FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

STRUCTURAL AND EVOLUTIONARY ANALYSIS OF THE ACTIVATION MECHANISM IN THE SK CHANNEL-CALMODULIN COMPLEX

An Undergraduate Honors Thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science

in

BIOLOGICAL SCIENCES

WITH HONORS

by

Brittany Montesino

2019

To: Dr. Steven Oberbauer, Chairperson

Department of Biological Sciences

This Undergraduate Honors Thesis in Biological Sciences, written by Brittany Montesino entitled "Structural and Evolutionary Analysis of the Activation Mechanism in the SK Channel-Calmodulin Complex" is submitted to you in partial fulfillment of the requirements for Undergraduate Honors in Biological Sciences. The Biological Sciences Undergraduate Honors Committee and the candidate's research supervisor have read this thesis. We recommend that it be approved.

_____

Dr. Jessica Siltberg-Liberles
Honors Research Supervisor

Dr. Walter M. Goldberg, Chairperson
Undergraduate Honors Committee

Date of Honors Research Presentation: April 15, 2019
This thesis by Brittany Montesino is approved.

_____

Dr. Steven Oberbauer, Chairperson
Department of Biological Sciences

Department of Biological Sciences
Florida International University
2019

DEDICATION

I dedicate this thesis to my mother and brother. Their support and encouragement throughout this process brought me great motivation. Without them, I would not have been able to accomplish all that I have achieved.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my mentor Dr. Jessica Siltberg-Liberles. Dr. Siltberg-Liberles has been an exceptional role model and I feel so fortunate to have worked alongside her for the past several semesters. Bioinformatics was an area I was unfamiliar with prior to this research, but she made it clear and understandable. She was there for me throughout all the stressful evenings and taught me that there are always solutions to every issue. I would also like to thank Janelle Nunez-Castilla, a bioinformatics graduate student in the Dr. Siltberg-Liberles lab, for her patience and help. She aided me on running programs and was always willing to explain something I did not understand.

In addition, I would like to acknowledge the FIU Honors College for providing me with resources so that I can exhibit my research in several conferences. Lastly, I would like to thank Dr. Walter Goldberg, who has allowed me to improve my scientific writing in ways that would be impossible had I not been a part of this program. His insights not only improved my thesis, but also provided me with knowledge I will be able to use during the rest of my career.

ABSTRACT

SK-channels (SKs) are small-conductance calcium-activated potassium channels that are involved in the hyperpolarization of neurons and other excitable cells. Calmodulin, a calcium-sensing protein, is complexed with these potassium channels. Upon calcium-binding, calmodulin undergoes a conformational change that ultimately opens the channel, allowing potassium through. The human genome encodes four different SK-channels (SK1-4). Each paralog is found in varying concentrations throughout the body, with varying effects on the brain and heart. Despite this, the activation mechanism has remained elusive until recently when it was shown how the human SK4 and calmodulin conformations change depending on the binding of calcium to calmodulin. In the inactive state, when calmodulin has not bound calcium, calmodulin is only partially bound to the channel. In the semi-active state, calmodulin is fully bound to the channel but the channel is closed. In the fully-active state, the channel is now open to transport potassium across the membrane. By mapping the interface residues found in each of the three conformations, I identified the residues important for the interactions between subunits. This was combined with an examination of the molecular evolution of this complex across vertebrates to elucidate the conservation of the activation mechanism for the four SK clades and their interactions with calmodulin in the different conformations. The channel-calmodulin interface was found to have more radical changes than within the channel, suggesting that the mechanism of activation by calmodulin may have diverged between the different SK channel paralogs, while the mechanism of channel activation is conserved. The divergence in activation by calmodulin can be explored as a way to therapeutically regulate the activation of the SK channels in a paralog specific manner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

SK-channels (SKs) are a family of small-conductance calcium-activated potassium channels found throughout neurons and other excitable cells. They regulate components of the afterhyperpolarization potential. This contributes to the refractory period after an excitable cell has fired, which is when the cell has departed from its resting potential. Currently, there are three components of the afterhyperpolarization potential (AHP): the fast, medium, and slow AHP. While fast AHPs are voltage-dependent and controlled by big-conductance potassium channels (BKs), medium and slow AHPs are voltage-independent and controlled by SK-channels (Berridge, 2014). The ability of these channels to regulate the medium AHP allows them to govern the interspike interval and as a result, sets the firing frequency (Maylie et al., 2004). This key role attributes this ion pore to various effects on the brain and heart, ranging from synaptic plasticity to the efficacy of T-lymphocytes.

The human genome encodes four different SK-channels (SK 1-4), encoded by the mammalian genes KCNN 1-4. These genes are paralogs, meaning that they are related by gene duplication. Paralogs tend to share some structural and functional characteristics, but often have functionally diverged following the gene duplication event. Each paralog forms a distinct clade on a phylogenetic tree. All SK-channels are composed of a tetramer of six-transmembrane alpha-helical domains (S1-S6), with the channel pore created by S5 and S6 (Weatherhall et al., 2010). Both the N and C-termini of the channel are found intracellularly (Nam et al., 2017). Calcium does not directly bind to the channel, but instead binds to a calcium messenger protein, calmodulin. This messenger is a highly

conserved protein found across all eukaryotes, binding to a wide range of targets due to its flexible linker region (Chin & Means, 2000). In addition, structural studies have indicated configurational variability at the binding interface of the calmodulin C-lobe, further increasing its ability to attach to proteins (Westerlund & Delemotte, 2018). Calmodulin constitutively binds to the SK-channel at its calmodulin binding domain, found on the C-terminus of the channel (Nam et al., 2017). In response to increasing calcium levels in the cell, calcium will bind to the four EF-hand domains of calmodulin. This causes a conformational change in calmodulin that in turn will open the channel pore. Potassium efflux occurs as a result.

SK 1-3 is found throughout electrically excitable cells in the central nervous system as well as the cardiovascular system. SK4 is minimally expressed in the central nervous system, and is instead found more prominently in erythrocytes, vascular endothelium, T-cells, and secretory epithelia (Cui et al., 2014). The SK3 channel is found in vascular smooth muscle, leading to vasodilation (Ledoux et al., 2006). Animal models have implicated these channels with conditions including ataxia and alcohol use disorders. Single nucleotide polymorphisms of SK channel genes have been associated with arrhythmias, hypertension, and other cardiovascular abnormalities (Cui et al., 2014). A study by Stackman et al. (2002) found that blockage of the SK channel enhanced hippocampal memory encoding and retention. Parkinson's disease is thought to stem from a loss of dopaminergic neurons, which synthesize dopamine, in the brain along with dysfunction of potassium channels. Therefore, blockers of the SK channels may be used for treatment (Liu, Wang, & Chen, 2010).

Regulation of the SK-channel occurs in neurons via posttranslational modification through phosphorylation (Adelman, Maylie & Sah, 2012). Phosphorylation of Thr79 in calmodulin results in decreased calcium sensitivity. Neurotransmitters such as acetylcholine and norepinephrine further regulate this phosphorylation (Nam et al., 2017).

Blockers of the SK-channel increase the rate at which neurons fire and sensitivity to excitatory stimuli. Current theories suggest that this is the result of a reduction in the post-spike after-hyperpolarization potential (Iyer et al., 2017). Since this potassium channel hyperpolarizes the membrane, a reduced potential due to inhibition allows for increased recovery of voltage-gated ion channels, such as those that carry sodium (Patlak, 1991). This enhanced activation allows for more available channels, leading the next round of depolarization to occur sooner. However, a disadvantage of these blockers is that they decrease firing regularity (Deister et al., 2009). By increasing the frequency of these spikes, there are moments where the sodium channels are not given enough time to reactivate, resulting in the cancellation of the action potential (Iyer et al., 2017). This results in inconsistent spikes.

One of the first discovered blockers of the SK 1-3 channels is apamin, which is a neurotoxin found in bee venom. Two others are TRAM-34, which binds to the inner pore of SK4, and UCL-1648, which binds to outer areas of the channel (Wulff et al., 2001). Various other compounds function by competing for the common channel pore in all subtypes. Negative allosteric modulators work by reducing the sensitivity of calmodulin for calcium, such as NS8593, which affects only SK1-3. More specific negative

modulators include RA-2, which affect SK3 and SK4 subtypes (Olivan-Viguera et al., 2013).

Activators, in turn, have the opposite effect and primarily function through positive allosteric modulation, which means that it indirectly changes the effect of the primary ligand that binds to the protein. This is done by attaching onto an area that is distinct from where the ligand binds. In this case, these positive allosteric modulators increase calcium sensitivity. Two known positive modulators of the SK channels are 1-ethyl-2-benzimidazolinone (1-EBIO) and NS309 (Nam et al., 2017). Positive modulators allow for the normalization of accelerated firing rates in cells by terminating repetitive firing. These two activators were found to bind to the same binding pocket at the interface between the calmodulin binding domain and the N-lobe of calmodulin (Nam et al., 2017).

Although the pharmacology has been extensively studied over the past few decades, the current blockers and modulators have poor specificity. Most target the same region, which is the common channel pore. Even the more specific modulators affect more than one SK channel subtype. As a result, these inhibitors are not therapeutically useful. The inability to target the subtypes individually also makes it difficult to differentiate between the functions of each SK. Moreover, the activation mechanism has remained elusive. To develop more specific drug targets for SK subtypes, a closer examination of the allosteric mechanism of the SK channel-Calmodulin complex is informative.

Recently, a study using cryogenic electron microscopy (CryoEM) by Lee and MacKinnon (2018) revealed the human SK4 channel bound to calmodulin in three different conformations. The first conformation represented an inactive state with no calcium bound to calmodulin. The second conformation illustrated an inactive state with calcium bound, but before the opening of the pore. The third showed the active state when potassium was able to pass through.

Here, I present a study that analyzes these conformations by identifying the interface residues involved in the conformational changes leading to the activation of the SK4 channel along with the interactions with calmodulin. The evolutionary conservation across all SK channels and within each clade will be examined. Radical amino acid substitutions in these clades, which reveal paralog-specific interfaces, will also be identified. These radical amino acid substitutions suggest functional divergence and may also be explored as binding regions for channel modulators. The identification of these novel sites can be used as potential subtype-specific drug targets for the multitude of ailments associated with SK-channels, including Parkinson's and hypertension.

METHODS

Sequence Retrieval

Three sequence datasets were constructed: (i) the SK protein family at the whole protein level in vertebrates, (ii) the calmodulin protein at the whole protein level in vertebrates and (iii) the calmodulin protein at the nucleotide level in vertebrates. For (i), a local metazoan database as created by Ahrens et al., (2016), was used to construct a representative protein family using UniProt reference sequence O15554 (KCNN4 in *Homo sapiens*) as the search query. This database contains complete, canonical proteomes for 24 metazoan taxa and one choanoflagellate (Ahrens et al., 2016). However, only ten of those species were examined for this analysis. Non-mammalian species used were the zebrafish *Danio rerio* (DR), the West Indian Ocean coelacanth *Latimeria chalmnae* (LC), the western clawed frog *Xenopus tropicalis* (XT), and the red junglefowl *Gallus gallus* (GG). Mammalian species used were the house mouse *Mus musculus* (MM), the gray short-tailed opossum *Monodelphis domestica* (MD), the giant panda *Ailuropoda melanoleuca* (AM), the horse *Equus caballus* (EC), and *Homo sapiens* (HS). Sequence identifiers for all vertebrate species in each dataset are given in Appendix A-C.

To maintain a high-quality multiple sequence alignment calmodulin chains were required to be 100-200 amino acids long. For the SK channel, sequences with insertions or gaps anywhere besides the N- and C- terminus that were longer than ten amino acids long were removed. Prior to removal, an NCBI BLAST search (Altschul et al., 1990) using the blastp algorithm against the species in question in the nr database from the National Center of Biotechnology Information (NCBI) was conducted. The e-value cutoff

used was 2e-30. If there was a more representative isoform for that protein, it was replaced. If not, it was removed.

For (ii), the local metazoan database was used with P62158 (Calmodulin in *Homo sapiens*) as the search query. Sequences were limited to a size between 100 to 200 amino acids long. Those that were not within these limits were searched on NCBI BLAST using the blastp algorithm against the species in question in the nr database with an e-value cutoff of 2e-30. If an equivalent sequence that was 100-200 AA in length was found, it was added to the dataset.

For (iii), the corresponding nucleotide sequences for the amino acid sequences obtained from (i) were retrieved from NCBI using TblastN, which searches the translated nucleotide database using a protein query. All of the protein sequences that did not have an equivalent nucleotide sequence were removed.

Sequence Alignment and Phylogenetic Reconstruction

The sequences for dataset (i) and (ii) were aligned with EMBL-EBI MUSCLE (Edgar, 2004; Chojnaki et al., 2017). Sequences in dataset (iii) were first checked for sequence concordance in (ii) using ExPASY (Gasteiger, 2003). Then, they were aligned using TranslatorX (Abascal, 2010), which maps the corresponding codons to the amino acid alignment from (ii).

Phylogenetic trees for dataset (i) were constructed using MrBayes v3.2.2 (Ronquist et al., 2012). A Bayesian MCMC analysis was performed using a mixed amino acid model with gamma distributed rate variation across sites. This was set to construct

possible trees for 1 million generations with a sampling frequency of 100 generations. A split frequency cutoff was set at 0.005. This is below the recommended cutoff by Ronquist et al. (2012), and should assure that the tree has reached convergence. The consensus tree was built using the default burn-in phase, which discards the first 25% of the trees and uses the 50% majority rule. For datasets (ii) and (iii), the phylogenetic trees were constructed using PhyML 3.0 (Guindon et al., 2010). Automatic substitution by Smart Model Selection based on the Akaike Information Criterion was used (Akaike, 1974). Branch support used the fast likelihood-based method using the SH-aLRT, the nonparametric version based on a procedure similar to the Shimodaira-Hasegawa tree selection (Anisimova et al., 2011). The resulting phylogenetic trees were all midpoint rooted and organized with increasing node order.

Interface Residue Mapping

The interface residues found within the channel and from the channel to calmodulin were determined by analyzing the cryo-electron microscopy structures of each SK4 channel conformation to identify the regions involved in the activation mechanism. These were extracted from the Protein Data Bank (PDB), which is a database for three-dimensional structural data of large biological molecules. The PDB IDs used for the analysis were 6CNM, 6CNN, and 6CNO. These represent the inactive state with no calcium bound, the first bound state, and the active state, respectively. The structures were visualized using Open-Source Pymol version 1.8.x and the contact sites were found using a Pymol Script created on python called “InterfaceResidues”. A dASA cutoff of 1Å was selected, as used by Jones & Thornton (1997). Accessible surface area (ASA) is the area of a protein that

is not buried. When two protein are bound to each other, their respective ASA is smaller than if they were unbound. The dASA is the accessible surface area that becomes inaccessible upon binding (Erijman et al., 2014). For a cutoff of 1Å the residues that make up the binding interface surface are those that decrease their ASA by more than 1 $Å^2$ upon forming the complex.

Once the interface residues were found, they were mapped directly onto a copy of the aligned sequence for *Homo sapiens* SK4 in order to easily identify the sites on the multiple sequence alignment (MSA) in Jalview. This was done by extracting the aligned sequence and replacing all sites that correspond to each specific interface interaction with a different non-amino acid symbol. The four SK channel chains were labeled Chains A-D and the four calmodulin proteins were labeled Chains E-H. For each conformation, the residues at the interfaces between all pairs of chains were identified. Contact sites were found between the four SK channel chains and where the four calmodulin chains interact with the SK.

Interface Residue Conservation

After this mapping was done, graphs were created to compare what amino acids were found at each interface site between each clade. SK1 was compared to SK2, then SK3, then SK4 and so forth. All residue divergences among the clades were recorded. These changes were then further grouped depending on whether the change represented a radical amino acid substitution, a variable change, or a rate shift. A change in residue was considered a radical amino acid substitution if it changed from one amino acid group to another. These groups are i) positively charged (K, R, and H), ii) negatively charged (D

and E), iii) hydrophilic (S, T, N, Q, C, G, and P), and iv) hydrophobic (A, I, L, M, F, W, V, and Y) (Gu, 2006). A rate shift is found on a site where one clade was completely conserved and the other had at least two different amino acids. A variable site included those where both clades had more than one type of amino acid. The percentage of the interface residues that were radical, variable, or a rate shift for each comparison were recorded and separated by conformation. A Mann-Whitney U-test (Mann and Whitney, 1947) was then conducted in order to determine whether the percentage changes per conformation were of statistical significance. A simplified Bonferroni correction was made to account to the established alpha value of 0.05 since multiple comparisons were made (Bonferroni, 1935).

Percent identity matrices were created using Clustal Omega (Chojnacki et al., 2017) to determine the conservation of the interfaces between the SK channels to calmodulin and between the channel subunits. This was done by extracting two sequence sets. The first included all the residues involved in the SK-calmodulin interface and the second included all the residues at the channel interface. Then, all sequences were compared to all other sequences within the same set. The resulting matrix allowed for a visual comparison of all the sampled sequences to each other.

RESULTS

Phylogenetic Reconstruction

Phylogenies for calmodulin and the SK channels were built to provide the evolutionary context for this study. Calmodulin was found to be part of a large protein family with several other calcium-binding proteins. For calmodulin itself, there are three gene copies in most vertebrates, but these sequences are nearly 100% identical on the amino acid level and therefore unresolved in the metazoan-wide phylogeny (not shown). The phylogenetic tree including the three calmodulin paralogs in vertebrates based on amino acid multiple sequence alignments was almost entirely unresolved due to the lack of information in the data (Figure 1). To further resolve this issue, a phylogeny based on the corresponding nucleotide sequences was constructed (Figure 2). Although calmodulin does not change at the amino acid level, it is still evolving at the nucleotide level, as indicated by Figure 2. This can be seen through the resolution of clades into clear dichotomies corresponding to Calmodulin-1, Calmodulin-2, and Calmodulin-3 and in agreement with the Uniprot annotation, which provides and standardizes names for the amino acid sequences. From these trees, it was concluded that the evolutionary context for all calmodulin sequences could be represented by any of the human calmodulin protein sequences since they are all identical. Therefore, all analyses use the human calmodulin sequence to represent the vertebrate calmodulin family.

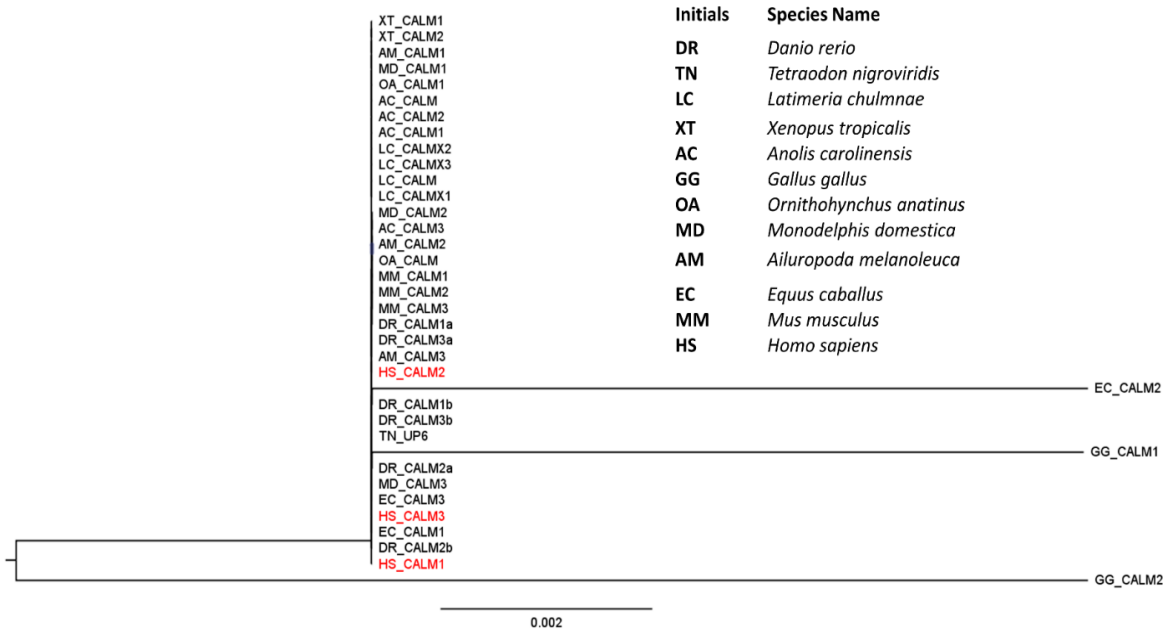

Figure 1. Phylogenetic reconstruction of calmodulin using a protein family sequence alignment. The horizontal axis is a distance scale representing the average number of amino acid substitutions per site. It shows the relative number of differences among the sequences as an indicator of evolutionary rate. The metazoan species names were abbreviated as shown in the legend. *Homo sapiens* sequences were highlighted in red for reference.

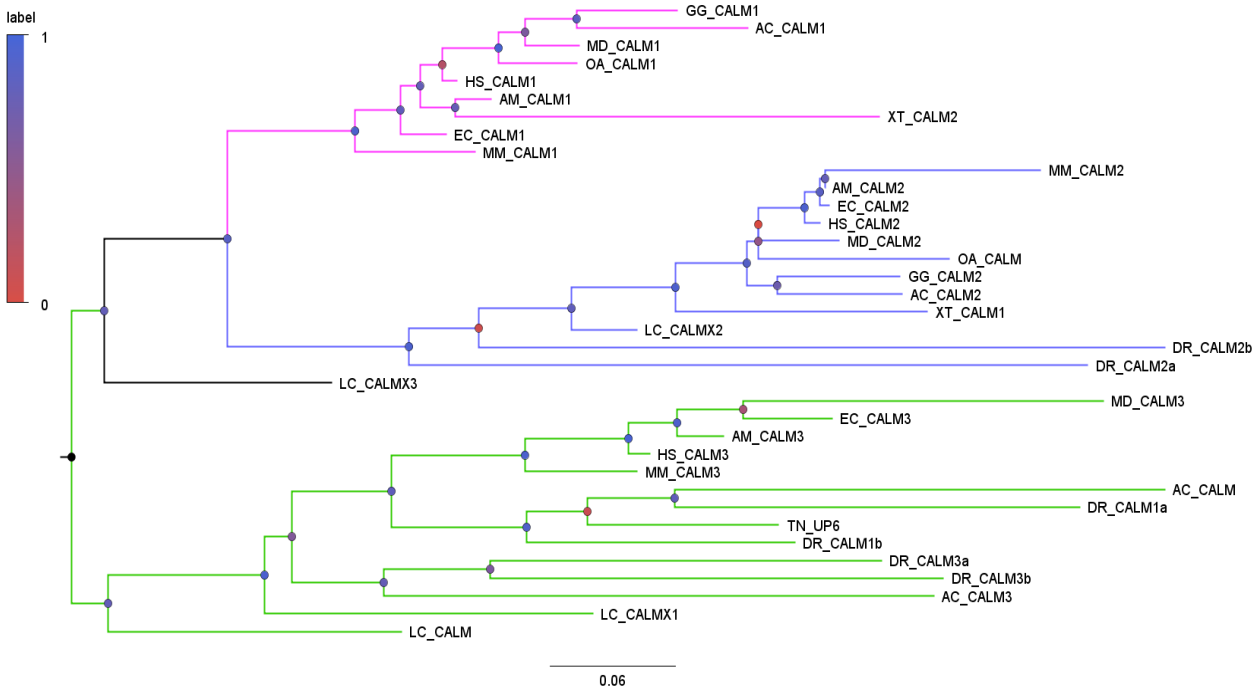

Figure 2. Phylogenetic reconstruction of the calmodulin using a nucleotide sequence alignment, based on the corresponding codons from the amino acid alignment for the tree in Figure 1. The distance scale on the horizontal represents the average number of nucleotide substitutions per site. The same species initials were used in Figure 1. Support values are indicated using a gradient ranging from 0 in red, to 0.96 in blue. Clades are separated by color, in which pink is calmodulin 1, blue is calmodulin 2, and green is calmodulin 3.

The SK channel phylogeny shows divergence at the amino acid level (Figure 3), unlike calmodulin. As a result, there was no need to create a phylogeny at the nucleotide level. The four clades formed were SK1, SK2, SK3, and SK4. It was found that *Homo sapiens* has four copies of the SK channel protein, while other vertebrates have varying amounts, ranging from two in AC to five in DR. Most mammals have four paralogs. The number of copies depends on whole genome duplication and all copies from some species are perhaps not represented due to incomplete proteome sequences. The phylogeny suggests that the rate of evolution is greatest for the SK4 clade, as shown by

the branch lengths. For conservation analysis, human SK4 was used, as it is the only SK channel protein that currently has an experimentally determined structure for the three conformations. Since all the SK channels are homologs and homologs tend to have similar structures, it can be assumed that the SK channel structures are similar. Therefore, the structure of the SK4 channel can be used as a representative of the entire SK channel family, although regulatory properties may have changed.

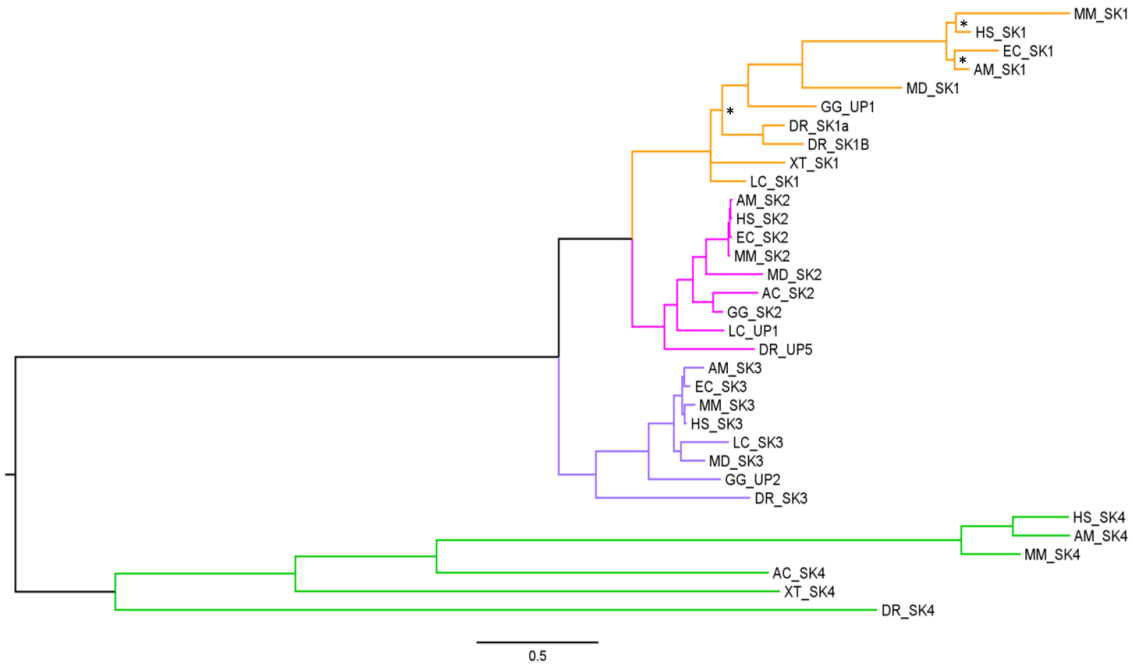

Figure 3. Phylogenetic reconstruction of the SK family using a protein sequence alignment. The horizontal axis is a distance scale which represents the average number of amino acid substitutions per site. Clades are separated by color. SK1 is orange, SK2 is in pink, SK3 is in purple and SK4 in green. Nodes indicated with "*" signify a support value of less than 0.8.

Identification of Interface Residues

The SK channel tetramer has several interfaces between its different subunits. In addition, four calmodulin chains interact with the SK channel tetramer. In this study, three different experimentally determined conformations of the SK channel-calmodulin complex were analyzed. The first conformation (PDB id: 6CNM) shows the inactive form when calmodulin does not bind calcium and the channel is closed. The second conformation (PDB id: 6CNO) shows the first bound state where calcium is bound to calmodulin, but the channel is not yet active. The third conformation (PDB id: 6CNO) shows the active form where calmodulin has bound to calcium and the channel is open. I identified all residues located at the different interfaces in the three different conformations using Pymol and the InterfaceResidues Python script. I found that each SK channel chain interacts with the three other SK chains. In total, there were 107 interface residues for each chain. In other words, Chain A of the SK channel interacted with the three other SK channel chains at 107 different residues as part of its activation mechanism. The three chains interacted with Chain A in different areas. While most of these interface residues remained constant throughout the different stages of activation, there was some variation in the number of interface residues and their locations for the different conformations. For conformation one, two and three, there were 95, 93, and 95 interface residues, respectively (Figure 4).

For the interface between calmodulin and the SK channel, it was found that each SK channel chain interacted with three different calmodulin chains. For conformation one, two and three, there were 24, 57, and 55 interface residues, respectively (Figure 5).

For the first conformation, fewer interface residues were found, since only one lobe of calmodulin was bound to the SK channel at that point.

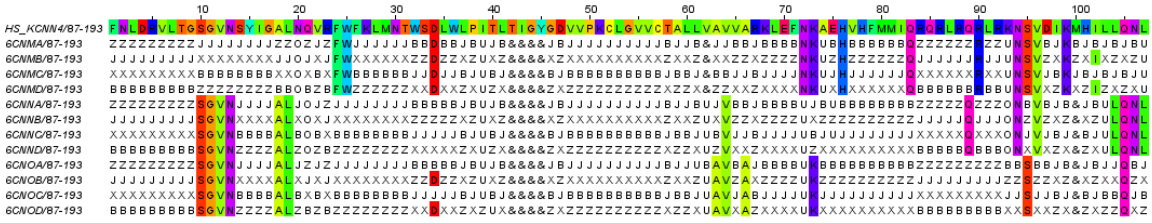

Figure 4. Interface residues between SK channel chains with all gaps and non-interface sites removed. The top line is the equivalent *Homo sapiens* sequence. Chains A-D for each conformation were mapped. 6CNM_A, 6CNN_A, and 6CNO_A map the interface residues with chain A for each SK state. This annotation was used for all four SK chain. Residues where there is contact with another chain are noted using non-amino acid characters. Residues interacting with Chain A were replaced in the corresponding sequence with the letter X, Chain B with B, Chain C with Z, and Chain D with J. Interfaces that involved two chains were labeled O and U. Interfaces where all four chains were in contact were labeled &.

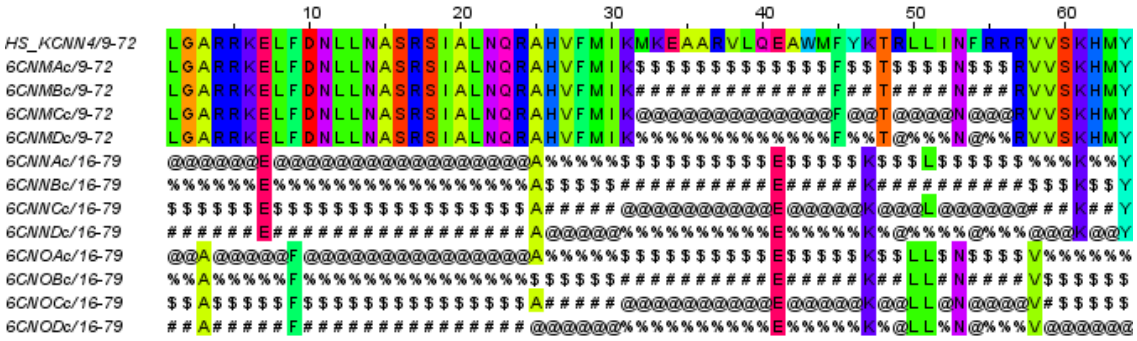

Figure 5. Interface residues between calmodulin and the SK channel chains. The top line is the equivalent *Homo sapiens* sequence. The four SK channel chains (A-D) were mapped and their interactions with the four calmodulin chains (E-H) were recorded. Residues that were in contact with residues from chain E were noted with $, F with #, G with @, and H with %.

Interface Residue Conservation

Although there are some differences in the interface residues in the different conformations, the evolutionary divergence of the interface residues for all channel-to-channel and all calmodulin-to-channel was analyzed in a conformation independent manner. In other words, the interface residues from the three SK channel conformations were compiled rather than separating them into three separate sets. The percent sequence identity for the residues at all channel-to-channel interfaces were calculated (Figure 6). A value of 100 indicates that all amino acids at the interface in human SK4 are identical across all sequences at the corresponding sites in the multiple sequence alignment. A value of 50 indicates that only 50% of the interface residues identified in human SK4 are conserved at the corresponding sites in the multiple sequence alignment. The percent identity calculation was also performed for the calmodulin-to-channel interface residues in a similar manner (Figure 6). The residues at the channel-to-channel interface are more conserved than those found between calmodulin and the channel, as indicated by a higher prevalence of blue in Figure 6B compared with Figure 6A. The level of interface residue conservation of SK1-3 is very similar across all clades, whereas SK4 shows more variance. Not only is SK4 dissimilar to SK1-3, it also appears to change rapidly within its own clade with a clear gap between mammals (MM, HS, and AM) and non-mammals (DR, XT, and AC). The maximum sequence identity for any set of interface residues between sequences within the SK4 clade is 76.64% and 82.48%, for the channel interfaces and for the calmodulin-to-channel interfaces, respectively. However, between mammalian SK4 interface residues are compared to each other; these are at least 92% identical for both calmodulin-to-channel and channel-to-channel interface (Figure 6A).

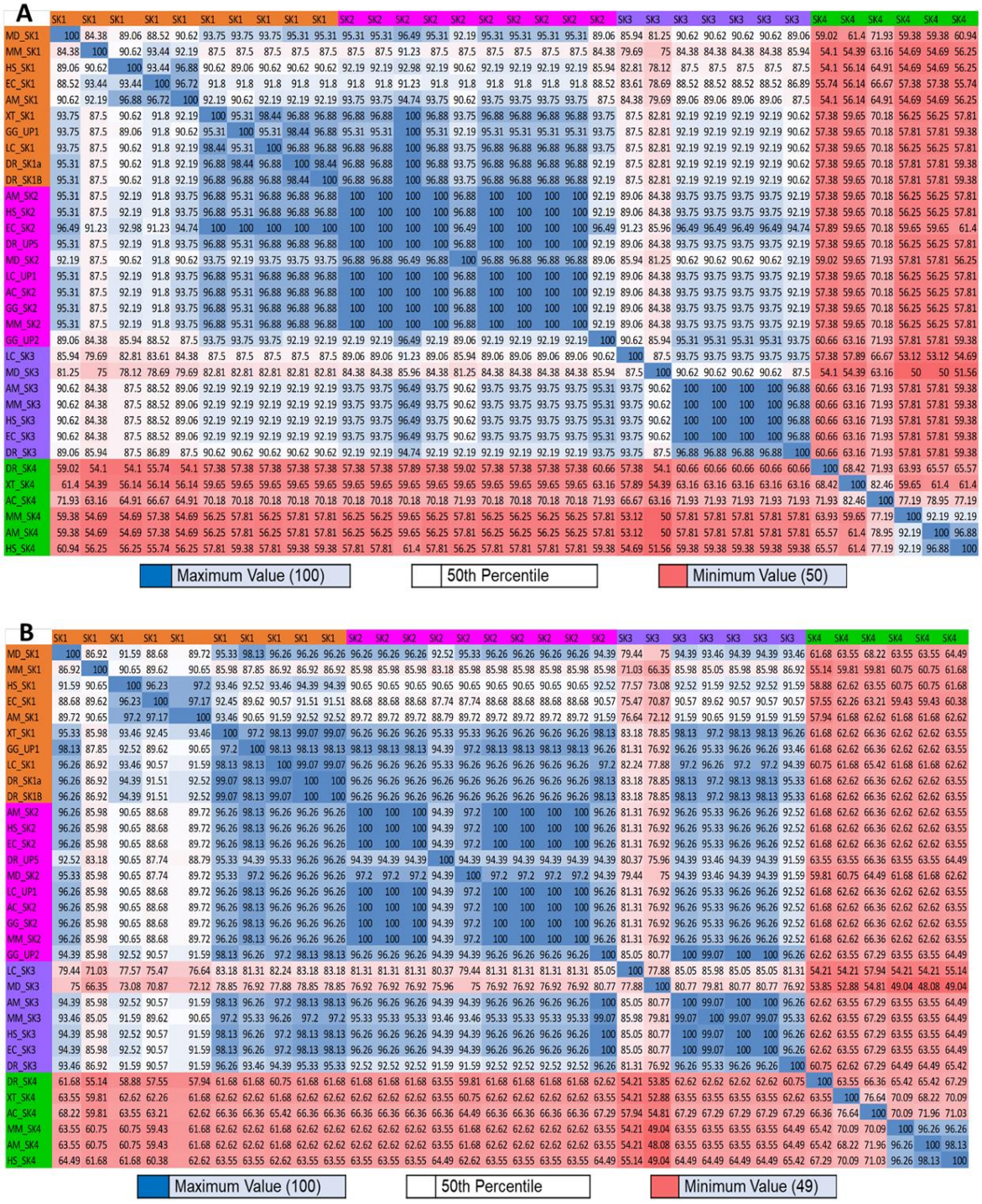

Figure 6. Percent identity matrices for residues located at the interface between channels to calmodulin (A) and channel subunits (B). Blue indicates the maximum value of 100, white is the 50th percentile, and red is the minimum value of either 50 or 49. Clades are colored in the same fashion as the phylogenetic tree (Figure 2), with orange for SK1, pink for SK2, purple for SK3, and green for SK4.

Diverging Sites

To follow-up on the divergences in the different clades, the number of different amino acids per alignment site and per clade were recorded based on the interface residues for the different conformations. The sites that were changing were classified as radical, variable, or rate-shifting. A site was classified as radical if the amino acid changed from one amino acid group to another. The amino acid groups were i) positively charged (K, R, and H), ii) negatively charged (D and E), iii) hydrophilic (S, T, N, Q, C, G, and P), and iv) hydrophobic (A, I, L, M, F, W, V, and Y) (Gu, 2006). For example, if the SK4 clade has a conserved A and the SK1 clade has a conserved D at the same site, this would be considered a radical change because A and D are not in the same amino acid group, therefore, a radical site. A site was classified as a variable site if both clades had more than one type of amino acid, regardless of amino acid group. A site was classified to have a rate shift if one clade was completely conserved and the other had at least three different amino acids.

This analysis indicates that there is high conservation within the channel-to-channel interfaces, but somewhat less conservation at the calmodulin-to-channel interface. There appears to be a difference between the different conformations. The first conformation, representing the closed channel, only experience an average of a 2% rate shift, with no radical substitutions or variable shifts. However, in the second conformation, once the N-lobe of calmodulin binds to the channel, it has the highest rate of divergence as shown with 4% of the interfaces observed to be radical sites, 2% variable sites and 9% rate shifts (Figure 7) For the third, fully active conformation, the interface between calmodulin and the SK channel has 3% radical sites, 2% variable sites,

and 8% rate shifts (Figure 7). The increase in the percentage of radical and variable sites from the inactive closed conformation to the other conformations are statistically significant (Table 1). The change in the percentage of rate-shifting sites is not significant (Table 1), although they appear to increase for all comparisons involving the SK4 channel. This was done using the Mann-Whitney U test. Since multiple comparisons were done, a multiple-comparison correction was performed using a simplified Bonferroni correction (Bonferroni, 1935). With this correction, the established alpha value must be divided by the number of tests that were run simultaneously, in this case it was six. This was because there were six clade comparisons made for each conformation, and the six percentages found for divergence category in conformation 1 (Figure 7) was compared to the six percentage of conformation 2 and 3 simultaneously. As a result, to have a significant result of less than 5% error rate, the p-value found must be less than .05/6, which is 0.0083. Within the channel, no significant differences are found for either class of sites (Table 1), even if there seems to be an increase in the percentage of radical sites for the active conformation.

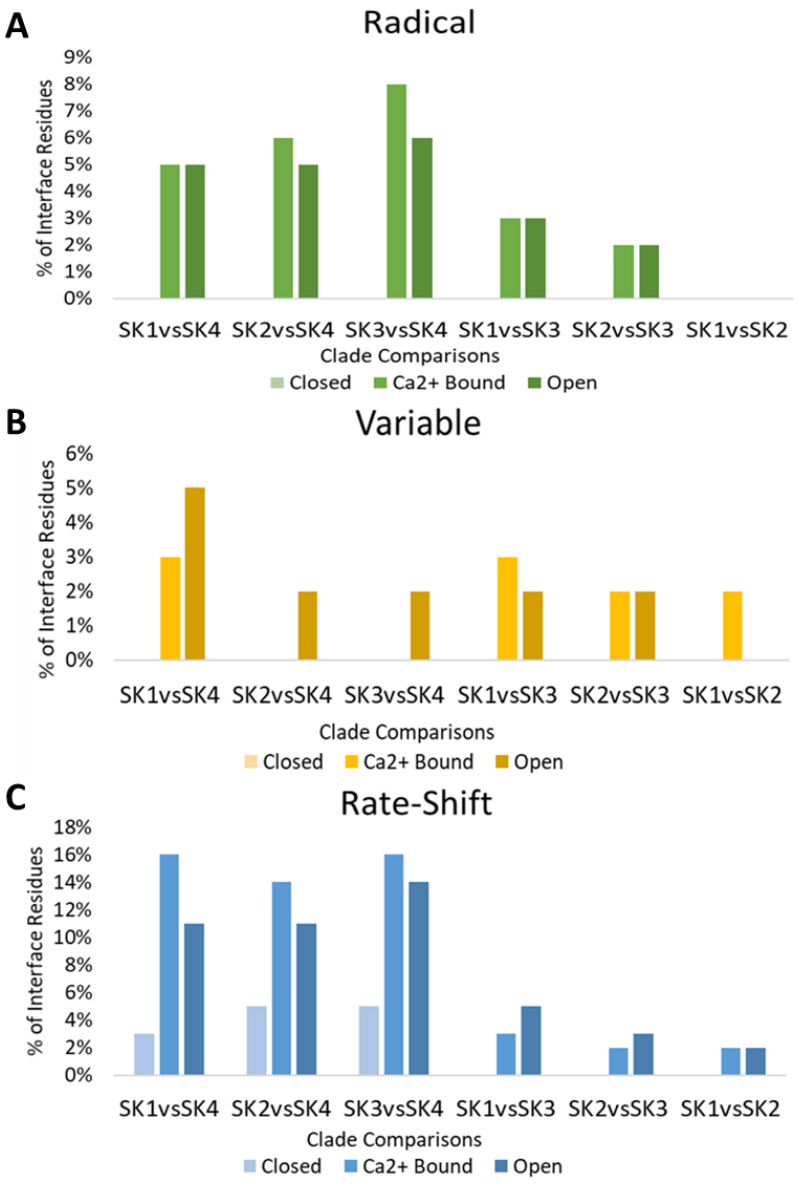

Figure 7. Percentage of interface sites between the channel and calmodulin with radical amino acid substitutions (A), variable sites (B), or rate shifts (C) in the closed, calcium bound, and open conformations for each clade comparison.

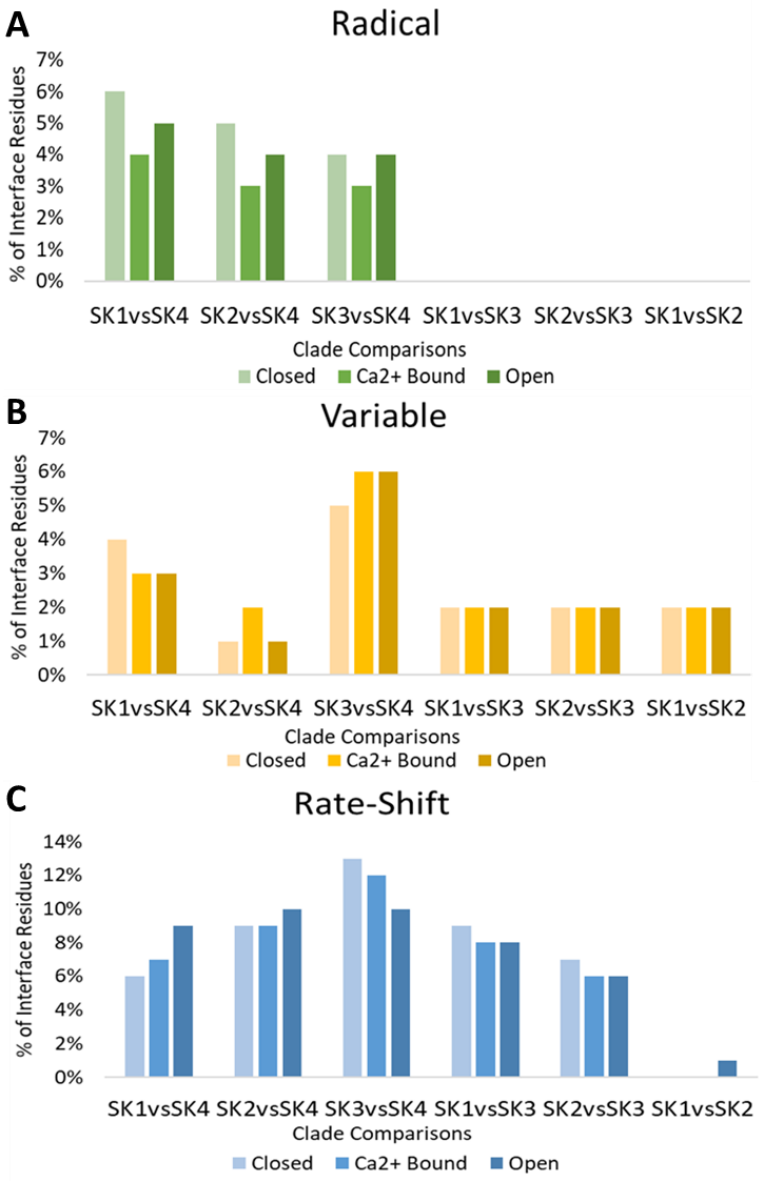

Figure 8. Percentage of interface sites amongst the channel chains with radical amino acid substitutions (A), variable sites (B), or rate shifts (C) in the closed, calcium bound, and open conformations.

Table 1. Tests of significance

p-values found for the comparison between the closed conformation and the other two conformations for the calmodulin-to-channel interface residues and channel to channel interface residues. Mann-Whitney U test; a *p*-value of 0.0083 or less was statistically significant.

	Calmodulin-to-channel Interface Residues		Channel-to-channel Interface Residues	
	6CNM vs 6CNN	6CNM vs 6CNO	6CNM vs 6CNN	6CNM vs 6CNO
Radical	0.00481	0.00473	0.274	0.398
Variable	0.01383	0.00406	0.394	0.466
Rate-Shift	0.07191	0.07191	0.436	0.436

Visualization of Interface Divergence

To better identify the areas where divergence is found, the diverging sites were mapped onto the 3D structures of the inactive, closed first conformation and the active, open conformation (Figure 9). The direction of potassium efflux is shown through the channel for the open and active conformation 3 (Figure 9B), while potassium efflux is prevented when the channel is closed and inactive in conformation 1 (Figure 9A). The view of these conformations from the cytosol was also added to visualize the changes that occur as the channel is activated (Figure 9C-9D). The location of the radical sites, variable sites, and rate shifts indicate that the majority of all diverging sites are found on the inside of the channel. Further, most of the radical sites at the calmodulin-to-channel interface are found at the N-terminus of the SK channel. Additionally, many radical amino acid substitutions are found within the channel pore, on the cytosol facing side (Figure 9). Of particular importance is the binding of the flexible N-lobe of calmodulin to the channel in order to facilitate the opening of the pore (Figure 9B).

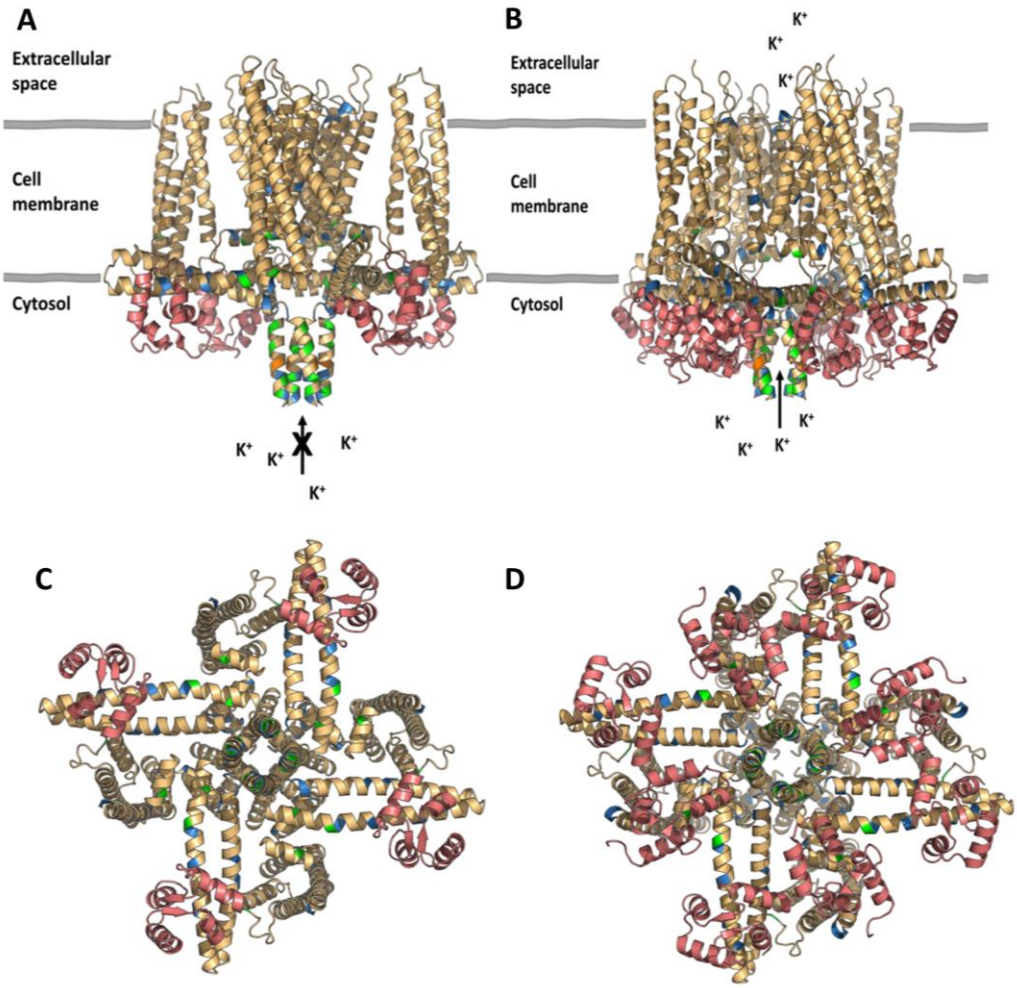

Figure 9. Visualization of radical amino acid substitutions (green), variable sites (orange) or rate shifts (blue). The interfaces between both the channel subunits and channel–calmodulin were included. B and D depicts the active third conformation (6CNO) from a side view and from the cytoplasm, respectively. A and C depicts the inactive first conformation (6CNM) from a side view and from the cytoplasm, respectively. The chains in pink are calmodulin. The chains in yellow are the SK channel. The direction of potassium efflux is also depicted in A and B.

DISCUSSION

The phylogenetic trees of both calmodulin and the SK channel provide evolutionary context to the study and supported previously reported information on these proteins (Chin & Means, 2000; Adelman, 2012). For calmodulin, the phylogenetic tree constructed from amino acid sequences supported the high conservation level of this protein across all vertebrate species (Chin & Means, 2000, Townsend et al., 2008). This may be due to calmodulin having a wide variety of specialized functions ranging from modulating the activities of enzymes, calcium transport, and binding to several structural proteins (Friedberg & Rhoads, 2001). In order to perform all these roles, calmodulin experiences strong selective pressure. However, although the three copies are identical, they have different concentrations in different tissues (Friedberg & Rhoads, 2001). This suggests that there must be distinct regions at the DNA level for each calmodulin copy, which allows for differential regulation. This was confirmed by the finding of three distinct clades and is supported by prior research (Friedberg & Rhoads, 2001). From the SK channel phylogeny, it was apparent that the SK4 clade was evolving at a faster rate than the other three copies of the channel indicated by the longer branch lengths (Figure 3). This suggests that there may be functional differences in this clade compared with the other clades. To determine if this difference involved the activation mechanism, a more in-depth analysis of the activation mechanism was performed.

The established activation mechanism described by Lee and MacKinnon (2018) for the SK channel involves the first conformation, 6CNM, which occurs when the channel is closed and only the C-lobe of calmodulin is bound. In the second

conformation, 6CNN, calcium binds the N-lobe of calmodulin that attaches onto the channel at the calmodulin binding domain. Finally, in the third conformation, 6CNO, the channel is opened and fully active.

Here, I analyzed the evolutionary conservation of all residues located at the different interfaces in the different activation conformations. After dividing the interface residues into those between calmodulin and the channel and within the channel itself, the percent identity of the interface residues was compared for channel-to-calmodulin and channel-to-channel, respectively (Figure 6). For sequences within the same clade, high similarity was seen. This aligns with prior findings that the different SK channels are highly similar. For example, human SK2 is found to be about 98% similar to mouse SK2 (Saqar, 2014). However, the interface for human SK2 and mouse SK2 are 100% conserved, implying the importance of these residues. When comparing across clades, the similarity percentage decreases and sharply declines when comparing SK1, 2, or 3 to SK4. Between the three first SK channels, they share approximately 70-75% identity, whereas SK4 shares only approximately 40% identity with the other three channels (Fanger et al., 1999; Begenisich et al., 2004; Saqar, 2014). However, in comparing the interface residues for SK4, the percentage identity does not fall below 49% (Figure 6). The matrix showed that the interface residues that are found within the channel were more conserved than those from the channel to calmodulin (Figure 6).

For a more detailed look into the divergence in the different clades, the type of divergence at the interface sites was investigated. This showed that from the inactive to active conformation, there is an increase in the percentage of radical sites, variable shifts,

and rate shifts (Figure 7). This increase in radical sites is due to most of them being found at the beginning of the SK channel sequence (N-terminus), which only encounters calmodulin during the first bound state. This increase was found to be statistically significant, implying that the interface residues in the first closed state are more conserved across all clades and sequences than the interface residues in the other two conformations. Consequently, the interface between the N-lobe of calmodulin and the SK channel has more radical amino acid substitutions amongst the clades. This is consistent with prior research which has found that although the SK channel subunits are highly similar across paralogs, they vary at the extreme N- and C-terminal domains, which encompasses where the N-lobe binds (Lee et al., 2003; Maylie et al., 2004; Saqar 2014). For the rate-shifting sites, the percentages were not significant between the conformations (Table 1). This is probably due to a small sample-size error, as there is still an increase in percentages from the first conformation to the second. The comparisons between SK4 and the other paralogs exhibited a pronounced increase in rate-shifting sites found from 3% for SK1 vs SK4, 5% for SK2 vs SK4, and 5% for SK3 vs SK4 in 6CNM to 16%, 14%, and 16%, respectively, in 6CNN (Figure 7). However, when comparing between the other three clades, the rate did not show this pronounced increase. As a result, the total results were not significant, although some portions may have been.

As for the other set of interface residues, which are those found within the SK channel, there is no statistically significant increase in the percentage of diverging sites (Table 1). This further supports the idea that the area that binds the calmodulin N-lobe, which is more variable across clades (Figure 7), should be targeted. This is noteworthy as the calmodulin sequence is not changing across vertebrate species.

The binding interface of the N-lobe at the calmodulin binding domain has been targeted in SK2 channels, as seen with current medications such as Riluzole, which is the only FDA-approved drug used to treat amyotrophic lateral sclerosis (Cui et al., 2014). This pharmaceutical functions as a positive modulator of the SK channel and increases channel sensitivity for calcium. More specifically, the medication binds to a calmodulin binding pocket at channel amino acid residues A477 and L480 (Cui et al., 2014; Nam et al., 2017; Orfali, 2018). This binding pocket is at the interface between the calmodulin binding domain and the calmodulin N-lobe. As a result, the firing speed of the affected neurons are regulated, showing positive outcomes (Zhang et al., 2012; Cui et al., 2014; Nam et al., 2017;). Amino acid residue A477 was found to be an interface residue for the first bound conformation that is conserved amongst clade SK1-3, but undergoes a substitution for V in SK4. Four other compounds also bind to this same site, where amino acids A477 within the binding pocket forms electrostatic interactions with M51 and K75 on calmodulin (Nam et al., 2017). Despite this, these modulators bind with varying potency, which reflects continuing need for SK channel modulators (Pedarzani, 2008; Cui et al., 2014). Since the alanine is conserved across the first three SK clades, targeting this residue would target them all. Lee & MacKinnon (2018) found that the N-lobe of calmodulin also binds to SK channel residues 176-187. All these residues were identified as both channel-to-channel interfaces or calmodulin-to-channel interfaces. Although these sites are conserved for SK1-3, they are variable across SK4, indicating a rate shift. When the calmodulin binding domain interfaces are compared to *Homo sapiens* SK4, there are three residues substitutions. Targeting multiple residues in this area may potentially increase the potency and binding strength of blockers.

The discovery that each conformation had presented various differences in the interface residues suggests the potential to inhibit these channels in a more specific manner than the current channel blockers. Interface residues with amino acids that are conserved in one clade but are variable or have a radically different amino acid in the others can bind drugs in a clade-specific manner. It should also be noted, that calmodulin binds to the SK channel on the cytosolic side, where the majority of the radical sites were found (Figure 9). Drugs being formulated must be able to cross the membrane in order to get into the cell. This has been taken into account previously, as seen with the modulator TRAM-30 and TRAM-34, which functions when bound to the cytoplasmic side (Wulff et al., 2001; Nam et al., 2017).

Drugs that target only SK4 appear to be the easiest ones to create, as this paralog is the most different. This channel presents the possibility for the most targets, as there are more radical sites than when comparing among SK1-SK3. This is probably the reason why there are already drugs being created that target only this clade (Nam et al., 2017). One of these is Senicapoc, which is used to prevent erythrocyte dehydration in patients with sickle cell disease (Ki et al., 2008; Tubman et al., 2015). It also can be used to treat hereditary xerocytosis (Rapetti-Mauss et al., 2016). These patients have abnormally active SK4 channels within their erythrocytes. As a result, this drug, which binds to the cytoplasmic ends of the pore close to V282, decreases the function of this channel. This residue occurs at a channel-to-channel interface that is conserved across all clades. In addition, all four chains are in contact at this amino acid. This is likely why it may be essential in SK4 gating (Lee & MacKinnon, 2018). There are sites with radical amino acid substitutions between the other clades as well and these could be explored as targets

for drugs aimed at either SK1, 2, or 3. The variable interfaces found from the clade comparisons do not provide specificity, but they indicate that these sites are less important for functions since they can vary. It also showed that there was high variation in the amino acids at that location within the clade, as they were only noted as variable if each clade had more than one type of amino acid. However, when developing a drug for humans, these sites may work if considering only human sequences, given that there is no residue variation among human populations.

If successful, the benefits to more clade-specific medications are numerous. Each of the potassium channel paralogs are found in varying concentrations in different parts of the body. For example, SK2 has been reported to exist and function in human cardiac myocytes (Tuteja et al., 2005). Therefore, its activity and regulation could affect the functionality of the heart and may improve function in cardiovascular disorders, especially those that pertain to arrhythmias or are caused by irregular excitability (Gu et al., 2018). Hypertension is associated with an overproduction of aldosterone, which has been seen to decrease using an SK channel agonist, 1-ethyl-2-benzimidazolinone (1-EBIO) (Gu et al., 2018). However, this is a non-discriminant agonist which also interacts at various levels with SK1-3 (Pedarzani et al., 2001). On the other hand, SK3 is expressed in endometrial cells and regulates cell migration; it also has implications in unsuccessful pregnancy outcomes (Lu et al., 2014). In addition, a better understanding of the binding mechanism of clade-specific drugs will improve the effectiveness and reduce side effects of medications that unintendedly block these SK channel paralogs. One of these is the antidepressant fluoxetine (Prozac) which is a selective serotonin reuptake inhibitor which also blocks SK1, SK2, and SK3 (Terstappen et al., 2003). While SK1 is

mainly found in the brain, SK2 and SK3 are more widely spread throughout the body (Rimini et al., 2000). This may be a contributing factor in the numerous side effects of these medications.

In summary, greater conservation within the channel suggests that the evolutionary mechanism in which the N-lobe of calmodulin binds to the channel may be unique to each clade. Once it is bound, the rest of the activation is more conserved across all clades. When looking at both matrices, all sequences being compared to those in the SK4 clade presented lower identity percentages compared with the rest of the comparisons in the matrix. This suggests that SK4 is evolving the fastest of all the SK family paralogs, which is also seen with SK4 sequences having the longest branch lengths in phylogenetic tree (Figure 3). The ability to target the calmodulin binding domain so that calmodulin cannot interact with and activate the channel must be further considered in pharmaceutical studies.

This research provides insights into the divergence of the SK channels and in particular the activation mechanism that can be examined in a molecular biology experimental setting. There is still a need for more information in order to further investigate possible drug target sites. A CryoEM structure of at least one of the three other paralogs should be investigated in order to confirm if the activation mechanism is actually the same amongst them. More mammalian species can be examined to improve the conservation analysis. The SK channel has implications in numerous disorders and understanding the evolution of the SK channel activation mechanism. Further, the work presented here also shows how protein evolution can contribute valuable insights to

potential translational applications. For the SK channels, this study offers a means to target one channel specifically rather than several generally.

LITERATURE CITED

Adelman JP, Maylie J, Sah P. 2012. Small-conductance Ca^{2+}-activated K^{+} channels: Form and function. Annual Review of Physiology 74:245–269 doi: 10.1146/annurev-physiol-020911-153336.

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations Nucleic Acids Res 38: W7-W13.

Ahrens J, Gomes Dos Santos H, Siltberg-Liberles J. 2016. The nuanced interplay of intrinsic disorder and other structural properties driving protein evolution. Molecular Biology and Evolution: 33: 2248-2256 doi: 10.1093/molbev/msw092.

Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19: 716-723 doi: 10.1109/TAC.1974.1100705

Altschul SF, Gish W, Miller W, et al. 1990. Basic local alignment search tool. Journal of Molecular Biology 245: 403–410

Anisimova M, Gil M, Dufayard JF et al. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Systematic Biology 60: 685-699 doi: 10.1093/sysbio/syr041

Begenisich T, Nakamoto T, Ovitt C. et al. 2004. Physiological roles of the intermediate conductance, Ca^{2+} -activated potassium channel Kcnn4. Journal of Biological Chemistry 279: 47681-47687. doi:10.1074/jbc.m409627200

Berridge MJ. 2014. Ion channels. Cell Signalling Biology 6:1-74 doi:10.1042/csb0001003

Bonferroni, CE. 1936 Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8: 3-62.

Chin D, Means, AR. 2000. Calmodulin: A prototypical calcium sensor. Trends in Cell Biology 10:322-328 doi:10.1016/s0962-8924(00)01800-6

Chojnacki S, Cowley A, Lee J, Foix A, Lopez R. 2017. Programmatic access to bionformatics tools from EMBL-EBI update: 2017. Nucleic Acid Research 45: W550-W553 doi: 10.1093/nar/gkx273

Cui M, Qin G, Yu K, et al. 2014. Targeting the small- and intermediate-conductance ca-activated potassium channels: the drug-binding pocket at the channel/calmodulin interface. Neuro-Signals 22: 65-78. doi: 10.1159/000367896

Deister CA, Chan CS, Surmeier, D. J. & Wilson, C. J. 2009. Calcium-activated SK channels influence voltage-gated ion channels to determine the precision of firing in *Globus pallidus* neurons. Journal of Neuroscience 29, 8452–8461, doi:10.1523/JNEUROSCI.0576-09.2009

Dosztányi Z, Csizmok V, Tompa P, Simon I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21: 3433–3434.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32: 1792–17 97. doi:10.1093/nar/gkh340.

Fanger CM., Ghanshani S, Logsdon NJ et al. 1999. Calmodulin mediates calcium-dependent activation of the intermediate conductance KCa channel, IKCa1. Journal of Biological Chemistry 274: 5746-5754. doi:10.1074/jbc.274.9.5746

Friedberg F, Rhoads A. 2001. Evolutionary aspects of calmodulin. Life 51:2015-221.

Gasteiger E, Gattiker A, Hoogland C, et al. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis Nucleic Acids Res. 31:3784-3788.

Gu M, Zhu Y, Yin X, et al. 2018. Small-conductance Ca^{2+}-activated K^{+} channels: insights into their roles in cardiovascular disease. Experimental and Molecular Medicine 50. DOI: https://doi.org/10.1038/s12276-018-0043-z

Gu X. 2006. A simple statistical method for estimating type-II (cluster-specific) functional divergence of protein sequences. 2006. Molecular Biology and Evolution 23: 1937-1945.

Guindon S, Dufayard JF, Lefort V, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Systematic Biology 59: 307-321

Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Computer Applied Bioscience 8: 275–282.

Jones S, Thornton JN. 1997. Analysis of protein-protein interaction sites using surface patches. Journal of Molecular Biology 272: 121-132.

Ki A, Smith WR, De Castro LM, et al. 2008. Efficacy and safety of the Gardos channel blocker, senicapoc (ICA-17043), in patients with sickle cell anemia. Blood 111: 3991-3997. doi: 10.1182/blood-2007-08-110098.

Ledoux J, Werner M, Brayden J, Nelson MT. 2006. Calcium-activated potassium channels and the regulation of vascular tone. Physiology 2: 69-78. doi: 10.1152/physiol.00040.2005

Lee C, MacKinnon R. 2018. Activation mechanism of a human SK-calmodulin channel complex elucidated by cryo-EM structures. Science 360: 508-513 doi:10.1126/science.aas9466

Lee WS, Ngo-Anh TJ, Bruening-Wright A, et al. 2003. Small conductance Ca^{2+}-activated K^+ channels and calmodulin. The Journal of Biological Chemistry 278: 25940-25946. doi: 10.1074/jbc.M302091200

Li W, Aldrich R. 2009. Activation of the SK potassium channel-calmodulin complex by nanomolar concentrations of terbium. Proceedings of the National Academy of Sciences 106:1075-1080. https://doi.org/10.1073/pnas.0812008106

Liu X, Wang G, Chen S. 2010. Modulation of the activity of dopaminergic neurons by SK channels: A potential target for the treatment of Parkinson's disease? Neuroscience Bulletin 26: 265-271 doi:10.1007/s12264-010-1217-4

Lu YC, Yang J, Ding GL et al. 2014. Small-conductance, calcium-activated potassium channel 3 (SK3) is a modulator of endometrial remodeling during endometrial growth. Journal of Clinical Endocrinology and Metabolism 99: 3800-3810 doi: 10.1210/jc.2013-3389.

Mann HB, Whitney DR. 1947. On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics 18:50-60

Maylie J, Bond CT, Herson, et al. 2003. Small conductance Ca^{2+}-activated K^+ channels and calmodulin. The Journal of Physiology 554: 255-261 doi: 10.1113/jphysiol.2003.049072

Nam Y, Orfali R, Liu T, et al. 2017. Structural insights into the potency of SK channel positive modulators. Scientific Reports 7: 1-10 doi:10.1038/s41598-017-16607-8

Oliván-Viguera A, Valero MS, Murillo M, et al. 2013. Novel phenolic inhibitors of small/intermediate-conductance Ca^{2+}-activated K^+ channels, KCa3.1 and KCa2.3. PLOS One 8: e58614 doi: 10.1371/journal.pone.0058614

Orfali R. 2018. SK channel modulators as drug candidates and pharmacological tools. Pharmaceutical Sciences (MS) Theses – Chapman University. 1.

Patlak J. 1991. Molecular kinetics of voltage-dependent Na+ channels. Physiological Reviews 71: 1047–1080.

Pedarzani P, Mosbacher J, Rivard A, et al. 2001. Control of electrical activity in central Neurons by modulating the gating of small conductance Ca^{2+}-activated K^+ channels. Journal of Biological Chemistry 276: 9762-9769. DOI: 10.1074/jbc.M010001200

Rapetti-Mauss R, Soriani O, Vinti H, et al. 2016. Senicapoc: a potent candidate for the treatment of a subset of hereditary xerocytosis caused by mutations in the Gardos channel. Haematologica, 101: e431–e435. doi:10.3324/haematol.2016.149104

Rimini R, Rimland JM, Terstappen GC. 2000. Quantitative expression analysis of the small conductance calcium-activated potassium channels, SK1, SK2 and SK3, in human brain. Molecular Brain Research 85:218-220 https://doi.org/10.1016/S0169-328X(00)00255-2.

Ronquist F, Teslenko M, van der Mark P, et al. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542

Saqar W. 2014. Characterization of small conductance calcium-activated potassium channels in a human lens epithelium cell line (B3*)*. B. Medicine and Surgery Wright State University:1-50.

Stackman RW, Hammond RS, Linardatos E, et al. 2002. Small conductance Ca^{2+}-activated K^+ channels modulate synaptic plasticity and memory encoding. The Journal of Neuroscience 22: 10163-10171 doi:10.1523/jneurosci.22-23-10163.2002

Tuteja D, Xu D, Timofeyev V et al. 2005. Differential expression of small-conductance Ca^{2+}-activated K^+ channels SK1, SK2, and SK3 in mouse atrial and ventricular myocytes. American Journal of Physiology-Heart and Circulatory Physiology 289:H2714-H2723.

Terstappen GC, Pellacani A, Aldegheri L et al. 2003. The antidepressant fluoxetine blocks the human small conductance calcium-activated potassium channels SK1, SK2 and SK3. Neuroscience Letters 346:85-88.

Tubman VN, Mejia P, Shmukler BE et al. 2015. The clinically tested gardos channel inhibitor senicapoc exhibits antimalarial activity. Antimicrobial Agents and Chemoteherapy 60: 613-616. DOI: 10.1128/AAC.01668-15

Townsend JP, Lopez-Giraldez F, Friedman R. 2008. The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. Journal of Molecular Evolution 67: 437-447. DOI: 10.1007/s00239-008-9142-0

Weatherall KL, Goodchild SJ, Jane DE, Marrion NV. 2010. Small conductance calcium-activated potassium channels: from structure to function. Progress in Neurobiology 91: 242-255 doi:10.1016/j.pneurobio.2010.03.002

Westerlund AM, Delemotte L. 2018. Effect of Ca^{2+} on the promiscuous target-protein binding of calmodulin. PLoS Comput Biol 14: e1006072. https://doi.org/10.1371/journal.pcbi.1006072

Wulff H, Gutman GA, Cahalan MD, Chandy KG. 2001. Delineation of the Clotrimazole/TRAM-34 binding site on the intermediate conductance calcium-activated potassium channel, IKCa1. Journal of Biological Chemistry 276: 32040–32045. doi: 10.1074/jbc.M105231200.

APPENDICES

APPENDIX A - Sequence Names and Accession Numbers - Calmodulin Protein Family

Sequence Name	Accession Number
XT_CALM1	XP_002932424.1
XT_CALM2	NP_001008160.1
AM_CALM	XP_002920827.1
MD_CALM1	XP_001371508.2
OA_CALM1	XP_001506524.1
AC_CALM	XP_008120909.1
AC_CALM2	XP_003217527.1
AC_CALM1	XP_008112670.1
LC_CALMX2	XP_006003477.1
LC_CALMX3	XP_005992764.1
LC_CALM	XP_006006588.1
LC_CALMX1	XP_005988862.1
MD_CALM2	XP_001375537.1
AC_CALM3	XP_016853235.1
AM_CALM2	G1LHZ6
OA_CALM	XP_001514069.4
MM_CALM1	P62204
MM_CALM2	NP_031615.1
MM_CALM3	NP_031616.1
DR_CALM1	Q6PI52
DR_CALM3a	AAH71404.1
AM_CALM3	XP_002923096.1
HS_CALM2	P0DP24
EC_CALM2	NP_001270011.1
DR_CALM1b	NP_956376.1
DR_CALM3b	AAH44434.1
TN_UP6	H3CQN4
GG_CALM1	NP_001103834.1
DR_CALM2a	NP_956290.1
MD_CALM3	F7CY56
EC_CALM3	XP_001500896.3
HS_CALM3	P0DP25
EC_CALM1	XP_023484032.1
DR_CALM2b	AAH54600.1
HS_CALM1	P62158
GG_CALM2	NP_990336.1

APPENDIX B - Sequence Names and Accession Numbers – Calmodulin Nucleotide Family

Sequence Name	Accession Number
GG_CALM1	NM_001110364.1
AC_CALM1	XM_008114463.2
MD_CALM1	XM_001371471.4
OA_CALM1	XM_001506474.2
HS_CALM1	BT006818.1
AM_CALM1	XM_002920781.3
XT_CALM2	NM_001008159.1
EC_CALM1	XM_023628264.1
MM_CALM1	AK151001.1
MM_CALM2	NM_007589.5
AM_CALM2	XM_011216806.2
EC_CALM2	NM_001283082.1
HS_CALM2	BT009916.1
MD_CALM2	XM_001375500.3
OA_CALM	XM_001514019.4
GG_CALM2	NM_205005.1
AC_CALM2	XM_003217479.2
XT_CALM1	XM_002932378.4
LC_CALMX2	XM_006003415.2
DR_CALM2b	NM_214736.1
DR_CALM2a	NM_199996.2
LC_CALMX3	XM_005992702.2
MD_CALM3	XM_007491977.2
EC_CALM3	XM_001500846.6
AM_CALM3	XM_002923050.3
HS_CALM3	BT006855.1
MM_CALM3	NM_007590.3
AC_CALM	XM_008122702.2
DR_CALM1a	BC097062.1
TN_UP6	CR715454.1
DR_CALM1b	NM_200082.1
DR_CALM3a	NM_182967.1
DR_CALM3b	NM_199570.1
AC_CALM3	XM_016997746.1
LC_CALMX1	XM_005988800.2
LC_CALM	XM_006006526.2

APPENDIX C - Sequence Names and Accession Numbers - SK Channel Protein Family

Sequence Name	Accession Number
MM_SK1	Q9EQR3
HS_SK1	Q92952
EC_SK1	F6Z4K2
AM_SK1	G1LWT0
MD_SK1	F7BJ86
GG_UP1	A0A1D5PV31
DR_SK1a	F1QND4
DR_SK1b	Q1LY44
XT_SK1	F6SM41
LC_SK1	H3AWP1
AM_SK2	G1L1N3
HS_SK2	Q9H2S1
EC_SK2	F7CKJ4
MM_SK2	P58390
MD_SK2	F7E0I4
AC_SK2	R4GCW1
GG_SK2	Q9PTS9
LC_UP1	H3AAN8
DR_UP5	A0A0R4IC13
AM_SK3	D2HJN4
EC_SK3	F7AWG9
MM_SK3	P58391
HS_SK3	Q9UGI6
LC_SK3	H3B7A2
MD_SK3	F6QE12
GG_UP2	A0A1D5PD82
DR_SK3	E7FAT2
HS_SK4	O15554
AM_SK4	G1LC04
MM_SK4	O89109
AC_SK4	H9G9A0
XT_SK4	F6TJW3
DR_SK4	X1WCL6