Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Cameron Hedden March 24, 2021

Determining the Role of Plasma Membrane Cholesterol in Modulating ClC Channel Function

By

Cameron T. Hedden

Nael A. McCarty, PhD
Adviser

Neuroscience and Behavioral Biology

Nael A. McCarty, PhD

Adviser

Ronald L. Calabrese, PhD

Committee Member

Michael D. Crutcher, PhD

Committee Member

Kirsten A. Cottrill

Committee Member

2021

Determining the Role of Plasma Membrane Cholesterol in Modulating ClC Channel Function

By

Cameron T. Hedden

Nael A. McCarty, PhD

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2021

Abstract

Determining the Role of Plasma Membrane Cholesterol in Modulating ClC Channel Function
By Cameron T. Hedden

Transmembrane ion channels facilitate the movement of charged species across plasma membranes in various cell types from all organisms, and channelopathies that impair or eliminate this ion conductance can result in lethal diseases like Cystic Fibrosis (CF). In recent years, new drugs have been synthesized that attempt to mitigate this defective ion conductance in CF, but there is mounting evidence that suggests that alterations of the plasma membrane impact the ability of these drugs to enhance patient outcomes. The plasma membrane in which ion channels exist and function is an incredibly complex mosaic of lipids, proteins and carbohydrates, yet how changes in this environment affect channel properties is not well understood. The present study sought to determine how depletion of plasma membrane cholesterol would impact the function of ion channels within the ClC family of voltage-gated transmembrane proteins, which are expressed widely throughout the body and are thought to be implicated in many neurological pathologies. Numerous cholesterol binding domains within the primary sequences of five ClC ion channels from two species were identified. Analysis of several primary amino acid sequences and computational modeling led to the identification of 10 potential cholesterol binding domains within human ClC-1 that were likely implicated in ClC protein-cholesterol interactions based on localization to the transmembrane domains of the protein. Of these 10 potential binding domains, sequence homology identified four domains that exhibited 100% conservation in all five of the ClCs studied. Two electrode voltage clamp (TEVC) studies of rabbit ClC-2 further supported an underlying dependence on membrane cholesterol for effective conductance of current across the plasma membrane. These results suggest an important role for plasma membrane cholesterol for effective function in ClC ion channels, and further suggest that depletion of plasma membrane cholesterol limits the ability of ClC ion channels to conduct current. With these results in mind, future work should focus on using site-directed mutagenesis to confirm the importance of the identified cholesterol binding domains and replicating TEVC studies with these mutant ClCs.

Determining the Role of Plasma Membrane Cholesterol in Modulating ClC Channel Function

By

Cameron T. Hedden

Nael A. McCarty, PhD

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2021

Acknowledgements

I am so grateful for my numerous mentors, peers and loved ones that have supported me and encouraged me over the years as I learn what it means to be a scientist. First, I am thankful for my parents for all of their sacrifice, love and support that has allowed me to get to where I am today. I am thankful for all of the faculty at Emory University for playing a role in my education and training, and providing me with the necessary tools to practice scientific inquiry. Drs. Sarah Fankhauser, LaTonia Taliaferro-Smith and Leah Roesch were especially instrumental in helping me achieve my academic goals, and for that I am deeply grateful. I would also like to thank Drs. Ron Calabrese and Mike Crutcher for taking the time out of their schedules to sit on my committee for this thesis, and for providing their input and expertise.

Since joining the McCarty laboratory two years ago, there have been numerous people that have shaped my experience who I would also like to thank. Kirsten Cottrill has been a constant source of support and guidance both within and outside of the scope of our laboratory work, and I am extremely grateful for everything that she has contributed to my education and my college career. Dr. Kerry Strickland also served as a key mentor during my time in the McCarty lab, and I am thankful for her mentorship and expertise in designing, conducting and interpreting experiments. I would also like to thank Dr. Guiying Cui, who has been very helpful in developing the methods for this thesis and has been so gracious and supportive during my time in the McCarty lab. Finally, I would like to thank my mentor Dr. Nael McCarty for seeing potential in me and allowing me to work and study in his lab over the last two years. So much of my time at Emory has been marked by my experiences with Dr. McCarty and the rest of his lab, and the invaluable experiences and skills that I have acquired under his mentorship have made me a better, more capable scientist.

Table of Contents

Abstract

Transmembrane ion channels facilitate the movement of charged species across plasma membranes in various cell types from all organisms, and channelopathies that impair or eliminate this ion conductance can result in lethal diseases like Cystic Fibrosis (CF). In recent years, new drugs have been synthesized that attempt to mitigate this defective ion conductance in CF, but there is mounting evidence that suggests that alterations of the plasma membrane impact the ability of these drugs to enhance patient outcomes. The plasma membrane in which ion channels exist and function is an incredibly complex mosaic of lipids, proteins and carbohydrates, yet how changes in this environment affect channel properties is not well understood. The present study sought to determine how depletion of plasma membrane cholesterol would impact the function of ion channels within the ClC family of voltage-gated transmembrane proteins, which are expressed widely throughout the body and are thought to be implicated in many neurological pathologies. Numerous cholesterol binding domains within the primary sequences of five ClC ion channels from two species were identified. Analysis of several primary amino acid sequences and computational modeling led to the identification of 10 potential cholesterol binding domains within human ClC-1 that were likely implicated in ClC protein-cholesterol interactions based on localization to the transmembrane domains of the protein. Of these 10 potential binding domains, sequence homology identified four domains that exhibited 100% conservation in all five of the ClCs studied. Two electrode voltage clamp (TEVC) studies of rabbit ClC-2 further supported an underlying dependence on membrane cholesterol for effective conductance of current across the plasma membrane. These results suggest an important role for plasma membrane cholesterol for effective function in ClC ion channels, and further suggest that depletion of plasma membrane cholesterol limits the ability of ClC ion channels to conduct current. With these results in mind, future work should focus on using site-directed mutagenesis to confirm the importance of the identified cholesterol binding domains and replicating TEVC studies with these mutant ClCs.

Introduction

Cystic Fibrosis and What it Taught Us

Cystic Fibrosis (CF) is an autosomal recessive disease that affects approximately 30,000 individuals in the United States and over 70,000 worldwide[1]. It was not until 1989 that the genetic basis of CF was localized to the long arm of human chromosome 7 and the associated gene and protein were named the Cystic Fibrosis Transmembrane conductance Regulator (CFTR)[2–4]. CFTR is a polytopic transmembrane protein composed of 1,480 amino acids that form two transmembrane domains, two intracellular nucleotide binding domains and an intracellular regulatory domain. Furthermore, CFTR is a unique member of the ATP-binding cassette (ABC) superfamily of protein transporters, as it is the only one known to also function as an ion channel[5]. Gating of CFTR consumes ATP through activity of its two nucleotide binding domains, and also requires phosphorylation of its regulatory domain via cAMP-dependent protein kinase (PKA)[6]. The flow of ions (predominantly chloride and bicarbonate) through CFTR during its gating cycle facilitates the movement of water across a variety of epithelia throughout the body.

CF is truly a multisystem disease, but the primary cause of premature death in people with CF is lung failure due to chronic inflammation, bacterial infection and mucus plugging of the lower airway. On the surface of lung epithelial cells, cilia project out from the membrane and are free to move in an aqueous layer called airway surface liquid (ASL). On top of the ASL there is also a thin layer of mucus that is responsible for trapping debris and pathogens that enter the lungs during respiration. The cilia sweep back and forth about 1,000 times per minute, all in a direction that pushes mucus and trapped particulate matter towards the pharynx where it can then be swallowed or expectorated[7].

In CF, failure of CFTR to facilitate ion movement across the lung epithelium results in the

inability to move water from the basolateral to apical side of the lung epithelium. Critically, without this movement of water the ASL becomes greatly reduced in height, and the thin, loose layer of mucus that normally coats the airway becomes thick, dehydrated, and viscous. Cilia that once projected out into an aqueous medium are consequently compressed and cemented in mucus, unable to effectively clear the lower airway. Without proper mucociliary clearance, patients with CF become increasingly susceptible to bacterial infections that are more frequent and more difficult to treat compared to patients without CF. The same issue of ion transport in pancreatic epithelia is known to cause damage to the exocrine pancreas where acidic digestive enzymes are not properly neutralized by bicarbonate moving through CFTR[8], and also causes increased levels of chloride in sweat that has been used for years as a diagnostic metric in screening for CF.

There are over 2,000 mutations that have been identified in CFTR, of which hundreds have been linked directly to disease[9]. Depending on the patient's unique allelic combination, functional CFTR could be completely absent, improperly trafficked to the plasma membrane, unstable during the gating cycle or present in insufficient quantities to achieve ideal ion conductance, all of which can lead to CF[10]. Prior to 2012, treatment of patients with CF focused on managing symptoms rather than addressing the underlying cause of disease. For most patients with CF, a complex treatment regimen to manage their symptoms begins soon after birth. Common treatments include inhaled medications such as hypertonic saline and/or dornase alfa (Pulmozyme®) that work to loosen mucus within the lungs, in addition to mechanical vests that physically shake the patient's chest in an effort to break up the mucus in the lower airway. Because recurrent bacterial infection is so frequent and of such great concern in people with CF, oral, inhaled and/or intravenous antibiotics are also commonly prescribed soon after initial signs of infection. When CF was first characterized, it was not typical for a child diagnosed with CF to survive past the first decade of

life. As management strategies have improved for patients with CF over the years, the life expectancy for a child born with CF today continues to rise[11,12], but it is important to note that the treatment burden of patients with CF has traditionally been intense, requiring strict adherence and numerous hours a day of care. In recent years, several small molecules known as Highly Effective Modulator Therapies (HEMTs) have been synthesized and prescribed in an effort to address - for the first time - the underlying defect in CF that causes disease.

The Advent of Highly Effective Modulator Therapy in CF

In 2012, the FDA approved Kalydeco® (Ivacaftor) as the first HEMT of its kind to treat patients with two copies of the G551D CFTR mutation, which causes CFTR in the plasma membrane to be unstable during its gating cycle. Kalydeco® is a CFTR potentiator, meaning it acts directly on the CFTR protein to stabilize it and keep it open longer during its gating cycle. For the first time, patients with CF had a therapeutic avenue that directly targeted the root cause of their disease. After just two weeks of treatment with Kalydeco®, patients showed significant clinical improvement indicated by increased forced expiratory volume in the first second (FEV_1) and reduced sweat chloride levels[13]. In 2014, the FDA expanded the label for Kalydeco®, approving its use in patients with at least one copy of the G551D mutation, and one copy of any of nine other specified CFTR mutations, making Kalydeco® a possible clinical avenue for about 8% of patients with CF in the United States.

In 2015, the FDA approved the first combination therapy, Orkambi® (Lumacaftor/Ivacaftor), which introduced Lumacaftor -- a CFTR corrector that allows more functional CFTR to be trafficked correctly to the plasma membrane where it could then be potentiated by Ivacaftor. Orkambi® is approved for patients with two copies of the most common CFTR mutation, ΔF508. After the introduction of Orkambi®, about 33% of patients in the United

States had a HEMT available to them based on their unique allelic combination. In 2018, the FDA approved another HEMT, Symdeko® (Tezacaftor/Ivacaftor and Ivacaftor), which expanded the label even further to include patients with one or more copies of 154 other specified mutations in CFTR with the introduction of Tezacaftor, another CFTR corrector similar to Lumacaftor. Most recently in 2019, the FDA approved Trikafta® (Elexacaftor/Tezacaftor/Ivacaftor and Ivacaftor) – which introduced another new corrector, Elexacaftor – for patients with at least one copy of the ΔF508 mutation and one copy of any of 177 other specified mutations. Between all of these HEMTs, about 90% of patients in the United States today have a drug that can be prescribed to them based on their unique combination of mutations in CFTR. A large focus of research today is focused on identifying therapeutics to treat the remaining 10% of patients that do not currently have a HEMT available based on their genotype.

Importantly and perhaps often overlooked, however, is that these drugs do not always provide the life-changing relief that many hope for and expect. For example, in some patients, these drugs are not well tolerated and require close monitoring of liver enzymes while on therapy. Furthermore, if a patient's liver enzymes deviate too far out of range, they may be precluded from staying on therapy. Additionally, none of these HEMTs currently have a generic equivalent available in the United States, and just one year of any of these HEMT drugs can cost a patient over $300,000 without insurance coverage. For this reason, there is an issue of access if the patient has low or no insurance coverage. Nevertheless, the large advances in our understanding and treatment of individuals with CF witnessed since the gene was cloned in 1989 is impressive, and serves as an exemplar of what precision medicine is and has the power to do[14].

Despite the general success that has been observed in patients prescribed HEMTs, little is known about the mechanisms of action that allow these drugs to mitigate the defect responsible

for causing CF. As a result, it is extremely difficult to refine these drugs to create better therapeutics. Furthermore, our understanding of potential endogenous molecules like cholesterol and how they might serve as natural ligands of CFTR or other transmembrane proteins is limited. Critically, how CFTR is impacted by cholesterol or other endogenous molecules may have severe consequences for how well HEMTs are able to rescue CFTR function. Experiments conducted in our lab have found that removal of cholesterol from the plasma membrane with methyl-β-cyclodextrin (MβCD) greatly increases the effect of Kalydeco® in potentiating non-mutant CFTR. While this may seem to suggest that MβCD treatment enhances Kalydeco® efficacy, what it actually suggests is that channels lacking the cholesterol ligand are in a less functional state before potentiation, and thus we observe a greater fold increase in CFTR mediated current upon treatment with Kalydeco®[15]. Regardless, Kalydeco® was unable to restore the full amount of transmembrane current after cells had been treated with MβCD. In order for therapeutics like Kalydeco® to continue to improve patient outcomes, it is clear that there must be a more robust understanding of how plasma membrane cholesterol interacts with transmembrane proteins.

The Plasma Membrane: Location is Everything

In order to understand issues of gating in CFTR or any other ion channel, it becomes important to understand the complex plasma membrane environment in which the protein exists. The animal cell plasma membrane is an impressive system that contains and separates one cell from another, allowing the smallest units of life to exist and carry out their functions while also separating the inside of cells from the extracellular environment. Aliphatic phospholipids self-assemble into a bilayer system due to hydrophobic interactions, with polar phosphate heads facing the aqueous intracellular and extracellular environments and hydrophobic fatty acids concealing themselves from water by pointing inwards towards the center of the two membrane leaflets. Major

contributing components of animal cell membranes include lipids such as phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, and sphingomyelin. Cholesterol is also a major component of the plasma membrane, and is thought to be present in the same molar amounts as phospholipids[16]. Cholesterol has many functions in the plasma membrane: it immobilizes the outer surface of the membrane and confers rigidity; it decreases membrane permeability to small water-soluble molecules; it separates phospholipid tails and prevents membrane crystallization; and critically, it secures membrane proteins and anchors regions of the membrane together, forming detergent insoluble microdomains also known as "lipid rafts".

The membrane lipid environment is also known to have important implications on the proteins that are embedded in the lipid bilayer. Furthermore, there are several ways in which the lipids in the plasma membrane can alter the properties of the protein. For example, lipids can bind directly to the protein and have a direct physical interaction with the proteins in the plasma membrane. Membrane lipids can also affect protein trafficking to the plasma membrane or cause proteins to localize to certain parts of the membrane. Additionally, lipids can initiate signaling cascades that cause post-translational modifications to the proteins in the plasma membrane. Further still, the lipid composition of the plasma membrane affects whole bilayer mechanics such as fluidity and curvature, which may have implications on how the protein is able to change its conformation as it carries out its biological function[17].

A New Protein with the Same Affliction

In addition to CFTR, there are numerous other transmembrane protein channels that facilitate the movement of ions across the plasma membrane. The ClC family of proteins includes several voltage-gated chloride channels that are expressed both in the plasma membrane as well as intracellular membranes of cells from most organisms. Proteins of the ClC family (hereafter

referred to as ClCs) are mostly impermeable to cations, and furthermore are quite selective for chloride ions. In fact, early studies of some ClCs suggested that they might serve as potential therapeutic targets to increase chloride conductance in the absence of functional CFTR[18]. There are nine known ClCs expressed in humans, of which four function as chloride channels and five function as chloride/proton exchangers. Regardless of this distinction, though, all nine of the ClCs expressed in humans have the same basic structure[19]. All ClCs are expressed in native cells as obligate homodimers.

ClCs are expressed in a variety of tissues and are thought to be implicated in a number of human diseases. For example, ClC-1 is highly expressed in skeletal muscle, and loss-of-function mutations in ClC-1 are known to cause myotonia, or the inability for skeletal muscle to relax after voluntary contraction[20,21]. ClC-2 has long been the topic of debate when it comes to its physiological role; however, it is widely expressed throughout the body. Furthermore, ClC-2 is highly expressed in both pyramidal hippocampal neurons[22] and astrocytes[23] of the brain, where it is thought to aid in postsynaptic inhibition and regulate cell volume[24]. Mutations in ClC-2 have also been observed in patients with certain forms of idiopathic generalized epilepsy. Mutations in ClC-Ka and/or ClC-Kb are known to cause Bartter Syndrome, a group of genetic disorders that affect salt reabsorption in the kidneys[25]. Mutations in another gene encoding barttin (a critical beta-subunit for ClC-Ka and ClC-Kb) leads to loss of function in both ClC-Ka and ClC-Kb, causing Bartter syndrome type IV with resulting K^+ accumulation in the inner ear and sensorineural deafness[26]. Like CFTR, these ClC transmembrane proteins are known to undergo significant conformational changes during their gating[27–29].

Despite only comprising approximately 2% of a human's body mass, the central nervous system contains about 25% of total unesterified cholesterol in the body[30]. As previously stated,

cholesterol is a critical part of the plasma membrane[16], and neurons are no exception to this phenomenon. Furthermore, alterations of membrane cholesterol have been observed in several diseases[31]. For example, Niemann-Pick type-C (NPC) is a neurovisceral disease that is characterized by defective trafficking of cholesterol within cells that leads to pathophysiologies of the bone marrow, liver, spleen, lungs and brain. Newborns with NPC may present with decreased muscle tone, delay in developmental milestones, and developmental regression compared to age matched controls. In childhood onset NPC, similar pathophysiologies are observed with the addition of slurred speech and seizures. In teenagers and adults with later onset NPC, psychiatric and other cognitive symptoms are more common[32]. Another example of a disease with imbalanced cholesterol is Smith-Lemli-Opitz Syndrome (SLOS), a developmental disorder that can cause microcephaly, cognitive impairment, moderate to severe physical disability and distinct facial features[33]. SLOS is caused by impairment of the cholesterol synthesis pathway, and many of the characteristic features of SLOS have been suggested to be due to impaired function of the sonic hedgehog (SHH) signaling pathway[34,35]. Similarly, several models of another neurodegenerative condition, Huntington's Disease, have shown that cholesterol synthesis and distribution are impaired and may serve as a potential therapeutic target and/or diagnostic metric[36].

Despite the implications of cholesterol in each of these neurological conditions, little is known about how altered cholesterol synthesis and trafficking impact neurological pathophysiologies. It is clear that the complex lipid environment in which a transmembrane protein exists has implications on how the protein is able to function. Furthermore, since cholesterol is so abundant in the plasma membrane, we sought to determine if alterations in plasma membrane cholesterol modulated ClC ion channel function through computational modeling, molecular docking simulations and electrophysiological methods.

Methods

Searching the Primary Sequences for Potential Cholesterol Binding Sites

The primary amino acid sequences of four human ClC ion channels (ClC-1, ClC-2, ClC-Kb and ClC-Ka) as well as rabbit ClC-2 were retrieved from UniProt[37]. The sequences were opened in Microsoft Word and converted to a continuous string of residues, removing any spaces or line breaks within each sequence. A function was built using RStudio which included a series of "for loops" and "if/then statements" to scan each primary amino acid sequence for CRAC and CARC domains (**Figure 1**). CRAC and CARC domains are thought to represent potential cholesterol binding sites, as they have been found to selectively bind cholesterol in myriad proteins of various organisms[38].

The CRAC domain is defined by the following relatively simple arrangement of amino acids beginning at the N-terminus and proceeding towards the C-terminus: a leucine or valine residue, followed by a sequence of 1-5 amino acids, followed by a tyrosine residue, followed by a sequence of 1-5 amino acids, and concluding with either lysine or arginine. The 1-letter amino acid algorithm that constitutes a CRAC domain is (L/V)-X_{1-5}-(Y)-X_{1-5}-(K/R), where X represents any amino acid.

In addition to the CRAC domain, the same sequence in the reverse direction (i.e., beginning at the C-terminus and proceeding towards the N-terminus) is known as a "CARC" domain, or an inverted "CRAC" domain. The CARC domain follows the same algorithm as the CRAC domain with the exception of the central residue which can be either tyrosine or phenylalanine. The 1-letter amino acid algorithm for a CARC domain is thus given as (L/V)-X_{1-5}-(Y/F)-X_{1-5}-(K/R).

The results of this R Script yielded the positions of three residues that constituted a CRAC or CARC domain. These three amino acid outputs were compared to one another, and any amino acids that were present in multiple potential CRAC or CARC binding domains were reported as

being one potential cholesterol binding site. As such, any of the identified domains may contain more than three amino acids. It should be understood that some combination of three amino acids within the domain that satisfy the definition of a CRAC or CARC cholesterol binding site would constitute the true CRAC or CARC binding site.

Mapping CRAC/CARC Domains onto Human ClC-1

With CRAC and CARC domains identified in the primary sequence, we proceeded to map these domains onto a solved structure of human ClC-1[39]. The .PDB file of the protein was obtained from the Protein Data Bank and opened in PyMOL[40]. The primary sequence of the protein was displayed in PyMOL, and each amino acid that was identified as belonging to a CRAC domain was highlighted and renamed as a new selection. CRAC domains were color coordinated to match the same domain on both monomers of the ClC-1 homodimer. This process was repeated to map CARC domains onto ClC-1, and the color-coded images were saved as separate PyMOL session files that could be exported as high-resolution images.

Multiple Sequence Alignment

In order to determine the conservation of these potential CRAC and CARC domains, and thus the potential biological significance, a sequence alignment of the five primary sequences discussed above was generated using Clustal Omega[41]. The primary sequences were uploaded to Clustal Omega along with their respective PathwayCommons identifier, origin species and associated gene name. The results of the sequence alignment were received via email shortly after the simulation was initiated. The output file was opened in Microsoft Word where conservation of specific residues was indicated by (*) to indicate 100% conservation across all sequences, (:) to indicate a high degree of conservation and (.) to indicate moderate conservation. Each CRAC and CARC domain was analyzed to determine the degree to which they were conserved, if at all.

Cholesterol Docking Simulations

The structure of cholesterol was created in ChemDraw, opened in PyMOL, and saved as a MOL2 file which was then uploaded to AutoDock Tools to add polar hydrogens and create a .PDBQT file. The Cryo-EM structure of human ClC-1 (**Figure 2)** was obtained from the Protein Data Bank as a .PDB file, which was edited in notepad to create a format that would be compatible with AutoDock Vina[42]. Headers and footers were removed from the plain text, and only information regarding the amino acid residues was retained. ClC-1 forms a homodimer consisting of two identical protein chains. To account for this, the edited .PDB file was split into two separate files, each containing one monomer. One of the new .PDB files for one monomer of ClC-1 was opened with AutoDock Tools, and polar hydrogens were added. The protein was saved as a .PDBQT file.

A grid box was generated with dimensions of 20 x 20 x 20 Å centered on the α-carbon of the central amino acid residue forming the potential CRAC or CARC domain of interest. This grid box is used by the docking simulation to determine the three-dimensional space within which the program searches for potential cholesterol binding states. The coordinates of the α-carbon on which the grid was centered were obtained by opening the .PDB file of the monomer in notepad and locating the corresponding amino acid. A configuration file was generated in notepad to interface with the AutoDock Vina executable software. This .TXT file included the file name of the protein, the file name of the cholesterol ligand, the coordinates and size (Å) of the grid box and an output file name (**Figure 3**).

The output files generated from AutoDock Vina included a .PDB file of the cholesterol in various possible binding states as well as a .TXT log file of the calculated binding energies for each possible state. Possible cholesterol binding states for each potential CRAC or CARC domain

were visualized and inspected by eye for viability based on several key factors: the position of the -OH group of the cholesterol relative to the lysine or arginine residue of the CRAC or CARC domain that could participate in hydrogen bonding; the position of the aromatic portion of cholesterol relative to the aromatic residue of the binding domain was considered for possible pi stacking; the position of the carbon chain of cholesterol relative to the hydrophobic leucine or valine residues were considered for possible Van der Waals interactions.

Constructs and Cell Maintenance

Xenopus laevis oocytes were acquired from EcoCyte Bioscience (Austin, TX). After the initial harvest, oocytes were stored and shipped on ice in Modified Barth's Saline (MBS) containing 89 mM NaCl, 1.0 mM KCl, 2.4 mM $NaHCO_3$, 10 mM HEPES and 0.82 mM $MgSO_4$ in deionized water (pH=7.4). Oocytes were expedited to arrive at our laboratory approximately 24 hours after harvest in order to ensure optimal cell health. Upon receipt of the oocytes, they were promptly transferred to half-strength Leibovitz's L-15 cell culture media containing L-glutamine (Invitrogen, Carlsbad, CA), 10 mM HEPES (pH=7.5), 50 I.U./mL penicillin, and 50 µg/mL streptomycin and were stored at 16°C. L-15 media was replaced daily.

Oocytes were injected with 5-10 ng of transcribed complementary RNA (cRNA) generated from the DNA construct of interest (mMessage mMachine Kit, Ambion, Inc., Austin, TX), in water that was treated with diethylpyrocarbonate (DEPC) to inactivate RNAses. Following cRNA injection, oocytes were incubated at 16°C and left to translate the cRNA into protein for 1-3 days. During this incubation, L-15 media was replaced daily.

Two Electrode Voltage Clamp

Two Electrode Voltage Clamp (TEVC) was used to monitor whole cell changes in current across the plasma membrane. TEVC electrode tips were made from borosilicate glass from Sutter

Instruments (Nocato, CA; catalog number BF150-86-10). Electrode tips were pulled using a Sutter Instruments Model P-2000 laser puller, and filled with 3M KCl. Electrodes containing 3M KCl were inserted into an electrode holder, which was screwed tight to form a seal. The electrodes were connected to a GeneClamp 500B amplifier and Clampex 10.2 software was used to initiate and record experiments. Current data were digitized at 2 kHz. The electrode tips were carefully positioned using micromanipulators into a plastic recording chamber filled with ND96 solution containing 96 mM NaCl, 2 mM KCl, 2 mM $MgCl_2$, 5 mM HEPES and 1 mM $BaCl_2$ in deionized water (pH=7.50). While in solution, the voltage readings of the electrodes were zeroed and the resistances in MΩ was recorded. Electrodes with resistance values between 0.3 and 2.0 MΩ were sufficient to be used with oocytes.

The electrodes were removed from the recording chamber and oocytes were carefully placed into the recording chamber one at a time and submerged in ND96 with the animal pole facing upwards. Fresh ND96 remained flowing and was suctioned by vacuum to remove waste and any debris that may accumulate in solution. The electrodes were carefully positioned back into solution and driven into the oocyte to puncture the plasma membrane. Oocytes were impaled bilaterally just below the cell surface, with each electrode piercing the membrane in the animal pole about 45º above the horizontal on the left and the right respectively. Upon impaling the cell, the resting membrane potential as well as holding current at a clamped voltage of -30 mV were recorded.

After entering voltage clamp mode, the system becomes charged and cannot be touched unless returned to setup mode. The oocyte was held at a potential of V_M = -30 mV, and the transmembrane current (holding current) at -30 mV was recorded. The voltage protocol that was used held the oocyte at a potential of -30 mV for 125 milliseconds and then stepped to a test

potential of -180 mV for 750 milliseconds. The potential then returned to -30 mV for 30 seconds to allow the transmembrane current to equilibrate. Each subsequent sweep followed the same general protocol, but the test potential increased by 20 mV in each sweep until a final test potential of 60 mV was recorded. These trials were conducted on four groups of oocytes: a) un-injected control oocytes, b) un-injected oocytes treated with 20 mM MβCD for 30 minutes at room temperature, c) ClC-2-injected control oocytes, and d) ClC-2-injected oocytes treated with 20 mM MβCD for 30 minutes at room temperature. The current-voltage relationships of these conditions were analyzed to determine if extraction of membrane cholesterol with MβCD affected ClC-2 channel function.

Current/Voltage Analysis

After collecting current/voltage data using TEVC, Clampfit 10.2 software was used to open the output files and observe trends in the current/voltage relationship. Although the experiments recorded a most hyperpolarized voltage of -180 mV, only currents between -160 mV and 60 mV were considered, as the currents at -180 mV occasionally were so large that the TEVC instrumentation could not sufficiently clamp the oocyte membrane to maintain the correct voltage. For the purposes of calculating the current at each voltage, the current was considered to be the final current at each voltage between -160 mV and 60 mV just before returning to a holding potential of -30 mV. Data was exported from Clampfit into Excel, and current-voltage plots were generated. Prism was also used to generate current-voltage plots, with data represented as means and standard deviations. Prism was also used to conduct statistical analyses. For the entire current-voltage plot, treatments were compared by multiple two-way ANOVAs, with repeated measures over the voltage range. For comparisons of the final current at -160 mV, multiple uncorrected, unpaired two-tailed *t* tests were performed.

Results

Previous literature has determined that cholesterol content of the lipid bilayer can affect ion channel activity. For example, our lab has determined that depleting cholesterol from the lipid bilayer can drastically affect the activity of CFTR[15]. While there are many mechanisms by which cholesterol modulation can affect ion channel activity, one such mechanism is by direct interactions with the protein at a specific binding domain. Therefore, we sought to understand the effects of changes in membrane cholesterol on ClC activity.

Identifying CRAC and CARC Domains in ClC Ion Channels

There are numerous potential cholesterol binding domains that can be identified from primary amino acid sequences that have been the focus of research on cholesterol-protein interactions, but one of the most popular is the Cholesterol Recognition/interaction Amino acid Consensus sequence, more commonly referred to as the CRAC domain[43]. CRAC domains have been found in diverse proteins that are known to bind cholesterol, and furthermore single mutations within the CRAC domain have been found to severely decrease or completely eliminate cholesterol-protein binding[38]. As described above, CARC domains are also commonly studied cholesterol binding domains in diverse proteins that have an inverted sequence algorithm compared to the CRAC domain, with the added possibility of having phenylalanine as the central residue of the domain.

Using the function generated in RStudio described in Methods (**Figure 1**), segments that code for CRAC and CARC domains were identified from the primary sequences of human ClC-1, ClC-2, ClC-Ka and ClC-Kb as well as rabbit ClC-2 (**Tables 1-10**). These sequences each yielded between six and nine unique CRAC domains and between 11 and 14 unique CARC domains in each primary sequence. However, while these represent potential cholesterol binding sites, more

information such as localization within the whole protein's structure is necessary to more confidently determine if these could truly be cholesterol binding sites.

Mapping CRAC/CARC Domains onto Human ClC-1

To determine tertiary localization, the CRAC and CARC domains for ClC-1 were mapped onto a solved Cryo-EM structure of the human ortholog[39]. **Figure 4** shows CRAC domains identified in the primary sequence, with each domain being color coded to match the same region on both monomers of human ClC-1. **Figure 5** shows CARC domains identified using the R function, again with the corresponding regions on each monomer shown in the same color. After reviewing the output of the R function and determining any overlap in residues that could contribute to a cholesterol binding domain, we determined there to be nine unique CRAC and thirteen unique CARC domains on human ClC-1. All future references to CRAC and CARC domains refer to domains in human ClC-1.

Evaluation of the tertiary localization of these CRAC and CARC domains revealed some that were localized to unresolved sections of the Cryo-EM structure. This suggests that this region of the protein is unstructured, and therefore is unlikely to bind cholesterol. Furthermore, some of the CRAC and CARC domains were localized to the intracellular domain rather than the transmembrane domain where cholesterol would be localized. Therefore, we were able to exclude these sites as potential cholesterol binding sites. Ultimately, these exclusions left us with five CRAC and five CARC domains of particular interest.

Multiple Sequence Alignment

To better understand the potential for these remaining five CRAC and five CARC domains to be true cholesterol binding sites, we determined their conservation by performing a sequence alignment on the primary sequences of human ClC-1, ClC-2, ClC-Ka and ClC-Kb and rabbit ClC-

2. A multiple sequence alignment was prepared using Clustal Omega[41]. The multiple sequence alignment showed high levels of conservation of the entire sequences. Interestingly, multiple of the cholesterol binding domains we identified to be of particular interest exhibited 100% conservation between the five sequences. On human ClC-1, these conserved sequences are denoted as CRAC_3, CRAC_5, CARC_1 and CARC_3. Cholesterol binding domains from ClC-1 that exhibit lesser or no degree of conservation may not actually bind cholesterol, or they may be unique to ClC-1. **Figure 6** shows the sequence alignment of the segments of protein that encoded CRAC and CARC domains of interest, along with their relative amino acid conservations.

Docking Cholesterol on Human ClC-1

After locating CRAC and CARC domains in the primary sequence and narrowing our focus to domains of particular interest by analyzing the tertiary structure and sequence alignment, docking simulations were performed on the solved Cryo-EM structure of ClC-1[39]. **Figure 7** shows singular binding states for five CRAC domains of interest. Cholesterol docking simulations using AutoDock Vina yielded multiple possible binding states for each CRAC domain, and a representative state was chosen based on the feasibility of true cholesterol binding. As described in Methods, feasibility was based on position of cholesterol's -OH group relative to lysine or arginine residues, the position of the aromatic portion of cholesterol relative to the aromatic residue of the binding domain, and the position of the carbon chain of cholesterol relative to the hydrophobic leucine or valine residues of the binding domain. **Figure 8** similarly shows single binding states for five CARC domains of interest.

These docking studies are preliminary, but provide at least an initial positioning for cholesterol in these potential binding sites. Interestingly, one CARC domain could be ruled out as a potential binding domain, as cholesterol could not access the amino acids of interest due to them

being buried in a series of α-helices (**Figure 8, Panel E**). Future studies using molecular dynamics simulations with the ClC-1 protein and the pre-docked cholesterol would be necessary to give the protein sufficient energy and flexibility to bind cholesterol more favorably. Following this, more docking studies could be performed to calculate binding energies.

Effects of MβCD on Rabbit ClC-2

We have identified many potential cholesterol binding sites on human ClC-1 that remain conserved in other ClCs. Therefore, it is very likely that cholesterol does indeed bind ClC channels. Thus, we were interested in determining the effects of cholesterol depletion on ClC-2 channel activity in particular, especially given the relevance of this channel to neuronal function. The *Xenopus laevis* oocyte is a common cell model used to study the properties of protein channels and transporters expressed in animal cell membranes. These cells have minimal expression of endogenous transmembrane proteins, and therefore present a useful model to study expression of exogenously inserted RNAs. Previous experiments from our lab have indicated that MβCD efficiently extracts plasma membrane cholesterol from *Xenopus* oocytes. Filipin III is a fluorescent molecule used to probe for cholesterol in the plasma membrane, as it binds to free (unesterified) cholesterol in the membrane[44]. After treating oocytes for 30 minutes in 20 mM MβCD at room temperature, filipin III intensity was reduced by half (*data not shown*), suggesting efficacious removal of plasma membrane cholesterol. We thus maintained these treatment conditions for our experiments.

TEVC is a common technique used to study ion conductance across cellular membranes *in vitro*. The TEVC system creates a circuit (**Figure 9**)[45] that allows the experimenter to monitor the amount of current injected into the cell that is needed to clamp the membrane potential to a given voltage. The amount of current that must be injected to keep the whole cell at a constant voltage

is the amount of current that is moving across the cell membrane; therefore, if the proteins of the cell membrane that conduct ions are limited or controlled, the experimenter can determine the activity of the channels of interest in a living cell. During TEVC experiments, it became evident that some oocytes were unhealthy or dying. In order to exclude cells that might confound our data, we established a healthy cell as having a resting membrane potential of between -10 and -30 mV. In some cases, cells clamped at a holding potential of -30 mV exhibited large amounts of unstable current which also indicated poor cell health, and in cases when a stable holding current at -30 mV could not be established, the cell was discarded.

Figure 10 shows the current-voltage relationship for the four treatment groups. Beginning with the most hyperpolarized potential of -160 mV, the amount of negative current elicited at each potential decreased until the cell's reversal potential was reached. At more positive, depolarizing potentials there was some increase in the amount of positive current elicited. Multiple two-way ANOVA analyses were used to compare the four treatment groups. There was no significant difference in net transmembrane current between un-injected control oocytes (i.e., not treated with MβCD) and un-injected oocytes treated with 20 mM MβCD for 30 minutes at room temperature ($p=0.9893$). This suggests that MβCD treatment did not significantly affect the amount of net transmembrane current in oocytes through a mechanism other than ClC-2. There was a significant difference between un-injected control oocytes and ClC-2-injected control oocytes (**$p=0.0076$). This suggests that oocytes injected with ClC-2 cRNA were appreciably expressing ClC-2 and conducting current across the cell membrane. Finally, there was a significant difference between ClC-2-injected control oocytes and ClC-2-injected oocytes treated with 20 mM MβCD (*$p=0.0120$), suggesting that MβCD treatment inhibits ClC-2 from conducting current.

Figure 11 shows the elicited current at the most hyperpolarized potential, -160 mV. This

very negative potential was evaluated due to the large magnitudes of current via ClC-2. Multiple unpaired, two-tailed *t* tests were used to compare between the four treatment groups. There was no significant difference between un-injected control oocytes and un-injected oocytes treated with 20 mM MβCD ($p=0.4352$). This further supports the notion that MβCD did not significantly affect the amount of net transmembrane current through a mechanism other than ClC-2. There was a significant difference between un-injected control oocytes and ClC-2-injected oocytes that were not treated with MβCD ($*p=0.0123$). This further supports the notion that oocytes injected with ClC-2 cRNA were appreciably expressing ClC-2 and conducting current across the cell membrane. Finally, there was a significant difference between ClC-2-injected control oocytes and ClC-2-injected oocytes treated with 20 mM MβCD ($**p=0.0099$), further supporting the notion that the removal of plasma membrane cholesterol inhibits ClC-2 from conducting current.

Discussion

The goal of this work was to determine if alterations in plasma membrane cholesterol would modulate ClC function by utilizing computational modeling, molecular docking simulations and electrophysiological methods. Numerous CRAC and CARC domains were identified in the primary sequences of five ClCs from *Homo sapiens* and *Oryctolagus cuniculus*. Using RStudio to build a function and scan the primary sequences of these proteins provided a quick method of searching proteins of any size for potential cholesterol binding motifs, and indeed similar methods could be used to scan primary sequences for other recognizable combinations of amino acid residues.

It is important to note that just because a sequence of amino acids constitutes a CRAC or CARC domain does not necessarily mean that the domain is likely to bind membrane cholesterol. Because there is a solved Cryo-EM structure of human ClC-1, it was possible to narrow our focus based on the likelihood that each domain would actually be implicated in any membrane cholesterol-protein interactions. Without solved structures of the other ClCs studied here, it is not possible to conduct docking experiments, but sequence homology could be a useful tool in determining potential cholesterol binding sites if they are highly conserved.

Numerous CRAC and CARC domains were identified in the primary sequence of human ClC-1 and mapped onto its Cryo-EM structure. Many of these domains could be ruled out as binding membrane cholesterol based on localization to the cytoplasmic domains and steric hindrance given the positioning of their α-helices. Using ClC-1 as a model, initial docking studies suggested that some of the CRAC and CARC domains could be true cholesterol binding domains. However, the docking simulations presented here should only be considered preliminary, and these potential interactions should be further analyzed using molecular dynamics simulations and

subsequent docking experiments. In the present docking simulations, the protein is presented in one conformation and not allowed to move around the cholesterol ligand. In reality, the binding of membrane cholesterol to ClC-1 or other transmembrane proteins would likely induce conformational changes[17] that would not be observed in these preliminary docking simulations. A more powerful molecular dynamics simulation would allow the protein and ligand to relax around one another to form a more viable binding. Future work to build homology models of the other ClCs based on the structure of ClC-1 will aid in performing similar screening methods for other ClCs of interest.

Results from TEVC studies further support an underlying dependence on membrane cholesterol for the effective gating of ClCs. This experimental data suggests that cholesterol is critical for rabbit ClC-2 activity, as the extraction of plasma membrane cholesterol with MβCD caused a reduction in the amount of whole cell net transmembrane current. In conducting these TEVC studies, there was some variability between and within batches of oocytes, which resulted in variability in the amount of current elicited at various voltages. This variability is especially visible in the ClC-2-injected, control group, due to the larger current densities. A possible reason for this variability is the difference of protein expression levels in oocytes based on how healthy the cell was at the time of recording. Furthermore, this variability is likely less evident in the ClC-2-injected, MβCD group because removal of plasma membrane cholesterol effectively inhibited any ClC-2 proteins in the plasma membrane, resulting in very little transmembrane current regardless of how much protein was expressed.

The dependence on plasma membrane cholesterol for effective ClC function has not been extensively evaluated, however there is one study in human embryonic kidney 293 (HEK293) cells that suggested that treatment with MβCD increased the open probability for ClC-2 at less

hyperpolarized potentials[46]. These results appear to be in contrast with our data in that they suggest an increased amount of transmembrane current after treatment with MβCD. A key difference of these studies is the model cell line used to evaluate ClC function, which necessitated different cell culture techniques. We also have noted on occasion that treatment of oocytes with MβCD yielded an increase in transmembrane current, but this was always accompanied by indications that the cell was unhealthy and were thus not considered. Useful metrics in gauging oocyte health include the resting membrane potential, and the holding current when the membrane voltage is clamped at -30 mV. Typically, we would expect our oocytes to have a resting potential between -10 and -30 mV to be considered healthy, with a stable holding current that typically fell between 0 and -100 nA at a clamped voltage of -30 mV.

Additionally, the authors noted that they were frequently unable to achieve a tight seal using their patch clamp method, which might further suggest poor cell health. In optimizing our experimental conditions, we found that treatment of oocytes with 20 mM MβCD for longer than one hour often resulted in massive amounts of transmembrane current at a holding potential of -30 mV, indicating an unhealthy/dying cell. Upon visual inspection of these cells, the plasma membrane virtually disintegrated in the recording chamber and the cells were much more likely to rupture upon impalement with the glass electrode tips used in TEVC. Therefore, it became evident to us that the concentration and duration of MβCD treatment had substantial impacts on the success of our recordings. In the Hinzpeter et al. study, HEK293 cells were treated with 10 mM MβCD for 30 minutes at 37°C. It is possible that this concentration of MβCD was too high for this type of cell, and at physiologic temperature extracted so much cholesterol from the plasma membrane that the increase in observed current was not actually due to increased ClC-2 activity, but was instead caused by decreased cell membrane integrity as a whole. We also note that those authors did not

use a ClC-2-selective inhibitor to confirm that the MβCD-induced currents were indeed attributable to ClC-2. Finally, our oocytes have a vitelline membrane that affords the cell durability but may lead to limited extraction of cholesterol, and that coupled with a lower temperature during MβCD treatment could explain our conflicting results.

Future directions should include site-directed mutagenesis to selectively disrupt potential cholesterol binding motifs within rabbit ClC-2 and repeat TEVC experiments to determine if the disruption of these domains would also disrupt channel function. Even though there is not a solved Cryo-EM structure of rabbit ClC-2, potential targets for mutagenesis could be gleaned from the sequence alignment in which CRAC and CARC domains were conserved across multiple ClC proteins. Based on the structure of human ClC-1 and sequence alignment, potential targets for mutagenesis in rabbit ClC-2 include the following sequences: 177-V to 186-K (CRAC_3); 275-V to 280-R (CRAC_5); 186-K to 192-V (CARC_4); 314-R to 324-L (CARC_6) which correspond to the ClC-1 domains CRAC_3, CRAC_5, CARC_1 and CARC_3 respectively.

Additionally, experiments should be conducted in which cholesterol is extracted from the oocyte membrane by MβCD, and epicholesterol is loaded back into the membrane in cholesterol's place. Epicholesterol is a stereoisomer of cholesterol and although it would not bind CRAC and CARC domains in ClCs, it would maintain whole cell membrane mechanics and make clearer how direct the effects of cholesterol extraction were on ClC function. Finally, based on the sequence and structural similarities between ClCs it is likely that TEVC studies on these ClCs would yield similar results. Additional TEVC experiments with different ClC cRNAs should be conducted to confirm if extraction of plasma membrane cholesterol using MβCD would also inhibit other ClCs in conducting current across the plasma membrane.

Conclusions

The present study identified numerous cholesterol binding domains within the primary sequences of five ClC ion channels from two species. Analysis of several primary amino acid sequences and computational modeling led to the identification of 10 potential cholesterol binding domains within human ClC-1 that were likely implicated in ClC protein-cholesterol interactions based on localization to the transmembrane domains of the protein. Of these 10 potential binding domains, sequence homology identified four domains that exhibited 100% conservation in all five of the ClCs studied. TEVC studies of rabbit ClC-2 further supported an underlying dependence on membrane cholesterol for effective conductance of current across the plasma membrane. It is likely that replicating these studies with other ClCs would yield similar results based on the similarities in sequence and in structure of other ClC ion channels. Future work should focus on using site-directed mutagenesis to confirm the importance of the identified cholesterol binding domains and replicating TEVC studies with these mutant ClCs.

Tables and Figures

Table 1. CRAC Domains in ClC-1 (Human)			
Site Name	**Residue 1**	**Residue 2**	**Residue 3**
CRAC_1	56-V	62-Y	66-K
CRAC_2	131-V	137-Y	141-K
CRAC_3	201-V	206-Y	210-K
	202-V		
	203-L		
CRAC_4	292-V	296-Y	300-R
CRAC_5	299-V	302-Y	304-R
CRAC_6	361-L	367-Y	370-R
	364-V		
CRAC_7	616-V	620-Y	626-R
		622-Y	
CRAC_8	681-L	686-Y	689-K
			691-R
	684-L		689-K
			691-R
CRAC_9	844-L	850-Y	856-K
	846-L		
	848-L		

Table 2. CARC Domains in ClC-1 (Human)

Site Name	Residue 1	Residue 2	Residue 3
CARC_1	210-K	212-F	216-V
			217-V
CARC_2	248-K	253-F	256-V
CARC_3	338-R	341-F	345-L
		343-F	
			348-L
CARC_4	380-K	385-F	391-L
CARC_5	388-K	393-Y	397-V
	390-R		
CARC_6	518-K	524-Y	526-V
CARC_7	585-K	589-Y	593-L
	586-K		
CARC_8	600-K	604-F	610-V
CARC_9	614-K	620-Y	625-L
CARC_10	680-K	686-Y	692-L
CARC_11	704-R	708-F	711-V
CARC_12	764-R	768-F	771-L
			772-L
CARC_13	836-K	841-F	843-L
			844-L
			846-L

Table 3. CRAC Domains in ClC-2 (Human)			
Site Name	**Residue 1**	**Residue 2**	**Residue 3**
CRAC_1	21-L	23-Y	25-R
CRAC_2	61-L	64-Y	66-R
			68-R
	62-L		66-R
			68-R
CRAC_3	174-V	179-Y	183-K
	175-V		
	176-L		
CRAC_4	225-L	230-Y	235-R
CRAC_5	272-V	275-Y	277-R
CRAC_6	337-L	340-Y	343-R
			344-K
CRAC_7	548-L	553-Y	558-R
CRAC_8	813-L	819-Y	825-R
	815-V		

Table 4. CARC Domains in ClC-2 (Human)

Site Name	Residue 1	Residue 2	Residue 3
CARC_1	13-R	17-Y	21-L
CARC_2	78-R	82-F	84-V
			87-V
CARC_3	177-K	179-Y	182-L
CARC_4	183-K	185-F	189-V
CARC_5	307-K	310-F	312-L
CARC_6	309-R	314-F	318-L
	311-R	316-R	
			321-L
CARC_7	353-K	358-F	364-L
CARC_8	361-R	366-F	369-6
			370-V
			372-L
	362-K		369-L
			370-V
			372-L
	363-R		369-L
			370-V
			372-L
CARC_9	428-R	432-F	435-L
			436-V
CARC_10	493-R	499-Y	501-V
			502-V
CARC_11	558-R	564-Y	568-L
	560-K		
	561-K		
CARC_12	572-R	576-Y	578-V
			580-V
CARC_13	789-K	795-F	797-L
			798-v
CARC_14	805-K	810-F	812-L
			813-L
			815-V

Table 5. CRAC Domains in ClC-Ka (Human)

Site Name	Residue 1	Residue 2	Residue 3
CRAC_1	133-V 135-L	138-Y	142-K
CRAC_2	174-V	179-Y	182-R 184-R
CRAC_3	233-V	236-Y	238-R
CRAC_4	294-V 295-L	299-Y	304-R
CRAC_5	387-L	392-Y	395-R
CRAC_6	526-V 529-L	531-Y	534-R

Table 6. CARC Domains in ClC-Ka (Human)			
Site Name	**Residue 1**	**Residue 2**	**Residue 3**
CARC_1	8-R	11-F	16-V
CARC_2	40-K	44-F	46-L
	42-K		
CARC_3	45-R	51-Y	53-L
			56-L
CARC_4	142-K	144-F	148-V
			149-V
CARC_5	165-K	169-F	172-L
			174-V
CARC_6	268-K	271-F	273-V
			275-V
CARC_7	272-R	277-F	279-L
CARC_8	311-K	315-Y	319-L
			320-L
CARC_9	324-K	327-Y	330-L
			333-L
CARC_10	395-R	399-F	402-L
CARC_11	527-K	531-Y	536-L
	528-K		

Table 7. CRAC Domains in ClC-Kb (Human)

Site Name	Residue 1	Residue 2	Residue 3
CRAC_1	133-V	138-Y	142-K
	134-V		
	135-L		
CRAC_2	174-V	179-Y	182-R
			184-R
CRAC_3	233-V	236-Y	238-R
CRAC_4	295-L	399-Y	304-R
CRAC_5	334-V	340-Y	346-R
	335-L		
CRAC_6	387-L	392-Y	395-R
CRAC_7	526-V	531-Y	534-R
	529-L		

Table 8. CARC Domains in ClC-Kb (Human)			
Site Name	**Residue 1**	**Residue 2**	**Residue 3**
CARC_1	40-K 42-K	44-F	46-L
CARC_2	45-R	51-Y	53-L 56-L
CARC_3	142-K	144-F	148-V 149-V
CARC_4	165-K	169-F	172-L 174-V
CARC_5	268-K	271-F	273-V 275-V
CARC_6	272-R	277-F	279-L
CARC_7	311R	315-F	319-L 320-L
CARC_8	324-K	327-Y	330-L 333-L
CARC_9	395-R	399-F	402-L
CARC_10	438-R	440-F	444-L
CARC_11	527-K 528-K	531-Y	536-L

Table 9. CRAC Domains in ClC-2 (Rabbit)			
Site Name	**Residue 1**	**Residue 2**	**Residue 3**
CRAC_1	24-L	26-Y	28-R
CRAC_2	64-L	67-Y	71-R
	65-L		
CRAC_3	177-V	182-Y	186-K
	178-V		
	179-L		
CRAC_4	228-L	233-Y	238-R
CRAC_5	275-V	278-Y	280-R
CRAC_6	340-L	343-Y	346-R
			347-K
CRAC_7	551-L	556-Y	561-R
CRAC_8	813-L	819-Y	825-R
	815-V		

Table 10. CARC Domains in ClC-2 (Rabbit)

Site Name	Residue 1	Residue 2	Residue 3
CARC_1	16-R	20-Y	24-L
CARC_2	81-R	85-F	87-V
			90-V
CARC_3	180-K	182-Y	192-V
CARC_4	186-K	188-F	192-V
CARC_5	310-K	313-F	315-L
CARC_6	312-R	317-F	321-L
	314-R		
		319-F	
			324-L
CARC_7	356-K	361-F	367-L
CARC_8	364-R	369-F	372-L
			373-V
			375-L
	365-K		372-L
			373-V
			375-L
	366-R		372-L
			373-V
			375-L
CARC_9	431-R	435-F	438-L
			439-V
CARC_10	496-R	502-Y	504-V
			505-V
CARC_11	561-R	567-Y	571-L
	563-K		
	564-K		
CARC_12	575-R	579-Y	581-V
			583-V
CARC_13	789-K	795-F	797-L
			798-V
CARC_14	805-K	810-F	812-L
			813-L
			815-V

```
CRAC_and_CARC_Domains <- function(aminoacids) {
  print("CRAC Domains")
  for (AA1 in 1:length(aminoacids)) {
    if(aminoacids[AA1]=="L" | aminoacids[AA1]=="V") {
      for(AA2 in 2:6) {
        if(aminoacids[AA1+AA2]=="Y") {
          for (AA3 in 2:6) {
            if (aminoacids[AA1+AA2+AA3]=="R" | aminoacids[AA1+AA2+AA3]=="K") {
              print(c(AA1, AA1+AA2, AA1+AA2+AA3))
              print(c(aminoacids[AA1],aminoacids[AA1+AA2], aminoacids[AA1+AA2+AA3]))
            }
          }
        }
      }
    }
  }
  print("CARC Domains")
  for (AA1 in 1:length(aminoacids)) {
    if(aminoacids[AA1]=="R" | aminoacids[AA1]=="K") {
      for(AA2 in 2:6) {
        if(aminoacids[AA1+AA2]=="Y"| aminoacids[AA1+AA2]=="F") {
          for (AA3 in 2:6) {
            if (aminoacids[AA1+AA2+AA3]=="L" | aminoacids[AA1+AA2+AA3]=="V") {
              print(c(AA1, AA1+AA2, AA1+AA2+AA3))
              print(c(aminoacids[AA1],aminoacids[AA1+AA2], aminoacids[AA1+AA2+AA3]))
            }
          }
        }
      }
    }
  }
}
```

Figure 1. Scanning primary amino acid sequences for CRAC and CARC domains using R. A function was built in R to search for potential cholesterol binding sites. By using a series of “for loops” and “if/then” statements, this code is able to move through the entire protein sequence and determine if a segment meets all of the criteria of either a CRAC or CARC domain.

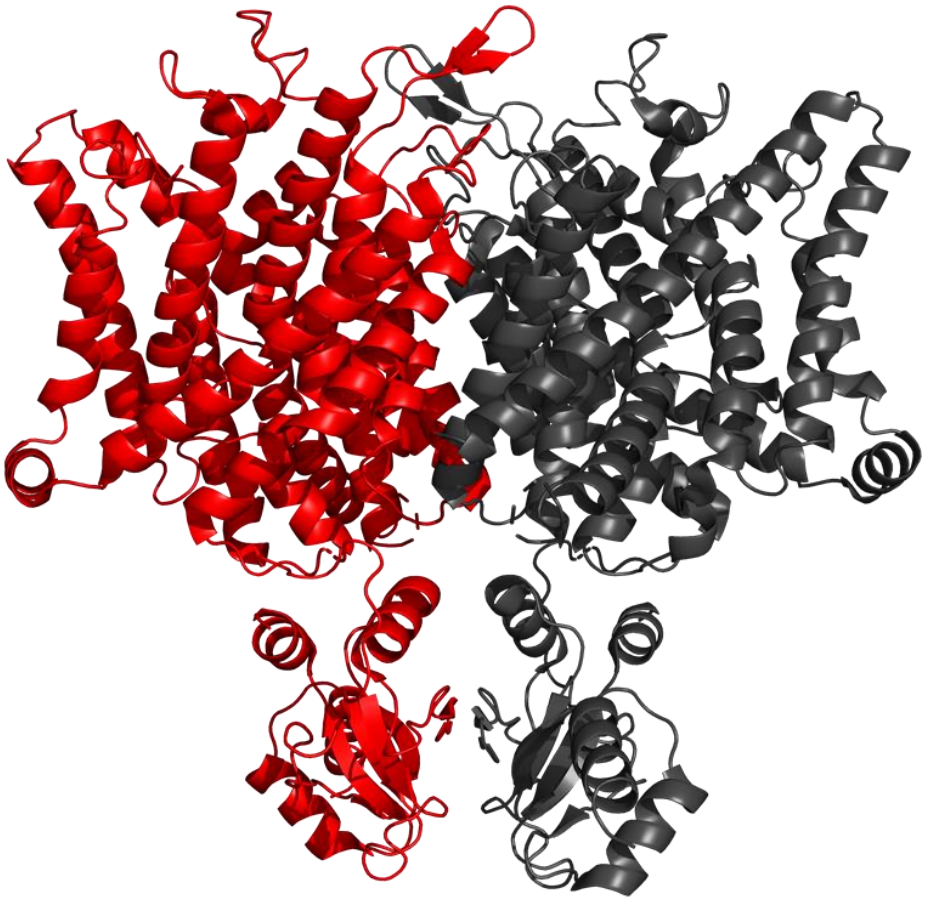

Figure 2. Solved Cryo-EM structure of human ClC-1. The two monomers that comprise the homodimer of human ClC-1 are shown with the two identical protein chains colored red and grey.

```
receptor = 6qvc_A_forADV.pdbqt
ligand = chol_2.pdbqt

center_x = 143.316
center_y = 178.364
center_z = 169.913

size_x = 20
size_y = 20
size_z = 20

log= log_clc1_crac2.txt
```

Figure 3. Representative configuration file for cholesterol docking studies. A representative image of a configuration file used to interface with AutoDock Vina executable software is shown. “Receptor” gives the file name of the protein of interest, and “ligand” gives the file name of the binding ligand of interest. The “center” coordinates are centered on the α-carbon of the central amino acid residue for the domain of interest. The “size” lengths represent the size of the grid box in angstroms. Finally, “log” provides a name for the output file generated by the executable software.

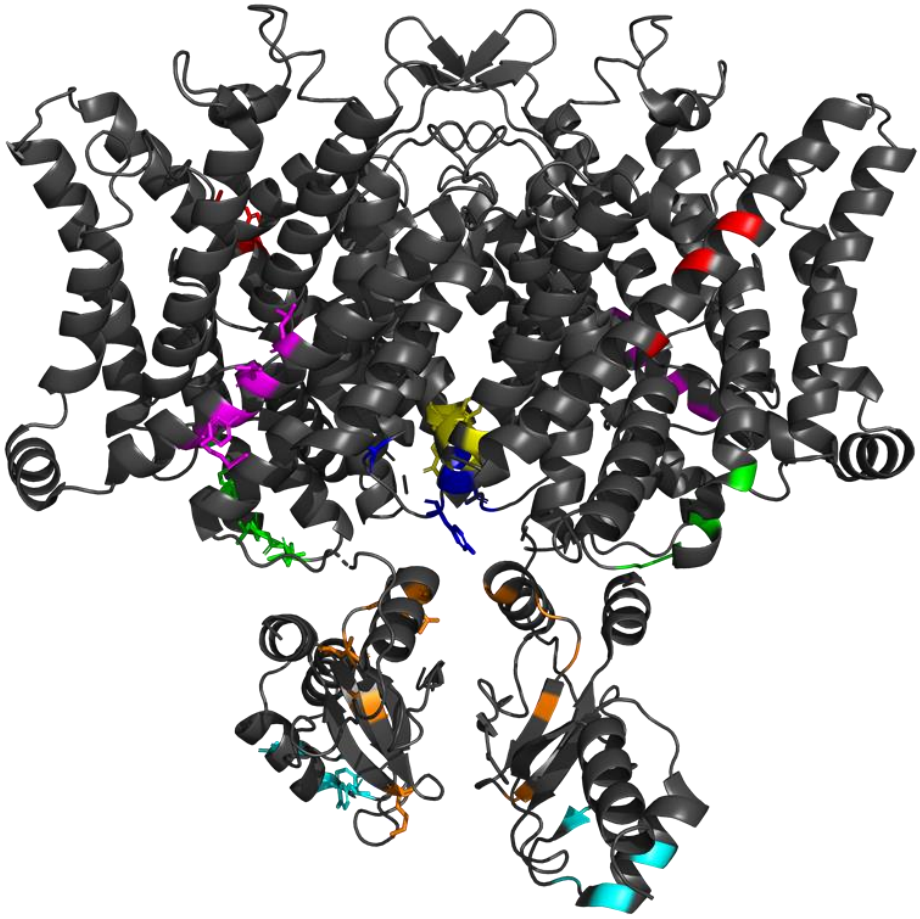

Figure 4. CRAC domains mapped onto the solved Cryo-EM structure of human CIC-1. CRAC domain sequences are shown in various colors, with the corresponding sequence on each protein monomer having the same color in each monomer. The remainder of the protein that has not been identified as constituting a CRAC domain is colored grey.

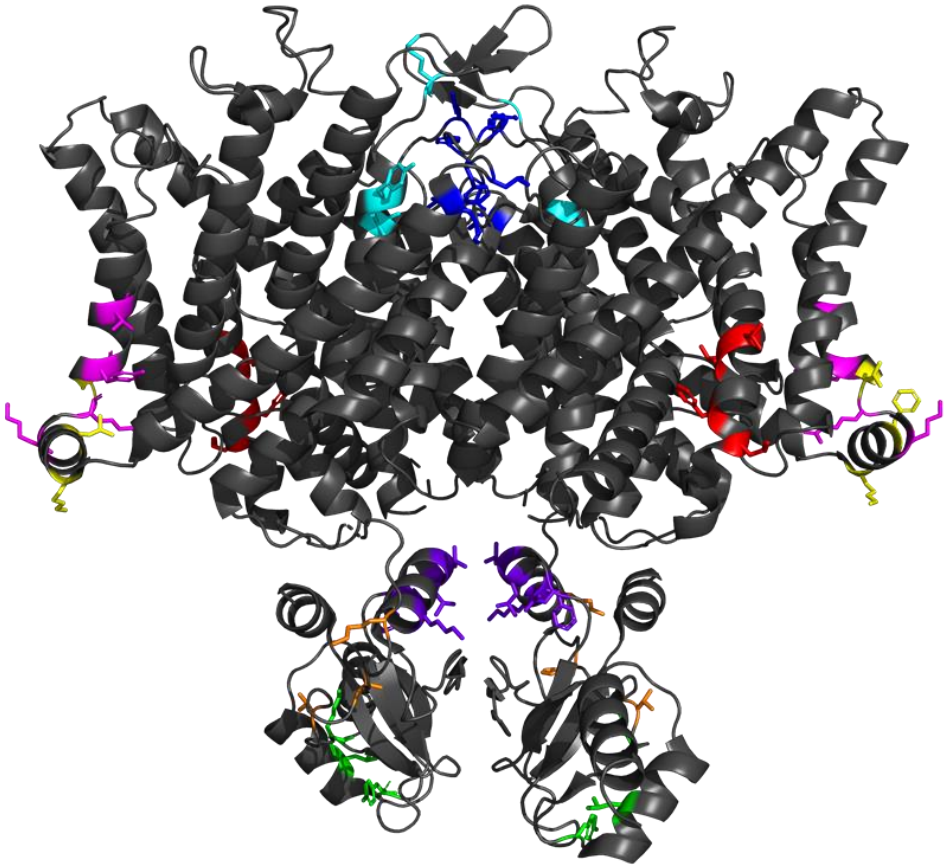

Figure 5. CARC domains mapped onto the solved Cryo-EM structure of human CIC-1. CARC domain sequences are shown in various colors, with the corresponding sequence on each protein monomer having the same color in each monomer. The remainder of the protein that has not been identified as constituting a CARC domain is colored grey.

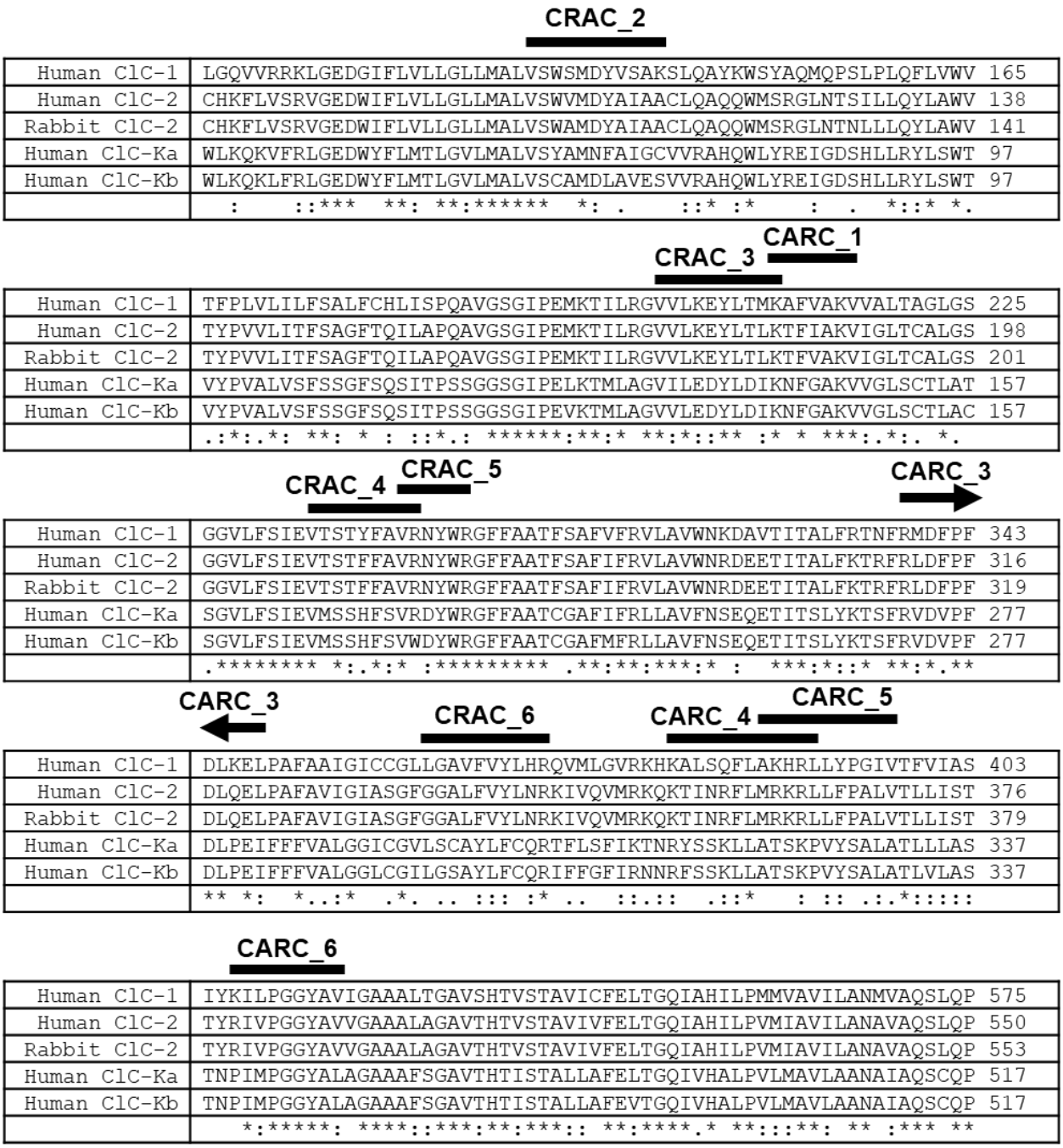

Figure 6. Multiple sequence alignment of five CIC proteins. A multiple sequence alignment was performed using four human CIC primary sequences and one rabbit CIC primary sequence. Each sequence is listed as a separate line, with the right most amino acid in each line having the position indicated by the number to its right within that primary sequence. CRAC and CARC sequences of interest are notated above each segment, and degree of conservation is indicated in the bottom line with (*) indicating 100% conservation, (:) indicating high level of conservation and (.) indicating moderate conservation at a given residue. Only segments of primary sequence encoding cholesterol binding domains of interest are shown.

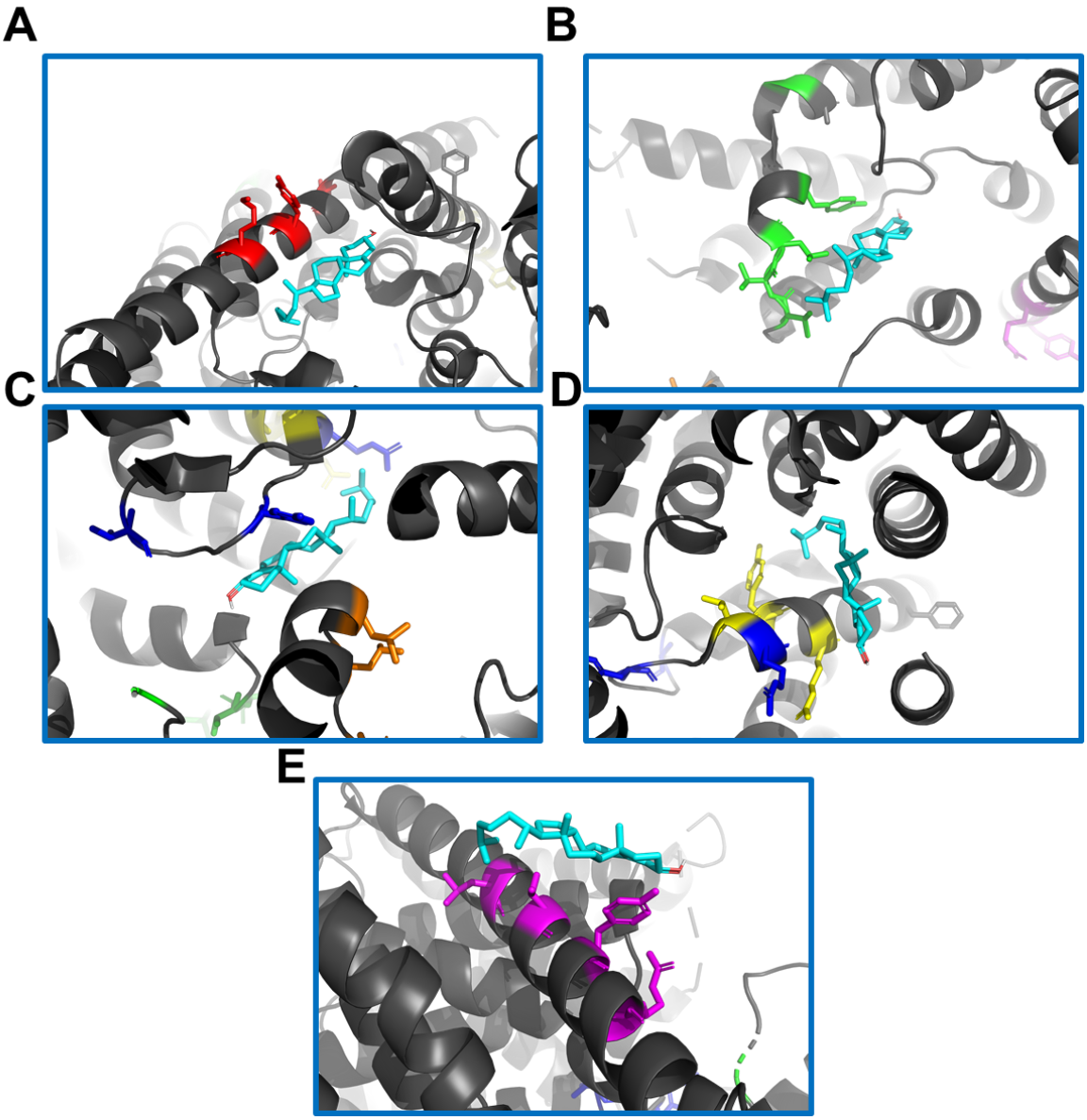

Figure 7. Representative docking positions of cholesterol at potential CRAC binding sites in human CIC-1. Cholesterol is shown in cyan with the oxygen colored red and the hydrogen on that oxygen colored white. The amino acids comprising the CRAC site of interest are colored **(A)** red (CRAC_2), **(B)** lime (CRAC_3), **(C)** blue (CRAC_4), **(D)** yellow (CRAC_5), **(E)** cyan (CRAC_6). The amino acid side chains on the residues of interest are shown, while the remainder of the protein is colored gray and side chains are not shown.

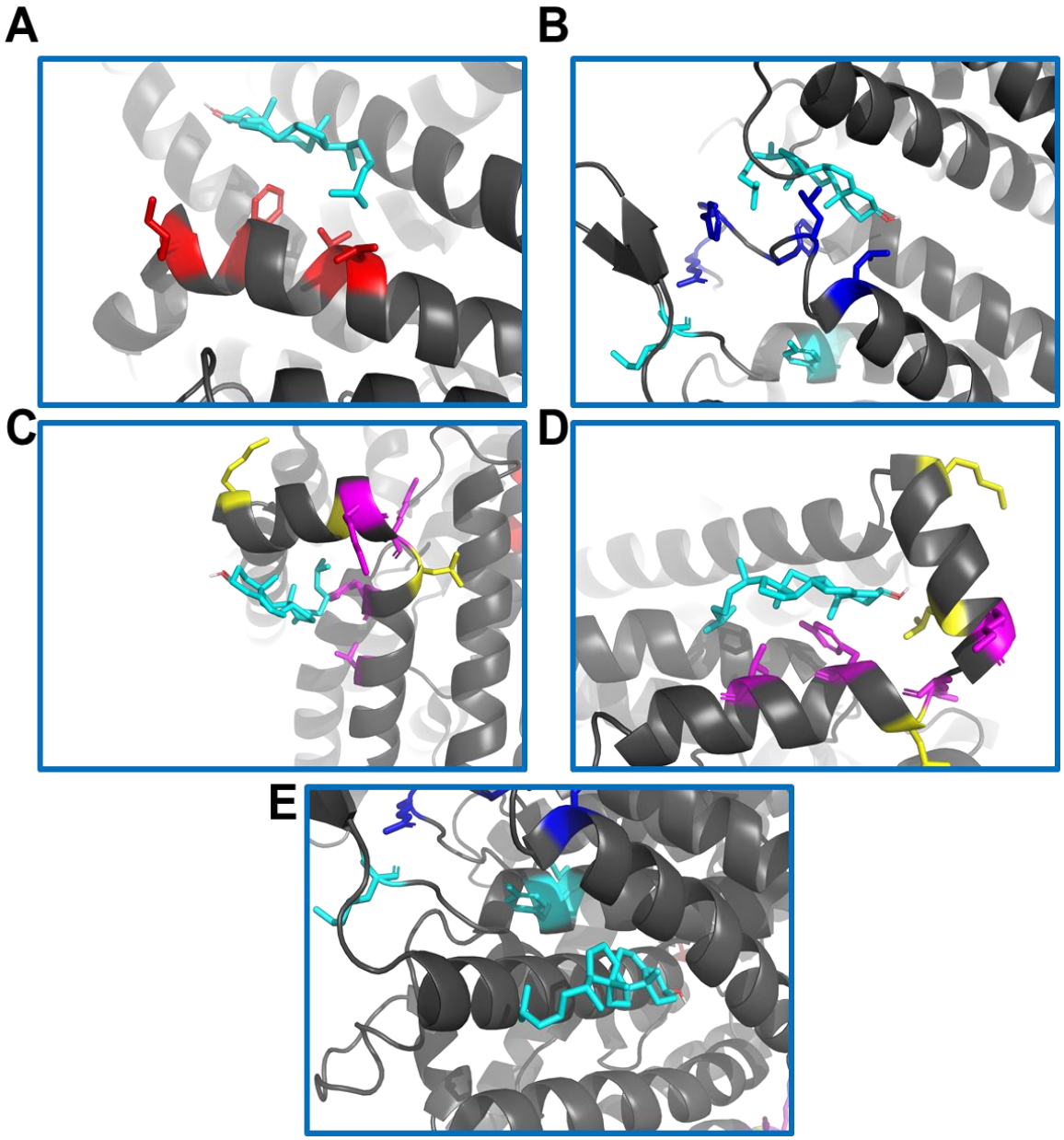

Figure 8. Representative docking positions of cholesterol at potential CARC binding sites. Cholesterol is shown in cyan with the oxygen colored red and the hydrogen on that oxygen colored white. The amino acids comprising the CARC site of interest are colored **(A)** red (CARC_1), **(B)** blue (CARC_3), **(C)** yellow (CARC_4), **(D)** pink (CARC_5), **(E)** cyan (CARC_6). The amino acid side chains on the residues of interest are shown, while the remainder of the protein is colored gray and side chains are not shown.

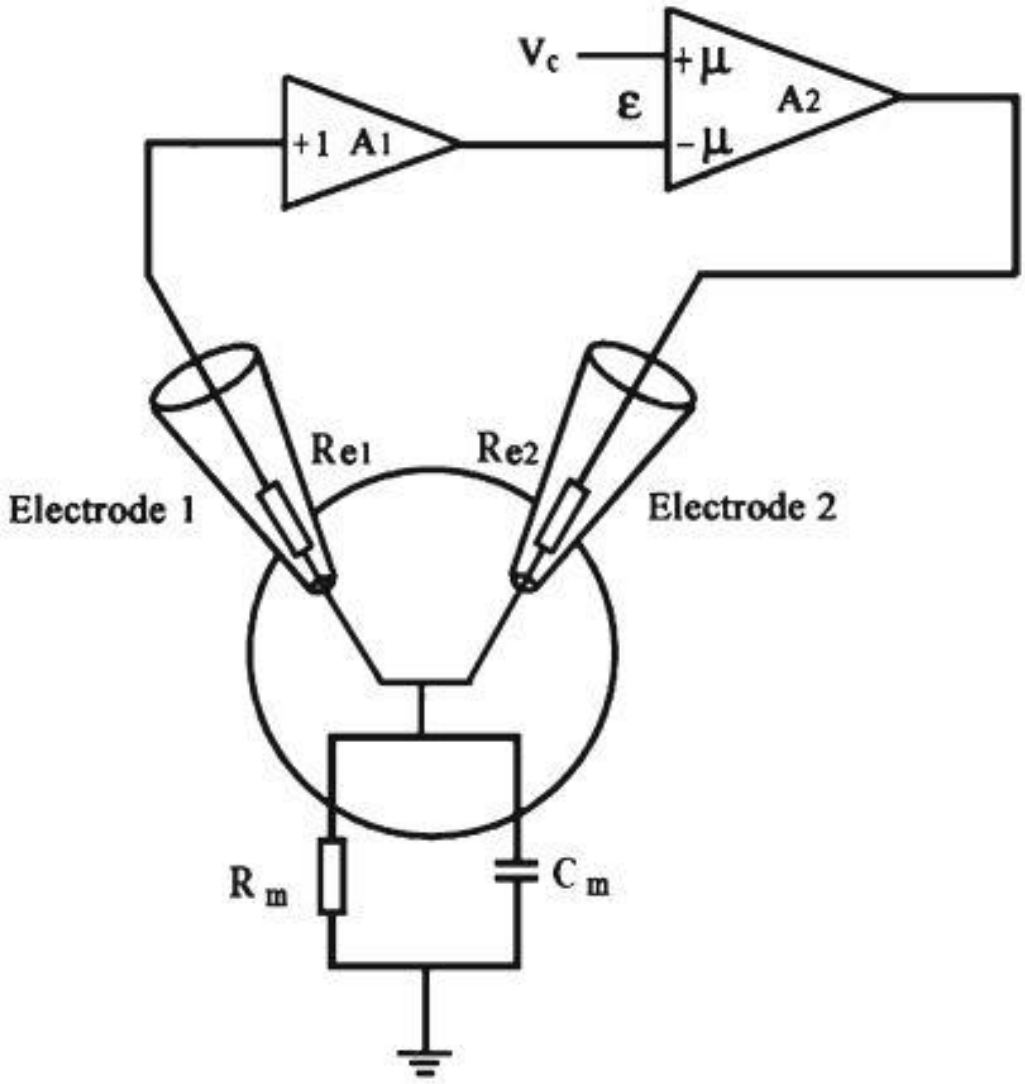

Figure 9. Circuit diagram of Two Electrode Voltage Clamp with a *Xenopus* oocyte. This figure was published by Guan et al. (2013) and depicts a simple circuit diagram of a two electrode voltage clamp apparatus. V_c denotes command voltage, which is the potential set by the experimenter using a computer interface. ε denotes the difference in voltage measured from the cell via Electrode 1 and the command voltage, which is used to inject an amount of current generated by amplifier 2 (A2) through Electrode 2 (with resistance Re_2) to bring ε to 0. C_m and R_m represent the capacitance and the resistance of the oocyte membrane, respectively. Electrode 1 (with resistance Re_1) records the membrane potential of the oocyte, and amplifier 1 (A1) relays this data to A2 where the loop repeats. This system allows for virtually instantaneous recording and adjustment of current-voltage relationships in living oocytes.

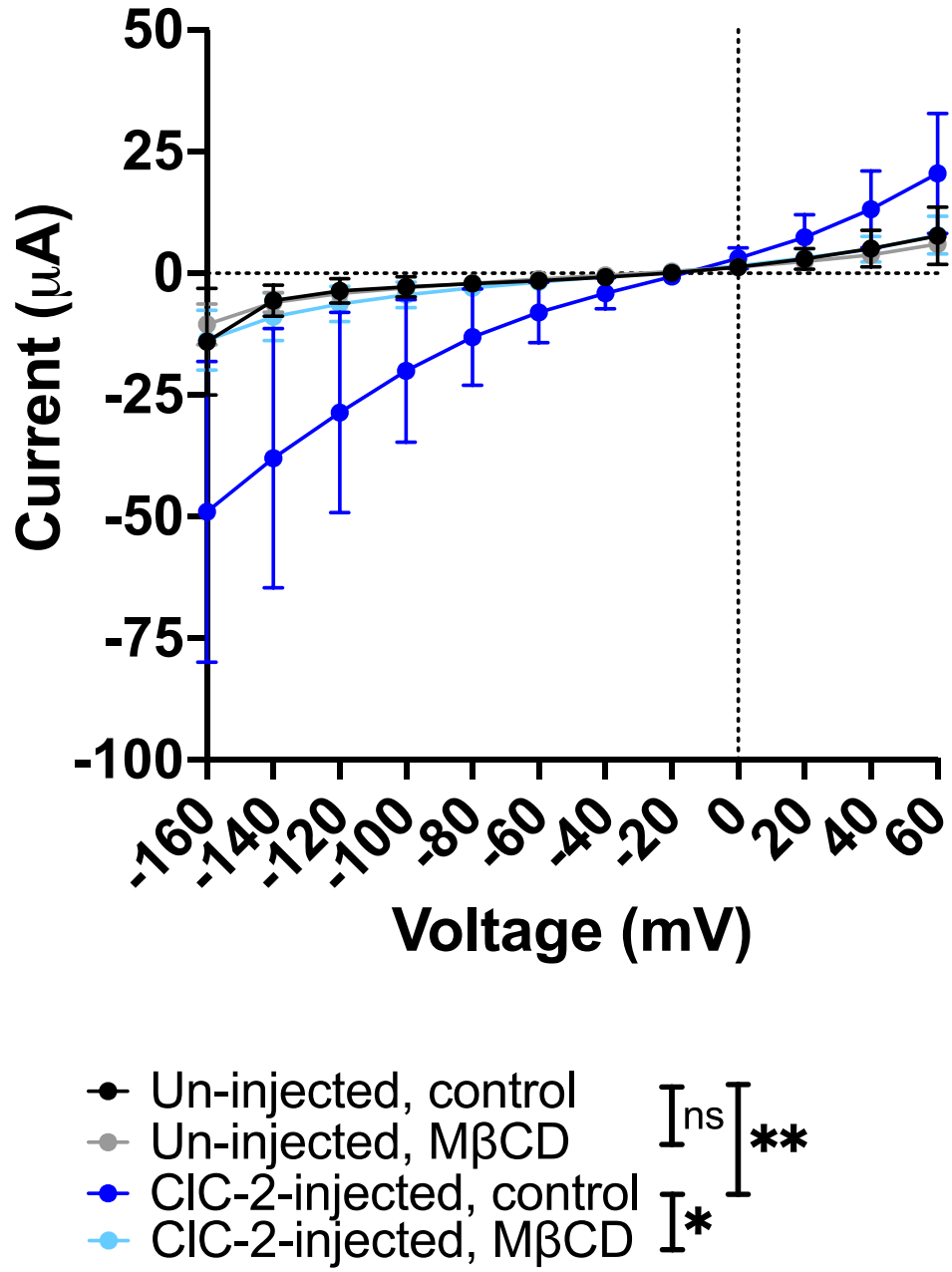

Figure 10. Current-voltage relationship of oocytes from four treatment groups. Average current-voltage curves for un-injected control cells (black, n=7), un-injected MβCD treated cells (grey, n=7), ClC-2-injected control cells (blue, n=10) and ClC-2-injected MβCD treated cells (cyan, n=7). Treatments were compared by multiple two-way ANOVAs, with repeated measures over the voltage range. There was no significant difference between un-injected control oocytes and un-injected oocytes treated with MβCD (p=0.9893). There was a significant difference between un-injected control oocytes and ClC-2-injected control oocytes (**p=0.0076). There was a significant difference between ClC-2-injected control oocytes and ClC-2-injected oocytes treated with MβCD (*p=0.0120).

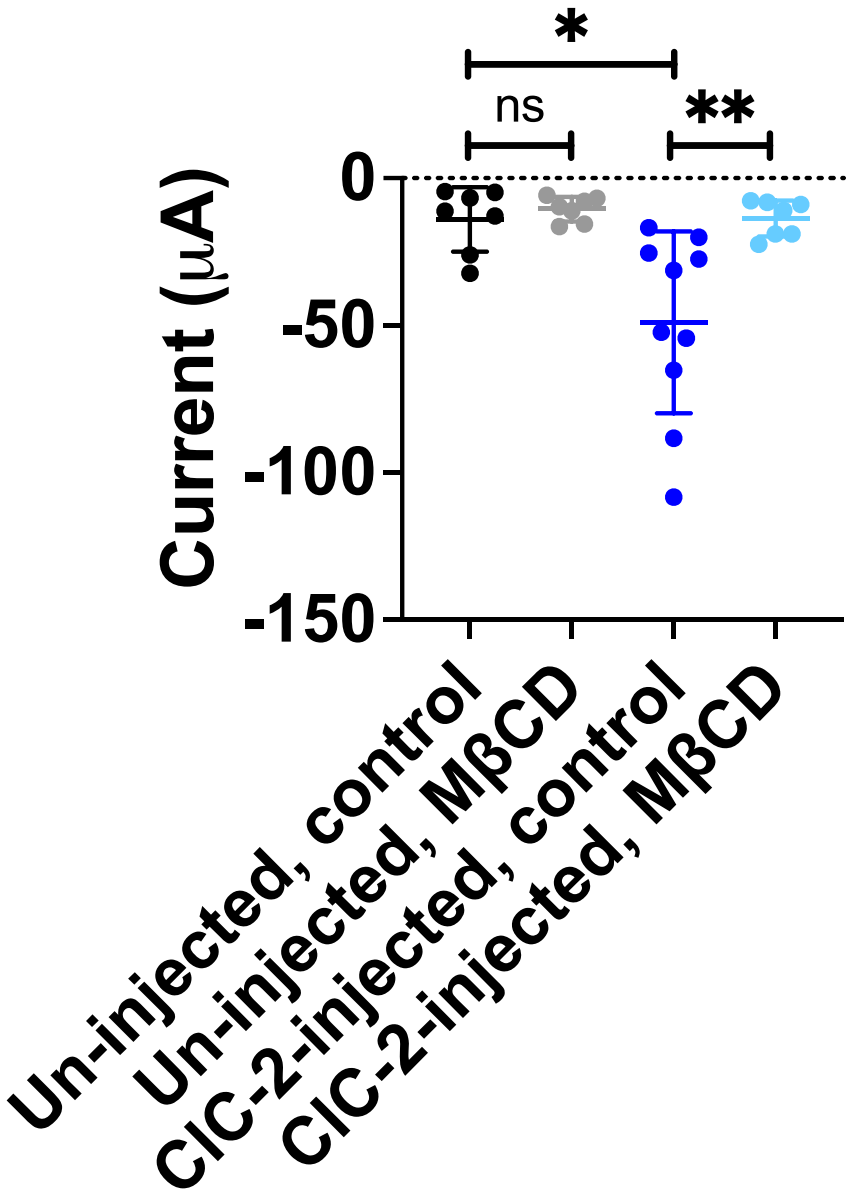

Figure 11. Whole cell net transmembrane current elicited at -160 mV for four treatment groups. ClC-2 elicited current at a voltage of -160 mV for un-injected control cells (black, n=7), un-injected MβCD treated cells (grey, n=7), ClC-2-injected control cells (blue, n=10) and ClC-2-injected MβCD treated cells (cyan, n=7). For comparisons of the final current at -160 mV, multiple unpaired two-tailed *t* tests were performed. There was no significant difference between un-injected control oocytes and un-injected oocytes treated with MβCD (p=0.4352). There was a significant difference between un-injected control oocytes and ClC-2-injected control oocytes (*p=0.0123). There was a significant difference between ClC-2-injected control oocytes and ClC-2-injected oocytes treated with MβCD (**p=0.0099).

References

1. Cystic Fibrosis Foundation. About Cystic Fibrosis. Accessed June 25, 2020. www.cff.org/What-is-CF/About-Cystic-Fibrosis/

2. Rommens JM, Iannuzzi MC, Kerem B, et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*. 1989;245(4922):1059-1065. doi:10.1126/science.2772657

3. Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989;245(4922):1066-1073. doi:10.1126/science.2475911

4. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science*. 1989;245(4922):1073-1080. doi:10.1126/science.2570460

5. Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. *J Lipid Res*. 2001;42(7):1007-1017.

6. Hwang T-C, Kirk KL. The CFTR ion channel: gating, regulation, and anion permeation. *Cold Spring Harb Perspect Med*. 2013;3(1):a009498. doi:10.1101/cshperspect.a009498

7. Breeze R, Turk M. Cellular Structure, Function and Organization in the Lower Respiratory Tract. *Environ Health Perspect*. 1984;55:3-24. doi:10.2307/3429688

8. Gibson-Corley KN, Meyerholz DK, Engelhardt JF. Pancreatic Pathophysiology in Cystic Fibrosis. *J Pathol*. 2016;238(2):311-320. doi:10.1002/path.4634

9. The Clinical and Functional TRanslation of CFTR (CFTR2). Published 2011. Accessed July 15, 2020. www.cftr2.org/

10. Cystic Fibrosis Foundation. Types of CFTR Mutations. Accessed March 2, 2021. www.cff.org/What-is-CF/Genetics/Types-of-CFTR-Mutations/

11. MacKenzie T, Gifford AH, Sabadosa KA, et al. Longevity of Patients With Cystic Fibrosis in 2000 to 2010 and Beyond: Survival Analysis of the Cystic Fibrosis Foundation Patient Registry. *Ann Intern Med*. 2014;161(4):233-241. doi:10.7326/M13-0636

12. Keogh RH, Szczesniak R, Taylor-Robinson D, Bilton D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data. *J Cyst Fibros Off J Eur Cyst Fibros Soc*. 2018;17(2):218-227. doi:10.1016/j.jcf.2017.11.019

13. Davies JC, Wainwright CE, Canny GJ, et al. Efficacy and Safety of Ivacaftor in Patients Aged 6 to 11 Years with Cystic Fibrosis with a G551D Mutation. *Am J Respir Crit Care Med*. 2013;187(11):1219-1225. doi:10.1164/rccm.201301-0153OC

14. Lopes-Pacheco M. CFTR Modulators: The Changing Face of Cystic Fibrosis in the Era of Precision Medicine. *Front Pharmacol*. 2020;10. doi:10.3389/fphar.2019.01662

15. Cui G, Cottrill KA, Strickland KM, Mashburn SA, Koval M, McCarty NA. Alteration of membrane cholesterol content plays a key role in regulation of CFTR channel activity. *Front Physiol. (Submitted).*

16. Cooper GM. Structure of the Plasma Membrane. *Cell Mol Approach 2nd Ed*. Published online 2000. Accessed January 25, 2021. http://www.ncbi.nlm.nih.gov/books/NBK9898/

17. Cottrill KA, Farinha CM, McCarty NA. The bidirectional relationship between CFTR and lipids. *Commun Biol*. 2020;3. doi:10.1038/s42003-020-0909-1

18. Schwiebert EM, Cid-Soto LP, Stafford D, et al. Analysis of ClC-2 channels as an alternative pathway for chloride conduction in cystic fibrosis airway cells. *Proc Natl Acad Sci*. 1998;95(7):3879-3884. doi:10.1073/pnas.95.7.3879

19. Zifarelli G, Pusch M. CLC Channels and Transporters. In: Roberts GCK, ed. *Encyclopedia of Biophysics*. Springer; 2013:320-326. doi:10.1007/978-3-642-16712-6_360

20. Imbrici P, Maggi L, Mangiatordi GF, et al. ClC-1 mutations in myotonia congenita patients: insights into molecular gating mechanisms and genotype-phenotype correlation. *J Physiol*. 2015;593(18):4181-4199. doi:10.1113/JP270358

21. Altamura C, Desaphy J-F, Conte D, De Luca A, Imbrici P. Skeletal muscle ClC-1 chloride channels in health and diseases. *Pflüg Arch - Eur J Physiol*. 2020;472(7):961-975. doi:10.1007/s00424-020-02376-3

22. Smith RL, Clayton GH, Wilcox CL, Escudero KW, Staley KJ. Differential expression of an inwardly rectifying chloride conductance in rat brain neurons: a potential mechanism for cell-specific modulation of postsynaptic inhibition. *J Neurosci Off J Soc Neurosci*. 1995;15(5 Pt 2):4057-4067.

23. Jeworutzki E, López-Hernández T, Capdevila-Nortes X, et al. GlialCAM, a protein defective in a leukodystrophy, serves as a ClC-2 Cl(-) channel auxiliary subunit. *Neuron*. 2012;73(5):951-961. doi:10.1016/j.neuron.2011.12.039

24. Okada Y. Volume expansion-sensing outward-rectifier Cl- channel: fresh start to the molecular identity and volume sensor. *Am J Physiol*. 1997;273(3 Pt 1):C755-789. doi:10.1152/ajpcell.1997.273.3.C755

25. Krämer BK, Bergler T, Stoelcker B, Waldegger S. Mechanisms of Disease: the kidney-specific chloride channels ClCKA and ClCKB, the Barttin subunit, and their clinical relevance. *Nat Clin Pract Nephrol*. 2008;4(1):38-46. doi:10.1038/ncpneph0689

26. Estévez R, Boettger T, Stein V, et al. Barttin is a Cl- channel beta-subunit crucial for renal Cl- reabsorption and inner ear K+ secretion. *Nature*. 2001;414(6863):558-561. doi:10.1038/35107099

27. Duffield M, Rychkov G, Bretag A, Roberts M. Involvement of Helices at the Dimer Interface in ClC-1 Common Gating. *J Gen Physiol*. 2003;121(2):149-161. doi:10.1085/jgp.20028741

28. Bennetts B, Parker MW. Molecular determinants of common gating of a ClC chloride channel. *Nat Commun*. 2013;4(1):2507. doi:10.1038/ncomms3507

29. McKiernan KA, Koster AK, Maduke M, Pande VS. Dynamical model of the CLC-2 ion channel reveals conformational changes associated with selectivity-filter gating. *PLOS Comput Biol*. 2020;16(3):e1007530. doi:10.1371/journal.pcbi.1007530

30. Dietschy JM, Turley SD. Cholesterol metabolism in the brain. *Curr Opin Lipidol*. 2001;12(2):105-112.

31. Vance JE. Dysregulation of cholesterol balance in the brain: contribution to neurodegenerative diseases. *Dis Model Mech*. 2012;5(6):746-755. doi:10.1242/dmm.010124

32. NORD (National Organization for Rare Disorders). Niemann Pick Disease Type C. NORD (National Organization for Rare Disorders). Accessed March 17, 2021. https://rarediseases.org/rare-diseases/niemann-pick-disease-type-c/

33. Porter FD, Herman GE. Malformation syndromes caused by disorders of cholesterol synthesis. *J Lipid Res*. 2011;52(1):6-34. doi:10.1194/jlr.R009548

34. Maity T, Fuse N, Beachy PA. Molecular mechanisms of Sonic hedgehog mutant effects in holoprosencephaly. *Proc Natl Acad Sci*. 2005;102(47):17026-17031. doi:10.1073/pnas.0507848102

35. Porter JA, Young KE, Beachy PA. Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science*. 1996;274(5285):255-259. doi:10.1126/science.274.5285.255

36. Valenza M, Rigamonti D, Goffredo D, et al. Dysfunction of the Cholesterol Biosynthetic Pathway in Huntington's Disease. *J Neurosci*. 2005;25(43):9932-9939. doi:10.1523/JNEUROSCI.3355-05.2005

37. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49(D1):D480-D489. doi:10.1093/nar/gkaa1100

38. Fantini J, Barrantes FJ. How cholesterol interacts with membrane proteins: an exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. *Front Physiol*. 2013;4. doi:10.3389/fphys.2013.00031

39. Wang K, Preisler SS, Zhang L, et al. Structure of the human ClC-1 chloride channel. *PLOS Biol*. 2019;17(4):e3000218. doi:10.1371/journal.pbio.3000218

40. *The PyMOL Molecular Graphics System*. Schrödinger, LLC. www.pymol.org/2/

41. F M, Ym P, J L, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019;47(W1):W636-W641. doi:10.1093/nar/gkz268

42. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31(2):455-461. doi:https://doi.org/10.1002/jcc.21334

43. Li H, Papadopoulos V. Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. *Endocrinology*. 1998;139(12):4991-4997. doi:10.1210/endo.139.12.6390

44. Gimpl G. Cholesterol–Protein Interaction: Methods and Cholesterol Reporter Molecules. In: Harris JR, ed. *Cholesterol Binding and Cholesterol Transport Proteins: Structure and Function in Health and Disease*. Subcellular Biochemistry. Springer Netherlands; 2010:1-45. doi:10.1007/978-90-481-8622-8_1

45. Guan B, Chen X, Zhang H. Two-Electrode Voltage Clamp. In: Gamper N, ed. *Ion Channels: Methods and Protocols*. Methods in Molecular Biology. Humana Press; 2013:79-89. doi:10.1007/978-1-62703-351-0_6

46. Hinzpeter A, Fritsch J, Borot F, et al. Membrane Cholesterol Content Modulates ClC-2 Gating and Sensitivity to Oxidative Stress. *J Biol Chem*. 2007;282(4):2423-2432. doi:10.1074/jbc.M608251200