Mid-Term Protein Bioinformatics

AS.410.639.81.SU20

Brian Wiley

Voltage-dependent P/Q-type calcium channel subunit alpha-1A (Cavα1)

Assigned Mutation or Natural Variant: A712T

BACKGROUND

Voltage-gated ion channels are transmembrane (TM) proteins that facilitate the transport of cation and anions into and out of the cell. The primary role of these voltage-gated ion channels is to propagate signaling in nerve cells to relay messages to various tissues and parts of the body (Alberts et al., 2014). They react to electrical potential, changes in membrane potential and switch from an inactive conformation to an active one that allows the influx or efflux of specific ions. Voltage-gated ion channels selectively permeable to Na^+, K^+, Ca^{2+}, and Cl^- have been discovered (Purves et al., 2001). The goal of this paper is to use protein bioinformatics techniques to learn about a particular subunit of the family of voltage-gated Ca^{2+} ion channels as well as to identify the role of a particular variant that has been associated with neurological disorders including but not limited to a variety of migraine headaches, ataxia, and epileptic encephalopathies (Campbell, Midterm Instructions, 2020; Simms & Zamponi, 2014).

Voltage-gated Ca^{2+} channels are especially important for mediating that actional potential propagated to the axon terminals of motor neurons and the release of neurotransmitters from the presynaptic motor neuron to serve as ligands for neurotransmitter receptors on the surface of muscle cells (Lodish et al., 2016; Few et al., 2005) or post synaptic dendrite of neighboring neurons. In doing so they are important in converting electrical signals into chemical signals that relay information in the nervous system and cells in the tissues that receive these signals. When neurons are in their resting state, extracellular (EC) calcium levels remain high and intracellular levels remain low; around 50-100 nM inside the cell (Grienberger & Konnerth, 2012; Simms & Zamponi, 2014). When propagated by depolarization at the end the synaptic terminal they open to allow the influx of Ca^{2+} ions inducing the release of neurotransmitters (Simms & Zamponi, 2014) which then cascades a variety of cellular processes including release of Ca^{2+} from the sarcoplasmic reticulum (SR) in muscle cells. The pore with which the calcium enters the ends of neuronal synapses is the alpha subunit of the voltage-gated Ca^{2+} channel complex known as Cavα1 (UniProt, 2020). There are two types of voltage-gated Ca^{2+} channels: high voltage-activated (HVA) channels and low voltage-activated (LVA) channels. In HVA channels Cavα1 associates as a heteromeric complex with $Cav\alpha_2\delta$ and Cavβ and possibly Cavγ subunits while LVA channels lack these additional subunit associations (Simms & Zamponi, 2014). There are three families of Cavα1 subunits in voltage-gated Ca^{2+} channels: Cav1, Cav2, and Cav3. The particular alpha subunit that is in the interest of this paper is in the second family Cav2 and is known as Cav2.1. Cav2.1 gives rise to P/Q type channels while the Cav1 family gives rise to L-type channels, Cav2.2 to N-type, Cav2.3 to R-type, and the family of Cav3 to T-type channels. We will refer to the alpha-1A subunit also known as “isoform alpha-1A” (UniProt, 2020) as Cav2.1 for the remainder of this paper.

THE ALPHA1 SUBUNIT AND CAV2.1

The Cavα1 subunit has the same basic form in all three families of Cavα1. It is a TM protein with four conserved TM spanning domains. Cav1 and Cav2 subfamily isoforms are all HVA channels (Simms & Zamponi, 2014; Andrade et al., 2019) and as such Cav2.1 associates with

Cav$\alpha_2\delta$ subunits which are post-translationally cleaved and then re-linked by disulfide bridge (Zhang et al., 2018) and Cavβ subunit (Figure 1). The Cav2.1 gene (CACNA1A) codes for many different spliced variant isoforms; 6 isoforms of Cav2.1 are listed by UniProtKB for accession ID O00555 while Ensembl lists 60 different spliced variants, 22 of them being protein coding (Ensembl, 2020). UniProtKB actually indicates there are 22 additional computationally mapped isoforms in which they map from Ensembl. The molecular mass of the canonical full-length isoform is around 282.5 kDa (UniProt 2020, Ensembl 2020). The length of this isoform listed in UniProt under Isoform 1 identifier O00555-8 is 2506 residues under accession ID O00555 however in UniProt under accession A0A1B0GTN7 it is listed at 2507 residues in length which includes an inserted glycine residue at G419 compared to accession O00555. This is important to reference that despite the assigned variant Ala712Thr, the SNP variant ID rs886037945 under National Center for Biotechnology Information (NCBI) documents this variant as p.Ala713Thr and most publications refer to this numbering of the variant from NCBI. We will be referencing the variant in accordance with NCBI as p.Ala713Thr (A713T) as well as the accession A0A1B0GTN7 sequence of 2507 residues for the remainder of this paper unless otherwise noted.

Cav2.1 expression profile indicates dominant expression in Purkinje neurons, glutamatergic neurons, spinal cord neurons and in the ganglia (Andrade et al., 2019). Diseases associated with Cav2.1 variant include Spinocerebellar ataxia 6, Migraine, familial hemiplegic, 1, Episodic ataxia 2, and in the case of variant A713, Epileptic encephalopathy, early infantile, 42 (EIEE42) (UniProt, 2020). One mouse study by Mallmann et al. indicates that complete ablation of Cav2.1 produces mice "with progressive neurological deficits" resulting in ataxia and dystonia and only survive up to 4 weeks (Mallmann et al., 2013). Two papers in reference to the variant A713T were both authored by various authors in conjunction with the EPI4k Consortium, an international group of scientists which characterize complex epilepsies. The first paper written in 2013 mapped the A713T mutation to the H19 human genome on chromosome 19 starting at position 13414398 with a C>T mutation on the non-coding strand that transcribes a GCU>ACU alanine to threonine mutation at protein position 713 and was linked to previously associated disorders of familial hemiplegic migraine and ataxia (Allen et al., 2013). UniProt and Ensembl both list the associated disease with variant A713T being Epileptic encephalopathy, early infantile, 42 (EIEE42) however the annotation source by ClinVar is no longer valid and the latest paper published by Epi4k in 2016 linked this variant to Early-Onset Epileptic Encephalopathy (EOEE) (Epi4K Consortium, 2016). Regardless both classifications present with epileptic seizures during the early infantile period and lead to severe neurological and motor development deficits.

A CLOSER LOOK AT STRUCTURE OF CAV2.1

As previously indicated all isoforms of Cavα1, in which there are 10 (Cav1.1, Cav1.2, Cav1.3, Cav1.4, Cav2.1, Cav2.2, Cav2.3, Cav3.1, Cav3.2, Cav3.3), are TM proteins of the plasma membrane in neurons and other cells. There are four TM spanning domains each consisting of six membrane spanning alpha helices (Simms & Zamponi, 2014). In each of these membrane

spanning repeat domains five of the six alpha helices indicated as S1-S6 are hydrophobic while the fourth segment, S4, in each of the four domains is marked by positively charged lysine or arginine every third residue and mostly likely represents the voltage sensor (UniProt, 2020). For instance, S4 of repeat domain 1 contains the residues LRTLRAVRVLRPLKLVSGI where each positive residue is underlined. The S5 and S6 helices of the TM domains form the central pore across the plasma membrane, "S6 lining the inner face" of the pore, where the calcium ions flow through from the EC matrix into the cytoplasm (Figure 2) (Cui, 2010). The EC linker between the S5 and S6 helices in each of TM repeats contains a characteristic P loop motif that reenters back into the plasma membrane and back out to the EC region and contain "highly conserved negatively charged amino acid residues," specifically glutamic acid residues in HVA channels (Figure 3) (Simms & Zamponi, 2014). For instance, the first P loop between the S5 and S6 of domain repeat III, residues 1398–1484, one would expect on average 4.3 of each amino acid and this sequence contains 9 glutamic acid residues. The negatively charge amino acids coordinate the pore of the channel formed by the S5 and S6 helices to accommodate the positively charged calcium cations (Figures 2 and 3). There are three cytoplasmic linker segments between the four TM repeat domains. The linker between domain repeat I-II is 127 residues long, between repeat II-III is the longest cytoplasmic linker segment at 528 residues long, and the shortest III-IV linker is 64 residues long. There is a long C-terminal cytoplasmic segment that is 706 residues long and includes a preIQ-IQ and IQ motif that are specific to Cav1 and Cav2 channels (Simms & Zamponi, 2014).

There has not been a significant portion of the 3-dimensional (3-D) structure of Cav2.1 currently completed. Two studies using X-ray diffraction have shown the interaction of the IQ domain with Ca^{2+} bound calmodulin. In each of the studies the length of the IQ domain in the structures were only 21 residues in Rat and 22 residues in European Rabbit and Human however only the last 20 of the 22 residues match the sequence in UniProt (Figure S1) (UniProt, 2020). A more in depth analysis of the IQ domain will be covered in later discussion in MOTIFS AND PROFILES section as well as SUBFAMILY ALIGNMENTS section. According to the crystal structure performed by depositors, this 21 residue IQ segment of Cav2.1 is an alpha helix structure (Figure 4) up until and including an alanine residue but not including next two lysine residues (Figure S1) (Mori et al., 2008; Kim et al., 2008). This is interesting, however, since lysine has the fifth highest helix propensity of the 20 amino acids (Pace & Scholtz, 1999).

Cav2.1 has twenty low complexity spanning subsequences (Pfam 33.1, El-Gebali et al., 2019). Using the complexity formula, equation (3) and equation (4) of Wootton & Federhen, the total complexity of the 2507 length Cav2.1 subunit is 4.204 while the complexity of all the low complexity sequences, ranging in length 9 to 33 and totaling 353 residues (~14%), range in complexity from 1.577 to 2.695.

As Cav2.1 is a TM protein it is natural to want to measure the hydrophobicity of the TM spanning helix domains and compare the TM helices to each other as well as to the ectodomains and cytoplasmic domains. Since the S5 and S6 helices form the central pore of the channel for transport of positively charged calcium ions, I hypothesize that similar to the pore of aquaporins, these helices (S5 and S6) will have a lower measurement of hydrophobicity compared to the

other four channels (S1-S4) in order to coordinate charged cations to enter the cell. To measure hydrophobicity of the 2507 residue Cav2.1 subunit I have utilized ProtScale from the ExPASy Bioinformatics Resource Portal (ExPASy) which allows you compute various statistics in addition to hydrophobicity profiles, including polarity, secondary structure α-helices and β-strands/sheets, and other physiochemical properties (Gasteiger et al., 2005). The results (Figure 5) using the Kyte and Doolittle scale (Kyte & Doolittle, 1982) with sliding window of 9 residues showed that all of the S1-S6 helical regions of all TM residues annotated by UniProt have higher hydrophobicity scores indicative that these are indeed TM domains (Campbell, Lecture 6, 2020). Interestingly Cav2.1 helices S5 and S6 actually have pretty high hydrophobicity peaks considering Cui indicates these helices form the central pore, on average being higher than helices S1-S4 of each TM domain (Figure 5). This is slightly unexpected as charged molecules such as ions do not pass easily through hydrophobic membranes and, as such, one will expect if hydrophobic helices make up the central pore this would prevent passage of charged cations. However, in fact, Hering et al. indicated that "hydrophobic residues in the lower third of" the S6 alpha helix TM repeat domain II are highly conserved in HVA calcium channels (see section SUBFAMILY ALIGNMENTS) and these residues "play an important role" in the activation of these gated channels and that mutation of these residues effect activation kinetics including slowing deactivation and decreased inactivation (Hering et al., 2008). The S6 helix in the second TM repeat domain (IIS6) from accession ID A0A1B0GTN7 spans residue 690-714 (689-713 in accession ID O00555) and mutation A713T is in the bottom third of this helix which may explain a role in the channelpathology of EIEE42 and EOEE and the role of this mutation in shifting voltage for activation (Hering et al., 2008). In fact, these 4 residues, LAIA in Cav1.2, indicated by Hering et al. aligns directly with residues 710-713 in Cav2.1. According to Kyte and Doolittle the hydropathy index for the Ala713Thr drops from 1.8 for an alanine residue to -0.7 for a threonine residue and using a sliding window of 9 residues in ProtScale the hydrophobicity score drops from 1.744 to 1.467 at position 713 (Figure 6).

It does make sense, because S4 has positively charged residues, lysines and arginines which have the lowest Kyte and Doolittle hydrophobicity score, at every third residue, that the S4 helices have the lowest hydrophobicity score than any other helix in the TM domains and that all helices S1-S6 have higher scores than the non-TM regions, i.e. the ectodomains and cytosolic domains (Figure 5). I also confirmed TM propensity with a prediction tool called TMHMM 2.0 (Krogh et al., 2001) since Hidden Markov Model (HMM) approach is also valuable in predicting TM proteins (Campbell, Lecture 6, 2020). This tool was pretty accurate (Figure S2) but missed 3 of 4 of the S3 and all four of the S4 TM helices in Cav2.1 (data not shown). This makes sense missing the S4 helices since they have conserved positively charge residues at every 3rd residue but should not have missed 3 of 4 of the S3 helices. I performed prediction with two other tools, PHDhtm (Rost et al., 1995) housed on NPS server (Combet et al., 2000) and TMpred (Hofmann & Stoffel, 1993) on the ExPASy server. Of these two tools, TMpred was more accurate than TMHMM 2.0 (data no shown) in that TMpred did not miss any of the S3 helices but like TMHMM 2.0 missed all of the S4 helices, most probable due to the positive residues. Interestingly, TMpred correctly predicted the reentrance helix in the P-loop between all four of the S5 and S6 TM helices (Figure 3) while TMHMM 2.0 considered these regions completely

extracellular or cytoplasmic. According to the crystal structure for Rabbit Cav1.1 PDB ID, the only HVA channel which has all of its TM domains crystallized, this reentrant TM P-loop is indeed an alpha helix (Figure S3). It is the conserved negatively charged glutamic acid residues of HVA channel in these P-loop helices that selectively coordinate the positively charged calcium cation (Simms & Zamponi, 2014) and this negative charge may be the reason why TMHMM 2.0 predicts this region to not be in the membrane. PHDhtm was the least accurate overall, despite predicting four groups of 6 TM helices, many of the spans of the predictions were longer, shorter, shifted or overlapping the annotated spans from UniProt (data not shown).

POST TRANSLATIONAL MODIFICATIONS

Post translational modifications (PTMs) of proteins are critical to a protein's stability, structure, interaction with other proteins and molecules, signaling, transport and regulation. PTMs are modifications that covalently add functional groups to proteins and include but not limited to phosphorylation (addition of phosphate group), glycosylation, ubiquitination (addition of ubiquitin protein), methylation, acetylation, and many types of lipidation which covalently attach lipids to proteins to recruit them to lipid bilayer of membranes or single lipid layer of micelles. Proteins that are imbedded plasma membrane as either TM (crossing from EC region into the cytosol) such as the Cavα1 subunit of voltage-gated calcium channels or lipid linked to the EC leaflet of the plasma membrane are often glycosylated on the ectodomains and/or phosphorylated on either ectodomains or cytosolic domains. Glycosylation PTM of proteins occurs shortly after the protein is synthesized and transported into the endoplasmic reticulum (ER) where sugar moieties are added to the protein on the ER lumen side and will ultimately be delivered to the plasma membrane through the secretory pathway where the glycan chains will face the outside of the cell (Colley et al., 2017). The two most common types of protein glycosylation are O-linked glycosylation where the sugars are added to oxygen on the side chains of serine or threonine residues, or N-linked glycosylation where sugars are attached as elaborate branched chains to the carboxamido nitrogen on the side chain of asparagine residues.

According to UniProt under accession ID O00555, the EC domain asparagine (Asn; N) residue N283 of Cav2.1 has an N-linked glycosylation PTM (UniProt, 2020). The consensus protein motif for N-linked glycosylation is Asn-x-Ser/Thr (N-x-S/T) where 'x' is any amino acid followed by a serine or threonine residue and the sequence of residues 283-285 for Cav2.1 is N-G-T so this indeed fits the motif. The motif for O-linked glycosylation is x(2)-S-T-x(2) where 'x' is any amino acid but highest propensity for proline, serine, threonine, or alanine (Hansen et al., 1995). It is possible to search through the entire sequence for this motif to determine other potential asparagine residues that can be post-translationally modified with N-linked glycosylation or motifs that are a match for possible O-linked glycosylation but it is important to be aware that only the ectodomains of membrane proteins are post-translationally modified with glycans and the cytoplasmic domains are never glycosylated (Colley et al., 2017). So, if we do not have a 3-D structure or alignment with a homologous protein whose structure and domain are verified searching for motifs for post-translation glycosylation is not going to be extremely accurate. We will go a little more in depth in motif searching and available tools besides using

regular expressions in the next section. I tried out the program NetOGlyc2.0 by DTU Health Tech to predict O-linked glycosylation however there is not functionality in the algorithm that includes modifying predictions for TM proteins and so the program erroneously predicts many threonine and serine residues within the long cytoplasmic linker domain between TM repeat domains II and III (Figure 3) which spans residues 715-1242 of Cav2.1 for accession ID A0A1B0GTN7 (Figure 7). The variant A713T would potentially be a target for O-linked glycosylation but being located on the cytoplasmic leaflet in alpha helix S6 of TM repeat domain II, this is extremely unlikely if not impossible.

Although phosphorylation of TM proteins it usually more critical to the regulation of TM receptor proteins involved in cell signaling pathways and not necessarily TM channel proteins like Cav2.1, UniProt also lists many phosphorylated serines and a couple threonine residues and links this PTM prediction through similarity to Rat and Mouse homologs of Cav2.1 (UniProt, 2020). The protein phosphorylation PTM on proteins is catalyzed by various protein kinases that target many different motifs and will be covered in the next section with profiles (Gooley & Packer, 1997).

MOTIFS AND PROFILES

In this section I have utilized the tool ScanProsite by ExPASy as well as the databases Pfam, SMART, and InterPro to confirm sequence spans for motif searching and profiles. Motif searching is a powerful tool which takes a consensus pattern either through scientific research and literature or through alignments of homologous proteins and those that share similar function and structure and converts into pattern syntax for searching programmatically. The pattern syntax, indicated by the PA (PAttern) line in PROSITE entries is one of the most widely used pattern syntax and is used extensively in scientific literature, although there are modifications made in publications on discretion of the authors (Hulo et al., 2006).

As previously indicated the UniProt database lists many phosphorylated residues for Accession ID O00555, mostly through similarity to Cav2.1 in other model organisms (Figure 8) and we can use ScanProsite (https://prosite.expasy.org/scanprosite/) with the "PROSITE collection of motifs" to validate these database entries. Phosphorylation of all three families of Cavα1 subunit, Cav1, Cav2, and Cav3 in some way provides regulation of these voltage-gated calcium channels such as modulating channel gating, responses to membrane depolarization, and increasing or preventing protein-protein interaction (Sang et al., 2016; Catterall, 2013; Blesneac, et al., 2015). Also, since the A713T mutant threonine residue is of the three residues with hydroxyl groups that are primarily phosphorylated, serine, threonine, and tyrosine, and because the location of residue 713 is at the bottom of the IIS6 helix near the cytoplasmic domain, this mutant A713T could potentially be phosphorylated. Ca^{2+}/Calmodulin-Dependent Protein Kinase II (CaMKII) is a serine/threonine kinase that is dependent on intracellular Ca^{2+} and as such its catalytic phosphorylation activity can induce calcium-dependent facilitation (CDF), a positive feedback mechanism (Bers et al., 2009). CaMKII does not have a motif entry in PROSITE, however studies by White et al. indicates "CaMKIIα preferentially phosphorylates substrates

with motifs" Hyd-X-Arg-X-X-Ser/Thr, and Hyd-X-Arg-NB-X-Ser/Thr-Hyd on either the serine or threonine residues where X is any residue, Hyd is any hydrophobic residue, and NB is any non-basic residue (White et al., 1998). In any case there must be an arginine residue 2 residues upstream of the phosphorylated residue and in the A713T mutant this residue is an alanine and so this mutant is not likely phosphorylated by CaMKII at A713T.

We can use ScanProsite to check their entire database of motifs and profiles simultaneously by providing a sequence or accession ID and therefore I have searched Cav2.1 using the accession ID O00555 as this is the entry in UniProt that contains the annotations for the phosphorylated residues. In addition to matching various enrichment profiles for a particular amino acid such as PS50323 arginine rich region profile in the cytoplasmic II-III linker domain, Cav2.1 has many positive motif hits for lipidation, glycosylation, amidation, and many hits for motifs corresponding to substrate sequences for a few different kinases; the greatest amount of matching motifs corresponding to substrates for Protein kinase C (PKC) which makes sense as elevated levels of cytosolic Ca^{2+} leads to activation of PKC (Lodish et al., 2016) as well as many hits for Casein kinase II substrate sequences (data not shown). Unfortunately, only 2 of the 16 indicated sites for phosphorylated residues for Cav2.1 in UniProt accession ID O00555 were matched with ScanProsite tool which could just mean the possibility that the mapping to other species by similarity in UniProt is incorrect or that PROSITE is missing some phosphorylation site motifs of other kinases or the current kinase motif entries in PROSITE might need to be updated or added upon.

PROSITE also has a profile matrix for the IQ domain. Its ACcession number (AC) is PS50096 and the consensus motif for this matrix leading to maximum score of 714 is **Q**<u>R</u>**RKAA**<u>V</u>**KIQA**<u>Y</u>**FRG**<u>H</u>**LAR**<u>R</u>**KYKK**<u>RRRERRS</u>. This is interesting, however, because it does not match the top row of the Sequence logo from PROSITE which I confirmed using WebLogo tool (Crooks et al., 2005) with the CLUSTAL format alignment file given with accession PS50096 in PROSITE (Figure 9). The bold residues above match the Sequence logo but the underlined residues with max matrix scores do not match the Sequence logo. Regardless Human Cav2.1 did not have a hit for this IQ profile in ScanProsite tool nor did any of the other 9 human isoforms for the Cavα1 subunits (data not shown), despite there being crystal structures for 5 of the isoforms (Cav1.1, Cav1.2, Cav2.1, Cav2.2, Cav2.3) in the Protein Data Bank (PDB) indicating an IQ domain bound by Ca^{2+}/calmodulin (CaM) (see section SUBFAMILY ALIGNMENTS) as well as plenty of citing articles that indicate the IQ domain of Cavα1 subunits is the binding site for Ca^{2+}/CaM in the calcium-dependent inactivation (CDI) of HVA channels (Simms & Zamponi, 2014).

SUBFAMILY ALIGNMENTS

Voltage-gated calcium channels have ways in regulating the response to extended depolarization and these channels have two well studied forms of channel inactivation, voltage-dependent inactivation (VDI) and that which we just mentioned CDI where, in HVA channels such as

Cav2.1 is dependent upon proper Ca^{2+}/CaM binding. The regulation mechanism that is of supreme interest to our IIS6 mutant A713T, is the VDI of HVA calcium channels.

Much research of VDI in HVA channels appear to have identified residues involved in VDI are in the S6 regions that line the inner face of the pore of Cavα1 subunits as well as the domain I-II linker (Simms & Zamponi, 2014; Tadross et al., 2010). As such it is natural to want to compare the S6 regions of TM domains Cav2.1 to the other Cavα1 isoforms to see how conserved the residues are in the S6 helices to determine what impact our A713T mutant may have. Therefore, I performed a multiple sequence alignment (MSA) with online version of Clustal Omega housed on the European Molecular Biology Laboratory – European Bioinformatics Institute (EMBL-EBI) server to align the S6 helices in all four TM repeat domains for all 10 isoforms as annotated by the topologies in UniProt (Figure 10) (Madeira et al., 2019; Sievers et al., 2011). To confirm we are aligning the correct residues for the S6 helices, I obtained crystal structures for almost complete TM components of Cavα1 subunits for Cav1.1 muscle isoform in Rabbit and Cav3.1 in Human and superimposed sequences -10 residues to cytosolic side of S5 to +10 residues to cytosolic side of S6, including the S5-S6 EC linker and reentrant P-loop for all four TM repeat domains (Figure 11). The MSA of the S6 helices for all four TM domains showed very similar or conserved residues totaling 20/25 (80%), 17/25 (68%), 16/25 (64%), 18/25 (72%) for TM repeat domains I to IV, respectively. In the bottom third (last 8 residues) of IIS6 these residues are completely conserved in HVA channels (Cav1 and Cav2 channels) having the sequence **N**VFL**AI**A**V** where residue 713 is the second alanine residue second from the bottom of the helix as well as completely conserved in Cav3 LVA channels having sequence **N**LLV**AI**L**V**. Three of the last four residues are completely conserved in all 10 isoforms: alanine, isoleucine, and valine. Overall in all four TM S6 helices it is the bottom half the helix that is more conserved across the HVA and LVA families compared to the top half of the helix, with the exception of TM domain 3 (TM3) S6 for Cav3 channels where an alanine in bottom half is replaced with a serine in Cav3.2 and Cav3.3.

As previously indicated in Hering et al., the residues in the lower third of IIS6 alpha helix are highly conserved and are thought to have "important role in activation gating…slowing of the activation kinetics…slowing of deactivation…and decreased inactivation" involved in functionality and regulation of calcium influx. Indeed a mutation at residue 713 at the bottom of the IIS6 helix from an alanine residue, with a small hydrophobic sidechain, which has the highest propensity for alpha helix to a threonine residue, with a medium sized polar sidechain, and much lower propensity for alpha helix might strain the bottom of the IIS6 helix enough to cause dysfunction in channel activation, deactivation, and depolarization that is observed in the phenotype of EIEE42 and EOEE.

The domain I-II linker region of HVA channels has also been linked to having a role in VDI and is thought to act as a "hinge-lid" and docks at the pore entrance of the cytoplasmic leaflet, possibly at the S6 helices, to physically block influx of calcium (Simms & Zamponi, 2014; Tadross et al., 2010). The Cav2.1 wild-type (WT) alanine residue at position 713 in Cav2.1 aligns with the alanine at position 660 in IIS6 region of muscle isoform Cav1.1 in Human as well as Cav1.1 in Rabbit, accession IDs Q13698 and P07293 in UniProt respectively. Indeed from

the crystal structure of Cav1.1 in Rabbit obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (Berman et al., 2000), there are some residues in the I-II cytosolic linker domain that are near this conserved alanine residue (Figure S4) and mutation of threonine at this location might also influence the role of this "hinge-lid"/S6 receptor site mechanism in the VDI of HVA calcium channels (Tadross et al., 2010). Although the two nearest residues in Figure S4, L333 and F337 for Cav1.1 are just below the IS6 helix these residues are not likely the residues of interest in the studies of linker I-II role in VDI since they are always near the pore, it is more likely we are interested in residues further in the cytosol of linker I-II region that change confirmation to act as a lid to inactivate the channel. It is these interactions between all the residues: the bottom third of IIS6 helix, the residues just below IS6 and those that make up the "hinge-lid" that are of interest. As such, I completed MSA using MAFFT version 7.429 (Katoh et al., 2002) for the I-II linker domain in Cav2.1 and all other Human isoforms of Cavα1 subunits in HVA families Cav1 and Cav2, as Cav3 channels have a different VDI mechanism (Simms & Zamponi, 2014), to see which residues may be conserved and thus might be important to VDI dysfunction of nearby mutant residues such as A713T during inactivation gating (Figure 12). The MSA in MAFFT shows a highly conserved region just after the IS6 TM helix in the cytoplasm. Studies indicate this region serves as region for binding β subunit interaction site and mutation in this region may serve in dysfunction of the kinetics of VDI (Berrou et al., 2001; Tadross et al., 2010). However, no known crystallization structure exists where the "hinge-lid" is in its conformation of inactivated Cavα1 subunit (Figure S5) so it would be hard to predict the interaction of these residues with our mutant A713T structurally and instead would have to perform kinetic experiments in mutant A713T vs wild-type (WT) as well as mutant residues in the conserved I-II linker to observe dysfunction and role in phenotype of EIEE42.

Studies of the other mechanism of voltage-gated calcium channel inactivation, CDI, place higher importance of Ca^{2+} dependent calmodulin associated binding to the IQ domain of the C-terminal region in HVA channels when channels are active and upon rise in intracellular calcium levels binds with the N-terminal region to inactivate the channel, a type of feedback inhibition. I performed MSA again with MAFFT on residues of the PreIQ domain (Figure 13A), labelled IQ-associated/GPHH_dom by InterPro database (InterPro 79.0, Finn et al., 2017), and IQ-associated/GPHH by Pfam (Pfam 33.1, El-Gebali et al., 2019), for HVA Cav1 and Cav2 channels as I am not aware of any evidence that Cav3 channels are regulated by CDI, as well as MSA on the IQ domain (Figure 13B), labelled IQ domain/VDCC_a1su_IQ in InterPro and IQ domain/Ca_chan_IQ in Pfam. I obtained the residue spans from the SMART database annotations for the IQ domain (SMART, Letunic et al., 2017) and cross referenced these with the sequences of the crystal structures of the Ca^{2+}/CaM with IQ domain of Cavα1 subunits that had entries in the RCSB PDB (Cav1.1, Cav1.2, Cav2.1, Cav2.2, Cav2.3). I also included Cav1.4 in the alignment despite this isoform lacking CDI, as it is able to also bind Ca^{2+}/CaM (Simms & Zamponi, 2014). The MSA for the IQ-associated or PreIQ domain had 31/60 (52%) evolutionarily conserved residues making it likely they have relatively good 3-D common structure and conserved secondary structure: strands, turns, loops, and helices (Campbell, Lecture 4, 2020). There were 4 conserved glycine residues and 5 conserved proline residues in the PreIQ domain making it possible for many secondary loop structures as these residues are

known to be helix breakers. In the next section we will look at secondary structure prediction of the PreIQ and IQ domains to confirm this, along with the key regions in VDI. The MSA had less conserved results for the IQ domain having only 7/35 (20%) residues conserved. Cav1 family channels have a true IQ (isoleucine-glutamine) motif while Cav2 have either a methionine or phenylalanine in place of a glutamine. Perhaps this could be one of the reasons why in Cav1 channels such as Cav1.2 and Cav1.3, individual channels can mediate their own CDI, that is a local increase in Ca^{2+} is only required to initiate CDI, while Cav2 type channels require global increase in Ca^{2+} in proximity to many channels to initiate CDI. Calmodulin might bind with higher affinity with true IQ containing sequences. I hypothesize that secondary structural predictions in the next section will align more closely for the PreIQ domains for these channels than the IQ domains.

Finally, I performed Smith-Waterman PSA for the P-loop to align the conserved S5-S6 P-loops in Rabbit Cav1.1 with Human Cav2.1 to obtain the location of the conserved glutamic acid residues (Figure S3C) which shows these P-loop regions are highly conserved in HVA channels. The alignment indicates that P-loop glutamic acid residues E292, E614, E1014, & E1323 of Rabbit Cav1.1 that coordinate the pore of HVA channels align to residues E318, E668, E1460, & E1756 respectively in Human Cav2.1.

STRUCTURAL PREDICTION TOOLS FOR SECONDARY STRUCTURE

Before doing any secondary structure prediction for Cav2.1 and other HVA isoforms, I was interested in performing pairwise alignments for Cav2.1 and the other 9 Human Cavα1 subunit isoforms including LVA channels as well as perform an MSA on all 10 isoforms using the secondary structural guide of Human Cav3.1 PDB IDs 6KZP and 6KZO (Figure S6). ClustalW and ClustalW2 are the two few MSA tools that can align sequences against the secondary structure of proteins with known 3-D structures (Campbell, Lecture 4, 2020).

I performed pairwise sequence alignment (PSA) Needleman-Wunsch algorithm (Needleman & Wunsch, 1970) which calculates optimal global alignments as well as Smith-Waterman algorithm (Smith & Waterman, 1981) which calculates local alignment for all 9 isoforms compared to Cav2.1 (Table 1). As expected the Cav2 family had the highest identity and similarity while Cav1 family had higher identity and similarity to Cav2.1 than Cav3 family did as Cav3 family are LVA channels.

I then performed structural MSA with the ClustalW2 algorithm using the secondary structures of Cav3.1 as a guide (Figure 14). This structural MSA indicated as suspected, since HVA channels associate with an ancillary β subunit as well as are regulated by VDI differently than LVA channels, that the I-II linker region that was highly conserved in Cav1 and Cav2 families did not have many residues that aligned from Cav3 family (Figure 14, middle) and likewise the PreIQ and IQ regions of Cav1 and Cav2 families also did not have many residues that aligned from Cav3 family (Figure 14, bottom). However, taking just one of the S5-S6 TM regions, TM domain 2, that form the central pore of the channel, these residues did align in all three Cavα1 subunit families (Figure 14, top). Given this information one would come to the conclusion that

the I-II linker regions of HVA channels as well as the C-terminal domains are very different from LVA channels in secondary and tertiary structure despite their TM helices being conserved and that in order to learn more about the A713T variant in Cav2.1 and its role in VDI and CDI, it would not add any value in studying the secondary and tertiary structures for Cav3 channels in the I-II linker and cytoplasmic C-terminal regions.

Despite the fact that Cav1 channels have low sequence similarity with respect to Cav2.1 (but higher than Cav3 channels), because they are both in HVA channel family, I still performed a PSA on two of the largest ectodomains as well as the I-II domain that were crystallized in Rabbit Cav1.1 with Cav2.1 to see if there were high similarity (Figure 15A,B,D,F,G). I then performed secondary structure prediction using PHD (Profile network from HeiDelberg) (Rost & Sander, 1993) on both Rabbit Cav1.1 and Human Cav2.1 to confirm with what level of certainty do the secondary structures of these regions in Cav2.1 being the same secondary structures of crystalized Rabbit Cav1.1 (Figure 15C,E,H). In the crystal structure for Rabbit Cav1.1, PDB ID 6BYO, most of the ectodomains are crystallized, where there are two substantial size segments, one between the IS5 and IS6 helices that is N-terminal to the entrance pore-forming P-loop and second region between IIIS5 and III6 also N-terminal to the P-loop (Figure 15A). There are others but these are the largest ones, greater than 20 residues required by PHD. The PSA between the first segment between IS5 and IS6 is extremely conserved between Rabbit Cav1.1 and Human Cav2.1 (Figure 15B). In the crystal structures for Rabbit Cav1.1 this entire region of around 60 residues is only loops and turns with no alpha helices or beta/extended beta strands. The secondary structure predicted by PHD for this loop segment for Rabbit Cav1.1 was pretty accurate predicting just over 72% random coil for loops and turns and of the 28% of the extended strand prediction, most of the prediction for these strands were not upper case and so were not predicted with high confidence (Figure 15C, left). The prediction for the aligned residues for this region from Cav2.1 in PHD also predicted a large portion to be random coil, around 88% (Figure 15C, right) and so we can be fairly confident the EC region between IS5 and IS6 of Cav2.1 is mostly random coil/loop structure. In addition the residues in the alignment for this region (Figure 15B,C) many of the residues for Cav2.1, 45/57 (79%), were either charged, polar, or special cases (glycine, proline, cysteine) and loops/coils usually have these hydrophilic residues and helix breakers that are exposed (Campbell, Lecture 6, 2020). PredictProtein (Yachdav, et al., 2014) also predicts this sequence between IS5/IS6 is mostly exposed (Figure S7A).

The second substantial EC region crystalized in Rabbit Cav1.1, the region between IIIS5 and IIIS6 has two sets of anti-parallel beta strands, each set containing a 3 and 5 residue strand forming a beta sheet characterized fold (Figure 15A) (Martinez-Ortiz & Cardozo, 2017). The PSA on this region with Human Cav2.1 was also extremely conserved (Figure 15D). The PHD prediction for Cav1.1 only missed the second set of antiparallel beta strands and was highly confident on the first set of antiparallel beta strands which align correctly to the crystal structure (Figure 15E, left). The PHD prediction for Human Cav2.1 was almost a mirror image of Rabbit Cav1.1 (Figure 15E, right) and so we can assume with fairly high confidence that this region between IIIS5 and IIIS6 of Cav2.1 has at least one set of antiparallel beta strands and is mostly random coil/loop. PredictProtein actually indicates that similar to Rabbit Cav1.1, this IIIS5-

IIIS6 linker of Cav2.1 does have 4 beta strand segments (Figure S7B) but only has prediction for exposed vs. buried residues for the first 25 of 47 residues.

I had already touched on the I-II linker region of Cav2.1 early when we discussed VDI and the "hinge-lid" mechanism of CDI, but we did not go into any detail on the secondary structure on this linker region. The crystal structure of Rabbit Cav1.1 for region I-II linker is entirely made up of alpha helices, except there is a gap of residues in the middle of the linker not included in the crystal structure (Figure 15F). The global PSA of Rabbit Cav1.1 I-II linker with full length Human Cav2.1 did align with residues of the I-II linker region annotated by UniProt, especially the beginning of the linker immediately C-terminal of IS6 as we already noted this region is highly conserved in HVA channels (Figure 12) however there is a gap in the aligment before and after the missing segment from the crystal structure of Rabbit Cav1.1 (Figure 15G, blue highlight). The PHD prediction of Rabbit Cav1.1 I-II linker incorrectly predicted the first 8 residues were random coil/loop when in fact these residues are an alpha helix but correctly predicts the rest of the most highly conserved region of the I-II linker being alpha helix with very high confidence (Figure 15H, left). Ironically, the segment missing from the crystal structure of Rabbit Cav1.1 shows very conserved alignment with Human Cav2.1 (Figure 15G, blue highlight) and the PHD prediction for Human Cav2.1 matches with Rabbit Cav1.1 up until the second helical region prediction (Figure 15H, right) so we can be fairly confident the highly conserved beginning region of cytoplasmic Cav2.1 I-II linker, and most likely all HVA Cavα1 subunits, are alpha helix, however because of gap in the alignment and gap in crystal structure it is hard to be confident on the I-II linker structure past this highly conserved beginning region of I-II linker.

Lastly, I performed secondary structure prediction with PHD of the PreIQ and IQ domains of HVA channels based on the fact that the PreIQ domains were more conserved than IQ to see if these secondary structures were more conserved in the prediction. There is only one crystal structure entry of the PreIQ domain for any HVA channel, which is for Cav1.2 however the PreIQ reference for this structure is just immediately upstream of the IQ domain and is not the IQ-associated domain indicated by InterPro and Pfam. Despite the MSA of the PreIQ showing higher percentage conserved residues in HVA channels than the IQ domain (Figure 14, bottom), the secondary structure predictions were pretty constant for both domains over the HVA Cav1 and Cav2 channels (Figure 16). PHD predicted three alpha helices for the PreIQ domain and correctly predicted the Ca^{2+}/CaM recognition helix in all HVA channel isoforms. One could hypothesize that the secondary structure of the IQ domain might not be the specific reason why only local increases of Ca^{2+} are needed to trigger CDI in Cav1 channels and global increases of Ca^{2+} are needed to trigger CDI in Cav2 channels and maybe the side chains of the specific residues in the helix causes higher affinity of Ca^{2+}/CaM for Cav1 over Cav2. We still do not know if the PreIQ and IQ domain bound by Ca^{2+}/CaM interacts at all with A713T mutant as we do not know if the CDI mechanism, which described by Simms and Zamponi, occurs when C-terminal bound Ca^{2+}/CaM with the IQ domain binds the N-terminal domain creating an inactive conformation, interacts with the channel pore at the cytoplasmic leaflet. More 3-D structure determination is needed especially with HVA channels in the inactivated conformation.

CAV2.1 ALIGNMENTS WITH MODEL ORGANISMS

Since the Cav2.1 channel does not have a high similarity to Cav1 channels, around 32% on average, more experimental information is needed to learn more about the 3-D structure of the pore of Cav2.1 P/Q type channels to determine the connection between the A173T mutant and phenotype EIEE42, not only in the active state when calcium influx occurs but during the initiation of inactivation as well as inactivated state. Therefore it would be valuable to see which model organism(s) have the highest sequence similarity in the full length Cav2.1 protein as well as highest sequence similarity in the regions of S5 and S6 helices that make up the central pore which includes our A713T variant at the bottom (cytoplasmic leaflet) of IIS6 and the I-II linker region for studying VDI in Cav2.1.

Currently in UniProt there are full length sequences determined for Cav2.1 for *Mus musculus* (Mouse), *Rattus norvegicus* (Rat), and *Oryctolagus cuniculus* (Rabbit). There are sequences for other species in UniProt, but they are not full length. Therefore, to obtain a few more species that may have high similarity for Cav2.1 or homolog, I performed a BLASTP (Altschul, et al., 1990) on NCBI server with Human Cav2.1 for each species individually filtering organisms for *Pan troglodytes* (Chimpanzee), *Gallus gallus* (Chicken), *Bos taurus* (Bovine/Cow), *Danio rerio* (Zebrafish), *Drosophila melanogaster* (Fruit Fly), *Equus caballus* (Horse), and *Saccharomyces cerevisiae* (Baker's Yeast). I ran BLASTP for full length Human Cav2.1 (accession ID O00555) with PAM250 substitution matrix as well as IIS5-IIS6 region, including linker, as this is where our mutant A713T (A712T in O00555) exists, with PAM30 substitution matrix (Table 2). Out of the seven species included in the BLASTs, all but Yeast showed high identity with Human Cav2.1, however there is only a partial sequence for Cav2.1 in Chicken and so the best match with highest identity and coverage is with Cav2.2 in Chicken. I then performed PSA on full length Human Cav2.1 with all but Chicken and Yeast, in addition to Rabbit, Mouse and Rat using the global Needleman-Wunsch algorithm, again with PAM250 substitution matrix, as well as PSA using local Smith-Waterman algorithm, again with PAM30 substitution matrix, for both IIS5-IIS6 region (residues 617-714 in A0A1B0GTN7) and the I-II linker (residues 361-487 in A0A1B0GTN7) in all organisms besides Yeast (Table 3). Chimpanzee and Bovine Cav2.1 had the highest global alignment with Human Cav2.1 as well as the highest local alignment score for IIS5-IIS6 and I-II linker regions. Zebrafish and Fruit Fly had the lowest global alignment similarity. All organisms in PSAs had high local alignment scores with the IIS5-IIS6 region as was expected with the transmembrane and pore domains being highly conserved in all Cavα1 isoforms, yet Zebrafish and Fruit Fly had only 79% and 46% local alignment similarity to the I-II linker. Thus, one can conclude from all the PSAs that the all organisms despite Zebrafish and Fruit Fly would be good model organisms to either test in vivo (Rabbit, Mouse, Rat, Chimpanzee) or isolate their Cav2.1 protein additionally for Bovine and Horse to perform experimental and 3-D X-ray/NMR structural studies in vitro in order to understand regulation and dysfunction in Cav2.1 VDI.

DISCUSSION

Progress for treating pathological disorders associated with mutations in voltage-gated calcium channels has been slow and difficult (McDonough, 2013). Obvious reasons for this are that experimental 3-D X-ray crystallography studies is difficult in transmembrane proteins. There are no known substantial structures of Cav2 channels and only a few for Cav1 and Cav3 channels. The experiments in this study encompass a collaborative of those that have been already been performed in Cav2 studies including alignments in the pore S5 and S6 region, the highly conserved P-loop with negatively charged residue in HVA channels that coordinate entry of calcium ions, and I-II linker region hypothesized to function in HVA VDI, hydrophobic studies to confirm transmembrane domains, alignments of the Pre-IQ and IQ domains of HVA channels with strong link to function in HVA CDI, as well as the use of 3-D crystal and electron microscopy structures of Cavα1 isoforms in Human and model organisms to align regions and predict homologous structure. There are no "new" experiments that have been performed in this paper; only I have applied techniques to link past research to find connections between sequence and structure of Cav2.1, other Cavα1 isoforms, and structural pieces that are possibly connected with this specific A713T mutation.

Most importantly we have learned that the Human Cav2.1 p.A713T mutation's role in the facilitation of epileptic seizures and impaired neuronal development phenotype of EIEE42. The S6 helices are extremely conserved in all Cavα1 channels precluding to their importance in the structure and function in facilitating calcium transport. The P-loop helices that also line the inner pore and coordinate the calcium ions are also conserved with negatively charged glutamic acid residues. According to homologous residues in Rabbit Cav1.1 structure, the mutant A713T residue is not in close proximity to the conserved glutamic acid residues (data not shown) so there would not be steric interaction causing dysfunction. The likely cause for dysfunction to gate activation/inactivation and shifts in polarization seen with mutations in the S6 region are attributable to A713T mutation along with rapid repolarization/hyperpolarization in cation voltage gated ion channels associated with the onset of epileptic seizures (Wei et al., 2017). The threonine residue likely disturbs IIS6 TM helix stability and interaction with nearby residues and hypothesized mechanism of interaction with the conserved I-II region in VDI regulation of Cav2.1 which leads to the channelpathology observed in EIEE42.

ADDITIONAL INFORMATION:

How to cite this paper: Wiley, B. (2020). Voltage-dependent P/Q-type calcium channel subunit alpha-1A, Assigned Mutation or Natural Variant: A712T. *Mid-Term Protein Bioinformatics AS.410.639.81.SU20*.

Any calculations, algorithms, or graphs used in this paper are cited and implemented in either Python 3.7.1. or R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night".

Supplementary figures are available in the file titled "A713T Supplemental Information".

TABLES AND FIGURES

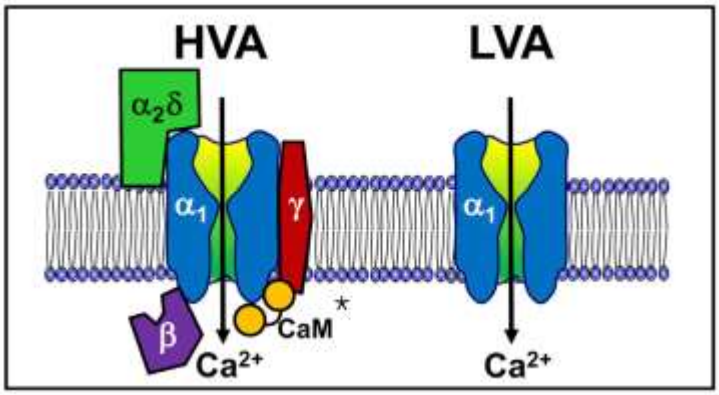


Figure 1. Overall cartoon structure of voltage-gated calcium channels. High-voltage activated (HVA) are mutimeric complexes that include Cav1 and Cav2 subfamilies. HVA channels associate with $\alpha_2\delta$ and β subunits. Calmodulin (CaM) is an important protein that regulates the activity of HVA channel inactivation. Figure taken from Simms and Zamponi, 2014, Neuronal Voltage-Gated Calcium Channels: Structure, Function, and Dysfunction.

A **B**

Figure 2. S5 and S6 alpha helices form pore of Cavα1 subunit in voltage-gated Ca^{2+} channels. Although this figure is from the Cavα1 subunit of the Cav3.1 T-type voltage-gated calcium channel created with electron microscopy by Zhao et al., as we will see later in the alignments section, the Cavα1 subunit of all 10 isoforms have the same overall channel structure. A) A view down the Cavα1 subunit from the exoplasmic leaflet. The four S6 helices lining the inner face of the central pore are shown in green while the S5 helices are shown in red. Calcium ions are in yellow. B) A view out to the EC region from the cytoplasmic leaflet of the TM spanning helices. Again, S6 helices are in green and S5 helices are in red. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. [Part (a) and (b) data from Y. Zhao et al., 2019, Nature 576: 492-497, PDB ID 6KZO.]

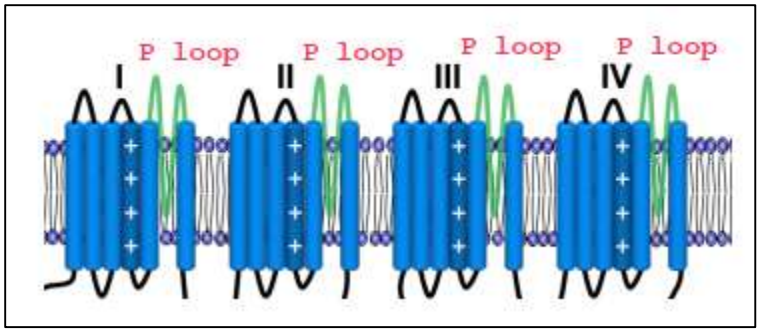


Figure 3. The P loop motifs of the Cavα1 subunit. The P loop motif, indicated in green, links between the S5 and S6 alpha helices in each the four membrane spanning repeat domains. In HVA channels, such as the Cav1 and Cav2 subfamilies, these motifs include negatively charged residues to coordinate the positively charged cations that enter through the pore created by S5 and S6 helices. The variant A713T is located at the bottom of the alpha helix S6 lining the pore of TM domain II (IIS6) (UniProt, 2020). Image taken from Simms & Zamponi, Neuronal Voltage-Gated Calcium Channels: Structure, Function, and Dysfunction. 2014.

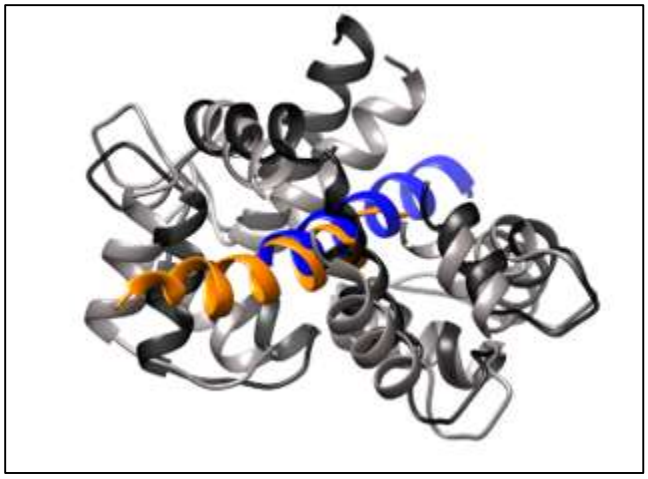


Figure 4. IQ Domain of Cav2.1 with calmodulin (CaM). This image contains the two crystal structures for Cav2.1 interacting with CaM. The black helices and blue helix are CaM and IQ domain of Cav2.1 respectively from PDB ID 3BXK in Rat. The grey helices and orange helix again are CaM and IQ domain of Cav2.1 respectively from PDB ID 3DVM in European Rabbit and Human. You can see that the alpha helices of IQ domains superimpose each other nicely in orange and blue showing homology across species. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. [Data from Kim et al., 2008, Structure 16: 1455-1467, PDB ID 3DVM and from Mori et al., 2008, PDB ID 3BXK]

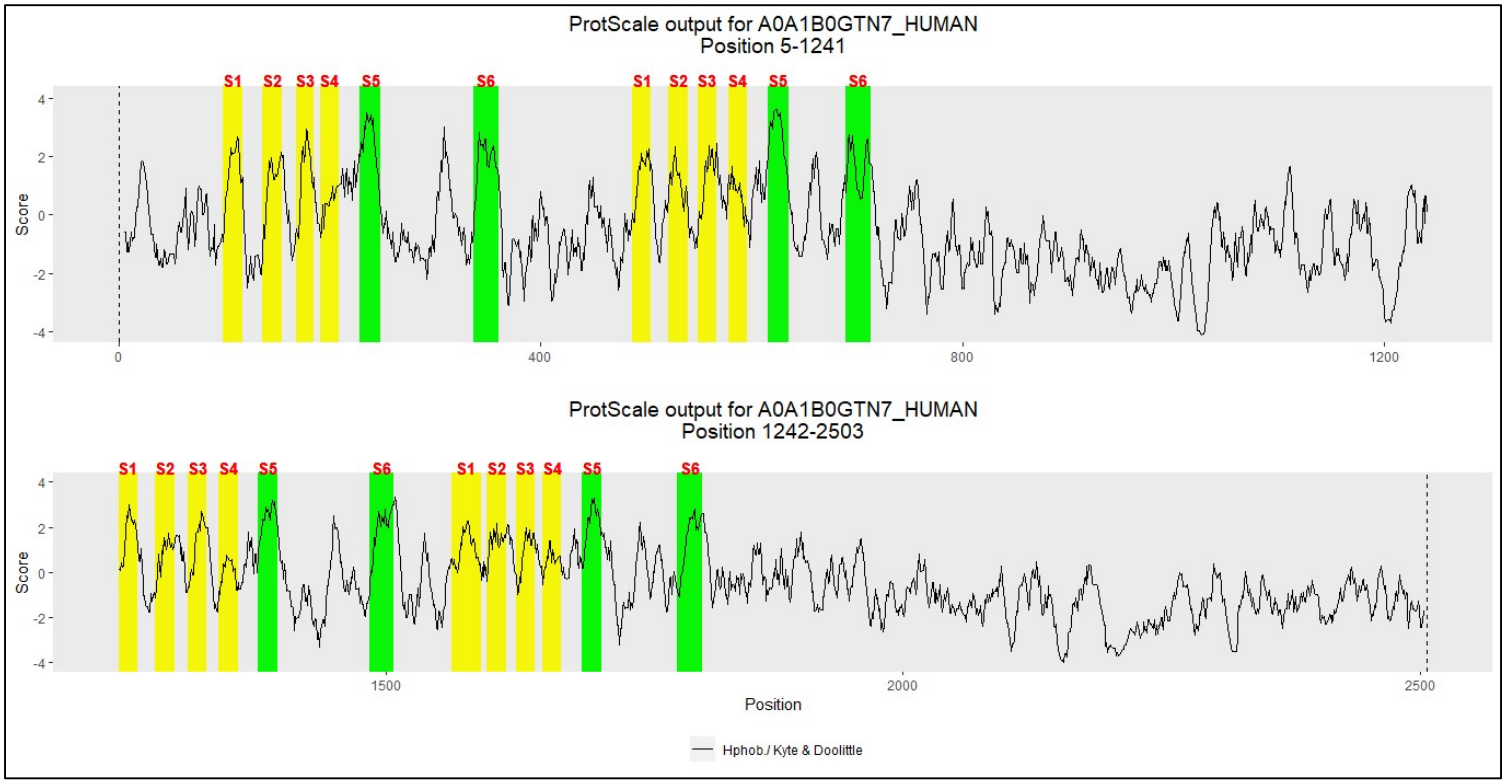


Figure 5. Hydrophobicity scores for Cav2.1 subunit. Scores were obtained using the ProtScale tool from ExPASy based off the Kyte & Doolittle hydrophobicity scale and a sliding window of 9 residues. The protein subunit was broken into two parts for better visibility of longer proteins. The TM domain helices are listed S1-S6 for all four repeats. Helices S1-S4 are highlighted yellow and S5-S6 helices that line the central pore are highlighted green to analyze if outside helices of the channel had different hydrophobicity than the inner helices that make up the pore. Plots were created with ggplot2 and gridExtra packages in R and source code is available upon request.

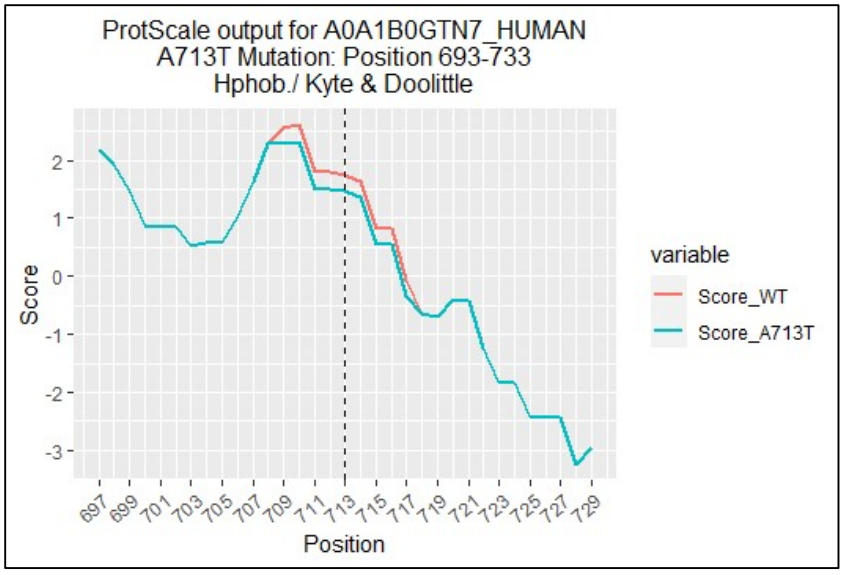


Figure 6. Hydrophobicity scores for Cav2.1 subunit showing drop in hydrophobicity for mutant A713T. Using the same Kyte & Doolittle hydrophobicity scale and window size of 9 residues as in figure 5, this figure is a closer look -20 to +20 residues around the site of our A713T mutant. The overall hydrophobicity drops in the range -5 residues to +5 residues around our mutant located at the bottom of IIS6 TM helix likely affecting central pore function. Plot was created with ggplot2 package in R and source code is available upon request.

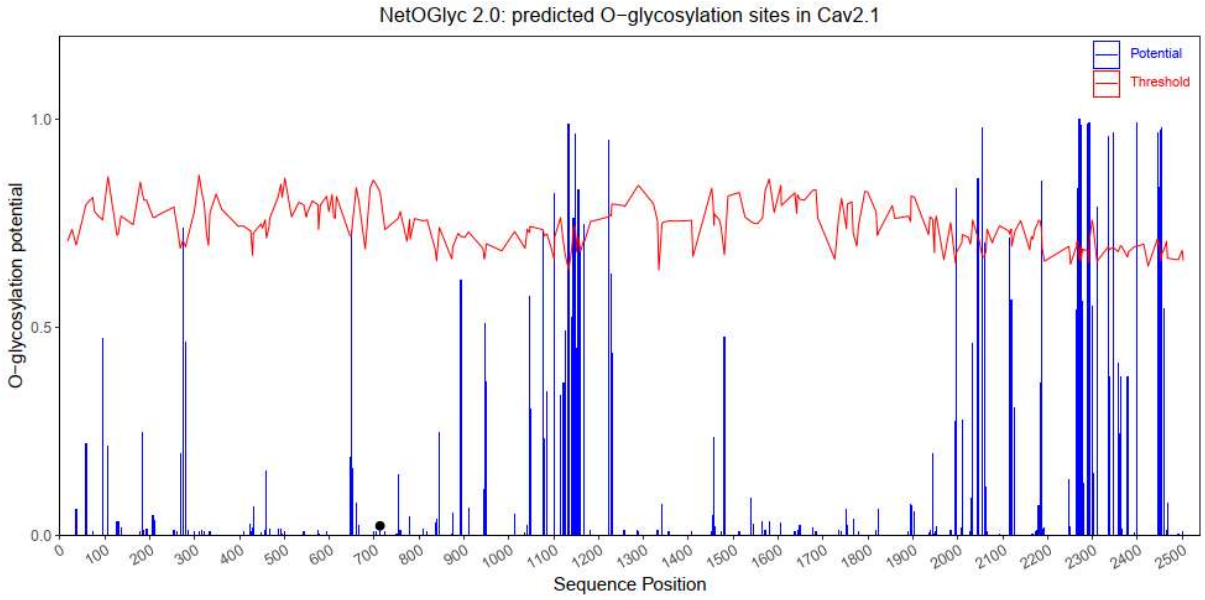


Figure 7. O-linked glycosylation prediction for Cav2.1 (accession ID A0A1B0GTN7) with mutant A713T substitution using NetOGlyc 2.0. The algorithm correctly indicates that glycosylation of 713T is low (black dot) which is accurate because this residue is near/at the bottom of IIS6 helix near the cytosolic leaflet. However the tool predicts many residues between position 1050-1200 and between 2000-2450 for O-linked glycosylation sites in regions that span the cytosolic domains between TM repeat domains II & III (residues 715-1242) as well as the cytosolic C-terminal domain respectively but glycosylation of TM proteins does not occur on cytosolic domains. Plot was created with ggplot2 package in R and source code is available upon request.

Modified residue	409	Phosphothreonine By similarity
Modified residue	447	Phosphoserine By similarity
Modified residue	450	Phosphoserine By similarity
Modified residue	749	Phosphoserine By similarity
Modified residue	752	Phosphoserine By similarity
Modified residue	789	Phosphoserine By similarity
Modified residue	1084	Phosphoserine By similarity
Modified residue	1093	Phosphoserine By similarity
Modified residue	1821	Phosphoserine; by PKA Sequence analysis

Figure 8. Some of the phosphorylated residues for Cav2.1. These are some, but not all, of the phosphorylated residues annotated on UniProtKB for accession ID O00555.

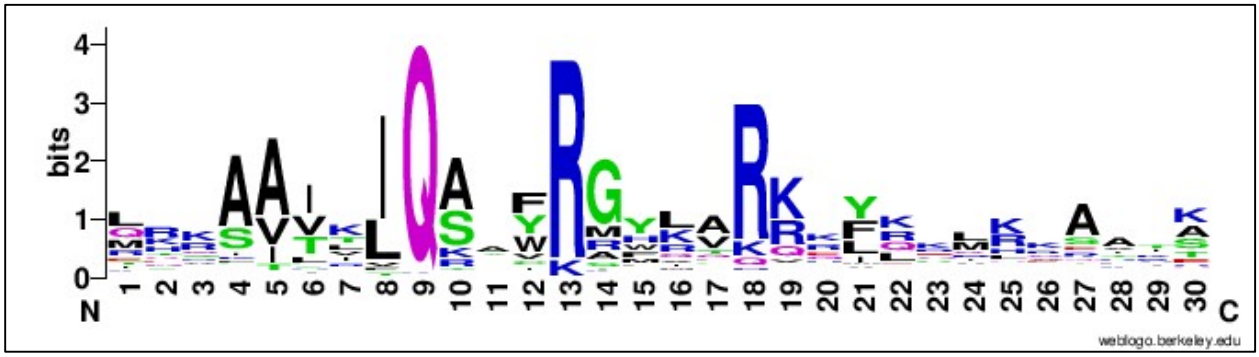


Figure 9. Sequence logo for PROSITE IQ matrix / profile. This consensus was created from an alignment of 1284 hits in 392 true positive sequences for the IQ domain which was confirmed with WebLogo (Crooks et al., 2005). However, the highest score from the matrix would indicate the sequence QRKAAVKIQAYFRGHLARRKYKKRRRERRS (score of 714) but the top level sequence from the Sequence logo based off the alignment, the sequence LRKAAIKIQAAFRGYLARKKYKKLKKAATK, would only score 602. These scores were confirm using the Bio.ExPASy.Prosite module in Python 3.7.1 and text processing to parse the matrix from matrix / profile PS50096 at PROSITE.

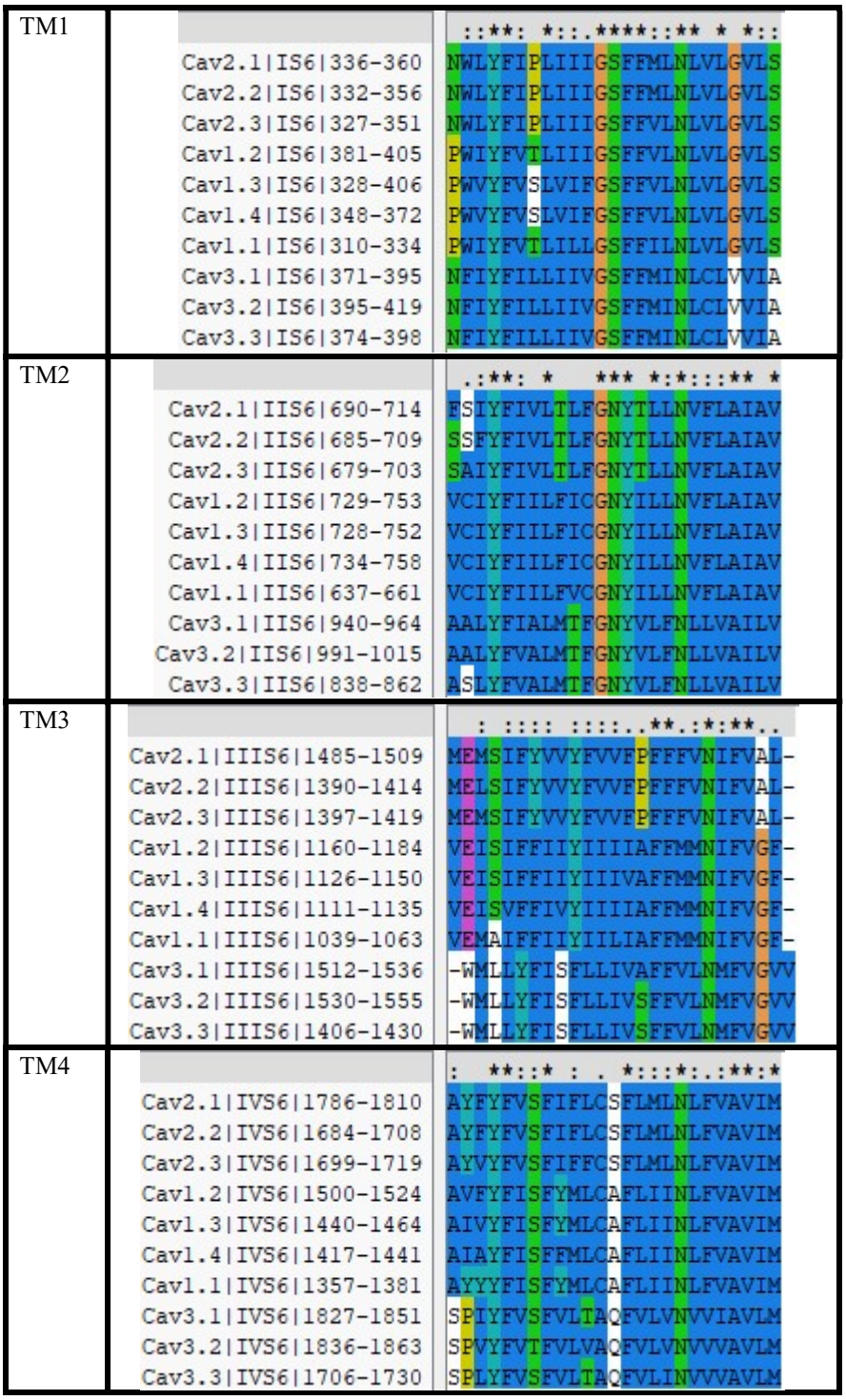


Figure 10. Multiple sequence alignment for S6 helices from all four TM repeat domains of Human Cavα1 subunit isoforms. S6 helix spans for each isoform were taken as annotated in UniProt and manually adjusted to perform the alignment. Each TM repeat domain is separated TM1-TM4. The residue locations for each isoform are indicated to the right of the isoform and domain number (Isoform|<TM#>S6|Residues). The residue numbers for Cav2.1 are for accession ID A0A1B0GTN7 sequence. Sequences were obtained from accession IDs: Cav2.2 (Q00975), Cav2.3 (Q15878), Cav1.2 (Q13936), Cav1.3 (Q01668), Cav1.4 (O60840), Cav1.1 (Q13698), Cav3.1 (O43497), Cav3.2 (O95180), and Cav3.3 (Q9P0X4). Visualization was done with ClustalX2 (Larkin et al., 2007). MSA in Clustal Omega was performed with default parameters.

A

B

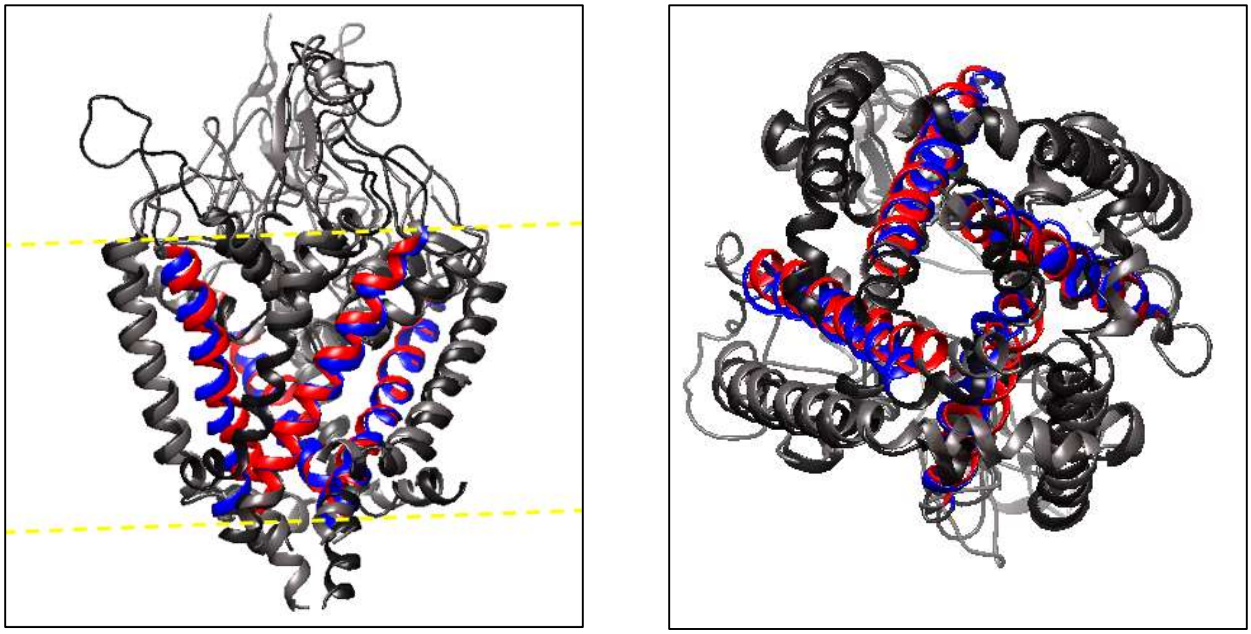
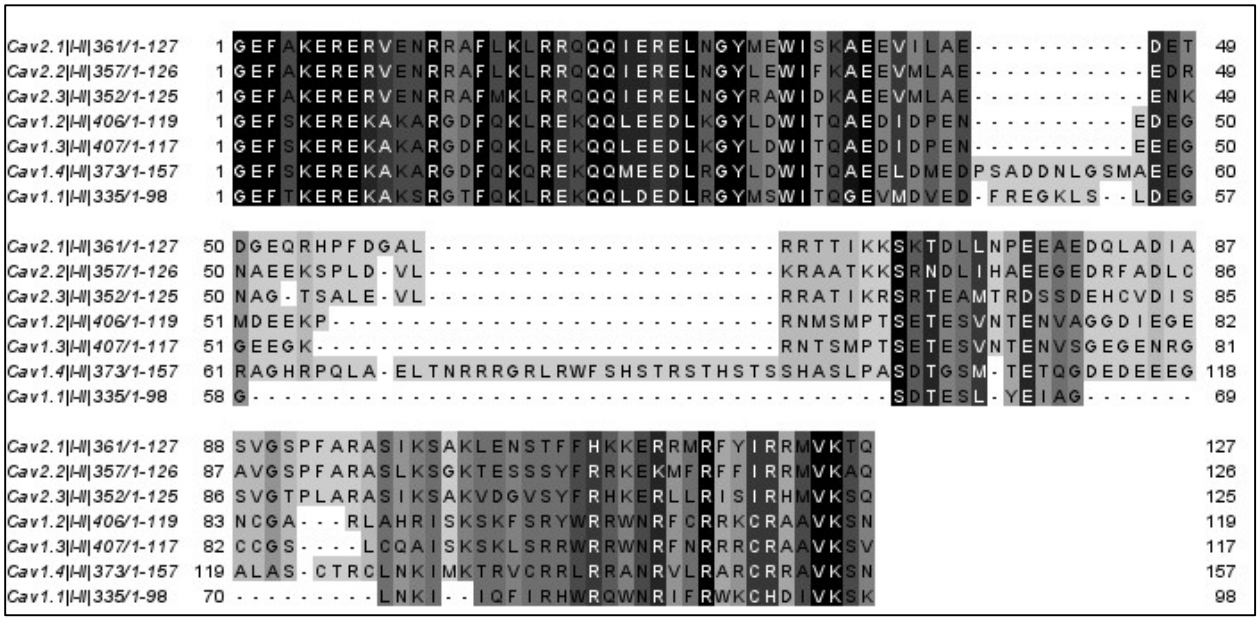


Figure 11. Confirming correct span of residues for S6 alignments of human Cavα1 subunit isoforms. A) This figure contains the crystal structures of Human Cav3.1 and Rabbit Cav1.1 superimposed from cytosolic -10 residues of S5 helix to +10 residues cytosolic of the S6 helix for all four TM domains. The grey residues contain the S5 and S5-S6 linker including EC P-loop of Human Cav3.1 (PDB ID 6KZO) with the twenty-five (25) S6 residues from the alignment of each TM domain in figure 10 colored in red. The black residues contain the S5 and S5-S6 linker including EC P-loop of Rabbit Cav1.1 (PDB ID 6BYO) with the twenty-five (25) S6 residues from the alignment of each TM domain in figure 10 (residues in Rabbit Cav1.1 are same as Human Cav1.1) colored in blue. The yellow dashed lines mark predicted EC and cytoplasmic leaflets of the plasma membrane. B) View from the bottom confirming we are using the correct span for S6 helix residues that lines the inner face of the central pore. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. [Data from Zhao et al., 2019, Nature 576: 492-497, PDB ID 6KZO and from Cardozo & Martinez-Ortiz 2017, PDB ID 6BYO]

A



B

Figure 12. MSA of the I-II linker domains of Human HVA Cavα1 subunit isoforms performed with MAFFT shows highly conserved beginning and ending segment. A) Shows full MSA of the entire I-II linker region where most of the conservation is in the beginning of the I-II linker just after IS6 as well as highly conserved residues toward end of the linker near IIS1. Visualization was done with Jalview (Waterhouse et al., 2009). B) Top: Highest conserved region of the I-II linker from A which is the cytoplasmic region directly after (C-terminal of) IS6. This is a span of 40 residues highlighted yellow (361-400 in both accession ID A0A1B0GTN7 and O00555 sequence for Cav2.1) of which 19 residues are conserved in HVA channels. Image is from MAFFT MSA ClustalW output housed on EMBL-EBI server. Bottom: Consensus sequence view from Jalview of the most conserved region of I-II linker from above. MAFFT MSA was done with BLOSUM62 matrix and default gap parameters.

A

```
Cav2.1|Pre-IQ|1 GPHHLDEYVRVWAEYDPAACGRIHYKDMYSLLRVISPPLGLGKKCPHRVACKRLLRMDLP
Cav2.2|Pre-IQ|1 GPHHLDEFIRVWAEYDPAACGRISYNDMFEMLKHMSPPLGLGKKCPARVAYKRLVRMNMP
Cav2.3|Pre-IQ|1 GPHHLDEFVRVWAEYDRAACGRIHYTEMYEMLTLMSPPLGLGKRCPSKVAYKRLVLMNMP
Cav1.2|Pre-IQ|1 GPHHLDEFKRIWAEYDPEAKGRIKHLDVVTLLRRIQPPLGFGKLCPHRVACKRLVSMNMP
Cav1.3|Pre-IQ|1 GPHHLDEFKRIWSEYDPEAKGRIKHLDVVTLLRRIQPPLGFGKLCPHRVACKRLVAMNMP
Cav1.4|Pre-IQ|1 GPHHLDEFKRIWSEYDPGAKGRIKHLDVVALLRRIQPPLGFGKLCPHRVACKRLVAMNMP
Cav1.1|Pre-IQ|1 GPHHLDEFKAIWAEYDPEAKGRIKHLDVVTLLRRIQPPLGFGKFCPHRVACKRLVGMNMP
                *******:  :*:***  * *** : ::  :*  :.****:** ** :** ***: *::*
```

B

```
Cav2.1|IQ|1948- STDLTVGKIYAAMMIMEYYRQSKAKKLQAMREEQD
Cav2.2|IQ|1846- PDEMTVGKVYAALMIFDFYKQNKTTRDQMQQAPGG
Cav2.3|IQ|1860- ASDLTVGKIYAAMMIMDYYKQSKVKKQRQQLEEQK
Cav1.2|IQ|1658- DDEVTVGKFYATFLIQEYFRKFKKRKEQGLVGKPS
Cav1.3|IQ|1598- DDEVTVGKFYATFLIQDYFRKFKKRKEQGLVGKYP
Cav1.4|IQ|1575- EEEVTVGKFYATFLIQDYFRKFRRRKEKGLLGNDA
Cav1.1|IQ|1515- DDEVTVGKFYATFLIQEHFRKFMKRQEE-YYGYRP
                  ::****.**:::* :.:::    : .
```

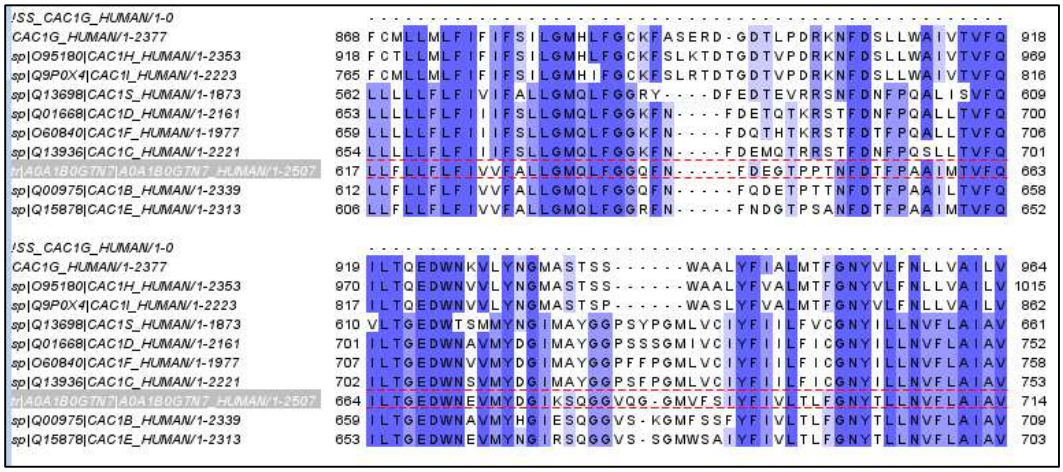
Figure 13. MSA of PreIQ and IQ domains of HVA calcium channels. A) MAFFT MSA of the PreIQ domains of HVA channels. This is known as the IQ-associated or GPHH domain in many databases. This domain is extremely conserved at 52% conservation between all 7 isoforms of HVA channels in Cav1 and Cav2 families. B) MAFFT MSA of the IQ domain of HVA channels which is not as well conserved as the PreIQ domain, only 20% of residues were conserved. MAFFT MSA was done with BLOSUM62 matrix and default gap parameters.

Needleman-Wunsch Global			
Isoform	Identity %	Similarity %	Gaps %
Cav1.1 (CAC1S_HUMAN)	31.1	43.5	34.9
Cav1.3 (CAC1D_HUMAN)	32.7	44.6	31.7
Cav1.4 (CAC1F_HUMAN)	32.2	44.0	31.6
Cav1.2 (CAC1C_HUMAN)	33.9	46.4	27.8
Cav2.2 (CAC1B_HUMAN)	**60.1**	**68.9**	**11.4**
Cav2.3 (CAC1E_HUMAN)	**55.4**	**65.4**	**13.8**
Cav3.1 (CAC1G_HUMAN)	22.2	35.8	33.5
Cav3.2 (CAC1H_HUMAN)	21.4	34.4	35.5
Cav3.3 (CAC1I_HUMAN)	22.3	35.1	34.7

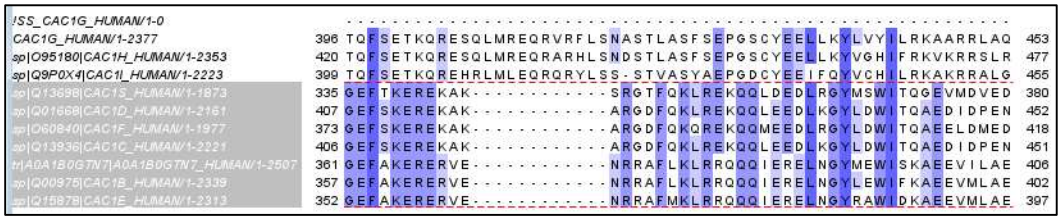
Smith-Waterman Local			
Isoform	Identity %	Similarity %	Gaps %
Cav1.1 (CAC1S_HUMAN)	36.0	49.7	27.4
Cav1.3 (CAC1D_HUMAN)	34.7	47.1	29.4
Cav1.4 (CAC1F_HUMAN)	33.6	45.8	29.7
Cav1.2 (CAC1C_HUMAN)	34.2	46.9	27.2
Cav2.2 (CAC1B_HUMAN)	**60.2**	**69.1**	**11.1**
Cav2.3 (CAC1E_HUMAN)	**55.4**	**65.4**	**13.7**
Cav3.1 (CAC1G_HUMAN)	22.5	36.5	32.6
Cav3.2 (CAC1H_HUMAN)	21.9	35.3	34.0
Cav3.3 (CAC1I_HUMAN)	23.0	36.5	32.9

Table 1. Cav3 LVA channels conservation versus Cav1 and Cav2 HVA channels conservation. Pairwise identity and similarity between Cav2.1 with the other 9 Human isoforms of Cavα1 subunits using global Needleman-Wunsch and local Smith-Waterman algorithms. Cav2.1 as expected has highest similarity with other Cav2 channels and higher similarity with Cav1 over Cav3 channels.

IIS5-IIS6



Domain I-II of HVA conserved region of linker



PreIQ and IQ Domains

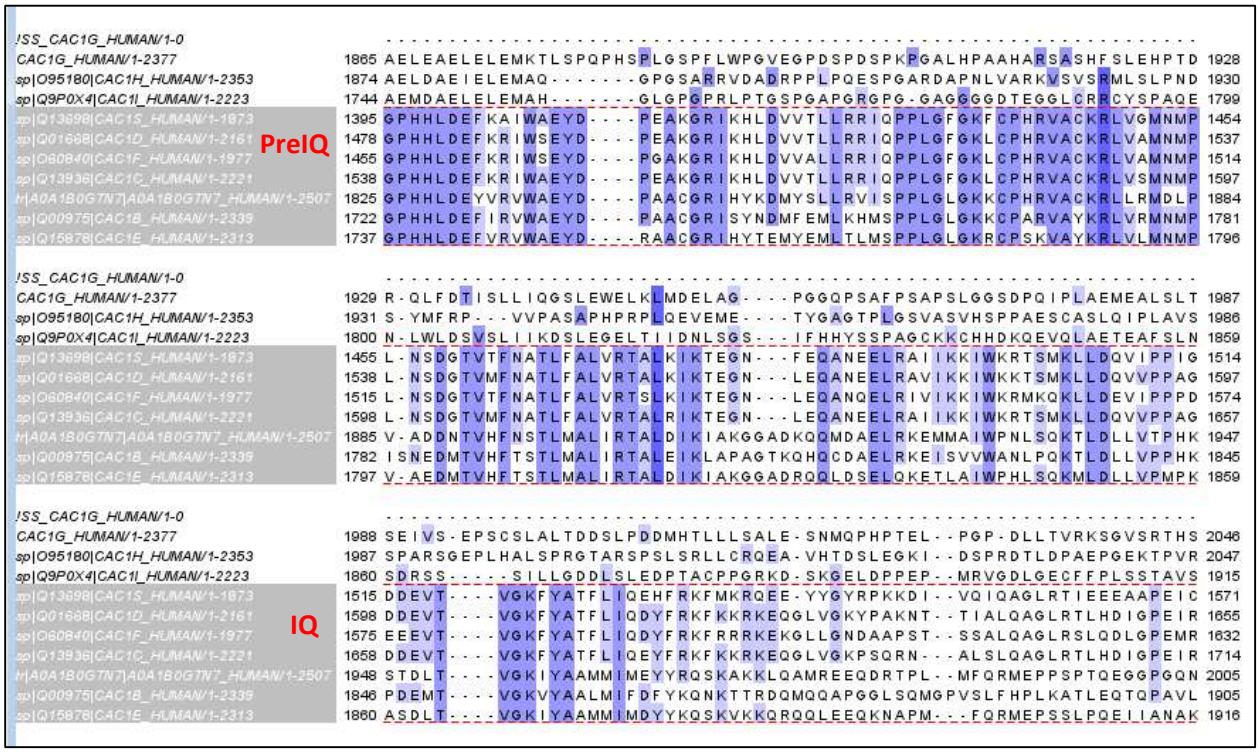
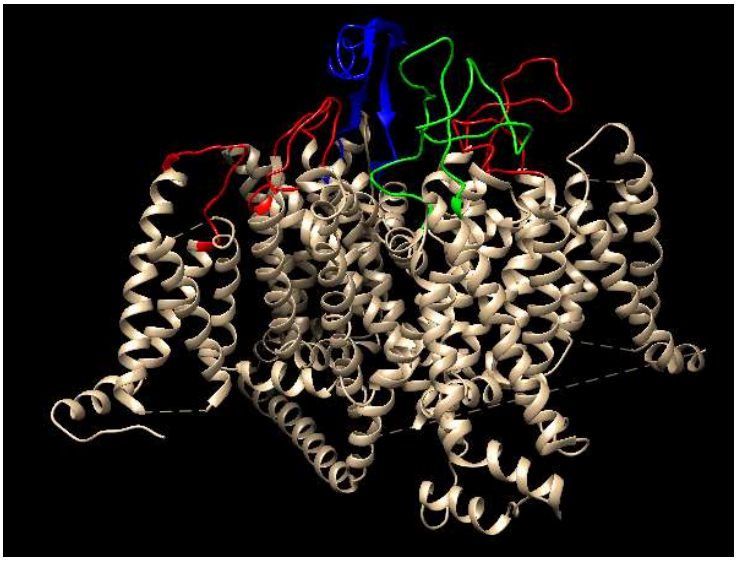


Figure 14. Cav3 LVA channels conservation versus Cav1 and Cav2 HVA channels conservation. Top: MSA with ClustalW2 based on the secondary structure from Cav3.1 PDB IDs 6KZP and 6KZO. The conserved TM helices of IIS5-IIS6 for all 10 isoforms indicates that the TM domains that form the pore of the channels are highly conserved in both HVA and LVA channels. Middle: The same MSA with ClustalW2 using secondary structure of Cav3.1 as a guide shows that the highly conserved region of the I-II linker, just C-terminal of IS6, that associates with the ancillary β subunit do not align with Cav3 channels. Bottom: Likewise the PreIQ-IQ domains are only conserved in Cav1 and Cav2 families (highlighted grey) in this region and the alignment of the C-terminal domain in Cav3 channels has very different residues indicating that the secondary and most likely tertiary structures of HVA and LVA channels differ drastically in the C-terminal domain. The order of the MSA is as follows: CAC1G/H/I are the top 3 rows for Cav3.1/2/3 respectively, CAC1S/D/F/C are middle 4 rows for Cav1.1/3/4/2 respectively and bottom 3 rows for A0A1B0GTN7(CAC1A)/B/E for Cav2.1/2/3 respectively. MSA performed with ClustalW2 secondary structure alignment was performed with default helix, strand, and loop gap penalty of 4, 4, and 1, respectively. Visualization was done with Jalview (Waterhouse et al., 2009).

A



B: Needleman optimal global alignment on the green linker with all of Cav2.1

```
EMBOSS_001         1 ------------------------------------------------KG      2
                                                                     .|
CAC1A_HUMAN      201 RPLKLVSGIPSLQVVLKSIMKAMIPLLQIGLLLFFAILIFAIIGLEFYMG    250

EMBOSS_001         3 KMHKTCYYIGTDIVATVENEKPSPCARTGSGRPCTINGSECRGGWPGPNH     52
                     |.|.||:..|||   .::.|.|:||......|.|. ||::|:..|.|||:
CAC1A_HUMAN      251 KFHTTCFEEGTD---DIQGESPAPCGTEEPARTCP-NGTKCQPYWEGPNN    296

EMBOSS_001        53 GITHFDNFG-------------------------------------------     61
                     |||.|||..
CAC1A_HUMAN      297 GITQFDNILFAVLTVFQCITMEGWTDLLYNSNDASGNTWNWLYFIPLIII    346
```

C

```
       10        20        30        40        50        60
        |         |         |         |         |         |
KGKMHKTCYYIGTDIVATVENEKPSPCARTGSGRPCTINGSECRGGWPGPNHGITHFDNFG
CCCeeceEecCCCCEeecccCCCCCCCCCCCCeeCCCCCeecCCCCCCCCceeeeecCC

Sequence length :    61

PHD :
   Alpha helix     (Hh) :     0 is   0.00%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :    17 is  27.87%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    44 is  72.13%
   Ambiguous states (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

```
       10        20        30        40        50
        |         |         |         |         |
MGKFHTTCFEEGTDDIQGESPAPCGTEEPARTCPNGTKCQPYWEGPNNGITQFDNIL
CCCeceeccCCCCcccCCCCCCCCcCCCCCCeeCCCCCeEeCCCCCCCccccCccccC

Sequence length :    57

PHD :
   Alpha helix     (Hh) :     0 is   0.00%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :     7 is  12.28%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    50 is  87.72%
   Ambiguous states (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

D: Needleman optimal global alignment on the blue linker with all of Cav2.1

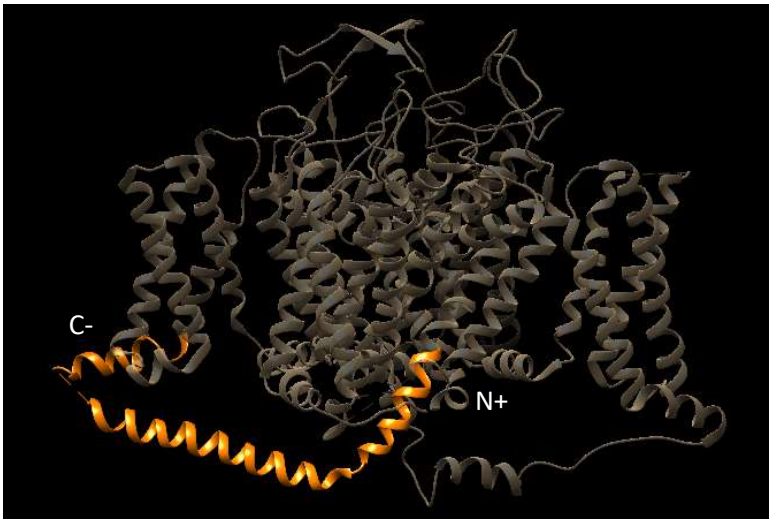
```
EMBOSS_001         1 ------------------------------------------------KG       2
                                                                     ||
CAC1A_HUMAN     1351 RPLKTIKRLPKLKAVFDCVVNSLKNVFNILIVYMLFMFIFAVVAVQLFKG    1400

EMBOSS_001         3 KFFSCNDLSKMTEEECRGYYYVYKDGDPTQMELRPRQWIHNDFHFDNV--      50
                     |||.|.|.||..|::||||.|.:|   :..:::.|.|:|...:||:|||
CAC1A_HUMAN     1401 KFFHCTDESKEFEKDCRGKYLLY---EKNEVKARDREWKKYEFHYDNVLW    1447
```

E

```
       10        20        30        40        50
        |         |         |         |         |
KGKFFSCNDLSKMTEEECRGYYYVYKDGDPTQMELRPRQWIHNDFHFDNV
CCceEEcCCCCCCCccccceEEEEEeCCCCCCCCccccecCCCcCcCCC

Sequence length :    50

PHD :
   Alpha helix     (Hh) :     0 is   0.00%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :    11 is  22.00%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    39 is  78.00%
   Ambiguous states (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

```
       10        20        30        40
        |         |         |         |
KGKFFHCTDESKEFEKDCRGKYLLYEKNEVKARDREWKKYEFHYDNV
CCCEEEcCCCcCCCcccccCcEEEEEeCCCCCccccccccCCCcCcCCC

Sequence length :    47

PHD :
   Alpha helix     (Hh) :     0 is   0.00%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :     9 is  19.15%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    38 is  80.85%
   Ambiguous states (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

F



G

```
EMBOSS_001        1 ------GVLSGEFTKEREKAKSRGTFQKLREKQQLEEDLRGYMSWITQGE     44
                        |||||||.||||:.::|..|.|||.:||:|.:|.|||.||::.|
CAC1A_HUMAN     351 MLNLVLGVLSGEFAKERERVENRRAFLKLRRQQQIERELNGYMEWISKAE    400

EMBOSS_001       45 VMDVEDLREGKLSLEEGGSDTESLYEIEGLNKIIQFIRHWRQWNRVFRWK     94
                          ::.|.|..:|.|..:..:.|.:.
CAC1A_HUMAN     401 ----------EVILAEDETDGEQRHPFDALRRT-----------------    423

EMBOSS_001       95 CHDLVKSR----------------------------------------     102
                      .:.||:
CAC1A_HUMAN     424 --TIKKSKTDLLNPEEAEDQLADIASVGSPFARASIKSAKLENSTFFHKK    471
```

H

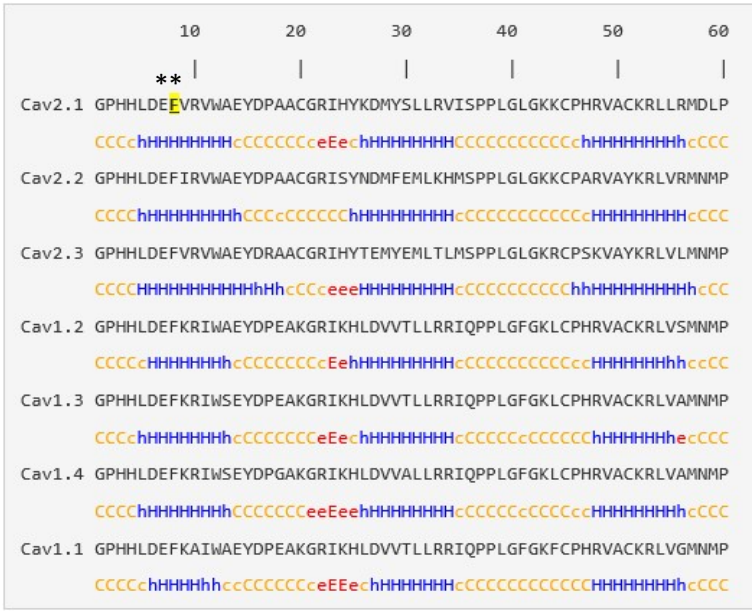
```
       10        20        30        40        50        60        70
        |         |         |         |         |         |         |
GVLSGEFTKEREKAKSRGTFQKLREKQQLEEDLRGYMSWITQGEVMDVEDLREGKLSLEEGGSDTESLYE
CCCCCccCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHhccCCCCCCcCccccCCCCCCCCCCCC
IEGLNKIIQFIRHWRQWNRVFRWKCHDLVKSR
CCCCCchHHHHHHHHHHHHHHHHHHHHHcCC

Sequence length :   102

PHD :
   Alpha helix     (Hh) :    59 is  57.84%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :     0 is   0.00%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    43 is  42.16%
   Ambiguous states (?)  :     0 is   0.00%
   Other states         :     0 is   0.00%
```

```
       10        20        30        40        50        60        70
        |         |         |         |         |         |         |
GVLSGEFAKERERVENRRAFLKLRRQQQIERELNGYMEWISKAEEVILAEDETDGEQRHPFDALRRTTIK
CCcCCcccHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcccccCCCCccCCCCccccccccce
KSK
cCC

Sequence length :    73

PHD :
   Alpha helix     (Hh) :    36 is  49.32%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :     1 is   1.37%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    36 is  49.32%
   Ambiguous states (?)  :     0 is   0.00%
   Other states         :     0 is   0.00%
```

Figure 15. Secondary Structure predictions of two EC domains and I-II linker region of Cav2.1 using Rabbit Cav1.1 crystal structure as guide. A) EC domains from crystal structure of Rabbit Cav1.1, PDB ID 6BYO. The IS5-IS6 linker is in green and the IIIS5-IIIS6 linker is in blue. Both of these linkers are N-terminal to the P-loop that

reenters the membrane between S5 and S6 helices. B) Needleman-Wunsch global alignment of EC linker between IS5 and IS6 for Rabbit Cav1.1 (EMBOSS_001) with Cav2.1 (accession ID O00555). C) Left: PHD prediction of linker between IS5 and IS6 for Rabbit Cav1.1. Right: PHD prediction of same alignment region of Human Cav2.1 between IS5 and IS6. D) Needleman-Wunsch global alignment of EC linker between IIIS5 and IIIS6 for Rabbit Cav1.1 (EMBOSS_001) with Cav2.1 (accession ID O00555). E) Left: PHD prediction of linker between IIIS5 and IIIS6 for Rabbit Cav1.1. Right: PHD prediction of same alignment region of Human Cav2.1 between IIIS5 and IIIS6. F) The I-II linker region of Rabbit Cav1.1. Same crystal structure as in A. G) Needleman-Wunsch global alignment of cytoplasmic domain I-II linker region for Rabbit Cav1.1 (EMBOSS_001) with Cav2.1 (accession ID O00555). The segment highlighted blue in Cav1.1 is missing from the crystal structure while the highlighted yellow indicates despite there being crystal structure for Cav1.1, there is gap in alignment with Human Cav2.1. H) Left: PHD prediction of I-II linker for Rabbit Cav1.1. Right: PHD prediction of same alignment region of Human Cav2.1 I-II linker. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. [Data from from Cardozo & Martinez-Ortiz 2017, PDB ID 6BYO]

A



B

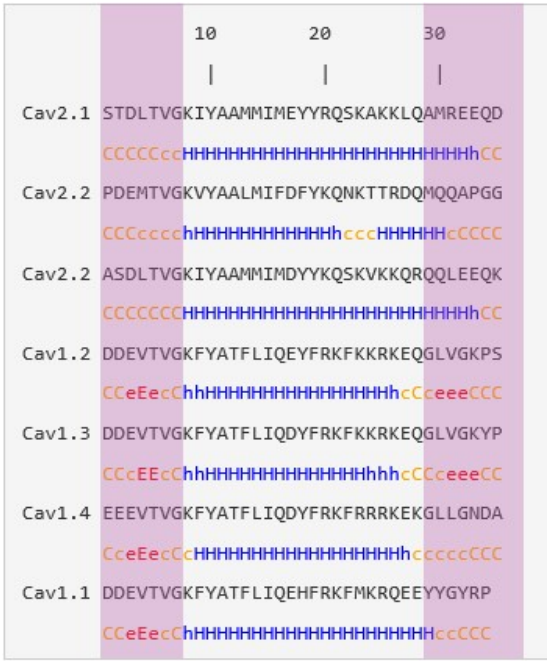


Figure 16. Secondary structure predictions of PreIQ and IQ domains of HVA channels. A) The PHD secondary structural predictions of all HVA channels for the PreIQ domain are all pretty close to each other. **Note there was an error for the input in Cav2.1 and PHD was not taking the sequence correctly so I had to mutate a tyrosine for a phenylalanine seen highlighted in yellow, bold, and underlined. The true sequence has a tyrosine (Y) residue at this location. B) The PHD secondary structure predictions of IQ domains are closer to each other than hypothesized. The predictions of the Ca^{2+}/CaM recognition helix region (the non-pink highlighted middle region) are consistently predicted correct as alpha helices which aligns with their crystal structures.

Organism	BLASTP: Full Length Cav2.1 % Identity best match	NCBI Accession\ (Full Length)	BLASTP: IIS5-IIS6 % Identity best match	NCBI Accession (best IIS5-IIS6 match)
Chimpanzee	99.6	XP_016790736.1	100.0	Same
Chicken	60.5*	XP_015134758.2	95.35	ATE62979.1
Bovine	95.2	NP_001068597.1	100.0	Same
Zebrafish	70.3	XP_021330116.1	90.70	Same
Fruit Fly**	59.8	NP_001356945.1	81.93	Same
Horse	92.2	XP_023501040.1	100.0	Same
Yeast***	23.2	PTN21583.1	No similarity	NA

Table 2. BLASTP with Cav2.1 for model organisms that do not have full length entries in UniProt. This table include results from NCBI's BLASTP. Full length Cav2.1 was blasted against each organism individually (second column) and again with residues 616-713 (accession ID O00555) for the IIS5-IIS6 region that contains the A>T variant (fourth column). BLASTP was performed with WT A712. *Cav2.2 N-type, P/Q Cav2.1 for Chicken was only partial in NCBI database. **Drosophila homolog for Cav2.1 is cacophony. ***Yeast homolog for Cav2.1 is Cch1p.

Organism	NW Full Length Similarity / Identity %	NCBI/UniProt Accession (Full Length)	SW IIS5-IIS6 Similarity / Identity %	SW I-II Linker Similarity / Identity %	NCBI/UniProt Accession (best IIS5-IIS6 match)
Rabbit	89.0 / 84.1	P27884	100.0 / 100.0	96.9 / 96.9	Same
Mouse	92.0 / 87.7	P97445	100.0 / 100.0	97.6 / 97.6	Same
Rat	86.3 / 82.9	P54282	99.0 / 99.0	97.6 / 97.6	Same
Chimpanzee	96.9 / 96.7	XP_016790736.1	100.0 / 100.0	99.2 / 98.4	Same
Chicken	NA*	XP_015134758.2	96.9 / 95.9	92.9 / 90.6	ATE62979.1
Bovine	98.4 / 96.2	NP_001068597.1	100.0 / 100.0	98.4 / 98.4	Same
Zebrafish	82.5 / 70.4	XP_021330116.1	94.9 / 91.8	78.7 / 70.9	Same
Fruit Fly	54.8 / 39.6	NP_001356945.1	87.8 / 82.7	45.7 / 43.8	Same
Horse	96.3 / 92.3	XP_023501040.1	100.0 / 100.0	96.1 / 96.1	Same

Table 3. PSA of Human Cav2.1 with Cav2.1 or homolog of model organisms. Global PSA were done with full length proteins using the Needle-Wunch (NW) algorithm (second column) while local PSA was done for IIS5-IIS6 (fourth column) (residues 617-714 in A0A1B0GTN7) and I-II linker (fifth column) (residues 361-487 in A0A1B0GTN7) were done with the Smith-Waterman (SW) algorithm. The first number in each cell is the similarity score and the second number is the identity score. *As indicated in Table 2 above, there was only partial for Cav2.1 for Chicken so did not complete global PSA. Best scores are highlighted yellow.

REFERENCES:

1. Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M. C., Roberts, K., Walter, P. (2014). Essential Cell Biology. 4th Edition. Taylor & Francis Group, LLC; New York, NY.
2. Purves D, Augustine GJ, Fitzpatrick D, et al., editors. Neuroscience. 2nd edition. Sunderland (MA): Sinauer Associates; 2001. Voltage-Gated Ion Channels. Available from: https://www.ncbi.nlm.nih.gov/books/NBK10883/
3. Campbell, C. Protein Bioinformatics: 410.639.81 SUMMER 2020 Mid-Term Exam. Johns Hopkins University. 2020.
4. Simms BA, Zamponi GW. Neuronal voltage-gated calcium channels: structure, function, and dysfunction. Neuron. 2014;82(1):24-45. doi:10.1016/j.neuron.2014.03.016
5. Lodish, H., Berk, A., Kaiser, C., Krieger, M., Bretscher, A., Ploegh, H., Amon, A. and Martin, K. (2016). Molecular Cell Biology. New York: W.H. Freeman and Company.
6. Grienberger C., Konnerth A. Imaging calcium in neurons. Neuron. 2012;73(5):862-885. doi:10.1016/j.neuron.2012.02.011
7. Few AP, Lautermilch NJ, Westenbroek RE, Scheuer T, Catterall WA. Differential regulation of CaV2.1 channels by calcium-binding protein 1 and visinin-like protein-2 requires N-terminal myristoylation. J Neurosci. 2005;25(30):7071-7080. doi:10.1523/JNEUROSCI.0452-05.2005
8. The UniProt Consortium, UniProt: a worldwide hub of protein knowledge, Nucleic Acids Research, Volume 47, Issue D1, 08 January 2019, Pages D506–D515, https://doi.org/10.1093/nar/gky1049
9. Andrade, A., Brennecke, A., Mallat, S., Brown, J., Gomez-Rivadeneira, J., Czepiel, N., & Londrigan, L. (2019). Genetic Associations between Voltage-Gated Calcium Channels and Psychiatric Disorders. International journal of molecular sciences, 20(14), 3537. https://doi.org/10.3390/ijms20143537
10. Zhang, F. X., Gadotti, V. M., Souza, I. A., Chen, L., & Zamponi, G. W. (2018). BK Potassium Channels Suppress Cavα2δ Subunit Function to Reduce Inflammatory and Neuropathic Pain. Cell reports, 22(8), 1956–1964. https://doi.org/10.1016/j.celrep.2018.01.073
11. Bronwen L. Aken, Sarah Ayling, Daniel Barrell1, Laura Clarke, Valery Curwen, Susan Fairley, Julio Fernandez Banet, Konstantinos Billis, Carlos Garcín Girón, Thibaut Hourlier, Kevin Howe, Andreas Kähäri, Felix Kokocinski, Fergal J. Martin, Daniel N. Murphy, Rishi Nag, Magali Ruffier, Michael Schuster, Y. Amy Tang, Jan-Hinnerk Vogel, Simon White, Amonida Zadissa, Paul Flicek and Stephen M. J. Searle. The Ensembl gene annotation system. Database 2016, baw093. doi: 10.1093/database/baw093
12. Mallmann, R. T., Elgueta, C., Sleman, F., Castonguay, J., Wilmes, T., van den Maagdenberg, A., & Klugbauer, N. (2013). Ablation of Ca(V)2.1 voltage-gated Ca^{2+} channels in mouse forebrain generates multiple cognitive impairments. PloS one, 8(10), e78598. https://doi.org/10.1371/journal.pone.0078598
13. Epi4K Consortium; Epilepsy Phenome/Genome Project, Allen AS, et al. De novo mutations in epileptic encephalopathies. Nature. 2013;501(7466):217-221. doi:10.1038/nature12439

14. Epi4K Consortium (2016). De Novo Mutations in SLC1A2 and CACNA1A Are Important Causes of Epileptic Encephalopathies. American journal of human genetics, 99(2), 287–298. https://doi.org/10.1016/j.ajhg.2016.06.003
15. Cui J. (2010). Reduction of CaV channel activities by Ca2+-CaM: inactivation or deactivation?. The Journal of general physiology, 135(4), 297–301. https://doi.org/10.1085/jgp.201010421
16. UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem.* 2004 Oct;25(13):1605-12.
17. PDB ID: 6KZO
Zhao Y, Huang G, Wu Q, et al. Cryo-EM structures of apo and antagonist-bound human Cav3.1. Nature. 2019;576(7787):492-497. doi:10.1038/s41586-019-1801-3
18. The Pfam protein families database in 2019: S. El-Gebali, J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Potter, M. Qureshi, L.J. Richardson, G.A. Salazar, A. Smart, E.L.L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, R.D. Finn. Nucleic Acids Research (2019) doi: 10.1093/nar/gky995
19. Wootton, J.C., Federhen, S. Statistics of local complexity in amino acid sequences and sequence databases, Comput. Chem. 17 (1993) 149–163, https://doi.org/10.1016/0097-8485(93)85006-X
20. PDB ID: 3BXK
Mori, M.X., Vander Kooi, C.W., Leahy, D.J., Yue, D.T. (2008). DOI: 10.2210/pdb3BXK/pdb
21. PDB ID: 3DVM
Kim E.Y., Rumpf C.H., Fujiwara Y., Cooley E.S., Van Petegem F., Minor Jr, D.L. Structures of CaV2 Ca2+/CaM-IQ domain complexes reveal binding modes that underlie calcium-dependent inactivation and facilitation. Structure. 2008;16(10):1455-1467. doi:10.1016/j.str.2008.07.010
22. Pace, C. N., & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. Biophysical journal, 75(1), 422–427. https://doi.org/10.1016/s0006-3495(98)77529-0
23. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.
24. Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of molecular biology, 157(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0
25. Campbell, C. Transmembrane (TM) identification methods. *Lecture_6 online*. Johns Hopkins University: Protein Bioinformatics, Summer 2020. Slide 14.
26. Hering, S., Beyl, S., Stary, A., Kudrnac, M., Hohaus, A., Guy, H. R., & Timin, E. (2008). Pore stability and gating in voltage-activated calcium channels. Channels (Austin, Tex.), 2(2), 61–69. https://doi.org/10.4161/chan.2.2.5999
27. Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete

genomes. Journal of molecular biology, 305(3), 567–580. https://doi.org/10.1006/jmbi.2000.4315
28. Rost, B., Casadio, R., Fariselli, P., & Sander, C. (1995). Transmembrane helices predicted at 95% accuracy. Protein science : a publication of the Protein Society, 4(3), 521–533. https://doi.org/10.1002/pro.5560040318
29. NPS@: Network Protein Sequence Analysis. TIBS 2000 March Vol. 25, No 3 [291]:147-150. Combet C., Blanchet C., Geourjon C. and Deléage G.
30. Hofmann K., Stoffel, W. (1993). TMbase - A database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler 374,166.
31. Colley K.J., Varki A., Kinoshita T. Cellular Organization of Glycosylation. (2017). In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 4. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453052/ doi: 10.1101/glycobiology.3e.004
32. Hansen, J. E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J. O., & Hansen, J. E. (1995). Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase. The Biochemical journal, 308 (Pt 3)(Pt 3), 801–813. https://doi.org/10.1042/bj3080801
33. Hulo N., Bairoch A., Bulliard V., Cerutti L., De Castro E.,Langendijk-Genevaux P.S., Pagni M., Sigrist C.J.A. The PROSITE database. Nucleic Acids Res. 34:D227-D230(2006).
34. Gooley A.A., Packer N.H. (1997) The Importance of Protein Co- and Post-Translational Modifications in Proteome Projects. In: Wilkins M.R., Williams K.L., Appel R.D., Hochstrasser D.F. (eds) Proteome Research: New Frontiers in Functional Genomics. Principles and Practice. Springer, Berlin, Heidelberg
35. Sang, L., Dick, I. E., & Yue, D. T. (2016). Protein kinase A modulation of CaV1.4 calcium channels. Nature communications, 7, 12239. https://doi.org/10.1038/ncomms12239
36. Catterall, W.A. Biochemical Studies of Voltage-Gated Ca2+ Channels. In: Madame Curie Bioscience Database [Internet]. Austin (TX): Landes Bioscience; 2000-2013. Available from: https://www.ncbi.nlm.nih.gov/books/NBK6526/
37. Blesneac, I., Chemin, J., Bidaud, I., Huc-Brandt, S., Vandermoere, F., & Lory, P. (2015). Phosphorylation of the Cav3.2 T-type calcium channel directly regulates its gating properties. Proceedings of the National Academy of Sciences of the United States of America, 112(44), 13705–13710. https://doi.org/10.1073/pnas.1511740112
38. Bers, D. M., & Grandi, E. (2009). Calcium/calmodulin-dependent kinase II regulation of cardiac ion channels. Journal of cardiovascular pharmacology, 54(3), 180–187. https://doi.org/10.1097/FJC.0b013e3181a25078
39. White, R. R., Kwon, Y. G., Taing, M., Lawrence, D. S., & Edelman, A. M. (1998). Definition of optimal substrate recognition motifs of Ca2+-calmodulin-dependent protein kinases IV and II reveals shared and distinctive features. The Journal of biological chemistry, 273(6), 3166–3172.

40. Crooks G.E., Hon G., Chandonia J.M., Brenner S.E. WebLogo: A sequence logo generator, Genome Research, 14:1188-1190, (2004).
41. Tadross, M.R., Ben Johny, M., and Yue, D.T. (2010). Molecular endpoints of Ca2+/calmodulin- and voltage-dependent inactivation of Ca(v)1.3 channels. J. Gen. Physiol. 135, 197–215.
42. Madeira F, Park YM, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research. 2019 Jul;47(W1):W636-W641. DOI: 10.1093/nar/gkz268.
43. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7:539 doi:10.1038/msb.2011.75
44. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.
45. PDB ID: 6BYO
Martinez-Ortiz, W., & Cardozo, T. J. (2018). An Improved Method for Modeling Voltage-Gated Ion Channels at Atomic Accuracy Applied to Human Cav Channels. Cell reports, 23(5), 1399–1408. https://doi.org/10.1016/j.celrep.2018.04.024
46. Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics, 23, 2947-2948.
47. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acid Res., 30:3059-3066
48. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics (Oxford, England), 25(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033
49. Berrou, L., Bernatchez, G., and Parent, L. (2001). Molecular determinants of inactivation within the I-II linker of alpha1E (CaV2.3) calcium channels. Biophys. J. 80, 215–228.
50. Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H. Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., … Mitchell, A. L. (2017). InterPro in 2017-beyond protein family and domain annotations. Nucleic acids research, 45(D1), D190–D199. https://doi.org/10.1093/nar/gkw1107
51. Letunic, I., Bork, P. 20 years of the SMART protein domain annotation resource, Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D493–D496, https://doi.org/10.1093/nar/gkx922.
52. Campbell, C. 3D structural alignment. *Lecture_4 online*. Johns Hopkins University: Protein Bioinformatics, Summer 2020. Slide 32.
53. Campbell, C. SECONDARY STRUCTURE OPTIONS in CLUSTALW. *Lecture_4 online*. Johns Hopkins University: Protein Bioinformatics, Summer 2020. Slide 26.

54. Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of molecular biology, 48(3), 443–453. https://doi.org/10.1016/0022-2836(70)90057-4
55. Smith, T. F., & Waterman, M. S. (1981). Identification of common molecular subsequences. Journal of molecular biology, 147(1), 195–197. https://doi.org/10.1016/0022-2836(81)90087-5
56. Rost, B., & Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. Journal of molecular biology, 232(2), 584–599. https://doi.org/10.1006/jmbi.1993.1413
57. Jones DT. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 292: 195-202.
58. Campbell, C. Loops (coils). *Lecture_6 online*. Johns Hopkins University: Protein Bioinformatics, Summer 2020. Slide 3.
59. PredictProtein–-an open resource for online prediction of protein structural and functional features. Yachdav, G.; Kloppmann, E.; Kajan, L.; Hecht, M.; Goldberg, T.; Hamp, T.; Hönigschmid, P.; Schafferhans, A.; Roos, M.; Bernhofer, M.; and others Nucleic acids research,gku366. 2014.
60. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410.
61. McDonough, S.I. (2013), Calcium ion channels: challenges and successes in drug discovery. WIREs Membr Transp Signal, 2: 85-104. doi:10.1002/wmts.71
62. Wei, F., Yan, L. M., Su, T., He, N., Lin, Z. J., Wang, J., Shi, Y. W., Yi, Y. H., & Liao, W. P. (2017). Ion Channel Genes and Epilepsy: Functional Alteration, Pathogenic Potential, and Mechanism of Epilepsy. Neuroscience bulletin, 33(4), 455–477. https://doi.org/10.1007/s12264-017-0134-1