RESEARCH ARTICLE

Three-dimensional Modelling of the Voltage-gated Sodium Ion Channel from *Anopheles gambiae* Reveals Spatial Clustering of Evolutionarily Conserved Acidic Residues at the Extracellular Sites

Rithvik S. Vinekar and Ramanathan Sowdhamini*

National Centre for Biological Sciences (TIFR), GKVK Campus, Bellary Road, Bangalore, India

Abstract: **Background**: The eukaryotic voltage-gated sodium channel(e-Nav) is a large asymmetric transmembrane protein with important functions concerning neurological function. No structure has been resolved at high resolution for this protein.

Methods: A homology model of the transmembrane and extracellular regions of an *Anopheles gambiae* para-like channel with emphasis on the pore entrance has been constructed, based upon the templates provided by a prokaryotic sodium channel and a potassium two-pore channel. The latter provides a template for the extracellular regions, which are located above the entrance to the pore, which is likely to open at a side of a dome formed by these loops.

Results: A model created with this arrangement shows a structure similar to low-resolution cryo-electron microscope images of a related structure. The pore entrance also shows favorable electrostatic interface.

Conclusion: Residues responsible for the negative charge around the pore have been traced in phylogeny to highlight their importance. This model is intended for the study of pore-blocking toxins.

1. INTRODUCTION

The eukaryotic voltage-gated sodium channel (e-Nav) is an essential component involved in the conduction of impulses in nerve tissue. The formation of action potential has been observed in *in vitro* patch-clamp experiments across membranes obtained from neurons. The Hodgkin-Huxley model [1-3] used electrical components to model electrical response across such a membrane. Besides membrane capacitance and voltage difference, voltage and time-dependent conductances were needed. These correspond to voltage-gated ion channels, consisted of two independent parts, *i.e.* potassium and sodium ion conductances respectively. Biochemical studies subsequently showed that two separate channel proteins, the eukaryotic voltage-gated potassium channel (Kv) and sodium channel (e-Nav) were responsible.

Many biochemical studies are devoted to understand the aspects of the structure of the Nav channel [4, 5] voltage sensing and gating [6, 7]. The structure of the simpler eukaryotic Kv channel [8] allowed detailed studies on the

e-Kv channel. The e-Nav consists of two subunits named alpha and beta. The alpha subunit is the main channel with the transmembrane region. Its secondary structure has been determined and shown to be similar to the e-Kv channel [8]. Prokaryotic Nav channel (p-Nav) structures were also determined [9, 10].

Both p-Nav and e-Kv channels form a homotetramer, with subunits around ~300-350 amino-acids (AA), with very similar conserved structure. The e-Nav and closely related eukaryotic voltage-gated calcium channels (e-Cav) [11] consist of one ~2000AA chain, with four duplicated and significantly diverged transmembrane domains (D1 to D4) corresponding to subunits of the homotetrameric p-Nav. Large protein regions in the extracellular and intracellular side between these repeats account for the significantly longer sequence than the total length of the p-Nav tetramer.

Atomic resolution structural determination of eukaryotic Nav channel is extremely challenging by virtue of its size, membrane-bound state and asymmetry. It falls out of scope for current x-ray crystallography and NMR techniques. Newer cryo-EM tomography methods are in their infancy and are better at resolving membrane associated molecules which exhibit symmetry [12], though these limitations are being challenged [13]. Low resolution 19Å cryo-EM density maps [14] of the structure of Nav from *Electrophorus*

*Address correspondence to this author at the National Centre for Biological Sciences (TIFR), GKVK Campus, Bellary Road, Bangalore, India; Tel: +91 80 23666250; Fax: +91 80 23636421; E-mail: mini@ncbs.res.in

neurons gives an approximate idea of the goal [15], while other studies add to the data [16]. The density maps show a bell-shaped molecule, with entrance to the channel present at the side of a large dome extending into the extracellular side, which is likely to be rich in helical structures [15]. The shape of the transmembrane region is, however, similar to the p-Nav channel. The importance of the Nav channel leads to the possibility that its structure may be determined in the near future. Currently however, homology modelling may be used to obtain a possible model.

Humans and mammals have greater diversity in sodium ion channels [11, 17, 18] compared to insects. Humans have around 10 paralogs of the sodium channel, with nomenclatures $Na_v1.1$ to $Na_v1.9$ and Na_x [19]. These Na_v nomenclatures are used mainly for mammalian homologs. They have been the target of drug research concerned with neurological disorders such as epilepsy. They are also a target for neurotoxins, both peptides and compounds, natural and artificial. Insecticides, such as pyrethroids, are lipophilic and target the channel from the lipid layer, preventing it from deactivating [20]. Pore-blocking toxins such as alpha-scorpion toxins [21] which are peptides, or toxins such as tetrodotoxin [TTX] [22] which are small molecules act by deactivating the channel. Different kinds of toxins and their actions on Nav channels have been classified [23] and is an active area of research. The extracellular loop regions play a major role in peptide toxin binding [24]. These may have evolved in an arms race between toxin producing organisms and target ion channels which evolved immunity to them: the evolution of immunity to toxins in garter snakes is one example [25].

This study concerns modelling the pore, transmembrane scaffold and the extracellular regions of an insect e-Nav channel – the malarial mosquito *Anopheles gambiae*. This model may then be subjected to theoretical studies such as molecular dynamics and docking to test scenarios and explain behaviour [26]. The human counterparts may also be modelled from this. The scope of this study currently does not consider aspects such as voltage gating or conductivity, though they are planned. Insect Nav channels have been extensively studied for the purpose of insecticide design against acquired resistance [28].

2. STRUCTURAL OVERVIEW

The voltage-gated sodium ion channel 'para' from mosquito *Anopheles gambiae* is represented by Uniprot [30] protein sequence A5I843_ANOGA [28] with a sequence length of 2139 residues. The sequence shows homology (~81.5% identity) with the annotated sequence SCNA_DROME corresponding to the *para* [31] channel of closely-related Diptera model organism *Drosophila melanogaster,* so named because knockout of this gene caused paralysis in the fly. There is no other known e-Nav paralog in *Drosophila*.

The closest homologues are the voltage-gated eukaryotic channel proteins with resolved complete structures, including both pore and voltage-sensing subdomains, are homo-tetrameric e-Kv channels represented by pdb: 3LUT [32], pdb:2A79 [33] and pdb:2R9R [34]. The p-Nav structures available are NavAb pdb:3RVY [9] from *Acrobacter butzleri* and its calcium-conducting mutant pdb:4MS2 [35] differing only by three residues near the selectivity filter region, NavRh pdb:4DXW [10] from *Rickettsiales sp.* HIMB114 representing a closed form, and the cryo-EM structure pdb:4BGN [36] from *Caldalkalibacillus thermarum*.

The e-Cav and e-Nav channels have four duplicated transmembrane domains each homologous to the e-Kv and p-Nav sequence fused into a single chain and are therefore pseudotetramers. Each domain consists of transmembrane helices S1-S6 and other structures in between, such as the intermembrane Pore helix and the filter loop. Substantial segments of the protein between these transmembrane domains may form supersecondary structures present on the cytosolic side, of currently unknown function. Only the cytosolic C-terminal region has been determined to have a calcium-binding (pdb:4JPZ) [27], but this is irrelevant to the current study.

The four domains (see Fig. **1**) are D1:133–463(330), D2:731–1059(328), D3:1287–1596(309) and D4:1605–1855
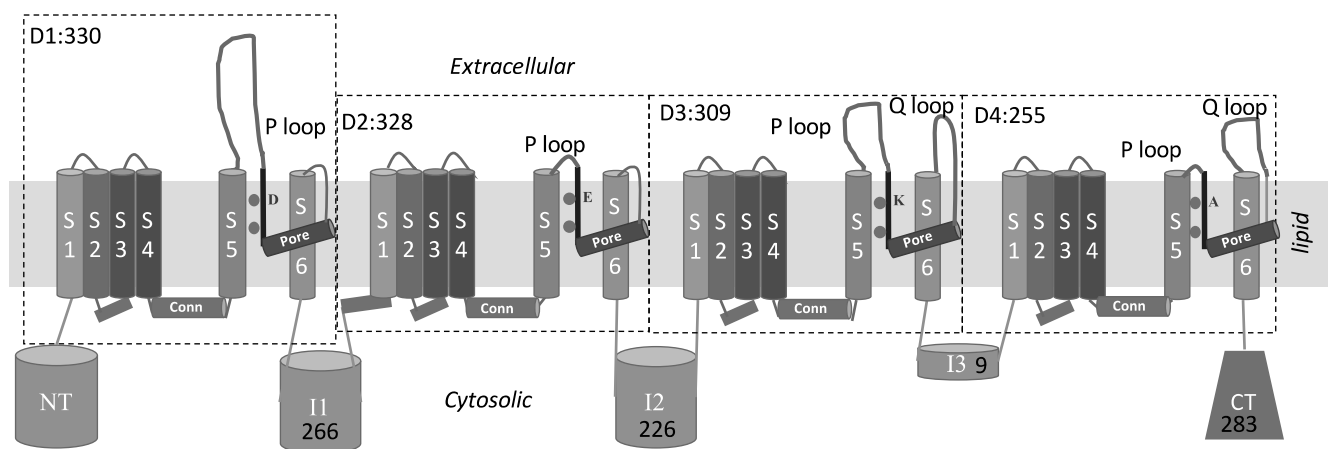

Fig. (1). Structure of Nav channel. The NT (N-terminal), I1, I2 I3 and CT (C-terminal) regions are represented as bins because their secondary structure is unknown and irrelevant to this study. The structure of the CT region however has been determined for human Nav channels, and is shown to have a calcium binding function, interacting with calmodulin and other factors. The circles indicate position of ion near the filter loop, which contribute to the DEKA tetrad.

(255) and intermediate regions I1:464–730(266), I2:1060–1286(226), I3:1596–1605(9) and CT:1856–2139(283). The I3 section is functionally important and called the 'Na channel gate' or IFM loop. The I1 and I2 regions are called the DUF3451 superfamily and the Na-trans_assoc (transport associated) subfamily domain respectively, which are both conserved across e-Nav but with unknown function and structure. Each of the duplicated domain (D1..D4) have diverged significantly from each other (Table **1**), and the core channel is composed of four related but very different units. The low sequence identity of <20% however is compensated by high structural similarity and compatibility to the channel of interest.

Table 1. **Sequence identity of the four duplicated domains in A51843_ANOGA.**

	D1	D2	D3	D4
D1	100			
D2	23.84	100		
D3	22.09	24.27	100	
D4	20.57	29.26	24.14	100
3RVY	19.8	19.9	20.9	24.3
KCNA2_rat	16.31	15.91	14.9	21.59

For each of these repeats, four helices (S1..S4) form the voltage sensing subdomain, with helix S4 characterized by alternating R (arginine) and sometimes K (lysine) residues whose protonation/deprotonation can affect the shape of the pore. The pore itself is defined by the pore subdomain, consisting of helices S5, S6, the pore helix and the filter region. The connecting helix, labelled 'Conn' transmits the force necessary for changes in orientation in the voltage-sensing subdomain to the pore subdomain.

The filter residues, shown by circles representing ions in Fig. (**1**), have a characteristic conserved DEKA tetrad motif or filter contributed by D1 to D4 respectively. In addition, the p-loop region extends into the extracellular part in domain D1 and D3. The D3 and D4 domains however have extracellular extension labelled q-loop by analogy with the p-loop, and, in the case of D4, a very short p-loop. D2 has a longer C-terminal side region (represented by additional helix), and a short p-loop, longer than that of D4. The transmembrane regions consist of four repeats, which have diverged from each other but still retain basic common structure.

Additional cytosolic side I1...I3 and CT regions of the protein are not modelled, and the transmembrane region is treated as a heterotetramer, and accessory subunits ignored. Around 30% of the structure is glycosylated [37] by N-linked glycans with negative charge, such as sialic acid. Removal of these glycans does not impair function or folding [37]. However, this adds a constraint to any model, as predicted glycosylated Asparagine residues must be exposed or surface-accessible for glycosylation.

Homology modelling has been used to model the protein [20] restricted to domains D2 and partially domain D3, as this region is known to interact with the lipophilic DDT (dichlorodiphenyltrichloroethane) and pyrethroid pesticides such as pyrethrin I. The goal of these studies were to understand the cause of immunity gained by insects to these insecticides. Other attempts at modelling these channels are seen on ModBase [38] and elsewhere [39], concentrate on one or another singular domain. A recent attempt at homology modelling looks at the pore region exclusively [40] and has been used for studying the permeability of the channel. The extracellular loop region has been ignored in this model, treated as loops, and the helices S5 and S6, the pore helix and filter region are the main focus of the study. An attempt to model the extracellular loops in the eel *electrophorus* Nav channel used a postulate that the structure may be an alpha-turn-beta structure [37, 41], but this hypothesis has not been verified. The low sequence identity to other membrane proteins of known structure and lack of secondary structure predicted by commonly used secondary structure prediction methods such as PSIPred [42] or JPred [43], adds to the challenge of modelling these portions of the protein.

The human two-pore potassium channel pdb:3UKM [44] is a dimer representing a channel protein with extracellular regions. Each subunit of the dimer may be split or cleaved into two parts, each correspond to parts of a repeat domain of the voltage gated channel. This part representing the core helices S5, pore helix and filter region, and helix S6. Only the first part has an extracellular extension, corresponding to that of D2 and D4.

3. METHODS

3.1. Full-length Alignment and Phylogeny

The target putative sequence of sodium channel protein A5I843_ANOGA [28] from *Anopheles gambiae* obtained from UNIPROT [30] was aligned first with other known sodium channel proteins. The Diptera para channel protein SCNA_DROME [14] from *Drosophila melanogaster* was included, as it is well-annotated. Two mammalian sodium channel proteins, one from human and one from rat were included in the alignment.

The detailed annotations in SCNA_DROME were used to determine boundaries of other sequences by alignment. The four transmembrane domain repeat boundaries D1, D2, D3 and D4 were marked and used to delineate similar boundaries on the other sequences. Transmembrane helices as well as the intermembrane helices which denote S1-S6 and the pore helix respectively were marked. The characteristic filter sequences are aligned.

This alignment with around 1500–2000+ residues was used as a reference, and not for modelling purposes, as there is no structure of a eukaryotic sodium or calcium channel resolved yet. The JalView [45] interface was used, and the alignment programs MUSCLE [46], CLUSTALW_Realign [47] and MAFFT [48] were used *via* this interface, to obtain the alignment. This alignment was extended by addition of total 242 e-Nav sequences obtained from UNIPROT [30], sampling deuterostomes such as amniotes, fishes, tunicates

etc. and protostomes such as annelids, molluscs and specifically insects (see Supplementary file 1). RAxML [49] was used with the maximum likelihood method to obtain phylogenetic tree with bootstrap 100 (Supplementary files 2 and 3).

4. PER-REPEAT ALIGNMENTS

Each repeat D1 - D4 were aligned to each subunit of the tetrameric template 4MS2 and e-Kv sequences. Repeats D1 and D3 were aligned to type A subunit of the template 3UKM and repeats D2 and D4 to type B subunit of the template 3UKM to yield a heterotetramer (3UKM_mod). Flexible structural superposition was used to obtain the initial sequence alignment using MATT [50] and the alignment was adjusted manually by visual inspection.

5. LOOP ALIGNMENTS

The extracellular p-loops of domains D1..D4 were separately aligned. It was observed that these domains were either conserved in length (D2, D4) or variable (D1) but have common conserved segments. These conserved segments were used to define putative secondary structure of the extracellular loops.

The conserved segments may represent helices, and the extracellular region may be similar to the structure seen in 3UKM. The loop alignments for D1, D3 and D4 are shown in Figs. (**2**, **3** and **4**).

6. MODELLING THE DOMAINS

The domain alignments were concatenated into a longer alignment. A short linker (I3) between D3 and D4 was also included with its corresponding secondary structure predictions. The program MODELLER [51] version 9.15 was used. The whole tetrameric structures of templates 4MS2 and 3UKM_

mod were used as templates. While pdb:3RVY represents a sodium channel and pdb:4MS2 a calcium mutant, there is very little root mean square difference (RMSD) between the two. Ions are not present in pdb:3RVY. Calcium ions present in pdb:4MS2were used as a structural reference.

The extracellular loops in D1 and D3 differ in size, unlike the template 3UKM with which they align. The loops from D1 seem to contribute three helices and that from D3 contributes one helix. The structure obtained is similar to the template, which has two helices from each subunit. One opening is left open to the channel, while the other side is blocked by the extra loops connecting the helices. Trial and error adjustments lead to the putative model shown in Fig. **5**.

The secondary structure predictions and disulphide connections were incorporated in the MODELLER script as restraint patches. Ions were treated as 'block residues' and distances from backbone atoms in interacting residues to stationary ions were added as restraint to aid modelling the filter region. To reproduce the putative arrangement shown in Fig. (**5**), the 3UKM_mod was included again, with most parts unaligned and outside the main alignment (Fig. **6**). Only the the second helix of the second subunit (h4 in Fig. **5**) extracellular region was aligned with the third helix position of D1 (For complete alignment, see Supplementary file 4).

The presence of several cysteines in the extracellular loops suggest that many disulphide bonds combinations are possible. Two such combinations were tested, one involving C207 and one involving C216 from D1 in the bond with C758 on D3. 20 models for each set were generated and ranked. Further refinement was done to understand the possibilities. Top-ranked models which showed unnatural knots and entanglements between the subunits were discarded. The current model discussed involves a bond involving C207. A conserved cysteine is also present on the D4 q-loop.
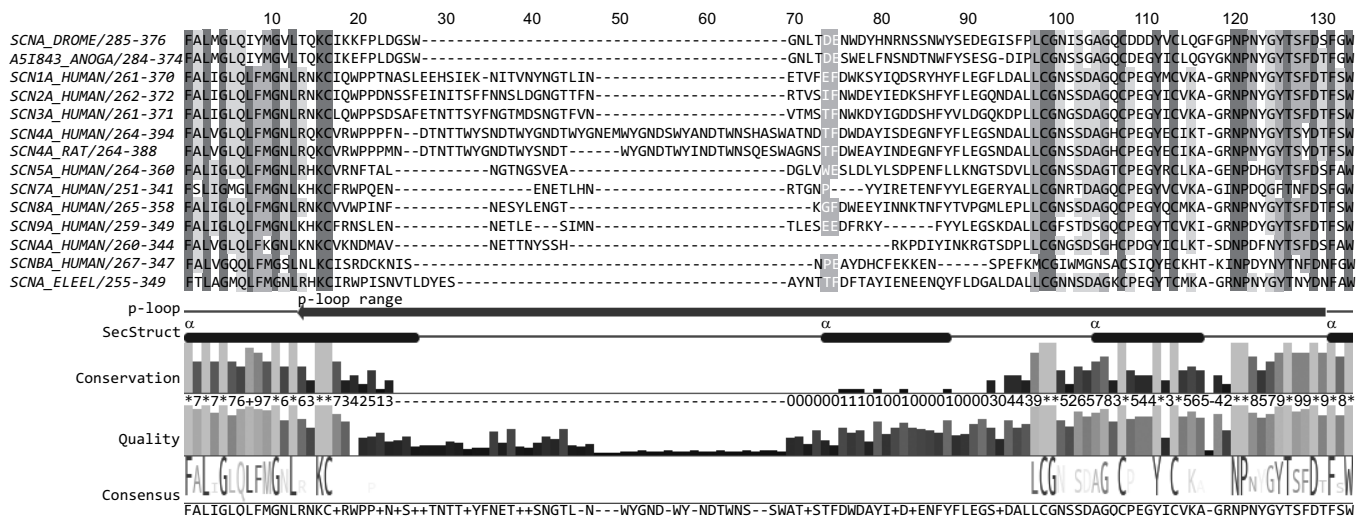
Fig. (2). Alignment of the D1 (domain 1) p-loop region and some flanking transmembrane regions. The *Anopheles* sequence is represented by A5I843_ANOGA and *Drosophila* by SCNA_DROME. The alignment is colored by sequence identity. The SecStruct annotation gives the secondary structure used in the modelling, which takes into account the conserved regions. Human SCNs have longer lengths, with SCN4A having the longest, despite clustering close to SCNA_DROME.
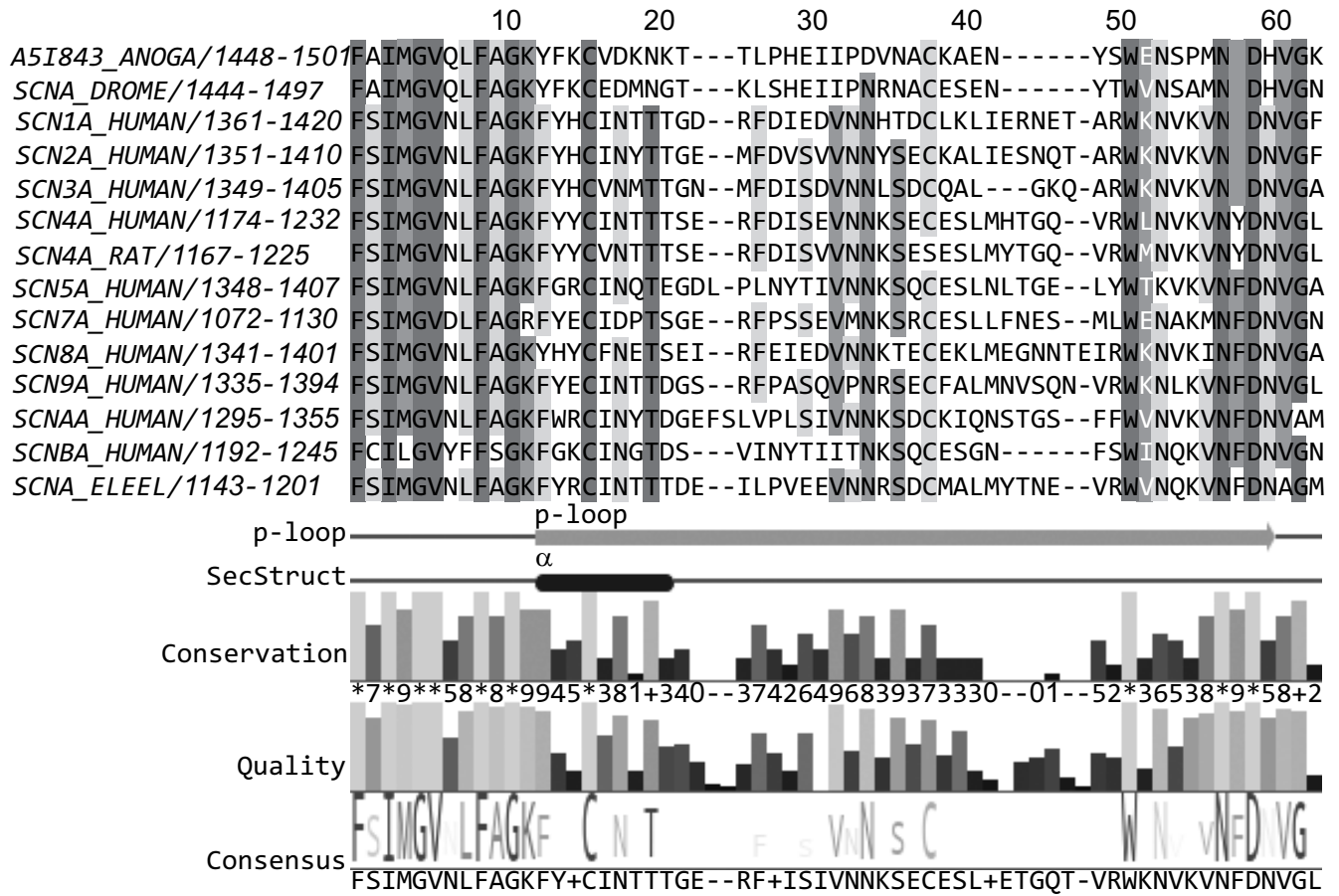
```
                                              10        20        30        40        50        60
A5I843_ANOGA/1448-1501  FAIMGVQLFAGKYFKCVDKNKT---TLPHEIIPDVNACKAEN------YSWENSPMN DHVGK
SCNA_DROME/1444-1497    FAIMGVQLFAGKYFKCEDMNGT---KLSHEIIPNRNACESEN------YTWVNSAMN DHVGN
SCN1A_HUMAN/1361-1420   FSIMGVNLFAGKFYHCINTTTGD--RFDIEDVNNHTDCLKLIERNET-ARWKNVKVN DNVGF
SCN2A_HUMAN/1351-1410   FSIMGVNLFAGKFYHCINYTTGE--MFDVSVVNNYSECKALIESNQT-ARWKNVKVN DNVGF
SCN3A_HUMAN/1349-1405   FSIMGVNLFAGKFYHCVNMTTGN--MFDISDVNNLSDCQAL---GKQ-ARWKNVKVN DNVGA
SCN4A_HUMAN/1174-1232   FSIMGVNLFAGKFYYCINTTTSE--RFDISEVNNKSECESLMHTGQ--VRWLNVKVNYDNVGL
SCN4A_RAT/1167-1225     FSIMGVNLFAGKFYYCVNTTTSE--RFDISVVNNKSESESLMYTGQ--VRWVNVKVNYDNVGL
SCN5A_HUMAN/1348-1407   FSIMGVNLFAGKFGRCINQTEGDL-PLNYTIVNNKSQCESLNLTGE--LYWTKVKVNFDNVGA
SCN7A_HUMAN/1072-1130   FSIMGVDLFAGRFYECIDPTSGE--RFPSSEVMNKSRCESLLFNES--MLWENAKMNFDNVGN
SCN8A_HUMAN/1341-1401   FSIMGVNLFAGKYHYCFNETSEI--RFEIEDVNNKTECEKLMEGNNTEIRWKNVKINFDNVGA
SCN9A_HUMAN/1335-1394   FSIMGVNLFAGKFYECINTTDGS--RFPASQVPNRSECFALMNVSQN-VRWKNLKVNFDNVGL
SCNAA_HUMAN/1295-1355   FSIMGVNLFAGKFWRCINYTDGEFSLVPLSIVNNKSDCKIQNSTGS--FFWVNVKVNFDNVAM
SCNBA_HUMAN/1192-1245   FCILGVYFFSGKFGKCINGTDS---VINYTIITNKSQCESGN------FSWINQKVNFDNVGN
SCNA_ELEEL/1143-1201    FSIMGVNLFAGKFYRCINTTTDE--ILPVEEVNNRSDCMALMYTNE--VRWVNQKVNFDNAGM
```

p-loop ──────────────── p-loop ──▶

SecStruct   α
              ━━━━━━━━━

Conservation
```
*7*9**58*8*9945*381+340--37426496839373330--01--52*36538*9*58+2
```

Quality

Consensus
```
FSIMGVNLFAGKFY+CINTTTGE--RF+ISIVNNKSECESL+ETGQT-VRWKNVKVNFDNVGL
```

Fig. (3). Alignment of the D3 p-loop. It shows the conserved cysteines, and highlights important acidic residues.

The resultant model shows possibility of more refinement and adjustment, such as the possibility of an anti-parallel beta sheet formed by the q-loop of D4 and these extra loops, as shown in Fig. (**5**). I-TASSER [52] predicted secondary structure predict the presence of beta strands which may form these. However, at this stage, the focus is on the opening to the pore. The resultant extracellular region in top view is shown in Fig. (**7**).

7. SIMULATION

A short simulation was carried out to check the stability of the model in a lipid-embedded environment. GROMACS 5.1 [53] was used. The membed [54] protocol was used to embed it in a dipalmitoylphosphatidylcholine(DPPC) pre-equilibrated and extended bilayer provided by Slipid [55]. The position of membrane was calculated using the PPM server [56, 57] and VMD [58] was used to align the bilayer to the protein. The Slipid parameters in conjunction with AMBER FF99SB [59] forcefield modified for Gromacs (ffamber) [60, 61]. Simulation was carried out for 10ns after a 1ns equilibration. 'Block' residues in the MODELLER generated model were replaced by Na+ and corresponding 4 neutralizing Cl- ions added. A total of 375468 particles, with 1101 protein residues, 1062 DPPC residues and 73146 water molecules was run on a GPGPU-based system. The RMSD and radius of gyration plots are provided in supplementary files 5 and 6 respectively.
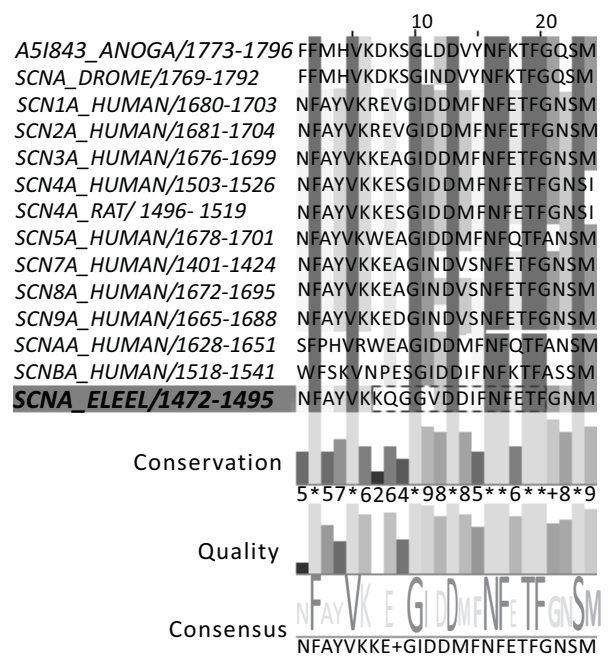
```
                                   10        20
A5I843_ANOGA/1773-1796  FFMHVKDKSGLDDVYNFKTFGQSM
SCNA_DROME/1769-1792    FFMHVKDKSGINDVYNFKTFGQSM
SCN1A_HUMAN/1680-1703   NFAYVKREVGIDDMFNFETFGNSM
SCN2A_HUMAN/1681-1704   NFAYVKREVGIDDMFNFETFGNSM
SCN3A_HUMAN/1676-1699   NFAYVKKEAGIDDMFNFETFGNSM
SCN4A_HUMAN/1503-1526   NFAYVKKESGIDDMFNFETFGNSI
SCN4A_RAT/1496-1519     NFAYVKKESGIDDMFNFETFGNSI
SCN5A_HUMAN/1678-1701   NFAYVKWEAGIDDMFNFQTFANSM
SCN7A_HUMAN/1401-1424   NFAYVKKEAGINDVSNFETFGNSM
SCN8A_HUMAN/1672-1695   NFAYVKKEAGINDVSNFETFGNSM
SCN9A_HUMAN/1665-1688   NFAYVKKEDGINDVSNFETFGNSM
SCNAA_HUMAN/1628-1651   SFPHVRWEAGIDDMFNFQTFANSM
SCNBA_HUMAN/1518-1541   WFSKVNPESGIDDIFNFKTFASSM
SCNA_ELEEL/1472-1495    NFAYVKKQGGVDDIFNFETFGNM
```

Conservation
```
5*57*6264*98*85**6***8*9
```

Quality

Consensus
```
NFAYVKKE+GIDDMFNFETFGNSM
```

Fig. (4). Alignment of the D4 p-loop. This loop forms the entrance to the pore in the model and has conserved DD or ND motif. Asparagine (N) residues serve as attachment points for glycans like Sialic Acid which are negatively charged and enhance the functioning of the channel.
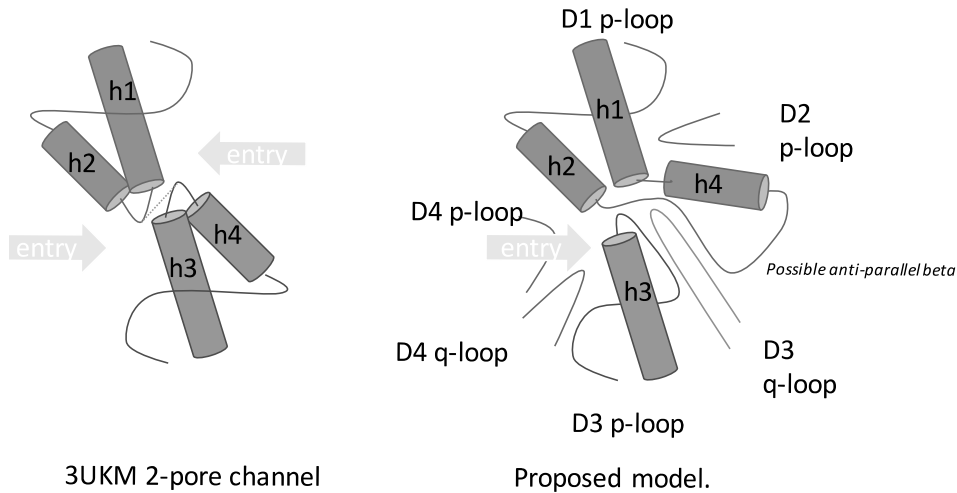
Fig. (5). Proposed secondary structure configuration of the extracellular region, fitted to the template provided by the human two-pore potassium channel structure pdb:3UKM. The proposed structure takes into account the length of D1 and D3 p-loops and their conservation in alignment. At least 4 cysteine residues are present, with the KC sequence common for D1 and D3 p-loops at around the same place, implying symmetry. Sequence-based structure prediction does not predict such a structure, due to absence of structural representatives in known structures. The D2 p-loop provides the negative electrostatics necessary at the pore by the presence of aspartates, with D1 and D3 providing additional aspartate and glutamates. Attachment points for N-linked glycans are also present around the pore entrance.
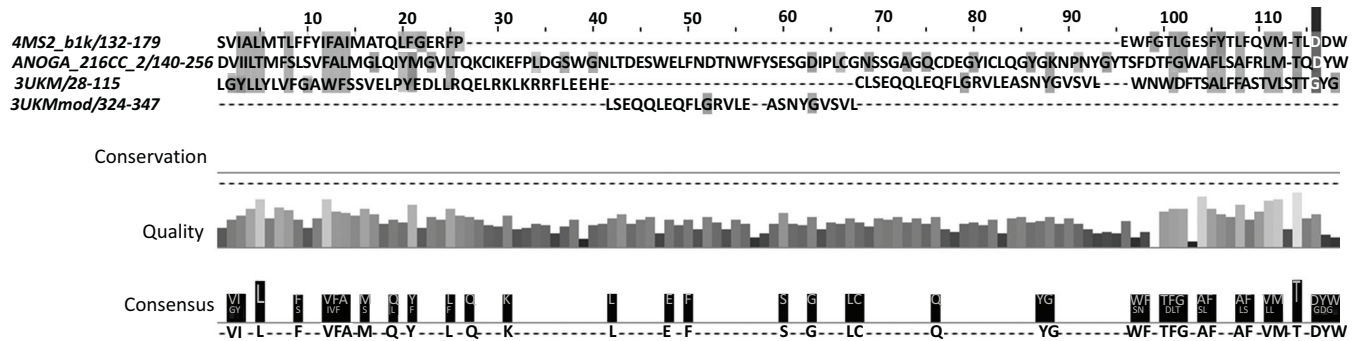
Fig. (6). The D1 p-loop region alignment used for generating the model using MODELLER. The 3UKM structure was included twice. One inclusion aligned the S5, S6 and filter residues with the corresponding sodium channel 4MS2 (with calcium ions renamed as BLK residues). Helix h1 and h2 aligned with this inclusion. The second 3UKM inclusion aligned only helix h4, with the rest of the sequence unaligned. Both inclusions of 3UKM were structurally aligned and essentially the same structure file. The filter residue of D1, which is the D of the DEKA tetrad, is marked in alignment at extreme right.

8. RESULTS

Whole sequence phylogeny analysis using eukaryotic channel Nav sequences shows that insect Nav sequences like *Anopheles gambiae* para-like and *Drosophila* para channel cluster close to the SCN4A sequence of human sodium channels.

The p-loops of the domains D1-D4 show variability in length in D1. Fig. **2** shows the alignment of the p-loop of the D1 domain, and some flanking transmembrane regions. *Anopheles gambiae* (ANOGA) sequence, the Drosophila homolog para (SCNA_DROME), human representatives, one rat representative and a channel from the electric eel *Electrophorus* have been included. The latter is performed because of the availability of its cryo-EM structure [15].

The sequence of SCN4A and the insect channels is particularly different. The SCN4A human has repeats of the sequence -WYGNDS- with point mutations. The SCN4A rat sequence is similar with one such repeat conspicuously absent (see Fig. **2**). These additional sequence lengths could contain more secondary structures or disordered regions or glycan binding sites that may prevent certain toxins from binding. The model in Fig. **5** can account for these, as the additions would be in the loop regions, conserving the overall scaffold. The *Electrophorus* channel sequence however has similar length to the insect channel sequence.

The model chosen here gave a favorable DOPE score, *i.e.* the lowest score model without showing structural abnormalities such as interchain knots. It also showed a well-formed pore entrance, similar to that predicted in the cryo-em data [15]. The cryo-EM data suggests a dome-shaped structure, with a pore entrance to one of the sides, not directly open vertically as in most resolved voltage-gated channel structures. Our model postulates a possible arrange-
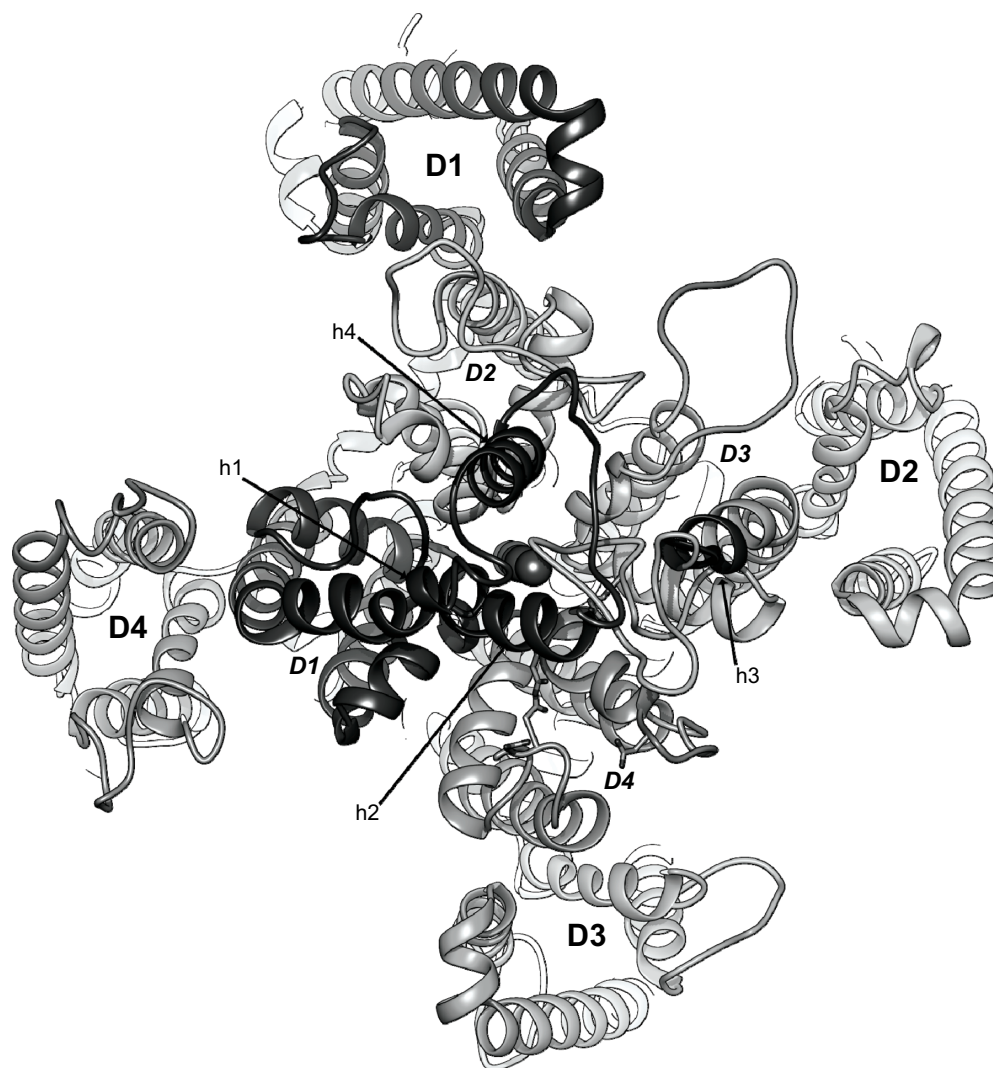
Fig. (7). Configuration of the modelled extracellular regions with lower areas clipped. The ions which were included as block residues are seen as spheres in center. This is comparable to the schematic in Fig. **5**. The domains are labelled D1 to D4, with italic labels denoting the pore subdomain and bold normal text denoting the voltage sensing subdomain. The helix positions are labeled h1 to h4.

ment of secondary structures forming this dome. The charge distribution around the entrance to the pore is, as expected in cation channels, with an overall negative charge. There are also prominent pockets with attachment points for ASN N-linked glycans. While the ASN residues themselves may contribute a net positive charge, negatively charged glycans like sialic acid can provide a conducive electrostatic environment for attracting cations to the mouth of the channel (Fig. **8**).

The residues responsible for providing this negative charge were traced in an alignment for the four domain p-loops. In particular, the domains D1, D3 and D4 are mainly responsible for contributing this charge. Five such residues have been noted, numbered GLU 218 (350 as in UNIPROT numbering), GLU 762 (1490), ASP 1056 (1784), ASP 1057 (1785) and ASP 1092 (1820) in the resultant model. The numbers in brackets are sequence numbers on single-chain Uniprot sequence A5I843_ANOGA. The first is present on D1, second on D3, the third and fourth on the floor of the

entrance on D4 p-loop and fifth on the D4 q-loop as seen in Fig. (**9**).

Conserved acidic residues are also present on D2. However, loop of D2 does not appear to be spatially proximate to the pore entrance in this model. There is, however, a possibility that the side chains of these residues may line the inner part of the pore within the "dome" as seen in the cryo-em image.

Residue 218 (350) (highlighted in Fig. **2**), along with a neighboring acidic aspartate residue, is present in insect ion channels, but not present in other ion channels. A toxin specific to binding these residues may be specific to insect channels.

The residue 762 (1490) is marked and is not conserved even in insects (Fig. **3**). The consensus sequence residue for these representative sequences is Lys. Residues 1056(1784), 1057(1785) are highly conserved on the D4 p-loop as seen in Fig. (**4**)-mostly as a sequence of two residues DD or sometimes
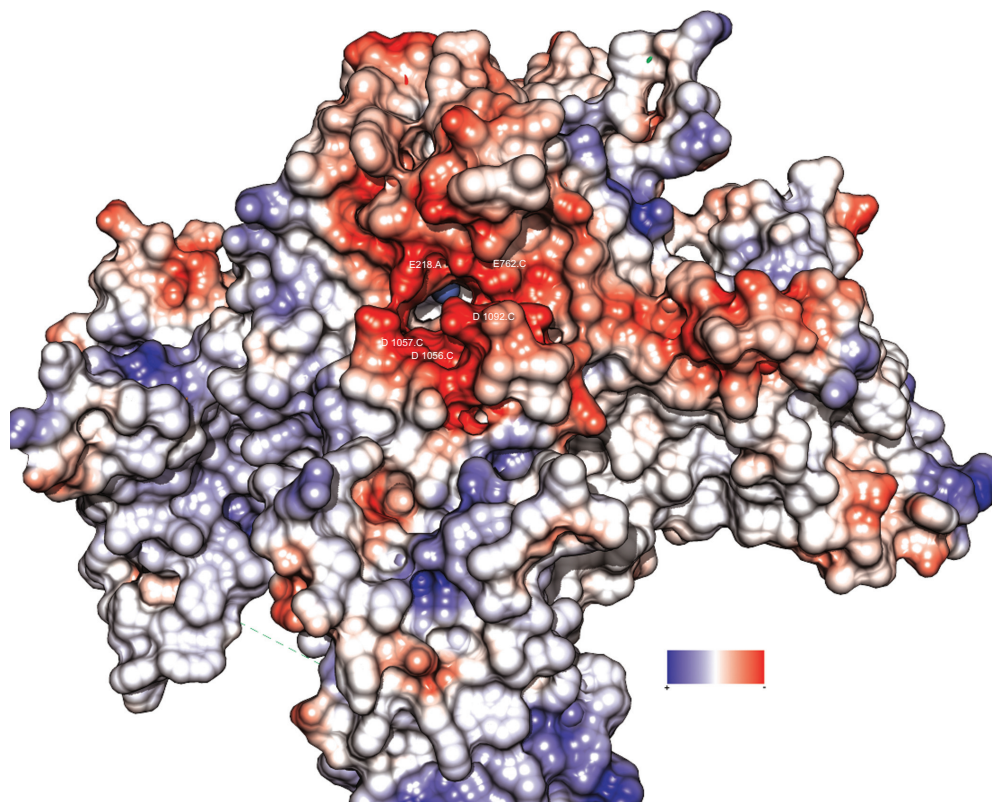
Fig. (8). Electrostatic surface map of the pore entrance. The contributing residues labelled E.218 is on D1 p-loop, E762 is on D3 p-loop, D1057 and D1058 are on D4 p-loop and D1092 is on D4 q-loop. The D4 p-loop DD or ND motif is conserved in Nav channels. Some ASN (N) residues present serve as attachment points for N-linked glycans such as sialic acid which contribute further to the negative charge in the entrance of the sodium channel and enhance its function.

Fig. (9). The ribbon view of structure shown in Fig. (**7**). The acidic residues surrounding the mouth of the pore are visible.

ND. Residue 1092 (1820) is present on the D4 q-loop. The D4 q-loop also has a highly conserved cysteine. The per-residue conservation and overall alignment with possible templates is shown in Supplementary File 9.

Minimization and simulation showed that the secondary structure of the extracellular loops remain stable. Minimization was done after placing the protein in the lipid bilayer. The ProSA–web interface [62] was used to assess the equilibrated model (Supplementary 7,8). It showed slightly lower quality in the transmembrane regions, which is expected as ProSA assumes a soluble molecule. However, the extracellular regions have a good score, with an overall Z-score of -4.58. A single bend near residue K755 in the D3 p-loop was the only part with unfavourable energy. Sodium ions move in and out of the channel during the simulation from the extracellular side, showing free movement in the absence of a voltage or ion concentration gradient. In the presence of a gradient, the ions may move into the channel. The negative electrostatic areas near the channel mouth prevent the ion from going too far and allow it to return.

CONCLUSIONS

This study shows a possible model of the sodium ion channel from an insect *Anopheles gambiae*. *Anopheles* is the vector for *Plasmodium falciparum* which causes malaria. The model shows a clearly formed pore with favorable electrostatics. The model mimics the two-pore channel in its extracellular regions. The two-pore channel and the Nav channel may be share homology and relationship of the extracellular regions of the pore channel may be similar to the e-Nav channels with their p-Nav homologs. While the basic architecture is similar, compensatory reshuffling where helix h4 (Fig. **2**) has been incorporated into D1 from D3 allows the pore helix to have a singular opening.

The favourable electrostatics seen near the entrance to the pore is contributed by acidic residues such as aspartates and glutamates (D, E) which surround the entrance. Some of these aspartates or glutamates are conserved, while some are very specific to insects. The DE motif in *Anopheles* Nav D1 p-loop is not present in mammalian channels, which have a proline in its place.

Previous knowledge on three-dimensional information of such transporters were limited to either low-resolution or not the entire transmembrane domain. To our knowledge, this is the first modelling attempt of the full transmembrane domain of sodium channel of *Anopheles gambie*, including all the extracellular loop regions. The current model has satisfied several structural compatibility and stability criteria. There is, however, scope for improvement and refinement.

The model will be employed for future studies. A large number of chemical compounds are known to interact with this protein, particularly pyrethroids, DDT and other insecticides [20]. A preliminary blind docking using Autodock Vina 1.1.2 [29], utilizing the lipophilic Pyrethrin-1 as ligand, showed the docking site in the transmembrane region between domains D2 and D3, similar to the case described by O'Reilly *et al*. [20]. The mode is shown in Supplementary file 10.

The current focus is to understand the specificity of pore-blocking toxins towards sodium channels. The interacting residues of the extracellular loops with these toxins and their conservation in phylogenetic clades may help us understand toxin specificity to an insect such as *Anopheles gambiae* and related clades, while not affecting other insects or mammals.

Blind docking using the soluble tetrodotoxin as ligand showed that this ligand is at least capable of entering through the extracellular loop region to interact with the filter residues as shown by Lukacs *et al*. [40]. However, it is still some distance away. This mode is shown in Supplementary Fig. (**11**).

Molecular dynamics simulations may further show interactions with the extracellular loops before the toxin moves into the pore. Apart from the filter residues, these regions may play an important role in toxin-channel interaction and cannot be ignored.

In future, such modelling studies could enable the detailed examination of human sodium channels, which are active targets for drugs treating epilepsy and other neurological disorders [63].

LIST OF ABBREVIATIONS

e-Nav = Eukaryotic Voltage gated sodium channel

p-Nav = Prokaryotic Voltage gated sodium channel

e-Kv = Eukaryotic Voltage gated potassium channel

e-Cav = Eukaryotic Voltage gated calcium channel

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

REFERENCES

[1]   Hodgkin, A.L.; Huxley, A.F.; Katz, B. Measurement of current-voltage relations in the membrane of the giant axon of Loligo. *J. Physiol.,* **1952**, *116*(4), 424-448. [http://dx.doi.org/10.1113/jphysiol.1952.sp004716] [PMID: 14946712]

[2]   Hodgkin, A.L.; Huxley, A.F. Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo. *J. Physiol.,* **1952**, *116*(4), 449-472. [http://dx.doi.org/10.1113/jphysiol.1952.sp004717] [PMID: 14946713]

[3]   Hodgkin, A.L.; Huxley, A.F. The dual effect of membrane potential on sodium conductance in the giant axon of Loligo. *J. Physiol.,* **1952**, *116*(4), 497-506. [http://dx.doi.org/10.1113/jphysiol.1952.sp004719] [PMID: 14946715]

[4]   Charalambous, K.; O'Reilly, A.O.; Bullough, P.A.; Wallace, B.A. Thermal and chemical unfolding and refolding of a eukaryotic sodium channel. *Biochim. Biophys. Acta - Biomembr.,* **2009**, *1788*, 1279-1286.

[5]   Duclohier, H. Structure-function studies on the voltage-gated sodium channel. *Biochim. Biophys. Acta,* **2009**, *1788*(11), 2374-2379. [http://dx.doi.org/10.1016/j.bbamem.2009.08.017] [PMID: 19747894]

[6]   Bagnéris, C.; Naylor, C.E.; McCusker, E.C.; Wallace, B.A. Structural model of the open-closed-inactivated cycle of prokaryotic voltage-gated sodium channels. *J. Gen. Physiol.,* **2015**, *145*(1), 5-16. [http://dx.doi.org/10.1085/jgp.201411242] [PMID: 25512599]

[7]   Bezanilla, F. How membrane proteins sense voltage. *Nat. Rev. Mol. Cell Biol.,* **2008**, *9*(4), 323-332. [http://dx.doi.org/10.1038/nrm2376] [PMID: 18354422]

[8]   Choe, S. Potassium channel structures. *Nat. Rev. Neurosci.,* **2002**, *3*(2), 115-121. [http://dx.doi.org/10.1038/nrn727] [PMID: 11836519]

[9]   Payandeh, J.; Scheuer, T.; Zheng, N.; Catterall, W.A. The crystal structure of a voltage-gated sodium channel. *Nature,* **2011**, *475*(7356), 353-358. [http://dx.doi.org/10.1038/nature10238] [PMID: 21743477]

[10]   Zhang, X.; Ren, W.; DeCaen, P.; Yan, C.; Tao, X.; Tang, L.; Wang, J.; Hasegawa, K.; Kumasaka, T.; He, J.; Wang, J.; Clapham, D.E.; Yan, N. Crystal structure of an orthologue of the NaChBac voltage-gated sodium channel. *Nature,* **2012**, *486*(7401), 130-134. [PMID: 22678295]

[11]   Goldin, A.L. Evolution of voltage-gated Na(+) channels. *J. Exp. Biol.,* **2002**, *205*(Pt 5), 575-584. [PMID: 11907047]

[12]   Lučič, V.; Rigort, A.; Baumeister, W. Cryo-electron tomography: the challenge of doing structural biology in situ. *J. Cell Biol.,* **2013**, *202*(3), 407-419. [http://dx.doi.org/10.1083/jcb.201304193] [PMID: 23918936]

[13]   Lu, P.; Bai, X.C.; Ma, D.; Xie, T.; Yan, C.; Sun, L.; Yang, G.; Zhao, Y.; Zhou, R.; Scheres, S.H.; Shi, Y. Three-dimensional structure of human γ-secretase. *Nature,* **2014**, *512*(7513), 166-170. [http://dx.doi.org/10.1038/nature13567] [PMID: 25043039]

[14]   Loughney, K.; Kreber, R.; Ganetzky, B. Molecular analysis of the para locus, a sodium channel gene in *Drosophila. Cell,* **1989**, *58*(6), 1143-1154. [http://dx.doi.org/10.1016/0092-8674(89)90512-6] [PMID: 2550145]

[15]   Sato, C.; Ueno, Y.; Asai, K.; Takahashi, K.; Sato, M.; Engel, A.; Fujiyoshi, Y. The voltage-sensitive sodium channel is a bell-shaped molecule with several cavities. *Nature,* **2001**, *409*(6823), 1047-1051. [http://dx.doi.org/10.1038/35059098] [PMID: 11234014]

[16]   Sato, C.; Sato, M.; Iwasaki, A.; Doi, T.; Engel, A. The sodium channel has four domains surrounding a central pore. *J. Struct. Biol.,* **1998**, *121*(3), 314-325. [http://dx.doi.org/10.1006/jsbi.1998.3990] [PMID: 9704503]

[17]   Catterall, W.A. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron,* **2000**, *26*(1), 13-25. [http://dx.doi.org/10.1016/S0896-6273(00)81133-2] [PMID: 10798388]

[18]   Wood, J.N.; Baker, M. Voltage-gated sodium channels. *Curr. Opin. Pharmacol.,* **2001**, *1*(1), 17-21. [http://dx.doi.org/10.1016/S1471-4892(01)00007-8] [PMID: 11712529]

[19]   Catterall, W.A.; Goldin, A.L.; Waxman, S.G. International Union of Pharmacology. XLVII. Nomenclature and structure-function relationships of voltage-gated sodium channels. *Pharmacol. Rev.,* **2005**, *57*(4), 397-409. [http://dx.doi.org/10.1124/pr.57.4.4] [PMID: 16382098]

[20]   O'Reilly, A.O.; Khambay, B.P.; Williamson, M.S.; Field, L.M.; Wallace, B.A.; Davies, T.G. Modelling insecticide-binding sites in the voltage-gated sodium channel. *Biochem. J.,* **2006**, *396*(2), 255-263. [http://dx.doi.org/10.1042/BJ20051925] [PMID: 16475981]

[21]   Campos, F.V.; Chanda, B.; Beirão, P.S.; Bezanilla, F. Alpha-scorpion toxin impairs a conformational change that leads to fast inactivation of muscle sodium channels. *J. Gen. Physiol.,* **2008**, *132*(2), 251-263. [http://dx.doi.org/10.1085/jgp.200809995] [PMID: 18663133]

[22]   Penzotti, J.L.; Fozzard, H.A.; Lipkind, G.M.; Dudley, S.C., Jr Differences in saxitoxin and tetrodotoxin binding revealed by mutagenesis of the Na+ channel outer vestibule. *Biophys. J.,* **1998**, *75*(6), 2647-2657. [http://dx.doi.org/10.1016/S0006-3495(98)77710-0] [PMID: 9826589]

[23]   Cestèle, S.; Catterall, W.A. Molecular mechanisms of neurotoxin action on voltage-gated sodium channels. *Biochimie,* **2000**, *82*(9-10), 883-892. [http://dx.doi.org/10.1016/S0300-9084(00)01174-3] [PMID: 11086218]

[24]   Stephens, R.F.; Guan, W.; Zhorov, B.S.; Spafford, J.D. Selectivity filters and cysteine-rich extracellular loops in voltage-gated sodium, calcium, and NALCN channels. *Front. Physiol.,* **2015**, *6*, 153. [http://dx.doi.org/10.3389/fphys.2015.00153] [PMID: 26042044]

[25]   Feldman, C.R.; Brodie, E.D., Jr; Brodie, E.D., III; Pfrender, M.E. The evolutionary origins of beneficial alleles during the repeated adaptation of garter snakes to deadly prey. *Proc. Natl. Acad. Sci. USA,* **2009**, *106*(32), 13415-13420. [http://dx.doi.org/10.1073/pnas.0901224106] [PMID: 19666534]

[26]   Li, Y.; Gong, H. Theoretical and simulation studies on voltage-gated sodium channels. *Protein Cell,* **2015**, *6*(6), 413-422. [http://dx.doi.org/10.1007/s13238-015-0152-6] [PMID: 25894089]

[27]   Wang, C.; Chung, B.C.; Yan, H.; Wang, H-G.; Lee, S-Y.; Pitt, G.S. Structural analyses of Ca^{2+}/CaM interaction with NaV channel C-termini reveal mechanisms of calcium-dependent regulation. *Nat. Commun.,* **2014**, *5*, 4896. [http://dx.doi.org/10.1038/ncomms5896] [PMID: 25232683]

[28]   Davies, T.G.; Field, L.M.; Usherwood, P.N.; Williamson, M.S. A comparative study of voltage-gated sodium channels in the Insecta: implications for pyrethroid resistance in Anopheline and other Neopteran species. *Insect Mol. Biol.,* **2007**, *16*(3), 361-375. [http://dx.doi.org/10.1111/j.1365-2583.2007.00733.x] [PMID: 17433068]

[29]   Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.,* **2010**, *31*(2), 455-461. [PMID: 19499576]

[30]   UniProt: a hub for protein information. *Nucleic Acids Res.,* **2015**, *43*(Database issue), D204-D212. [PMID: 25348405]

[31]   Thackeray, J.R.; Ganetzky, B. Developmentally regulated alternative splicing generates a complex array of Drosophila para sodium channel isoforms. *J. Neurosci.,* **1994**, *14*(5 Pt 1), 2569-2578. [PMID: 8182428]

[32]   Chen, X.; Wang, Q.; Ni, F.; Ma, J. Structure of the full-length Shaker potassium channel Kv1.2 by normal-mode-based X-ray crystallographic refinement. *Proc. Natl. Acad. Sci. USA,* **2010**, *107*(25), 11352-11357. [http://dx.doi.org/10.1073/pnas.1000142107] [PMID: 20534430]

[33]   Long, S.B.; Campbell, E.B.; Mackinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K^+ channel. *Science,* **2005**, *309*(5736), 897-903. [http://dx.doi.org/10.1126/science.1116269] [PMID: 16002581]

[34]   Long, S.B.; Tao, X.; Campbell, E.B.; MacKinnon, R. Atomic structure of a voltage-dependent K^+ channel in a lipid membrane-like environment. *Nature,* **2007**, *450*(7168), 376-382. [http://dx.doi.org/10.1038/nature06265] [PMID: 18004376]

[35]   Tang, L.; Gamal El-Din, T.M.; Payandeh, J.; Martinez, G.Q.; Heard, T.M.; Scheuer, T.; Zheng, N.; Catterall, W.A. Structural basis for Ca2+ selectivity of a voltage-gated calcium channel. *Nature,* **2014**, *505*(7481), 56-61. [http://dx.doi.org/10.1038/nature12775] [PMID: 24270805]

[36]   Tsai, C.J.; Tani, K.; Irie, K.; Hiroaki, Y.; Shimomura, T.; McMillan, D.G.; Cook, G.M.; Schertler, G.F.; Fujiyoshi, Y.; Li, X.D. Two alternative conformations of a voltage-gated sodium channel. *J. Mol. Biol.,* **2013**, *425*(22), 4074-4088. [http://dx.doi.org/10.1016/j.jmb.2013.06.036] [PMID: 23831224]

[37]   Cronin, N.B.; O'Reilly, A.; Duclohier, H.; Wallace, B.A. Effects of deglycosylation of sodium channels on their structure and function.

Biochemistry, **2005**, *44*(2), 441-449. [http://dx.doi.org/10.1021/bi048741q] [PMID: 15641768]

[38] Pieper, U.; Webb, B.M.; Barkan, D.T.; Schneidman-Duhovny, D.; Schlessinger, A.; Braberg, H.; Yang, Z.; Meng, E.C.; Pettersen, E.F.; Huang, C.C.; Datta, R.S.; Sampathkumar, P.; Madhusudhan, M.S.; Sjölander, K.; Ferrin, T.E.; Burley, S.K.; Sali, A. ModBase, a database of annotated comparative protein structure models, and associated resources. *Nucleic Acids Res.,* **2011**, *39*(Database issue), D465-D474. [http://dx.doi.org/10.1093/nar/gkq1091] [PMID: 21097780]

[39] Bruhova, I.; Tikhonov, D.B.; Zhorov, B.S. Access and binding of local anesthetics in the closed sodium channel. *Mol. Pharmacol.,* **2008**, *74*(4), 1033-1045. [http://dx.doi.org/10.1124/mol.108.049759] [PMID: 18653802]

[40] Lukacs, P.; Gawali, V.S.; Cervenka, R.; Ke, S.; Koenig, X.; Rubi, L.; Zarrabi, T.; Hilber, K.; Stary-Weinzinger, A.; Todt, H. Exploring the structure of the voltage-gated Na^+ channel by an engineered drug access pathway to the receptor site for local anesthetics. *J. Biol. Chem.,* **2014**, *289*(31), 21770-21781. [http://dx.doi.org/10.1074/jbc.M113.541763] [PMID: 24947510]

[41] Lipkind, G.M.; Fozzard, H.A. KcsA crystal structure as framework for a molecular model of the $Na^{(+)}$ channel pore. *Biochemistry,* **2000**, *39*(28), 8161-8170. [http://dx.doi.org/10.1021/bi000486w] [PMID: 10889022]

[42] Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.,* **1999**, *292*(2), 195-202. [http://dx.doi.org/10.1006/jmbi.1999.3091] [PMID: 10493868]

[43] Drozdetskiy, A.; Cole, C.; Procter, J.; Barton, G.J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.,* **2015**, *43*(W1), W389-94. [http://dx.doi.org/10.1093/nar/gkv332] [PMID: 25883141]

[44] Miller, A.; Long, S. Crystal structure of the human two - pore domain potassium channel K2P1. *Science,* **2012**, *335*, 432.

[45] Russell, B.R.; Barton, G.J.; Rbr, R.B. Structural Alignment of Multiple Proteins. *Proteins,* **1999**, *335*, 432-6436.

[46] Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics,* **2004**, *5*, 113. [http://dx.doi.org/10.1186/1471-2105-5-113] [PMID: 15318951]

[47] 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.; Clustal, W. Clustal W and Clustal X version 2.0. *Bioinformatics,* **2007**, *23*(21), 2947-2948. [http://dx.doi.org/10.1093/bioinformatics/btm404] [PMID: 17846036]

[48] Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.,* **2013**, *30*(4), 772-780. [http://dx.doi.org/10.1093/molbev/mst010] [PMID: 23329690]

[49] Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics,* **2014**, *30*(9), 1312-1313. [http://dx.doi.org/10.1093/bioinformatics/btu033] [PMID: 24451623]

[50] Menke, M.; Berger, B.; Cowen, L. Matt: local flexibility aids protein multiple structure alignment. *PLOS Comput. Biol.,* **2008**, *4*(1), e10. [http://dx.doi.org/10.1371/journal.pcbi.0040010] [PMID: 18193941]

[51] Eswar, N.; Webb, B.; Marti-Renom, M.A.; Madhusudhan, M.S.; Eramian, D.; Shen, M-Y.; Pieper, U.; Sali, A. Comparative protein structure modeling using Modeller, **2007**. [http://dx.doi.org/10.1002/0471140864.ps0209s50]

[52] Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics,* **2008**, *9*, 40. [http://dx.doi.org/10.1186/1471-2105-9-40] [PMID: 18215316]

[53] Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX,* **2015**, *1-2*, 19-25.

[54] Wolf, M.G.; Hoefling, M.; Aponte-Santamaría, C.; Grubmüller, H.; Groenhof, G. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J. Comput. Chem.,* **2010**, *31*(11), 2169-2174. [http://dx.doi.org/10.1002/jcc.21507] [PMID: 20336801]

[55] Jämbeck, J.P.; Lyubartsev, A.P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B,* **2012**, *116*(10), 3164-3179. [http://dx.doi.org/10.1021/jp212503e] [PMID: 22352995]

[56] Lomize, M.A.; Lomize, A.L.; Pogozheva, I.D.; Mosberg, H.I. OPM: orientations of proteins in membranes database. *Bioinformatics,* **2006**, *22*(5), 623-625. [http://dx.doi.org/10.1093/bioinformatics/btk023] [PMID: 16397007]

[57] Lomize, M.A.; Pogozheva, I.D.; Joo, H.; Mosberg, H.I.; Lomize, A.L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.,* **2012**, *40*(Database issue), D370-D376. [http://dx.doi.org/10.1093/nar/gkr703] [PMID: 21890895]

[58] Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.,* **1996**, *14*(1), 33-38, 27-28. [http://dx.doi.org/10.1016/0263-7855(96)00018-5] [PMID: 8744570]

[59] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins,* **2006**, *65*(3), 712-725. [http://dx.doi.org/10.1002/prot.21123] [PMID: 16981200]

[60] Sorin, E.J.; Pande, V.S. Exploring the helix-coil transition *via* all-atom equilibrium ensemble simulations. *Biophys. J.,* **2005**, *88*(4), 2472-2493. [http://dx.doi.org/10.1529/biophysj.104.051938] [PMID: 15665128]

[61] DePaul, A.J.; Thompson, E.J.; Patel, S.S.; Haldeman, K.; Sorin, E.J. Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics. *Nucleic Acids Res.,* **2010**, *38*(14), 4856-4867. [http://dx.doi.org/10.1093/nar/gkq134] [PMID: 20223768]

[62] Wiederstein, M.; Sippl, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.,* **2007**, *35*(Web Server issue), W407-10. [PMID: 17517781]

[63] Catterall, W.A. Sodium channels, inherited epilepsy, and antiepileptic drugs. *Annu. Rev. Pharmacol. Toxicol.,* **2014**, *54*, 317-338. [http://dx.doi.org/10.1146/annurev-pharmtox-011112-140232] [PMID: 24392695]

[64] Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.,* **2004**, *25*(13), 1605-1612. [http://dx.doi.org/10.1002/jcc.20084] [PMID: 15264254]