G_{K+}/G_{Na+}, but the relative lithium permeability was maintained in comparison with wild type ASIC1a. The Leu440 substitution, which we expected to reduce selectivity by influencing the

Lewis basicity of the Gly439 carbonyl, had a minimal effect, only slightly increasing the permeability of lithium.

These data are somewhat more consistent with the recent structures that include what is likely to be the true positioning of the GAS motif (**Figure 4.2C,D**). The I446Iah mutation eliminates a hydrogen bonding interaction with the carbonyl of Leu440 that helps maintain the conformation of the selectivity filter. This likely results in suboptimal positioning of the GAS ring, which would disrupt selectivity for sodium.

The Lewis basicity of the backbone carbonyl at Ala444, which points into the channel pore near the level of an observed Cs^+ ion, is weakened by an S445Sah mutation. The observed increases in the permeabilities of both Li^+ and K^+ accompanying this substitution are consistent with an important role for the Ala444 carbonyl in preferentially stabilizing Na^+. However, the effect could also be explained by disruption of what may be a structurally important hydrogen bonding interaction between the backbone N-H of Ser445 and the side chain of Thr448 (N-O distance 3.1 Å).

In contrast, neither of the backbone mutations influencing Lewis basicities of Gly443 or Gly439 backbone carbonyls appeared to impact ion selectivity (although surface expression levels were too low to observe K^+ currents). These two sets of carbonyls were proposed to be the primary coordinating groups in the selectivity filter.[19] Our observations may indicate that this is not the case; however, it is also possible that preferential stabilization of Na^+ is largely contingent on the positioning of the coordinating groups, but somewhat independent of their Lewis basicities.

Together with the available structures, these data support critical structural roles for the backbone amides of Ile446 and possibly Ser445, as well as potential involvement of the backbone carbonyl at Ala444 in ion selectivity. However, these data are preliminary, having

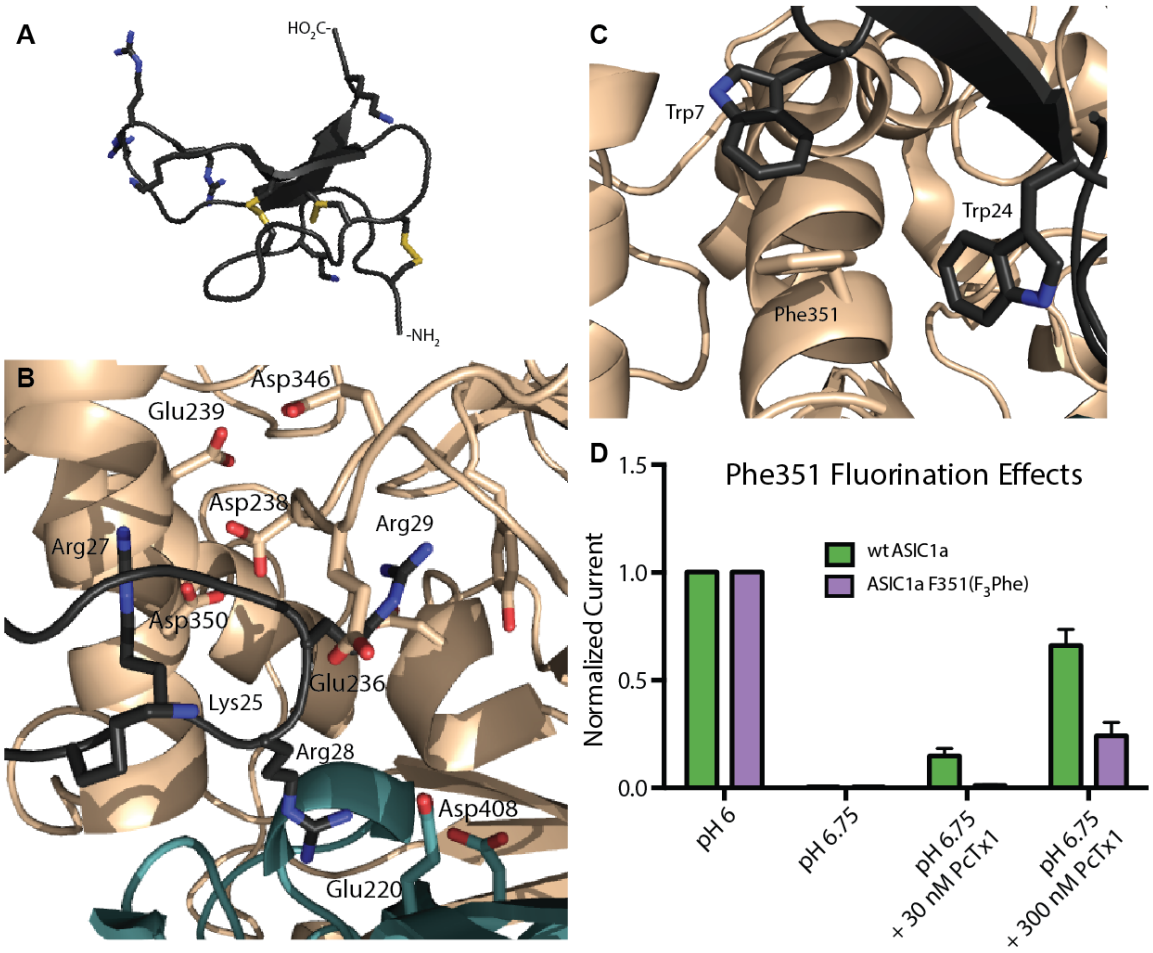


Figure 4.9. The binding interaction between PcTx1 and ASIC1a. (A) Structure of PcTx1. Disulfide bridges are highlighted in yellow. Cationic residues residues are shown as sticks (PDB 4FZ0). (B) Electrostatic interactions between cationic residues on PcTx1 and anionic side chains in the ASIC1a acid pocket. (C) Aromatic-aromatic interactions involving Phe351 and two trypophan residues on PcTx1 thought to be involved in toxin selectivity for ASIC1a. (D) Effects of fluorinating Phe351 on current responses in the presence of PcTx1. Mean normalized currents from 5 oocytes are plotted, with standard errors indicated.

been performed on only one batch of oocytes, and will need to be reproduced before final conclusions may be drawn.

*4.3.2 A noncovalent interaction with Phe351 influences psalmotoxin selectivity for ASIC1**

Peptide and protein inhibitors of ASICs offer potential routes for the development of pain and stroke therapies, as well as probes for understanding the fundamental mechanism of action of the receptors. Here, we have examined the interaction of ASIC1a with psalmotoxin (PcTx1), the active component of the venom from *Psalmopoeus cambridgei,*

the Trinidad chevron tarantula (**Figure 4.3A**). PcTx1 is an inhibitor cysteine knot protein, about 40 amino acids in length, and containing three disulfide linkages. There are several available X-ray structures of PcTx1 in complex with ASIC1a, in which it appears that a cluster of cationic residues reach into the electrostatically negative interface between subunits to make contacts with anionic and aromatic side chains at the extracellular subunit interface.[17,18] This interaction apparently increases the affinity of the receptor for protons, allowing it to enter the desensitized state at pH ~7.5, resulting in an inactive receptor under physiological conditions.[11]

PcTx1 specifically influences the function of ASIC1. Prior to the aforementioned structural studies, sequence alignments between ASIC1 and ASIC2 led a different group to investigate the specific amino acids that confer this selectivity. Mutagenesis of Phe351 to leucine (which replaces Phe351 in ASIC2) significantly reduced the pH_{50} of ASIC1a (~6.2) and ablated any effect of PcTx1 on the channel's activation curve.[33] This observation, along with the large number of cationic residues present in PcTx1, suggested that a cation-π interaction could potentially be responsible for determining the selectivity of the toxin for ASIC1. We therefore proceeded by using nonsense suppression to substitute Phe351 with a fluorinated phenylalanine analog, a strategy that we have frequently used to investigate the energetic importance of electrostatic aromatic interactions.[34-36]

In this experiment, we used 3,4,5-trifluorophenylalanine (F_3Phe), an analog that should have a dramatically reduced electrostatic potential at the center of the aromatic ring. In the absence of PcTx1, the F351(F_3Phe) mutant gave dose-response curves that were very similar to those of the wild type receptor (pH_{50} ~ 6.4). We then expressed wild type ASIC1a and the F351(F_3Phe) mutant in *Xenopus laevis* oocytes, and then used the two-electrode voltage clamp to monitor peak whole-cell currents under a number of different conditions.

For each oocyte, we normalized responses to the current observed upon application of pH 6 buffer. Then each cell was treated with buffer at pH 6.75, which gave negligible responses, followed by coapplication of pH 6.75 buffer along with two concentrations of PcTx1 (**Figure 4.3D**).

As shown by these responses, both receptors increase in their acid sensitivity at pH 6.75 when PcTx1 concentration increases, implying that fluorination of Phe351 does not completely ablate the ability of PcTx1 to bind or modulate function. The extent to which PcTx1 sensitized the wild type receptor was consistently greater than the effect observed on the mutant, indicating that an electrostatic attraction at Phe351 is likely an important factor for binding selectivity.

The structures showing PcTx1 bound to ASIC1a do indeed suggest that Phe351 is involved in binding interactions with the toxin.[17,18] Rather than a cation-π interaction, two edge-face interactions with Trp7 and Trp24 appear to help anchor PcTx1 in place, along with numerous other Coulombic interactions between toxin arginine residues and receptor aspartic or glutamic acid residues at the subunit interface (**Figure 4.3B,C**).

Our data are again only preliminary, conducted on a single batch of cells. A more rigorous analysis of the interaction would involve creation of a Schild plot showing the relationship between the concentration of toxin and the magnitude of shift in the steady-state desensitization curve. However, due to the cost of the toxin and the availability of detailed structural data on the interaction, we opted not to quantitatively characterize the effect or pursue studying the interaction further.

4.3.3 Efforts toward probing titratable residues with fluorinated glutamate and aspartate analogs

The ASIC1a structure reveals a clustering of anionic residues at the extracellular subunit interface, dubbed the "acid pocket."[15] This pocket is situated between the thumb and finger of one subunit and the palm of the neighboring subunit. Aspartic and glutamic acid residues are usually found deprotonated at physiological pH, resulting in anionic side chains that can participate in salt bridges and contribute to protein structural stability. However, the proximity of several pairs of acidic side chains in this structure (~3-4 Å) suggests that at least one of the carboxylates in each pair must be protonated, mediating association via a hydrogen bond (**Figure 4.4**). Indeed, it is well documented that small molecules containing proximal carboxylic acids can demonstrate significant shifts in pK_a, which could enable titration of these pairs at pH ~6.5 in the protein microenvironment.[37]

In the first structural study, the authors proposed that the carboxylic acids may individually coordinate calcium in the resting state, then dimerize in a cooperative fashion as they are protonated and the receptor undergoes a conformational shift into the open state.[15] This hypothesis was tested by a number of studies in which mutation of these residues modulates the proton sensitivity of the receptor.[38] However, most of these mutations are

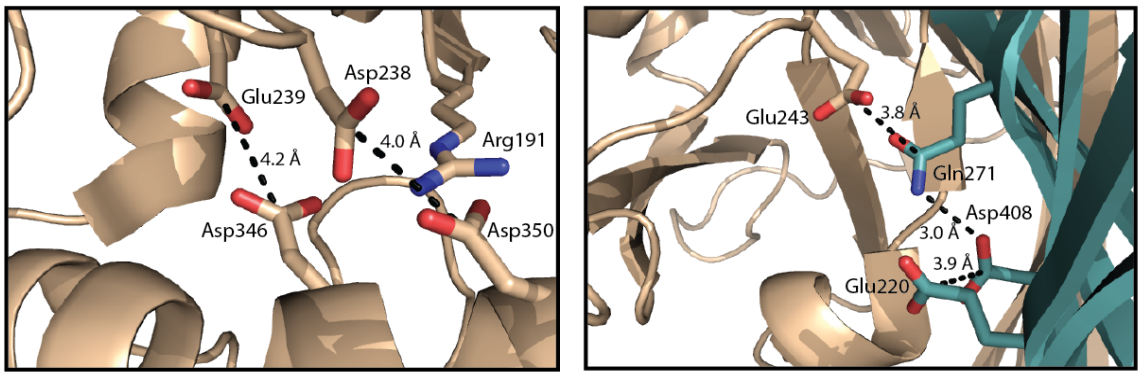


Figure 10.4. The ASIC1a acid pocket. Shown are several carboxylic acid residues that have been proposed as putative proton binding sites that are titrated during receptor activation (PDB 2QTS).

fairly dramatic, removing the negative charge from the side chain entirely. We sought to develop amino acid analogs with quantifiable shifts in acidity. Ideally, these analogs would minimally perturb the steric parameters of aspartic or glutamic acid. Such subtle control over electrostatics could be useful not only in building a model for the activation of ASICs, but also for tweaking reactivity in metalloprotein active sites or exploring factors governing protein-protein interactions, such as PcTx1 binding.

Analogous to our strategy for perturbing aromatic side chains, we chose to attempt incorporation of (2*S*,3*S*)-3-fluoroaspartic acid (**3**, FAsp) and (2*S*,4*R*)-4-fluoroglutamic acid (**11**, FGlu), due to the known shifts in the pK_a values caused by α-fluorination of carboxylates,[39] the small difference between the atomic radii of hydrogen and fluorine, and the relative accessibility of these stereoisomers.[40,41] After synthesizing the free amino acids, they must be appropriately derivatized for attachment to suppressor tRNAs. This sequence is particularly challenging for amino acids with carboxylic acid side chains, since the backbone carboxylic acid must be selectively activated for coupling with the pdCpA dinucleotide (abbreviated here as dCA. See Chapter 1 for more details on chemical tRNA acylation and nonsense suppression). We therefore chose to selectively derivatize the side chain carboxylate with a photocleavable protecting group, allowing activation of the backbone carboxylic acid as a cyanomethyl ester, as previously described.[42]

Our strategy is illustrated in **Figure 4.5** and **Figure 4.6**. In each case, we prepared the free amino acid as described by the literature precedent[40,41,43] and then protected the backbone amine under standard conditions by reacting it with nitroveratryl chloroformate for ease of handling and to slow hydrolysis from the tRNA during storage. At this stage, we chose to condense the two carboxylic acids into a cyclic anhydride, which could then be opened via nucleophilic attack with an alcohol, masking one carboxylate as an ester. We

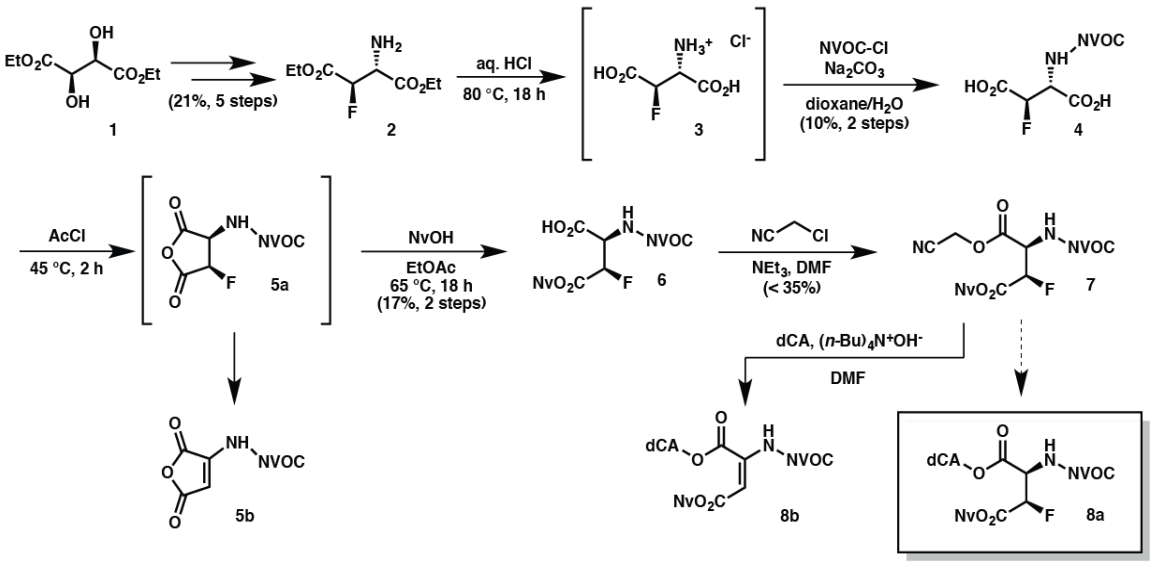


Figure 4.11. Attempted synthesis of photocaged dCA-FAsp. Elimination of hydrofluoric acid under dCA coupling conditions prevented access to the desired product.

hoped that selective nucleophilic attack with nitroveratryl alcohol on the carbonyl closest to the fluoro substituent would enable us to preferentially form a side chain ester, leaving the backbone carboxylate free for activation. Prior to injection into oocytes, both the nitroveratryl ester and the backbone carbamate would then be photochemically cleaved.

We began by optimizing the anhydride formation and alcoholysis steps on FAsp. After trying a number of dehydration reagents, including phosphorus trichloride, oxallyl chloride, thionyl chloride, and acetic anhydride, we determined that stirring a suspension of **4** in neat acetyl chloride gave fairly clean conversion to anhydride **5a** in under 5 hours, along with a minor product (**5b**, <10%) that appears to result from elimination of hydrofluoric acid. After treatment of the anhydride with nitroveratryl alcohol for 16 hours, we were able to isolate a product with the expected mass of acid **6**, although in about equal proportion to the elimination product **5b**. Analysis by HMBC revealed long-range couplings that confirmed the product as having the desired connectivity. We were able to proceed with this material, activating the carboxylic acid as a cyanomethyl ester (**7**). Isolation of **7** was

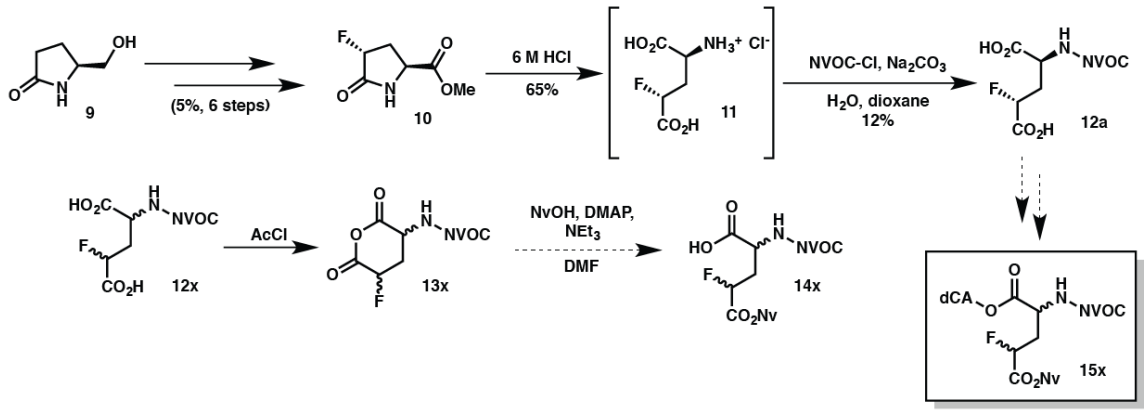


Figure 4.12. Attempted synthesis of photocaged dCA-FGlu. Although significant elimination was not observed, the overall synthetic route was prohibitively long and low-yielding.

nontrivial; as we were concerned about the potential for elimination of hydrogen fluoride, we submitted the partially purified product to dCA coupling conditions. However, upon HPLC purification of the product mixture, we were only able to assign masses corresponding to elimination product **8b**.

Considering the previously observed elimination products, it is no surprise that the basic conditions of the dCA coupling reaction resulted in decomposition. It is possible that further optimization of these conditions could allow for the intact amino acid to be coupled to dCA. However, the reactivity of the amino acid side chain could easily cause decomposition during the tRNA ligation or at any time after injection into the oocyte cytoplasm. As it is difficult to assess the quality of the amino acid after it has been coupled to a biomacromolecule, we decided not to continue work with this analog.

We expect that the reactivity of FAsp is due to the formation of a large conjugated π-system upon elimination of hydrofluoric acid, making decomposition highly favorable. Since a similar elimination from FGlu would generate a smaller π-system, we were hopeful that it would serve as a more stable probe. Unfortunately, due to several low-yielding steps in the synthesis of FGlu, we were only able to access a very small mass of diacid **12a**. In an effort to assess the viability of our route, we opted to make use of a commercially available mixture

of stereoisomers of 4-fluoroglutamic acid, which we were similarly able to protect as a nitroveratryl carbamate (**12x**). We then used acetyl chloride to effect intramolecular condensation into a mixture of anhydrides (**13x**). Notably, no significant elimination product was observed. Treatment of the resulting crude mass with nitroveratryl alcohol, dimethylaminopyridine, and triethylamine yielded a new mixture of compounds, including some with masses corresponding to the desired monoacid product (**14x**). However, due to the very low crude yield and the complexity of the NMR obtained from the diastereomeric mixture, we were unable to determine the regioselectivity of the alcoholysis reaction.

Despite our failure to incorporate these fluorinated probes, these synthetic efforts have yielded a potentially general strategy for regioselectively protecting amino acids with carboxylic acid side chains. It is likely that such regioselectivity is limited to cases where the electrophilicity of the side chain carboxylate is activated by an electron-withdrawing group at the α-position, since alcoholysis of anhydrides derived from the parent amino acids is usually selective for the backbone position.[44] Further studies will be required to determine the scope of this selectivity and the overall utility of the strategy for the purposes of chemical tRNA acylation with other glutamic and aspartic acid analogs.

4.4 Conclusions

To our knowledge, these studies represent the first applications of the nonsense suppression methodology to the investigation of the acid-sensing ion channels. In the context of new structural data, the ion selectivity experiments bring up interesting questions as to which backbone carbonyl groups are involved in cation stabilization at the selectivity filter. Future computational modeling studies could resolve whether the reductions in Lewis basicity that we introduce via backbone ester mutations would be expected to influence selectivity in a significant way.

We have shown that Phe351 is likely to play a role in the selectivity of PcTx1 activity on ASIC1a via noncovalent aromatic-aromatic interactions. Although there are clearly countless other important electrostatic interactions at the binding interface, it is a testament to the power of non-canonical amino acid mutagenesis that such a complex association can be energetically probed through such a subtle perturbation.

Finally, our unsuccessful efforts to incorporate FAsp are an indication that introduction of strongly electronegative groups at the 3-position of aspartic acid is likely to lead to decomposition. Before additional work is committed to protection and activation of a non-canonical amino acid analog containing a functionalized side chain, it is advised that the limits of its stability are vigorously tested. Apart from this caveat, it is likely that anhydride alcoholysis is a generalizable strategy for producing activated amino acids protected by photocleavable groups.

4.5 Experimental procedures

4.5.1 Molecular biology

pCR2.1-hASIC1 was a gift from Cecilia Canessa (Addgene plasmid # 32790). The region containing the ASIC1a gene was amplified by PCR using primers with overhangs containing 5' HindIII and 3' XhoI restriction sites. These restriction enzymes were then used to subclone the gene into the pGEMhe vector for expression in *Xenopus laevis* oocytes. Site-directed mutagenesis was performed using the Stratagene QuikChange protocol to generate the appropriate codon. For non-canonical amino acid mutants and conventional mutants generated by nonsense suppression, the site of interest was mutated to the TAG stop codon. Plasmids were linearized with the SbfI restriction enzyme, and receptor mRNA was then prepared by in vitro runoff transcription using the Ambion T7 mMessage mMachine kit.

Hydroxy or amino acid-dCA conjugates were enzymatically ligated to truncated 74mer THG73 tRNA as previously described.[27,45] The 74mer tRNA was prepared using the Ambion T7MEGAshortscript kit by transcription from a DNA oligonucleotide template modified to enhance RNA transcript homogeneity, as described in the literature.[46] Crude tRNA-amino acid or tRNA-hydroxy acid product was used without desalting, and the product was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry on a 3-hydroxypicolinic acid matrix. Deprotection of the NVOC groups on the tRNA-amino acids was carried out by photolysis for 5 minutes on a 300-watt high-pressure mercury arc lamp with WG-335 and UG-11 filters immediately prior to injection.

PcTx1 was purchased as a lyophilized powder from Peptide Institute, Inc. The toxin was dissolved in Millipore water to a stock concentration of 0.1 m*M*, and then stored at -80 °C until use.

4.5.2 Oocyte preparation and RNA injection

Stage V-VI oocytes of *Xenopus laevis* were harvested and injected with RNAs as described previously.[45] For nonsense suppression experiments, each cell was injected with 1 - 5 ng of mRNA and 20 ng of appropriate tRNA approximately 24 h before recording.

For wild type experiments and conventional mutants, each cell received a single injection of 0.05-0.5 ng of ASIC1a mRNA approximately 18 h before recording. Injection volumes for each injection session were 50 nL per cell.

As a negative control for suppression experiments at each site, unacylated full-length tRNA was co-injected with mRNA in the same manner as charged tRNA. Wild-type recovery conditions (injecting tRNA charged with the appropriate amino acid to regenerate a wild type channel via nonsense suppression at a TAG stop codon) were injected alongside mutant nonsense suppression conditions as a positive control.

4.5.3 Electrophysiology

Oocyte recordings were made in two-electrode voltage clamp mode using the OpusXpress 6000A (Axon Instruments). Unless otherwise noted, oocyte equilibration and washes were performed with ND96 (96 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 1.8 mM $CaCl_2$, 5 mM HEPES) adjusted to pH 7.8 with 1 N NaOH. The pH of buffers for concentration-response curves were adjusted accordingly with NaOH or HCl, and buffers with pH below 6.8 contained 5 mM MES in place of HEPES. Initial holding potential was -60 mV. Data were sampled at 125 Hz and filtered at 50 Hz. Oocytes were equilibrated for 30 s at 1 mL/min before each pH application. The pH buffer applications lasted for 15 s at 4 mL/min. Cells were then washed for 2 min at 3 mL/min before the subsequent equilibration. For concentration-response experiments, responses from each cell were normalized to the maximum response observed from that cell, and the normalized currents were then averaged.

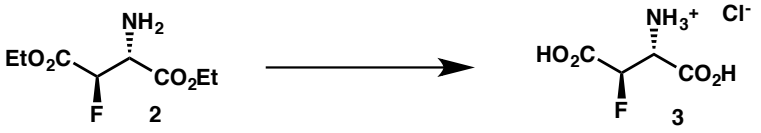
4.5.4 Chemical synthesis and characterization

4.5.4.1 Materials. Unless otherwise specified, all reactions were performed under an argon atmosphere in a flame-dried round-bottom flask sealed with a rubber septum, and in freshly dried solvents. Air- and moisture-sensitive liquids were transferred using stainless steel needles or cannulae. Commercial reagents and solvents, purchased from Sigma Aldrich and VWR, were used as received. Triethylamine was freshly distilled from CaH_2 prior to use. Deuterated solvents for NMR experiments were purchased from Cambridge Isotope Laboratories and used without further purification. Compounds **2**[43] and **10**,[40] as well as pdCpA[47] were synthesized as previously reported. Flash chromatography was performed using silica gel from Merck (230-400 mesh). Analytical TLC was performed using plates

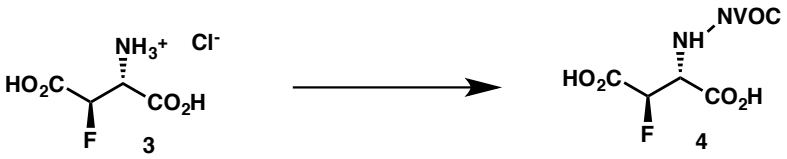
from Sigma-Aldrich (60 Å pore size, F_{254}, 0.25 mm) and was visualized by UV irradiation (254 nm) or staining with potassium permanganate.

4.5.4.2 Instrumentation. ^{1}H and ^{13}C NMR spectra were obtained using Varian NMR instruments. Chemical shifts are reported in parts per million (ppm, δ) referenced to the residual ^{1}H resonance of the solvent. ^{13}C NMR spectra were referenced to the residual ^{13}C resonance of the solvent. Splitting patterns are designated as follows: s, singlet; br, broad; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet. Mass spectroscopic data were acquired using an Agilent 1100 Series LC/MSD equipped with an ESI-APCI source. HPLC preparative purification was performed using a Waters 1525 binary pump, a Waters 2996 photodiode array detector, and an Atlantis Prep T3 OBD 5μm 19x150mm column.

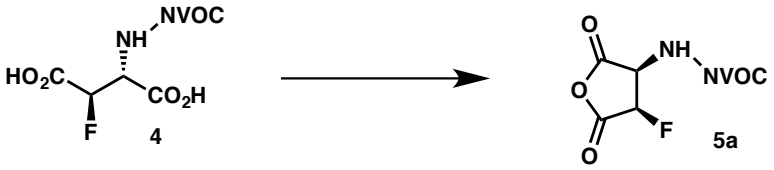
4.5.4.3 Reaction conditions.



(1*R*,2*R*)-1,2-dicarboxy-2-fluoroethan-1-aminium chloride (3). Diester **2** (800 mg) was dissolved in an aqueous 5 *N* solution of hydrochloric acid (30 mL) in a 100-mL round bottom flask fitted with a reflux condenser. The mixture was heated to 80 °C and stirred overnight. The solvent and hydrochloric acid were then removed by concentration in vacuo, followed by iterative dissolution and lyophilization from Millipore water. The resulting crude white solid (198 mg) was carried on without further purification.

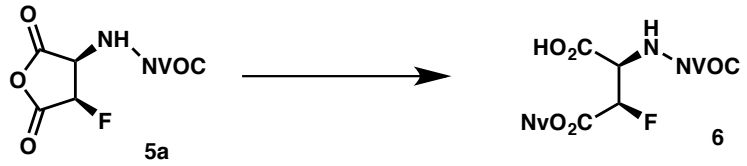


(2R,3R)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-fluorosuccinic acid (4). Crude amino acid hydrochloride **3** (198 mg) was dissolved in water (30 mL) and dioxane (30 mL). Sodium carbonate (694 mg) was added as a solid in one portion, followed by 6-nitroveratryl chloroformate (360 mg). The resulting mixture was stirred for 48 h under an argon atmosphere. The reaction mixture was then washed with ethyl acetate (3 x 50 mL) to remove residual 6-nitroveratryl chloride. The aqueous layer was then cooled in an ice-water bath and acidified to a pH of 2 by slow addition of 6 N aqueous HCl. At this time, the aqueous solution was quickly extracted with ethyl acetate (3 x 100 mL) until no UV-active material was observable by TLC in the aqueous layer. The combined organic fractions were then dried over anhydrous sodium sulfate, filtered, and concentrated to give a yellow-orange residue. This residue was triturated with diethyl ether and dried in vacuo to give **4** as a tan solid (96 mg, 16% yield over 2 steps). 1H NMR (300 MHz; CD_3CN) δ: 10.01 (s, br, 1H), 7.74 (s, 1H), 7.18 (s, 1H), 6.57 (d, br, J = 9.0, 1H), 5.49 (s, 2H), 5.34 (dd, J_F = 47.0, J_H = 2.5, 1H), 4.99 (ddd, J_F = 27.1, J_H = 8.7, 2.3, 1H), 3.97 (s, 3H), 3.91 (s, 3H); LRMS (ESI) 389.0 $[M-H]^-$.

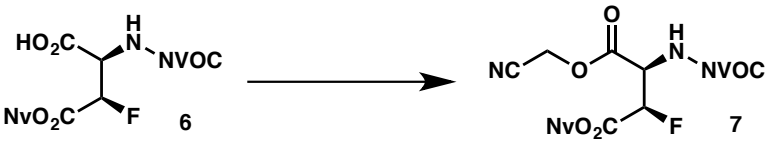


4,5-dimethoxy-2-nitrobenzyl ((3R,4R)-4-fluoro-2,5-dioxotetrahydrofuran-3-yl)carbamate (5a). Under air, diacid **4** (30 mg) was added to a 2-dram scintillation vial equipped with a stir bar and suspended in acetyl chloride (1.5 mL). The vial was fitted with a

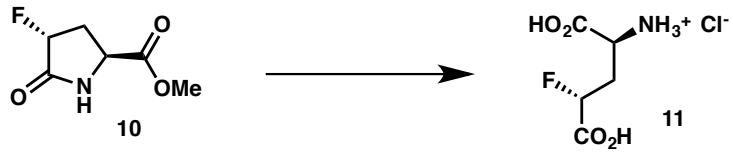
rubber septum vented through an oil bubbler, and the suspension was heated to 45 °C with stirring. After 2 hours, or when the mixture appears homogenous, the reaction progress was assessed by NMR. Upon complete consumption of the diacid, the reaction was returned to room temperature and the solvent removed in vacuo. Iterative suspension in anhydrous benzene and evaporation in vacuo gave anhydride **5a** as a crude tan solid (27 mg), which was carried on without further purification. LRMS (ESI) 353.0 $[M-H]^-$.



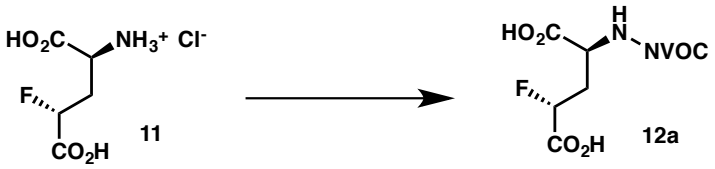
(2R,3R)-4-((4,5-dimethoxy-2-nitrobenzyl)oxy)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-fluoro-4-oxobutanoic acid (6). A flame-dried 2-dram vial equipped with a stir bar was charged with crude anhydride **5a** (27 mg) and 6-nitroveratryl alcohol (82 mg) under argon. Freshly distilled ethyl acetate (2 mL) was added by syringe, and the reaction was heated to 65 °C with vigorous stirring for 16 h. The solvent was then removed in vacuo, and the resulting residue was purified on a short plug of silica gel (EtOAc to elute residual alcohol and anhydride, followed by 2% AcOH/EtOAc to elute monoacid product) to yield **6** as an orange solid with some small impurities (12 mg, 17% yield over 2 steps). HSQC and HMBC were used to assign the reaction regioselectivity. ^{1}H NMR (600 MHz; CD_3CN) δ: 7.71 (s, 1H), 7.70 (s, 1H), 7.17 (s, 1H), 7.14 (s, 1H), 6.58 (d, J = 8.2, 1H), 5.57 (s, 2H), 5.45 (s, 2H), 5.50-5.42 (m, 1H), 5.04-4.98 (m, 1H), 5.00 (ddd, J_F = 27.2, J_H = 8.6, 1.9, 1H), 3.93 (s, 3H), 3.93 (s, 3H), 3.89 (s, 3H), 3.89 (s, 3H); ^{13}C NMR (126 MHz; d_6-DMSO) δ: 169.05, 169.00, 166.78, 166.58, 156.2, 154.0, 153.8, 148.3, 148.1, 139.7, 139.3, 128.3, 126.3, 111.2, 110.4, 108.50, 108.47, 89.4 (d, J_F = 188.57), 64.0, 63.4, 61.0 (d, J_F = 217.96), 56.7, 56.5; LRMS (ESI) 584.1 $[M-H]^-$.



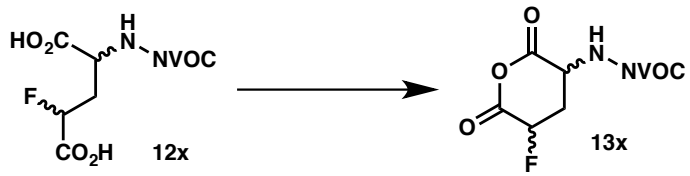
(2R,3R)-1-(cyanomethyl) 4-(4,5-dimethoxy-2-nitrobenzyl) 2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-3-fluorosuccinate (7). A flame-dried 2-dram vial equipped with a stir bar was charged with acid **6** (8 mg). The solid was dissolved in chloroacetonitrile (1 mL) and DMF (1 mL) under argon. Triethylamine (1.9 μL) was then added slowly as a solution in DMF, and the mixture was stirred under argon for 18 h at room temperature. The solvents were then removed in vacuo, and the resulting solid was partially purified by silica gel chromatography, giving an orange-tan solid (~3 mg, < 35% yield). Purification of product on this scale was difficult, and overlapping signals in the mixed ^{1}H NMR made peak assignments impossible. LRMS (ESI) 603.1 [M-H_2F]$^-$.



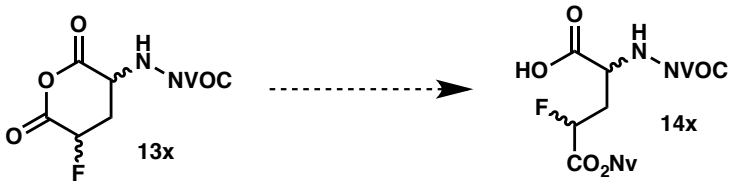
(1*S*,3*R*)-1,3-dicarboxy-3-fluoropropan-1-aminium chloride (11). Lactam **10** (15 mg) was dissolved in an aqueous 6 *N* hydrochloric acid solution (1 mL) in a scintillation vial vented through an argon bubbler. The mixture was heated at 90 °C for 4 hours, then concentrated in vacuo. The resulting residue was twice iteratively redissolved in Millipore water and concentrated by lyophilization, giving a crude yellow residue (12 mg), which was carried on without further purification.



(2*S*,4*R*)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-4-fluoropentanedioic acid (12a). Crude amino acid **11** (12 mg) was subjected to the conditions described for preparation and purification of NVOC carbamate **4.** Following chromatography, fractions were concentrated, giving **12a** as a yellow solid (3 mg, 8% over 2 steps). 1-H NMR (500 MHz; $CDCl_3$) δ: 7.74 (s, 1H), 7.20 (s, 1H), 4.99 (d, J = 3.7, 2H), 4.03 (s, 3H), 3.99 (s, 3H), 2.62-2.59 (m, 1H), 1.57 (m, 3H); LRMS (ESI) 403.1 $[M-H]^-$.



4,5-dimethoxy-2-nitrobenzyl (5-fluoro-2,6-dioxotetrahydro-2*H*-pyran-3-yl)carbamate (13x). A mixture of stereoisomers of NVOC carbamate **12x** (20. mg, derived from 4-fluoro-DL-glutamic acid purchased from Sigma-Aldrich) was added to a pear-shaped flask under air, and acetyl chloride (2 mL) was added. The flask was capped with a rubber septum with a vent needle, and the suspension was stirred at 55 °C for 2 hours. The mixture was cooled to room temperature and stirred overnight, and then concentrated in vacuo. The resulting crude mixture was carried on without further purification. LRMS (ESI) 385.1 $[M-H]^-$.



Attempted synthesis of 5-((4,5-dimethoxy-2-nitrobenzyl)oxy)-2-((((4,5-dimethoxy-2-nitrobenzyl)oxy)carbonyl)amino)-4-fluoro-5-oxopentanoic acid (14x). Crude NVOC carbamate **13x** was dissolved in dry DMF (1 mL) in a scintillation vial. Nitroveratryl alcohol (11.5 mg), DMAP (0.6 mg), and triethylamine (6.8 μL) were added under argon. The mixture was stirred at 60 °C overnight, and the solvent was removed in vacuo. The correct mass of the desired product was observed by LCMS. However, attempts to remove the remaining impurities by silica gel chromatography (0-5% acetic acid/ethyl acetate) yielded too little product to obtain a ^{1}H NMR or assign the regioisomer by HMBC. LRMS (ESI) 598.1 $[M-H]^-$.

4.6 References

1. Waldmann, R. *et al.* A proton-gated cation channel involved in acid-sensing. *Nature* **386,** 173–177 (1997).
2. Lingueglia, E. Acid-sensing Ion Channels in Sensory Perception. *J. Biol. Chem.* **282,** 17325–17329 (2007).
3. Wemmie, J. A., Taugher, R. J. & Kreple, C. J. Acid-sensing ion channels in pain and disease. *Nat. Rev. Neurosci.* **14,** 461–471 (2013).
4. Sluka, K. A., Winter, O. C. & Wemmie, J. A. Acid-sensing ion channels: A new target for pain and CNS diseases. *Curr. Opin. Drug Di. De.* **12,** 693–1394 (2009).
5. Xiong, Z.-G. *et al.* Neuroprotection in ischemia: blocking calcium-permeable acid-sensing ion channels. *Cell* **118,** 687–698 (2004).
6. Kellenberger, S. & Schild, L. Epithelial sodium channel/degenerin family of ion channels: a variety of functions for a shared structure. *Physiol. Rev.* **82,** 735–767 (2002).
7. Lingueglia, E. *et al.* A Modulatory Subunit of Acid Sensing Ion Channels in Brain and Dorsal Root Ganglion Cells. *J. Biol. Chem.* **272,** 29778–29783 (1997).
8. Grunder, S., Geissler, H.-S., Bässler, E.-L. & Ruppersberg, J. P. A new member of acid-sensing ion channels from pituitary gland. *Neuroreport* **11,** 1607 (2000).
9. Akopian, A. N. *et al.* A new member of the acid-sensing ion channel family. *Neuroreport* **11,** 2217–2222 (2000).
10. Hesselager, M., Timmermann, D. B. & Ahring, P. K. pH Dependency and desensitization kinetics of heterologously expressed combinations of acid-sensing ion channel subunits. *J. Biol. Chem.* **279,** 11006–11015 (2004).
11. Chen, X., Kalbacher, H. & Grunder, S. The Tarantula Toxin Psalmotoxin 1 Inhibits Acid-sensing Ion Channel (ASIC) 1a by Increasing Its Apparent H^+ Affinity. *J. Gen. Physiol.* **126,** 71–79 (2005).
12. Diochot, S. *et al.* Black mamba venom peptides target acid-sensing ion channels to abolish pain. *Nature* **490,** 552–555 (2012).

13. Bohlen, C. J. *et al.* A heteromeric Texas coral snake toxin targets acid-sensing ion channels to produce pain. *Nature* **479,** 410–414 (2011).
14. Yang, Z.-J. *et al.* Neuroprotective effect of acid-sensing ion channel inhibitor psalmotoxin-1 after hypoxia–ischemia in newborn piglet striatum. *Neurobiol. Dis.* **43,** 446–454 (2011).
15. Jasti, J., Furukawa, H., Gonzales, E. B. & Gouaux, E. Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH. *Nature* **449,** 316–323 (2007).
16. Gonzales, E. B., Kawate, T. & Gouaux, E. Pore architecture and ion sites in acid-sensing ion channels and P2X receptors. *Nature* **460,** 599–604 (2009).
17. Dawson, R. J. P. *et al.* Structure of the Acid-sensing ion channel 1 in complex with the gating modifier Psalmotoxin 1. *Nat. Comms.* **3,** 936 (2012).
18. Baconguis, I. & Gouaux, E. Structural plasticity and dynamic selectivity of acid-sensing ion channel–spider toxin complexes. *Nature* **489,** 400–405 (2012).
19. Baconguis, I. *et al.* X-ray structure of acid-sensing ion channel 1-snake toxin complex reveals open state of a Na^+-selective channel. *Cell* **156,** 717–729 (2014).
20. Baron, A., Waldmann, R. & Lazdunski, M. ASIC-like, proton-activated currents in rat hippocampal neurons. *J. Physiol. (Lond.)* **539,** 485–494 (2002).
21. Chen, C. *et al.* A sensory neuron-specific, proton-gated ion channel. *Proc. Natl. Acad. Sci. USA* **95,** 10240–10245 (1998).
22. Askwith, C. C. *et al.* Acid-sensing ion channel 2 (ASIC2) modulates ASIC1 H^+-activated currents in hippocampal neurons. *J. Biol. Chem.* **279,** 18296–18305 (2004).
23. Babini, E. *et al.* Alternative splicing and interaction with di- and polyvalent cations control the dynamic range of acid-sensing ion channel 1 (ASIC1). *J. Biol. Chem.* **277,** 41597–41603 (2002).
24. Canessa, C. M., Merillat, A. M. & Rossier, B. C. Membrane topology of the epithelial sodium channel in intact cells. *Am. J. Physiol.* **267,** C1682–90 (1994).
25. Dudev, T. & Lim, C. Ion selectivity strategies of sodium channel selectivity filters. *Acc. Chem. Res.* **47,** 3580–3587 (2014).
26. England, P. M., Zhang, Y., Dougherty, D. A. & Lester, H. A. Backbone Mutations in Transmembrane Domains of a Ligand-Gated Ion Channel. *Cell* **96,** 89–98 (1999).
27. England, P. M., Lester, H. A. & Dougherty, D. A. Incorporation of esters into proteins: Improved synthesis of hydroxyacyl tRNAs. *Tetrahedron Lett.* **40,** 6189–6192 (1999).
28. Rienzo, M. *et al.* Perturbation of Critical Prolines in Gloeobacter violaceus Ligand-gated Ion Channel (GLIC) Supports Conserved Gating Motions among Cys-loop Receptors. *J. Biol. Chem.* **291,** 6272-6280 (2016).
29. Rienzo, M., Lummis, S. C. R. & Dougherty, D. A. Structural Requirements in the Transmembrane Domain of GLIC Revealed by Incorporation of Noncanonical Histidine Analogs. *Chem. Biol.* **21,** 1700–1706 (2014).
30. Van Arnam, E. B., Lester, H. A. & Dougherty, D. A. Dissecting the Functions of Conserved Prolines within Transmembrane Helices of the D2 Dopamine Receptor. *ACS Chem. Biol.* **6,** 1063–1068 (2011).
31. Yang, L. & Palmer, L. G. Ion conduction and selectivity in acid-sensing ion channel 1. *J. Gen. Physiol.* **144,** 245–255 (2014).
32. Kellenberger, S., Gautschi, I. & Schild, L. A single point mutation in the pore region of the epithelial Na^+ channel changes ion selectivity by modifying molecular sieving. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 4170–4175 (1999).
33. Sherwood, T. *et al.* Identification of protein domains that control proton and calcium sensitivity of ASIC1a. *J. Biol. Chem.* **284,** 27899–27907 (2009).
34. Van Arnam, E. B. & Dougherty, D. A. Functional probes of drug-receptor interactions implicated by structural studies: Cys-loop receptors provide a fertile testing ground. *J. Med. Chem.* **57,** 6289–6300 (2014).
35. Xiu, X. *et al.* Nicotine binding to brain receptors requires a strong cation-pi interaction. *Nature* **458,** 534–537 (2009).
36. Daeffler, K. N.-M., Lester, H. A. & Dougherty, D. A. Functionally important aromatic-aromatic and sulfur-π interactions in the D2 dopamine receptor. *J. Am. Chem. Soc.* **134,** 14890–14896 (2012).
37. Brown, H.C. et al., in Braude, E.A. and F.C. Nachod Determination of Organic Structures by Physical Methods, Academic Press, New York, 1955.
38. Sherwood, T. W., Frey, E. N. & Askwith, C. C. Structure and activity of the acid-sensing ion channels. *Am. J. Physiol.-Cell Physiol.* **303,** C699–710 (2012).
39. Dippy, J., Hughes, S. & Bray, L. G. Chemical constitution and the dissociation constants of

monocarboxylic acids. Part XIX. Dihydroxy-and dichloro-benzoic acids. *J. Chem. Soc.* (1959).

40. Konas, D. W. & Coward, J. K. Electrophilic Fluorination of Pyroglutamic Acid Derivatives: Application of Substrate-Dependent Reactivity and Diastereoselectivity to the Synthesis of Optically Active 4-Fluoroglutamic Acids. *J. Org. Chem.* **66,** 8831–8842 (2001).
41. Charvillon, F. B. & Amouroux, R. Synthesis of nonracemic 3-fluoro-aspartic acids. *Tetrahedron Lett.* **37,** 5103–5106 (1996).
42. Noren, C., Anthony-Cahill, S., Griffith, M. & Schultz, P. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244,** 182–188 (1989).
43. Saito, S., Komada, K. & Moriwake, T. Diethyl (2*S*,3*R*)-2-(N-*tert*-butoxycarbonyl) amino-3-hydroxysuccinate. *Org. synth.* **73,** 184–200 (1996).
44. Isidro-Llobet, A., Álvarez, M. & Albericio, F. Amino Acid-Protecting Groups. *Chem. Rev.* **109,** 2455–2504 (2009).
45. Nowak, M. W. *et al.* in *Method. Enzymol.* **293,** 504–529 (Elsevier, 1998).
46. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5,** 1268–1272 (1999).
47. Robertson, S. A. *et al.* The use of 5'-phospho-2 deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA. *Nucleic Acids Res.* **17,** 9649–9660 (1989).

Chapter 5: Progress Toward Mapping Protein-Protein Interfaces With Photocrosslinking Non-Canonical Amino Acids

5.1 Abstract

Pentameric ligand-gated ion channels are known to participate in interactions with other proteins that regulate their function in vivo. In this chapter, we discuss our efforts to establish methods using photochemical crosslinking amino acid analogs to explore these protein-protein interfaces. Using the Schultz methodology, we have begun to optimize expression and crosslinking conditions in a soluble protein model system: the transcription factor estrogen receptor α. We have also successfully incorporated two crosslinking amino acids into several pentameric ligand-gated ion channels, including functional serotonin receptors. Progress toward development of a mass spectrometry-based readout for crosslinking is discussed.

5.2 Introduction

5.2.1 Covalent crosslinking with non-canonical amino acids

Extensive studies of ligand-gated ion channels (LGICs) in model systems have revealed detailed pharmacological profiles of many receptor types, contributing to a molecular-level understanding of fast synaptic transmission.[1,2] However, signal transduction at neuronal synapses is mediated by a complex mixture of interacting extracellular, membrane-associated, and cytoplasmic molecules; occasionally profiles of individual LGICs overlook physiologically relevant aspects of their function. Numerous reports of interactions among ligand-gated ion channels have underscored the non-independent nature of receptor function – pairs of receptors have been shown to colocalize using fluorescent microscopy, and simultaneous activation of different LGICs has often resulted in non-additivity of total currents.[3-19] It is likely that the same interplay observed in these systems is involved in

integrating complex input from a presynaptic cell, but there has been little headway in elucidating molecular mechanisms of such crosstalk.

Direct assessment of interactions among membrane proteins is a challenge, since the conformations of the proteins involved and the strength of any interactions are likely dependent upon the presence of the lipid bilayer.[20] X-ray crystallography and cryo-EM are capable of providing very high resolution depictions of these interfaces, but especially in the cases of membrane proteins and relatively weak interactions, the observed conformations fail to provide the whole story. Coimmunoprecipitation, which is heavily relied upon by protein biochemists for low-resolution information about protein complexes, is generally performed from a lysate mixture and thus may miss important interactions or reveal spurious ones.[21] The situation is further complicated by the possibility that receptor-receptor interactions are mediated by additional integral membrane and/or cytoplasmic proteins.

Chemical crosslinking is a potential workaround for these problems, as it allows one to preserve information about interactions of interest even after destruction of the native cellular environment. Countless bifunctional crosslinking molecules have been developed for different applications, including some with isotope labels for easy detection by mass spectrometry and/or affinity handles to facilitate purification or imaging.[22] Advantages of such reagents include their ease of introduction into the system and the potential for targeting specific amino acid side chain functionalities. However, they are limited to crosslinking solvent-accessible sites at the periphery of binding interfaces, and usually result in a highly complex mixture of products. To increase the scope of accessible sites and decrease the number of different photoadducts, we have focused on using a site-specific crosslinking strategy employing non-canonical amino acid mutagenesis.[23,24] Specifically, a tyrosine-based tRNA/synthetase has been optimized for incorporation of 4-substituted

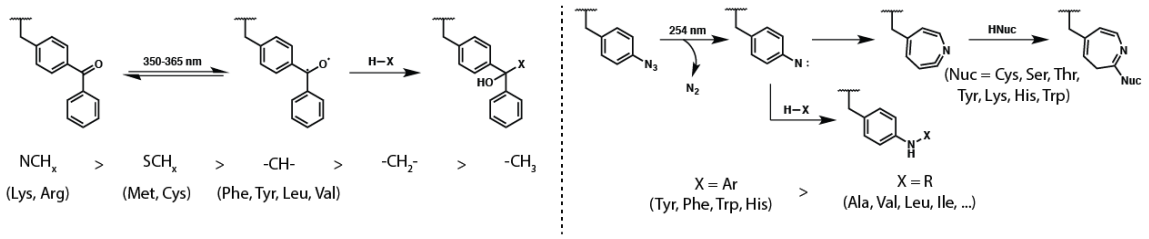


Figure 5.13. Reactivity of photocrosslinking non-canonical amino acids used in this study. (Left) Benzoylphenylalanine (Bpa) is reversibly excited by UV light, and preferentially inserts into C-H bonds adjacent to heteroatoms or at tertiary carbons. (Right) 4-Azidophenylalanine (N_3Phe) releases dinitrogen after UV excitation, and the resulting nitrene can either react via C-H insertion at aryl or alkyl positions, or via rearrangement to an electrophilic species.

phenylalanine derivatives, including the commonly used photoreactive crosslinkers 4-azidophenylalanine (N_3Phe) and 4-benzoylphenylalanine (Bpa). This approach uses site-selective introduction of the crosslinking reagent, which is reactive toward various X-H bonds (X = C, N, O, S) or nucleophilic groups (**Figure 5.1**).[25] Furthermore, because N_3Phe and Bpa are photoreactive, they enable temporal control over the crosslinking reaction.

Following crosslinking, a number of techniques can be used to analyze products. On the simplest level, SDS-PAGE allows detection of higher molecular weight products, with immunoprecipitation and/or immunoblotting as options for enrichment or selective detection of adducts containing proteins of interest. For higher resolution data, mixtures can be subjected to proteolytic digestion, and the resulting fragments identified by tandem mass spectrometry. In theory, this procedure could allow for identification of the specific residues on the binding partner in contact with the site of incorporation chosen for the crosslinker. In practice, low abundance and incomplete fragmentation of crosslinked peptides can often make it difficult to achieve the necessary levels of confidence to identify meaningful interaction sites. Still, there have been a number of reports using crosslinking and mass spectrometry in conjunction.[26,27] However, only four have used amino acids incorporated via nonsense suppression,[28-31] three of which were performed on purified proteins, and none of which involved membrane proteins.

Of course, we aim to use crosslinking to characterize associations involving LGICs in an intact membrane, potentially in a number of different functional states. This goal faces a concurrence of several complicating factors, and after initially struggling with our problem of primary interest, we chose to optimize our basic experimental design in a simpler model system: estrogen receptor α.

5.2.2 Estrogen receptor α

There are a number of pathways by which estrogen influences gene expression, involving both membrane receptors and internal soluble receptors. Estrogen receptor α (ERα) belongs to the latter group, and in its resting state exists in the nucleus as a complex with heat shock protein 90 (Hsp90) and other chaperone proteins (**Figure 1.4**).[32] When receptor agonists, such as β-estradiol (E2), make their way into the nucleus, they bind to the receptor and induce its dissociation from Hsp90. ERα then dimerizes and binds to the estrogen response element, a regulatory region of DNA that allows transcription to be controlled by the presence of E2 and other agonists.

The three-dimensional structure and mechanism of action of ERα are partially understood, thanks to numerous crystal structures of the receptor in the presence of various ligands. Dimerization is facilitated by two domains: that which is responsible for interaction with DNA (the DNA-binding domain, DBD) and that which recognizes and binds to agonists or antagonists (the ligand-binding domain, LBD). The structures of the dimeric states of these domains have been determined;[33,34] the remainder of the protein is disordered (**Figure 5.2A**).

The activity of ERα is heavily influenced by its association with coactivator or corepressor proteins. The propensity for the receptor to bind different proteins is governed

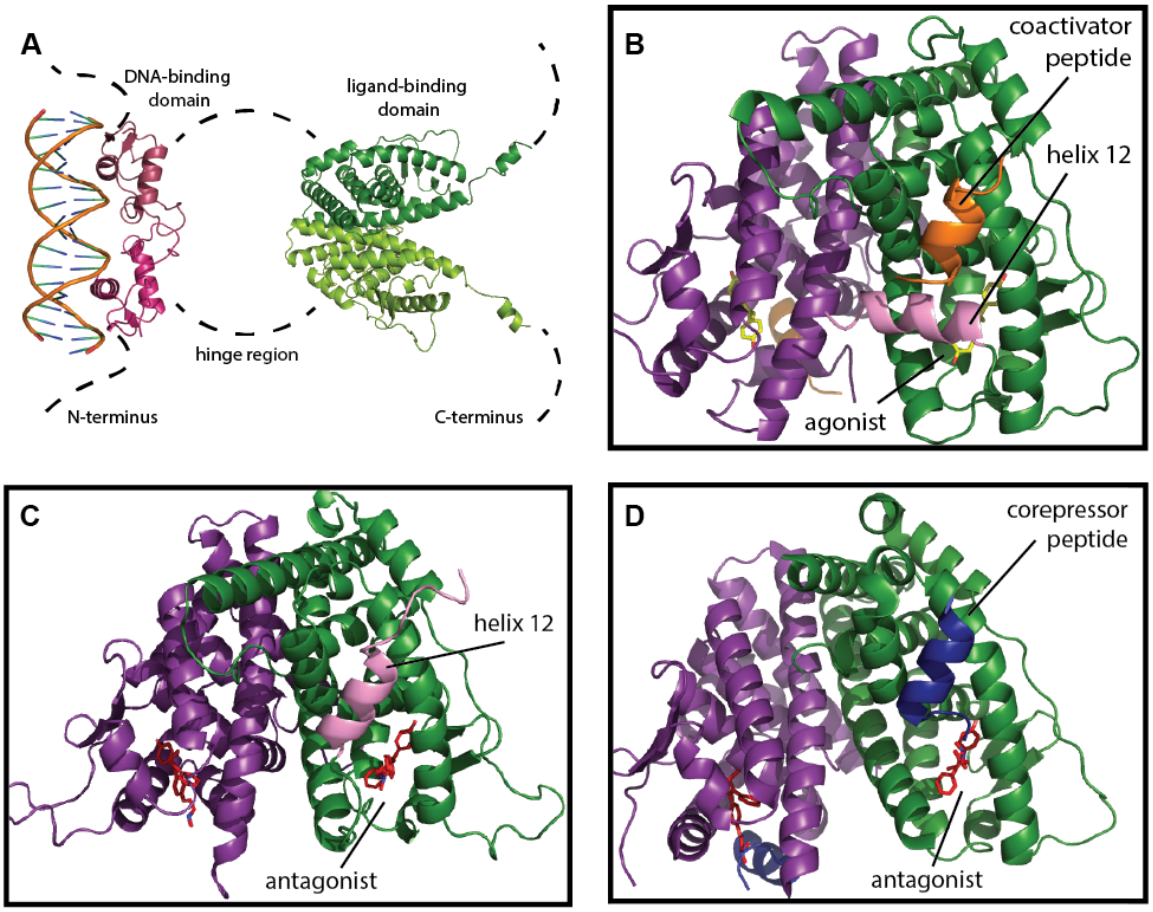


Figure 5.14. ERα structure and cofactor interactions. (A) Structural overview of the ERα dimer, showing crystal structures for the DBD (PDB 4AA6) and LBD (1A52). Other domains are disordered and shown by dashed lines. (B) When agonist is bound, C-terminal helix 12 of the LBD adopts a conformation stabilizing coactivator helical domain binding (PDB 3ERD). (C and D) Antagonist binding can cause helix 12 to block the coactivator helix binding site (C, PDB 3ERT) or allow corepressor helical domain binding (D, PDB 2JF9).

by a peptide recognition domain, activation function 2 (AF2), which forms a hydrophobic cleft in the LBD. When an agonist is bound at the LBD, C-terminal helix 12 adopts a conformation that allows the AF2 cleft to bind helical peptides from coactivator proteins with the consensus sequence L*XX*LL (**Figure 5.2B**).[35] Antagonists block activity by causing helix 12 to either block the AF2 cleft, mimicking the peptides of coactivators (**Figure 5.2C**);[35] or to dissociate from the receptor in a fashion that favors the binding of helices from corepressors, with consensus sequence L*XXX*L*XXX*L (**Figure 5.2D**).[36] Current efforts to tune the transcriptional activity of ERα and other nuclear receptors focus on the

design of small molecules that can induce these conformational changes or peptides that can stabilize them.

As our main motivation in studying ERα was to become comfortable with the incorporation and reactivity of photochemical crosslinkers, we have used this system as a positive control for the viability of our assays, using currently available structures to guide our mutagenesis. However, the same methodology could likely be applied as direct readout of changes in tertiary structure or recruitment of cofactors in vivo in future work. Indeed, two of the aforementioned examples of genetically encoded crosslinkers used in conjunction with mass spectrometry explored protein conformation in related nuclear receptors, the peroxisome proliferator-activated receptors.[28,31] However, these studies used purified receptors, and focused exclusively on mapping the conformation of a single subunit, rather than protein-protein interactions.

5.2.3 Pentameric ligand-gated ion channels (pLGICs)

As the pLGICs have been discussed in depth in Chapters 1-3, we will forgo a detailed introduction and give only a brief overview of the specific receptors relevant to the following sections.

There are several classes of serotonin (5-hydroxytryptamine, 5-HT) receptors expressed in humans, named 5-HT_1 through 5-HT_7. Of these, all are 7TM receptors except for the 5-HT_3 group, which are pLGICs (**Figure 1.2**).[37] There are 5 subunits expressed in the human brain, labeled A through E. Functional 5-HT_3 receptors can be homopentameric, with five A subunits, or heteropentameric, with at least 2 A subunits along with one other type of subunit.[38] Because the 5-HT_3A homopentamer expresses robustly in cultured cells, produces strong signals in functional assays, and does not require subunit stoichiometry ratio

optimization, much of our preliminary work on crosslinking in pLGICs has been carried out in this receptor.

The nicotinic acetylcholine receptor (nAChR) is a major focus of study for our group.[39,40] In humans, 16 different genes code for nicotinic receptor subunits (α1-7 and α9-10, β1-4, γ, δ, and ε), which can be arranged into pentameric arrangements with diverse pharmacological properties. Subunits are expressed differentially through regions of the central and peripheral nervous system, while signaling at the neuromuscular junction is mediated specifically by the muscle-type nAChRs, with stoichiometry $(\alpha 1)_2\beta 1\gamma\delta$.

The nAChRs containing the α6 subunit are of particular interest because of the recent evidence that their coexpression with the purinergic P2X receptors in dorsal root ganglia provides protection against exaggerated pain responses in mice.[6] In the same study, we have also shown that coactivation of α6-containing nAChRs and P2X receptors results in reduction of the total current in the *Xenopus laevis* oocyte model system. A molecular model for this phenomenon has not yet been formulated, and so application of crosslinking to a cell culture-based model system could yield very useful information toward this goal. Thus, we aimed to establish the technique in the α6β4 and α6β2 receptor subtypes. Because α4-containing receptors have also been shown to interact with P2X receptors[3] and have historically been much more amenable to functional expression in model systems, we have also attempted to use the α4β4 and α4β2 subtypes.

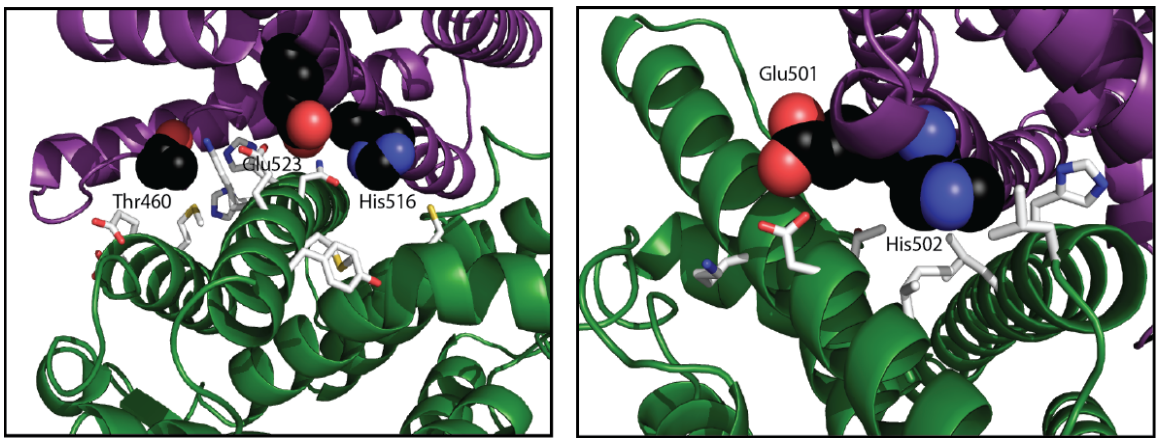


Figure 5.15. LBD residues selected for crosslinking of ERα dimers. Positions used for nonsense suppression are shown as black spheres. Side chains on the opposing subunit most likely to react during crosslinking are shown as white sticks (PDB 1A52).

5.3 Results and discussion

5.3.1 Exploring the ERα ligand-binding domain dimer interface

5.3.1.1 Nonsense suppression and crosslinking. We began our nonsense suppression experiments with an ERα construct containing an N-terminal enhanced GFP (eGFP) fusion. Without boiling lysate samples, some fraction of the eGFP remains intact and fluoresces even after SDS-PAGE, allowing for rapid evaluation of nonsense suppression and crosslinking without the need for western blotting. This is a crude assay, as the proportion of fluorescent eGFP need not be identical among all mutants and adducts; however, it has dramatically reduced the time required for condition optimization, and a more rigorous immunoblotting approach is under way.

For the purposes of a positive control, an ideal position for the incorporation of a photocrosslinking amino acid analog satisfies a number of conditions. (1) The positioning of the probe should be such that irradiation will likely result in crosslinking. The particular mechanism of reactivity of the selected analog should be taken into account, including functional group preferences for linking partners and the specific atom(s) within the side chain that form new bonds. (2) Care should be taken not to disrupt any structurally

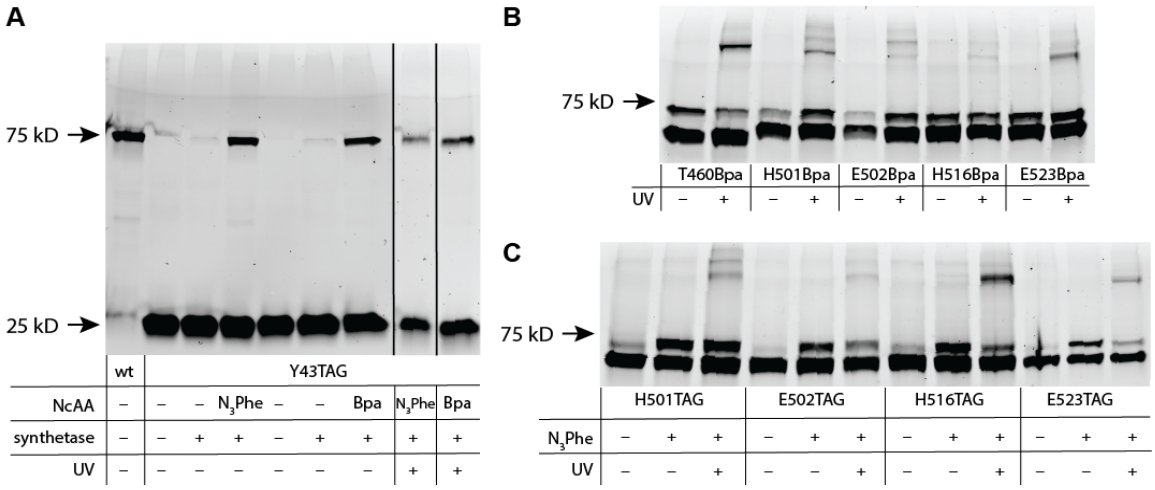


Figure 5.16. Nonsense suppression in ERα visualized by GFP fluorescence. (A) Results of N_3Phe and Bpa incorporation at a negative control site. Irradiation of these cells appears to yield no covalent dimer. (B and C) Irradiation of cells expressing ERα variants with Bpa (B) or N_3Phe (C) incorporated at the LBD, with negative controls.

or functionally important noncovalent interactions. (3) Additional substituents (i.e. N_3 or Bz for N_3Phe or Bpa) should not create steric clashes that disfavor complex formation.

X-ray crystallography data for the dimeric ERα ligand-binding domain show a number of residues that fit most or all of these criteria (**Figure 5.3**).[33] We performed site-directed mutagenesis to insert TAG mutations at 5 positions: Thr460, His501, Glu502, His516, and Glu523. We also inserted a TAG mutation at an unrelated position in the unstructured N-terminal domain, Tyr43, as a negative control for the crosslinking experiment.

We cotransfected HEK293T cells with the mutant constructs, along with either the Bpa or N_3Phe synthetase plasmid. Following transfection, we supplemented the cell growth medium with the appropriate photoreactive amino acid. After allowing protein expression, we introduced 10 μM agonist (E2) to induce receptor dimerization, and incubated for another 24 h. We then lysed the cells and analyzed them by electrophoresis, imaging via GFP fluorescence (**Figure 5.4**). We observe strong production of a protein at ~75 kD, the same apparent molecular weight as the ERα construct. Because nonsense suppression is not

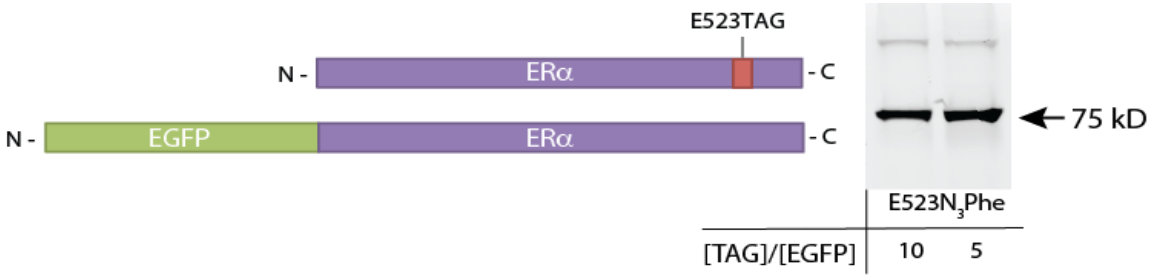


Figure 5.17. Confirmation of homodimer crosslinking. To ensure that the observed high molecular weight band corresponds to crosslinking of two ERα subunits, we cotransfected receptors containing either nonsense mutations or eGFP fusions in ratios of 10:1 and 5:1 (w/w).

perfectly efficient, a second band also appears at a lower molecular weight, corresponding to the truncated product. In negative control lanes lacking the synthetase plasmid, no full-length protein is observed (**Figure 5.4A**). However, if the synthetase plasmid is present, weak bands do still appear at the molecular weight of the full-length fusion protein even in the absence of the non-canonical amino acid supplement (**Figure 5.4A,C**). This species is most likely produced by nonspecific acylation of the suppressor tRNA with tyrosine by the mutant synthetase, which was derived from an *E. coli* tyrosine synthetase.

We next subjected the transfected cells to 45 minutes of irradiation at 354 nm with a 190-mW LED. Analysis of the lysates revealed new bands at high molecular weights, which varied in intensity and mobility based on the mutation site and the identity of the crosslinker (**Figure 5.4B,C**). These bands did not appear upon irradiation of cells expressing the Y43TAG mutants, which we would not expect to be at the interface of dimerization. Moreover, no new products were observed when cells lacking either the amino acid or the synthetase were irradiated (controls performed only for the E523N_3Phe mutant; data not shown).

This evidence strongly suggests that the newly formed species are covalently linked dimers including at least one ERα subunit. Although the dimerization affinity of the receptor should be quite high in the presence of E2,[41] it is also possible that the receptor is crosslinking with chaperone proteins, transcription cofactors, or other proteins that interact

with the ligand-binding domain. To show that we are capable of forming an adduct between two authentic ERα subunits, we used the following experimental approach. Two constructs were cotransfected – one containing an N-terminal eGFP fusion without a nonsense mutation, and one containing a nonsense mutation, but lacking the fluorescent protein fusion (**Figure 5.5**). After irradiation of these cells and SDS-PAGE, two bands were observed – one at the usual size for the full-length eGFP-ERα fusion, and one band at a higher molecular weight. Because the eGFP-containing protein lacks a photoreactive amino acid, the new band is presumably created specifically by the crosslinking of the invisible ERα with the fusion protein. The relative weakness of this band is likely explained by the presence of either only one or zero eGFP molecules per adduct, rather than two, as in the previous experiments. However, its presence is indicative that the homodimer exists among the new bands formed from irradiation, and suggests that the strong band at ~75 kD is mostly of this composition.

5.3.1.2 Time-dependence of photocrosslinking. Presumably, the fraction of dimeric ERα observed at completion of the photoreaction should depend upon the thermodynamic stability of the noncovalent dimer, providing a potential readout for the effects of mutations or ligands on the dimerization equilibrium. Thus, we were curious whether the rate of formation of the photoadduct was dependent upon the position or identity of the crosslinker. We proceeded with a set of time-course experiments (Figure 5.6). In each test, we irradiated dishes of identically transfected and treated HEK293T cells for varying durations, followed by immediate lysis and SDS-PAGE. The observed bands corresponding to the photoadduct displayed clear time-dependence, which we roughly quantified through densitometric analysis. It is worth noting that photobleaching of eGFP is possible during

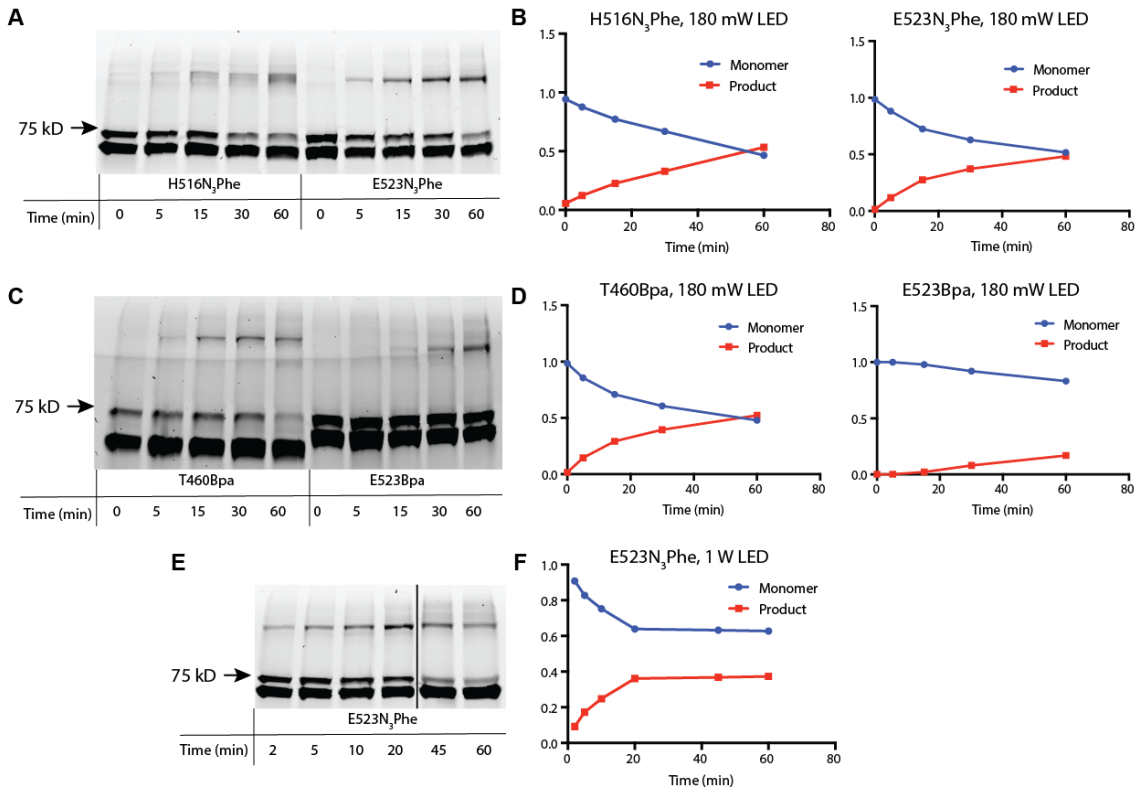


Figure 5.18. Time dependence of photocrosslinking. Time-course experiments for crosslinking of two N_3Phe mutations (A and B), for two Bpa mutations (C and D), and using a higher-power LED (E and F). All plots show integrated band intensities normalized to the sum of the intensities for the monomer and product bands.

the crosslinking reaction. In an attempt to correct for an overall rate of decay of fluorescence, we chose to normalize the densitometry measurements to the sum of the integrations for the monomer and adduct bands. At all sites examined, adduct intensity increased monotonically over time relative to the full-length monomer band. Interestingly, N_3Phe and Bpa formed significantly different final proportions of crosslinked dimer when incorporated at Glu523. This difference could be due to unfavorable steric clashes introduced by the bulky Bpa side chain during dimerization, as the region surrounding this residue appears somewhat congested in crystal structures (**Figure 5.3**).

We were curious whether the rate of crosslinking was limited by the photon flux of the light source, so we also subjected cells expressing E523N_3Phe to irradiation with a more powerful 1-W LED (**Figure 5.6E,F**). Interestingly, the crosslinking reaction appeared

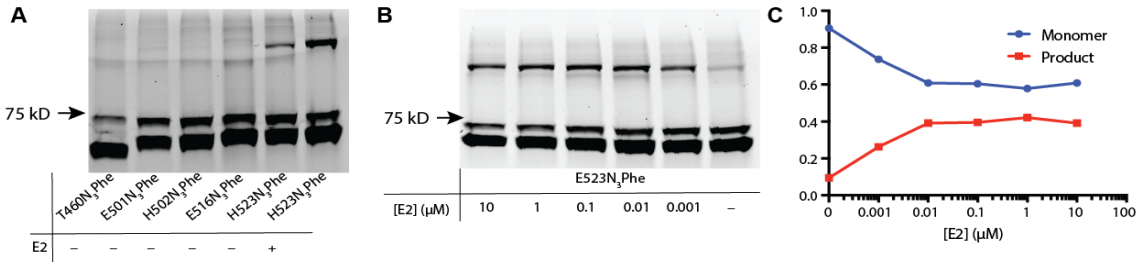


Figure 5.19. Ligand dependence of photocrosslinking. (A) In the absence of E2, irradiation results in weaker dimer bands than observed in the presence of the ligand. H523N_3Phe still produces significant dimer, but noticeably less than with 10 μM E2. (B) Results of crosslinking E532N_3Phe in the presence of various concentrations of E2. (C) Integrated band intensities from (B), normalized to the sum of the values determined for the monomer and dimer bands.

to reach completion within 20 minutes, but the final proportion of photoadduct was somewhat lower than that produced with the 180-mW LED. There was also a significant overall decrease in fluorescence in both the monomer and adduct bands after prolonged irradiation with the stronger LED, likely due to exacerbated photobleaching of the eGFP moiety.

5.3.1.3 Ligand-dependence of crosslinking. An ideal assay for the thermodynamics of receptor activation should be able to corroborate the transcriptional output of the system with dimerization over a relevant dynamic range of agonist concentrations. Following expression of ERα crosslinking mutants, we treated cells with between 1 nM and 10 μM E2 before analysis. No significant effect was observed on the final integration of the photoadduct band until E2 concentrations dropped to 1 nM (**Figure 5.7B,C**). This came as a surprise, since a previous in vitro monomer exchange study has shown that ERα-LBD dimer affinity varies over an E2 concentration range from 10 nM to 1 μM.[41] In contrast, the agonist response in our system seems to be already saturated once extracellular agonist concentrations have reached a threshold of 10 nM. The dynamic range could be modulated by the presence of coactivator or corepressor species, which might shift the binding equilibrium by sequestering ERα dimers in larger protein complexes. This explanation is in line with a recent in vivo fluorescence complementation assay that has produced E2 dose-

response curves closer to what we observe.[42] Even in the absence of added E2, some extent of covalent dimerization was still observed. This could potentially be due to adventitious estrogens present in the cellular growth medium, including components of fetal bovine serum or the phenol red pH indicator dye. Efforts to obtain lower baseline signal in the absence of added agonist are in progress.

5.3.2 Progress toward intersubunit crosslinking in the human 5-HT$_3$A receptor

5.3.2.1 Nonsense suppression and photocrosslinking. Our initial efforts to carry out nonsense suppression in 5-HT$_3$A were aimed at finding a positive control for intermolecular photocrosslinking. Because a crystal structure of the homopentameric receptor is currently available,[43] the simplest target was to attempt crosslinking between subunits. We again used a fluorescent protein fusion to aid in detection of the receptor following SDS/PAGE. To minimize the impact of the fusion on protein structure or oligomerization, we used a construct with a CFP inserted into the flexible intracellular loop (see experimental section for details). In this case, cell lysis conditions were somewhat more challenging to optimize than with the soluble ERα. Despite attempting receptor denaturation with various reducing reagents (TCEP, BME, DTT) and solubilization with numerous detergents (NP-40, Triton-100, Tween-20, SDS, CHAPS, DDM), all lysis conditions resulted in gels with broad, hazy bands and what appear to be protein aggregates at high apparent molecular weights. We settled on the use of a standard RIPA buffer containing NP-40, SDS, and deoxycholate as surfactants. Although this buffer did not successfully disaggregate the high molecular weight species, it seemed to be the most effective of those we tried.

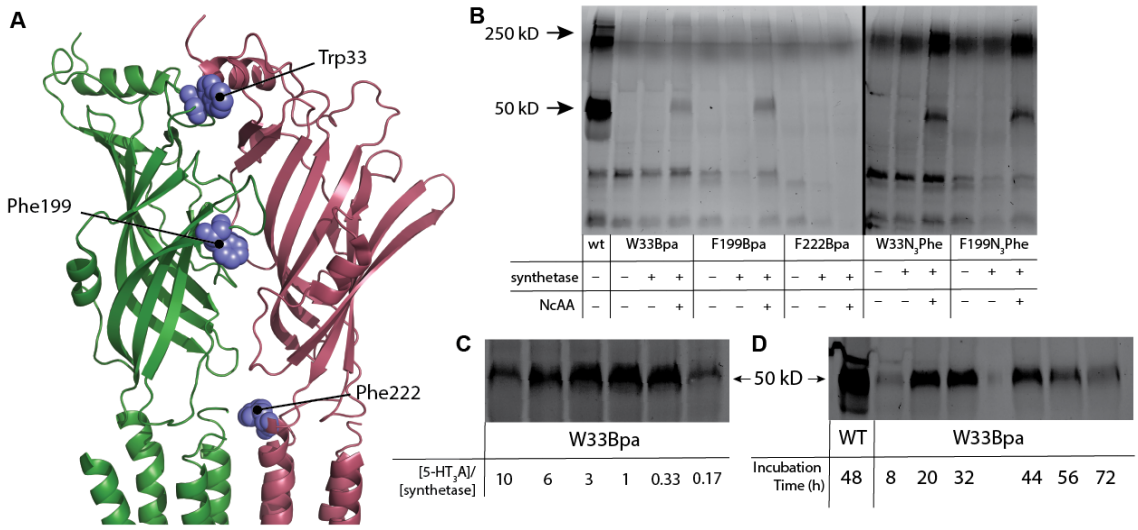


Figure 5.20. Nonsense suppression in the 5-HT$_3$A receptor. (A) Crosslinking sites selected at the subunit interfaces of the 5-HT$_3$A homopentamer (PDB 4PIR). (B) Full-length subunits containing Bpa or N$_3$Phe at Trp33 and Phe199. (C and D) Optimization of transfection ratios (C) and incubation time (D) for maximum expression of mutants. Receptors contained a CFP fusion in the M3-M4 loop; images show CFP fluorescence.

Based on the structure of the mouse 5-HT$_3$A receptor, we selected 3 corresponding positions in the human receptor that followed the aforementioned criteria for probable crosslinking sites: Trp33, Phe199, and Phe222 (**Figure 5.8A**). We used the same tRNA/synthetase pairs as in the ERα system, again transfecting the receptor and synthetase plasmids in a 3:1 ratio into HEK293T cells. Although the gels were not as crisp as in the ERα model system, we were able to discern a distinct band at the expected molecular weight for the 5-HT$_3$A monomer in two of the three mutants (**Figure 5.8B**). The band was not observed in lanes from cells lacking either the non-canonical amino acid or the synthetase plasmid, supporting our assignment of the species as full-length monomer. After some optimization, we concluded that a 3:1 or 1:1 transfection ratio with ~24-36 h incubation produced the greatest intensity of the monomer band (**Figure 5.8C,D**). A range of concentrations of non-canonical amino acid were also tested, but did not seem to significantly affect suppression efficiency.

As in the ERα model system, we then attempted irradiating cells expressing the mutant receptors in an effort to crosslink subunits. Although some slight differences in

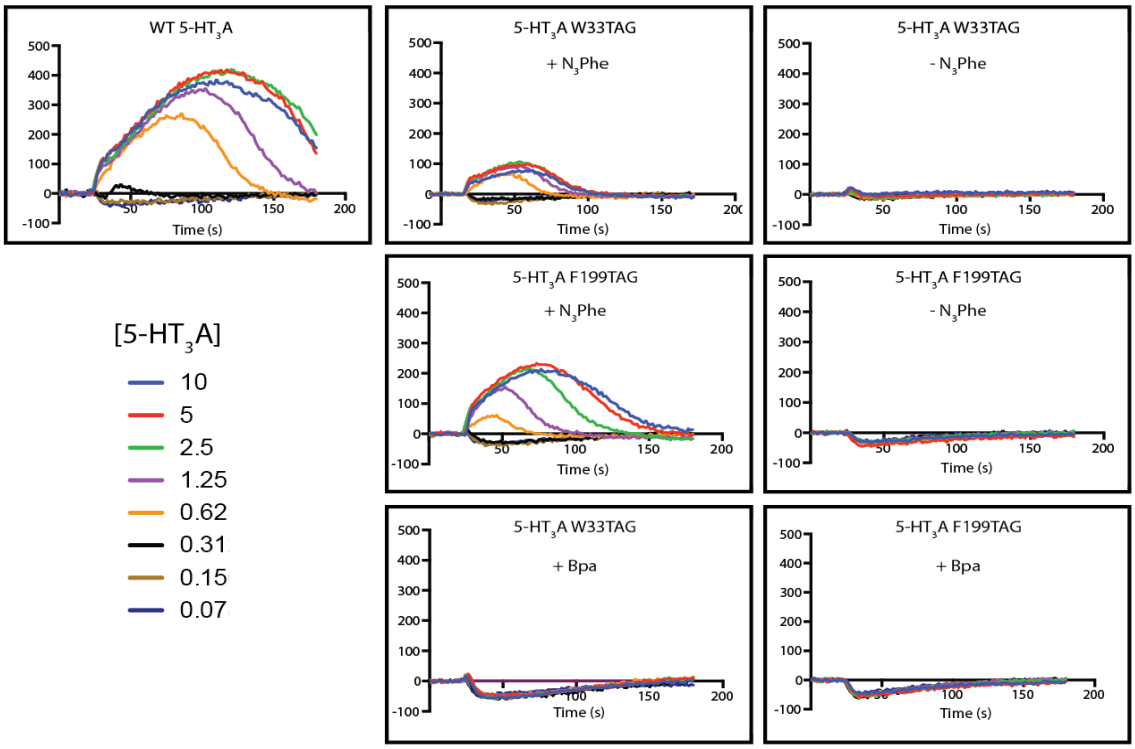


Figure 5.21. Functional validation of mutant 5-HT$_3$A receptors. Each box shows representative FlexStation3 fluorescence measurements from FLIPR membrane potential dye-loaded cells in 8 wells of a 96-well plate. Each well received a different dose of serotonin after ~25 s of recording.

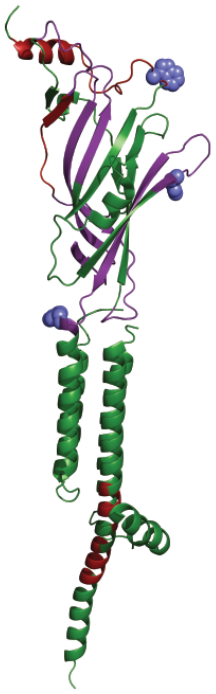
intensity were observable, the background signal in every gel lane made it impossible to discern any distinct bands corresponding to a dimeric or higher order photoadduct. In hopes that an immunoblotting detection method might provide less background, we also generated mutant receptors containing FLAG epitope tags in the M3-M4 loop and at the N- and C-termini. Unfortunately, none of the experiments with these mutants yielded convincing evidence of nonsense suppression.

5.3.2.2 Functional validation. Using these conditions, we prepared cells expressing mutant receptors for functional validation. For both N_3Phe mutants that gave visible monomer expression, we observed serotonin-dependent depolarization of plasma membranes (**Figure 5.9**). These currents are not as robust as those seen in cells expressing wild type receptors, consistent with the relatively low expression levels of the mutants. Interestingly, we saw no

membrane depolarization in cells expressing Bpa-containing receptors. Because monomer expression for the two was comparable, this suggests that either receptor assembly or function was compromised by the bulky benzoyl substitution introduced by the Bpa group at Phe222.

5.3.2.3 Mass spectrometric analysis.* Although confirming crosslinking by way of SDS/PAGE is beyond our present capabilities, we also began exploring the feasibility of analyzing 5-HT_3A receptors through HPLC-MS^2. The ultimate goal of this strategy would be to find crosslinked peptide fragments, identify which residues the unnatural crosslinker has reacted with, and then use these data to constrain a model of the protein-protein binding interface. However, membrane proteins are notoriously difficult to express, purify, digest, and ionize. There are several examples of the use of mass spectrometry to gain structural information about pLGICs, including the glycine receptor[44,45] and the muscle-type nAChR.[46] This section represents our first basic efforts toward obtaining sequence coverage of the wild type 5-HT_3A receptor. We attempted two methods for sample preparation: in-gel digestion and filter-aided sample preparation (FASP). For each approach, we enriched the receptor via immunoprecipitation before proteolytic digestion by targeting an N-terminal FLAG epitope. After elution from the beads, the sample was subjected to fractionation by SDS-PAGE. Even after immunoprecipitation, the protein content was still not high enough for bands to be visible from staining with Coomassie blue, so protein content was located by running a second gel in parallel and immunoblotting against the FLAG epitope. Both the gel regions corresponding to the protein aggregate and the monomer were excised, and in-gel reduction, alkylation, and tryptic digestion were performed (see experimental section for details). After

* These experiments were carried out with the kind assistance of Caltech's Proteome Exploration Laboratory.



		Peptide Sequence	Intensity[a]	Posterior Error Probability
Run #1	1	RDEIREVAR	125100000	1.98 E-2
	2	VGSVLDK	954460	1.94 E-2
	3	SPNIPYVYIR	1828300	1.97 E-2
	4	GVRPVRDWR	1435400	1.05 E-3
	5	LSDYLLTNYRK	6627600	1.82 E-2
	6	LSDYLLTNYR	-	3.67 E-3
	7	NTTRPALLR	679600	3.08 E-3
	8	SRNTTRPALLR	9362400	1.93 E-3
	9	CSPPPPPR	-	1.46 E-2
Run #2	10	AILNVDEKNQVLTTY	3050300	0
	11	FREFSMESSNYYAEMKF	63073	1.06 E-2
	12	KPLQVVTACSLDIY	836270	2.38 E-6
	13	KPLQVVTACSLDIYNFPF	156270	1.42 E-2
	14	SMESSNYYAEMKFY	89474	2.67 E-70
	15	VDVGKSPNIPYVY	2503400	5.70 E-50
	16	VYIRHQGEVQNY	105470	5.93 E-226
	17	YRQYWTDEFLQWNPEDF	96678	1.40 E-2

Table 5.3. Results of LC-MS² experiments on the 5-HT$_3$A receptor. Sequence coverage is depicted by regions highlighted in red (Run #1, in-gel tryptic digestion) and purple (Run #2, filter-aided chymotryptic digestion). Sites selected for mutagenesis are shown as spheres (PDB 4PIR). [a]Peptides without listed intensities were identified by MS-MS spectra triggered by precursor ions of other peptides, so a precursor ion with the matching m/z was not detected or its intensity was not recorded.

extracting peptides and desalting, the peptides were submitted to HPLC-MS^2 using a Thermo Orbitrap Fusion instrument.

The mass spectrometry analysis returned evidence for seven distinct peptide fragments covering 13% of the receptor (**Table 5.1**). Not surprisingly, all of the peptides detected were derived from the more hydrophilic extracellular and intracellular domains. Although detected peptides did span one of our chosen nonsense suppression sites (Trp33), most peptide intensities were quite low, ~10^6-10^7 compared with 10^9 for the most intense contaminant detected. This is somewhat worrisome, as nonsense suppression usually yields less protein than cells expressing the wild type receptor, and crosslinking will generate a number of different species, further reducing the abundance of each peptide product.

In hopes of minimizing the protein lost during gel extraction, we proceeded to attempt FASP. Receptors were similarly expressed and enriched using immunoprecipitation,

then loaded onto a 10 kD microcentrifuge filter. Following buffer exchange and removal of detergents, the adsorbed protein was reduced, alkylated, and digested. In this round of sample preparation, we opted to use chymotrypsin rather than trypsin, due to the increased number of potential cleavage sites in the transmembrane domain. After the product peptides were eluted, dried, and desalted, we again submitted them for HPLC-MS^2 analysis, this time using a Thermo Fisher Orbitrap Elite instrument. The analysis returned eight distinct peptide fragments covering 17% of the protein sequence (**Table 5.1**). Again, only peptides derived from the hydrophilic portions of the receptor were detected. Possibly because of the high promiscuity of the chymotryptic digestion, ion intensities were even lower than in the results obtained from the in-gel digestion (10^4-10^6, while the top 100 ions were 10^7-10^{10}). However, combining the coverage from both preparation procedures, we were able to observe peptides covering 59% of the receptor's extracellular domain and 27% of the entire receptor.

5.3.3 Nonsense suppression in nAChRs

Because of our interest in interactions between α6-containing nAChRs and P2X receptors, we began by testing nonsense suppression in the α6 subunit. There are no available structures of this subunit, so we used the web-based homology-modeling server SWISS-MODEL[47] to predict the structure of the subunit based upon alignment with the 5-HT_3A receptor structure and sequence (**Figure 5.10A**). We introduced a C-terminal HA tag into the α6 subunit, and then incorporated three TAG mutations at positions that could potentially be involved in interactions with other receptors: Tyr142 and His216 in the

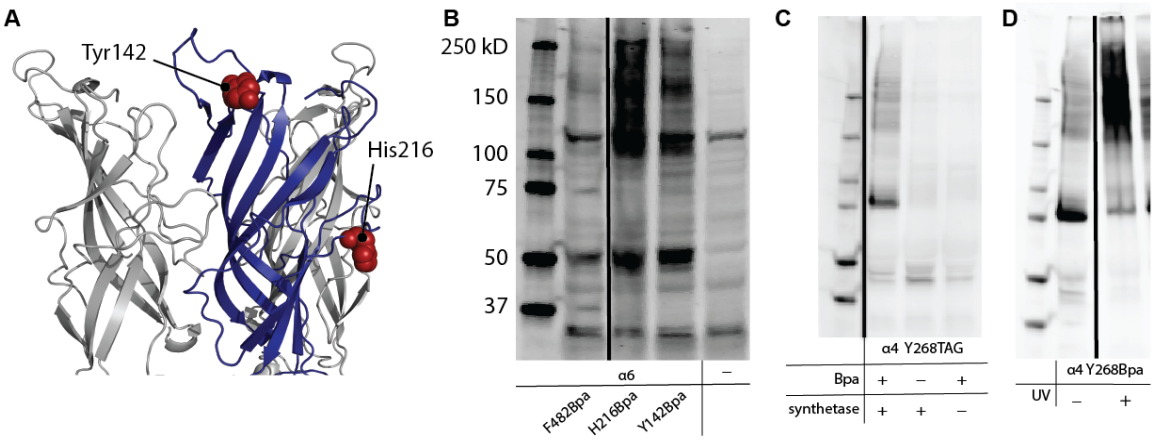


Figure 5.22. Nonsense suppression in nAChRs. (A) Homology model of α6 subunits based upon the structure of the mouse 5-HT$_3$A receptor. Two solvent-exposed sites selected for nonsense suppression in the ECD are highlighted as red spheres; the third site, Phe482, was chosen by sequence alignment of the M4 helix. (B) Western blot against C-terminally HA-tagged rat α6L9'S subunits, suppressing with Bpa at each position. Negative control cells were treated only with transfection reagent. (C and D) Nonsense suppression with Bpa at Tyr268 in the mouse α4 receptor, containing a GFP-HA fusion in the M3-M4 loop, and an L9'A mutation (C), and subsequent attempts at photocrosslinking in cells expressing the same mutant (D). The α4 mutant was cotransfected with a β2‡L9'S subunit. Blots were imaged with a mouse primary antibody against the HA epitope and a secondary Alexa Fluor 680-conjugated goat antibody.

extracellular domain, and Phe482 in the M4 transmembrane helix (facing out into the lipid environment).

We cotransfected each of the α6 mutants with the β4 construct and the Bpa synthetase plasmid in a ratio of 9:1:2.7, then introduced a solution of Bpa. After 48 h of incubation, we lysed the cells, subjected the lysates to SDS-PAGE, and blotted against the HA epitope tag. As with the 5-HT$_3$A receptor, the resulting lanes were somewhat smeared. Nevertheless, a definite signal was observed at the expected molecular weight of the monomer, aligning well with a band in the parent α6 sample (**Figure 5.10B**). Unfortunately, membrane potential dye experiments to validate the function of even the wild type α6β4 receptor in HEK293T cells revealed no acetylcholine-dependent receptor. Despite clear expression of both subunits in the cells, surface protein isolation experiments indicated relatively low membrane expression levels of either subunit, possibly causing the lack of function we observed.

In the past, our group has had success in elevating the membrane expression levels of α6β2 receptors in *Xenopus laevis* oocytes by introducing gain-of-function L9'S pore mutations in both subunits, along with an ER export motif in the intracellular domain of β2 (notated as $\beta2^{\ddagger}$).[48] We subcloned the $\beta2^{\ddagger}$ gene into our mammalian expression vector and cotransfected it along with the α6 plasmid, but again, we observed no acetylcholine-dependent activity.

In contrast, functional expression of α4β4 receptors in HEK293T cells had previously been achieved in our group, suggesting that the α6-containing receptors were not optimal for these studies. We were indeed able to reproduce acetylcholine-dependent membrane depolarization by cotransfecting an α4 construct containing a GFP fusion in the M3-M4 loop, along with a wild type β4 construct. Proceeding with the α4-GFP construct, we performed a sequence alignment with the 5-HT_3A receptor. Based on sequence alignment with the 5-HT_3A receptor, we selected Tyr268 in the M4 helix, which we expect to be lipid-exposed. We were able to easily confirm successful nonsense suppression with both N_3Phe and Bpa (**Figure 5.10C**). However, we once again did not detect acetylcholine-induced membrane depolarization in cells expressing the mutant receptors.

Interestingly, irradiation of cells expressing the Y268Bpa mutant did appear to induce a significant difference in the band pattern observed: the monomer band was reduced in intensity, and instead the hazy smear at the top of the lane dramatically increased in intensity (**Figure 5.10D**). This change could possibly be explained by crosslinking of the receptor to lipids, resulting in reduced electrophoretic resolution, but further identification of any new species was not pursued.

5.4 Conclusions

We have succeeded in using the nonsense suppression approach to incorporate non-canonical photoreactive amino acids into several proteins in mammalian cells. In the ERα system, this has enabled us to crosslink receptor dimers in a fashion that depends upon the concentration of activating ligand present. This sets the foundation for informative in vivo experiments that could reveal the molecular factors governing interactions with coactivators or corepressors that have not yet been crystallized. Optimization of the concentration-dependent dimerization experiments could also eventually yield a thermodynamic assay for receptor activation equilibria. Important steps toward making this type of assay possible will include background signal reduction, increasing the dynamic range of the assay, improvement of current functional validation tests, and removal of the reliance upon fluorescent protein fusions.

Our nonsense suppression results in pLGICs have also been somewhat promising. Although discouraged by the challenges of handling and analyzing highly hydrophobic proteins, we have succeeded in introducing crosslinking probes into sites in three different pLGIC subunits, including functional 5-HT_3A homopentamers. In order for such probes to be useful in the investigation of protein-protein interfaces, significant optimization of cell lysis, protein denaturation, and electrophoresis conditions may be required. Furthermore, prior to extending mass spectrometry techniques to analysis of crosslinker-containing receptors, extensive work may be needed to improve the digestion and ionization of receptor domains. We advise that future efforts focus first on validation of structurally characterized protein-protein interactions with these techniques, ideally involving soluble binding partners of the pLGICs.

5.5 Experimental procedures

5.5.1 Molecular biology

The pcDNA3.1(+) expression vector was used for expression of proteins in mammalian cells, and was purchased from Invitrogen (Carlsbad, CA). Genes for human wild type 5-HT3A and eGFP-ERα; rat α6L9'S and β2‡L9'S; and mouse α4-GFP and wild type β4 were subcloned into the vector by restriction digest and subsequent ligation. Small modifications, such as introduction of epitope tags or point mutations, were accomplished using the QuikChange method (Stratagene). Introduction and removal of GFP fusions were carried out using the Q5 mutagenesis kit (New England Biosciences). In the eGFP-ERα construct, eGFP was connected to the N-terminus of ERα by way of a 15-residue linker, SGRSRAASNSAVDGT. In the mouse α4-GFP construct, the insertion point for the HA-GFP fusion was between Pro428 and Arg429.

For nonsense suppression experiments, PU6 plasmids[49] encoding the tRNA/synthetase pairs for Bpa and N_3Phe were a kind gift from the lab of Professor Michelle Krogsgaard. These plasmids were amplified and used directly in experiments without modification.

5.5.2 Mammalian cell culture and protein expression

All incubations were performed at 37 °C under 5% CO_2 unless otherwise indicated. All manipulations requiring opening of culture flasks or dishes was carried out in a class II biosafety cabinet. The growth medium used for human embryonic kidney cells (HEK293T) contained 91% DMEM/F12 (Gibco) and 9% heat-inactivated FBS (Gibco), with 100 U/mL penicillin and 100 μg/mL streptomycin.

HEK293T cells purchased from ATCC were stored as a DMSO suspension at -80 °C. Freezedowns were immersed in a 37 °C water bath for 2 minutes, until 80% thawed, then resuspended with a serological pipet in prewarmed cellular growth medium. The cells were then pelleted by centrifuging at 1000 g in a 15-mL conical tube, and the supernatant was removed and discarded. Cells were resuspended by gently rocking in 14 mL prewarmed growth medium, and then transferred to a T75 flask. Cells were incubated for 24 h, at which point the medium was exchanged for 10 mL fresh medium.

Cells were propagated every 3 days (or whenever they reached 80-90% confluence). The old medium was removed, the cells were washed briefly with 10 mL cold phosphate-buffered saline (PBS), and 1 mL TrypLE dissociation reagent (Gibco) was added. The cells were incubated for 5 min to allow for dissociation, then TrypLE was deactivated by addition of 9 mL fresh growth medium and cells were resuspended by pipetting forcefully five times with a serological pipet. A new T75 flask was filled with 9 mL fresh growth medium, and 1 mL of the cell suspension was added. The new flask was then gently rocked to ensure even distribution of cells on the growth surface, and then returned to incubation.

Proteins of interest were expressed using the Xfect reagent (Clontech). At the time of propagation, 35-mm petri dishes or 6-well plates were seeded by diluting 300 μL of the cell suspension into 2.7 mL fresh growth medium. For larger experiments, a 10-cm petri dish was seeded by diluting 1 mL of cell suspension into 9 mL fresh growth medium. Cells were then incubated until reaching ~50-60% confluence. Before transfecting, the old medium was removed and replaced with 1 mL of fresh growth medium (10 mL for a 10-cm plate).

Transfection mixtures were made by premixing plasmid DNA in the desired stoichiometry, diluting with Xfect buffer to a total volume of 100 μL (600 μL for a 10-cm dish), and then adding Xfect polymer (0.3 μL/μg plasmid DNA). Typical total DNA loading

was 10 μg for a 35-mm dish or 30 μg for a 10-cm dish. Following incubation at room temperature for 10 min, the entire transfection mixture was added dropwise to the dish, which was then rocked gently to ensure homogeneity. The cells were then incubated for 4 hours before the transfection reagent was removed and replaced with 3 mL fresh growth medium. For nonsense suppression experiments, the media added after transfection contained 0.05 mM Bpa or 0.5 mM N_3Phe.

5.5.3 Photocrosslinking

Photocrosslinking experiments were performed only on cells seeded in individual 35-mm dishes to allow for independent UV irradiation. After incubation for 24-48 h, medium was removed, the cells were rinsed once gently with fresh growth medium to remove residual extracellular photoreactive amino acids, and then 1 mL of fresh medium added. The dishes were then irradiated from below with a 190-mW 365 nm LED for the desired crosslinking duration, using a ring stand with two clamps to ensure consistent positioning among experiments. For ERα experiments, E2 was added directly to the growth medium 24 h following transfection to induce receptor dimerization, and irradiation was carried out after another 24 h.

5.5.4 Lysis, immunoprecipitation, SDS-PAGE, and immunoblotting

Following protein expression (and, if desired, crosslinking), media were removed and discarded. Lysis buffers were selected based on the application. For ERα experiments, 1% SDS in PBS was used. For most pLGIC experiments, a RIPA buffer (1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 50 mM Tris HCl, pH 8.0) was used. In experiments requiring enrichment by immunoprecipitation, 1% Triton X-100 in PBS was used to avoid denaturation of antibodies. The desired lysis buffer (typically 300 μL for a 35-

mm dish or 1 mL for a 10-cm dish) was then added, and the cells were gathered using a 1-mL pipette tip or disposable cell scraper, then transferred to a 1.5-mL Eppendorf tube. Lysates were homogenized with an immersion sonicator for 10 s, then stored at -80 °C or used immediately.

Immunoprecipitations were carried out using immobilized protein A/G agarose (Pierce). For each 10 cm plate of HEK293T cells, 100 µL resin was washed 3 times with lysis buffer. 300 µL of Triton X-100 was used for lysis of each plate. Following sonication, samples were centrifuged and supernatant was transferred to the resin, where it was incubated overnight at 4 °C with rocking. The resin was then washed 3 times for 5 minutes each with lysis buffer, before eluting with 250 µL of 3xFLAG peptide (150 ng/µL) for 1 h at 4 °C. A second identical elution was then performed to remove remaining protein, and samples were used directly for SDS-PAGE fractionation or FASP.

For most experiments, precast Biorad Mini-Protean TGX 15-well gels with an AnykD gradient were used, running samples for ~1 h at 150 V. Fluorescent protein fusions were imaged immediately using a Typhoon FLA 9000 with 473 nm laser excitation and an LPB (510LP) detection filter. For immunoblotting, gels were transferred to LF-PVDF membranes, blocked using 5% milk in PBS with 0.1% Tween-20 (PBST) for 1 hour, and then blotted overnight with a solution of the appropriate primary antibody in the same blocking medium. Membranes were then washed 4 x 10 min with PBST, treated with a solution of an Alexa Fluor 680-conjugated secondary antibody in blocking medium for 1 hour, and washed again 4 x 10 min before imaging on a LI-COR Odyssey with 685 nm laser excitation and 700 nm detection. Densitometric analysis was performed using ImageJ.

5.5.6 Membrane potential dye assay

Immediately following transfection, cells in 6-well plates or 35-cm dishes were dissociated with 300 μL TrypLE per well (or 1 mL, for 10-cm dishes). After incubation for 5 min, the cells were resuspended in 2.7 mL fresh growth medium (containing Bpa or N_3Phe, if necessary) with a 1-mL pipette tip (or 9 mL, for 10-cm dishes). The suspension was then transferred to a V-shaped basin. A multichannel pipet was used to plate 100 μL of the suspension into each well in three rows of a clear-bottom 96-well culture plate (or into all 96 wells, for a 10-cm dish). The cells were then incubated for 24-48 h before performing the functional assay.

Thirty minutes prior to the experiment, growth medium was removed from all wells by a multichannel pipet and discarded. The cells were then loaded with the commercially available FLIPR Blue Assay Kit membrane potential dye (Molecular Devices). The solid dye was dissolved as recommended by the manufacturer, and then diluted 1:1 in Flex buffer (115 mM NaCl, 1 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM D-glucose, and 10 mM HEPES at pH 7.5)[50] and transferred to a V-shaped basin. A multichannel pipette was used to add 100 μL to each well, and the plate was incubated for 30 min. Data collection was performed using the FlexStation3 in Flex mode. Fluorescence readings were taken from each well every 1.52 s for a total of 180 s. After ~25 s, 100 μL of ligand solution was added to each well via the liquid handler arm. Data were processed using Microsoft Excel, at which point a constant baseline correction corresponding to the mean signal observed prior to ligand addition was subtracted from the trace for each well.

5.5.7 Proteolysis and HPLC-MS analysis

For in-gel digestions, work was performed in a laminar-flow workspace. Gel pieces were excised and placed into Eppendorf tubes. Sample processing was performed as optimized by Caltech's Proteome Exploration Laboratory (PEL). Briefly, gel pieces were destained with acetontrile, reduced with DTT, alkylated with iodoacetamide, digested with trypsin, extracted with acetonitrile and formic acid, and then desalted using a C18 Ziptip (Millipore). For a detailed procedure, see the PEL website (pel.caltech.edu).

For FASP, eluted proteins from immunoprecipitation were applied to a 10 kD size-exclusion spin column. Reduction, alkylation, digestion, and elution were carried out as previously described.[51] Peptides were desalted by flowing through a C18 plug.

Peptide samples were fractionated by HPLC and analyzed using Thermo Scientific Orbitrap Elite or Orbitrap Fusion instruments, according to protocols optimized by the PEL.[52] Specific procedures are available upon request.

Data were analyzed using the MaxQuant software, searching against the sequence for the human 5-HT_3A receptor and a contaminant database (246 entries). The digestion enzyme was specified as either trypsin or chymotrypsin, depending on the sample preparation, using up to 2 missed cleavages. Variable modifications included methionine oxidation (+15.9949) and N-terminal protein acylation (+42.0106), with a fixed modification of cysteine carbamidomethylation (+57.0215). Precursor mass tolerance was less than 4.5 ppm. False discovery rates were fixed at 1% and estimated using a decoy database search performed by MaxQuant.

5.6 References

1. Dacosta, C. J. B. & Baenziger, J. E. Gating of pentameric ligand-gated ion channels: structural insights and ambiguities. *Structure* **21,** 1271–1283 (2013).
2. Nys, M., Kesters, D. & Ulens, C. Structural insights into Cys-loop receptor function and ligand

recognition. *Biochem. Pharmacol.* **86,** 1042–1053 (2013).
3. Khakh, B. S. *et al.* An angstrom scale interaction between plasma membrane ATP-gated P2X2 and α4β2 nicotinic channels measured with fluorescence resonance energy transfer and total internal reflection fluorescence microscopy. *J. Neurosci.* **25,** 6911–6920 (2005).
4. Shrivastava, A. N., Triller, A., Sieghart, W. & Sarto-Jackson, I. Regulation of GABAA Receptor Dynamics by Interaction with Purinergic P2X2 Receptors. *J. Biol. Chem.* **286,** 14455–14468 (2011).
5. Birdsong, W. T. *et al.* Sensing Muscle Ischemia: Coincident Detection of Acid and ATP via Interplay of Two Ion Channels. *Neuron* **68,** 739–749 (2010).
6. Wieskopf, J. S. *et al.* The nicotinic α6 subunit gene determines variability in chronic pain sensitivity via cross-inhibition of P2X2/3 receptors. *Sci. Transl. Med.* **7,** 287ra72–287ra72 (2015).
7. Toulmé, E. *et al.* An intracellular motif of P2X3 receptors is required for functional cross-talk with $GABA_A$ receptors in nociceptive DRG neurons. **102,** 1357–1368 (2007).
8. Boué-Grabot, E., Toulmé, E., Emerit, M. B. & Garret, M. Subunit-specific coupling between gamma-aminobutyric acid type A and P2X2 receptor channels. *J. Biol. Chem.* **279,** 52517–52525 (2004).
9. Boué-Grabot, E. *et al.* Cross-talk and co-trafficking between ρ1/GABA receptors and ATP-gated channels. *J. Biol. Chem.* **279,** 6967–6975 (2004).
10. Boué-Grabot, E. *et al.* Intracellular cross talk and physical interaction between two classes of neurotransmitter-gated channels. *J. Neurosci.* **23,** 1246-1253 (2003).
11. Li, Y. & Xu, T.-L. State-dependent cross-inhibition between anionic GABAA and glycine ionotropic receptors in rat hippocampal CA1 neurons. *Neuroreport* **13,** 223 (2002).
12. Nakazawa, K. ATP-activated current and its interaction with acetylcholine-activated current in rat sympathetic neurons. *J. Neurosci.* **14,** 740-750 (1994).
13. Sokolova, E., Nistri, A. & Giniatullin, R. Negative cross talk between anionic $GABA_A$ and cationic P2X ionotropic receptors of rat dorsal root ganglion neurons. *J. Neurosci.* **21,** 4958-4968 (2001).
14. Decker, D. A. & Galligan, J. J. Cross-inhibition between nicotinic acetylcholine receptors and P2X receptors in myenteric neurons and HEK-293 cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **296,** G1267–76 (2009).
15. Decker, D. A. & Galligan, J. J. Molecular mechanisms of cross-inhibition between nicotinic acetylcholine receptors and P2X receptors in myenteric neurons and HEK-293 cells. *Neurogastroenterol. Motil.* **22,** 901–8– e235 (2010).
16. Karanjia, R. *et al.* Cross-inhibitory interactions between GABAA and P2X channels in myenteric neurones. *Eur. J. Neurosci.* **23,** 3259–3268 (2006).
17. Barajas-López, C., Montaño, L. M. & Espinosa-Luna, R. Inhibitory interactions between 5-HT3 and P2X channels in submucosal neurons. *Am. J. Physiol. Gastrointest. Liver Physiol.* **283,** G1238–48 (2002).
18. Barajas-López, C., Espinosa-Luna, R. & Zhu, Y. Functional interactions between nicotinic and P2X channels in short-term cultures of guinea-pig submucosal neurons. *J. Physiol. (Lond.)* **513,** 671–683 (1998).
19. Zhou, X. & Galligan, J. J. Non-additive interaction between nicotinic cholinergic and P2X purine receptors in guinea-pig enteric neurons in culture. *J. Physiol. (Lond.)* **513,** 685–697 (1998).
20. Skach, W. R. Cellular mechanisms of membrane protein folding. *Nat. Struct. Mol. Biol.* **16,** 606–612 (2009).
21. Avila, J. R., Lee, J. S. & Toriia, K. U. Co-Immunoprecipitation of Membrane-Bound Receptors. *The Arabidopsis Book* **13,** e0180 (2015).
22. Paramelle, D., Miralles, G., Subra, G. & Martinez, J. Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* **13,** 438–456 (2013).
23. Chin, J. W. & Schultz, P. G. In vivo photocrosslinking with unnatural amino Acid mutagenesis. *Chembiochem* **3,** 1135–1137 (2002).
24. Farrell, I. S., Toroney, R., Hazen, J. L., Mehl, R. A. & Chin, J. W. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat Meth* **2,** 377–384 (2005).
25. Tanaka, Y., Bond, M. R. & Kohler, J. J. Photocrosslinkers illuminate interactions in living cells. *Mol. Biosyst.* **4,** 473–480 (2008).
26. Sinz, A., Arlt, C., Chorev, D. & Sharon, M. Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. *Protein Sci.* **24,** 1193–1209 (2015).
27. Sinz, A. Investigation of protein–protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal. Bioanal. Chem.* **397,** 3433–3440 (2010).
28. Schwarz, R. *et al.* Monitoring Conformational Changes in Peroxisome Proliferator-Activated Receptor α by a Genetically Encoded Photoamino Acid, Cross-Linking, and Mass Spectrometry. *J. Med. Chem.*

56, 4252–4263 (2013).
29. Forné, I. *et al.* Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. *Mol. Cell Proteomics* **11,** M111.012088–M111.012088 (2012).
30. Berg, M. *et al.* An In Vivo Photo-Cross-Linking Approach Reveals a Homodimerization Domain of Aha1 in S. cerevisiae. *PLoS ONE* **9,** e89436 (2014).
31. Schwarz, R., Tänzler, D., Ihling, C. H. & Sinz, A. Monitoring Solution Structures of Peroxisome Proliferator-Activated Receptor β/δ upon Ligand Binding. *PLoS ONE* **11,** e0151412 (2016).
32. Fliss, A. E., Benzeno, S., Rao, J. & Caplan, A. J. Control of estrogen receptor ligand binding by Hsp90. *J. Steroid Biochem.* **72,** 223-230 (2000).
33. Tanenbaum, D. M., Wang, Y., Williams, S. P. & Sigler, P. B. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 5998–6003 (1998).
34. Schwabe, J. W., Chapman, L. & Rhodes, D. The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. *Structure* **3,** 201–213 (1995).
35. Shiau, A. K. *et al.* The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95,** 927–937 (1998).
36. Heldring, N. *et al.* Structural insights into corepressor recognition by antagonist-bound estrogen receptors. *J. Biol. Chem.* **282,** 10449–10455 (2007).
37. Thompson, A. J. & Lummis, S. C. R. 5-HT_3 Receptors. *Curr. Pharm. Design* **12,** 3615-3630 (2006).
38. Niesler, B. *et al.* Characterization of the novel human serotonin receptor subunits 5-HT_3C, 5-HT_3D, and 5-HT_3E. *Mol. Pharmacol.* **72,** 8–17 (2007).
39. Van Arnam, E. B. & Dougherty, D. A. Functional probes of drug-receptor interactions implicated by structural studies: Cys-loop receptors provide a fertile testing ground. *J. Med. Chem.* **57,** 6289–6300 (2014).
40. Albuquerque, E. X., Pereira, E. F. R., Alkondon, M. & Rogers, S. W. Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiol. Rev.* **89,** 73–120 (2009).
41. Tamrazi, A. *et al.* Estrogen Receptor Dimerization: Ligand Binding Regulates Dimer Affinity and DimerDissociation Rate. *Mol. Endocrinol.* **16,** 2706–2719 (2013).
42. McLachlan, M. J., Katzenellenbogen, J. A. & Zhao, H. A new fluorescence complementation biosensor for detection of estrogenic compounds. *Biotechnol. Bioeng.* **108,** 2794–2803 (2011).
43. Hassaine, G. *et al.* X-ray structure of the mouse serotonin 5-HT_3 receptor. *Nature* **512,** 276–281 (2014).
44. Leite, J. F., Amoscato, A. A. & Cascio, M. Coupled proteolytic and mass spectrometry studies indicate a novel topology for the glycine receptor. *J. Biol. Chem.* **275,** 13683–13689 (2000).
45. Liu, Z. *et al.* Crosslinking constraints and computational models as complementary tools in modeling the extracellular domain of the glycine receptor. *PLoS ONE* **9,** e102571 (2014).
46. Leite, J. F. *et al.* Conformation-dependent hydrophobic photolabeling of the nicotinic receptor: Electrophysiology-coordinated photochemistry and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 13054–13059 (2003).
47. Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42,** gku340–W258 (2014).
48. Post, M. R., Limapichat, W., Lester, H. A. & Dougherty, D. A. Heterologous expression and nonsense suppression provide insights into agonist behavior at α6β2 nicotinic acetylcholine receptors. *Neuropharmacology* **97,** 376–382 (2015).
49. Wang, W. *et al.* Quantitative analysis of T cell receptor complex interaction sites using genetically encoded photo-cross-linkers. *ACS Chem. Biol.* **9,** 2165–2172 (2014).
50. Lummis, S. C. R., Thompson, A. J., Bencherif, M. & Lester, H. A. Varenicline Is a Potent Agonist of the Human 5-Hydroxytryptamine3 Receptor. *J. Pharmacol. Exp. Ther.* **339,** 125–131 (2011).
51. Erde, J., Loo, R. R. O. & Loo, J. A. Enhanced FASP (eFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments. *J. Proteome Res.* **13,** 1885-1895 (2014).
52. Kalli, A., Smith, G. T., Sweredoski, M. J. & Hess, S. Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. *J. Proteome Res.* **12,** 3071–3086 (2013).