

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**INSTITUTO DE CIÊNCIAS BIOLÓGICAS**  
**PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA**

**MODELAGEM ESTRUTURAL DO CANAL DE SÓDIO 1.4 E ESTUDO TEÓRICO**  
**DA LIGAÇÃO DE ALFA E BETA TOXINAS DE ESCORPIÃO**

**JOÃO RONALDO CLEMENTE FERNANDES**

**Orientador: Prof. Dr. PAULO SÉRGIO LACERDA BEIRÃO**

**Co-Orientador: Prof. Dr. LUCAS BLEICHER**

**BELO HORIZONTE - MG**

**JANEIRO/2019**

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

MODELAGEM ESTRUTURAL DO CANAL DE SÓDIO 1.4 E ESTUDO TEÓRICO DA  
LIGAÇÃO DE ALFA E BETA TOXINAS DE ESCORPIÃO

JOÃO RONALDO CLEMENTE FERNANDES

Tese apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais para a obtenção do título de Doutor em Bioinformática.

Orientador: Prof. Dr. Paulo Sérgio Lacerda  
Beirão

Co-Orientador: Prof. Dr. Lucas Bleicher

Belo Horizonte - MG

Janeiro - 2019



## "MODELAGEM ESTRUTURAL DO CANAL DE SÓDIO 1.4 E ESTUDO TEÓRICO DA LIGAÇÃO DE ALFA E BETA TOXINAS DE ESCORPIÃO"

**Joao Ronaldo Clemente Fernandes**

Tese aprovada pela banca examinadora constituída pelos Professores:

Prof. Paulo Sérgio Lacerda Beirão - Orientador  
UFMG

Prof. Lucas Bleicher - Coorientador  
UFMG

Profa Raquel Cardoso de Melo Minardi  
UFMG

Prof. José Miguel Ortega  
UFMG

Prof. Werner Treptow  
UNB

Prof. Pedro Geraldo Pascutti  
UFRJ

Belo Horizonte, 30 de janeiro de 2019.

## AGRADECIMENTOS

A meus pais (*in memorian*).

Ao Professor Doutor Paulo Sérgio Lacerda Beirão por ter me aceito como orientando, colaborando para meu crescimento pessoal e, principalmente, como cientista algo que tanto amo.

Aos colegas do Laboratório de Membranas Excitáveis (Lamex) que, de uma forma ou outra, participaram deste trabalho.

À Universidade Federal de Minas Gerais, ao Programa Interunidades de Pós-graduação em Bioinformática, sempre presentes neste trabalho.

Ao Professor Doutor Lucas Bleicher pelo apoio, auxílio e grande paciência com as minhas inúmeras dúvidas.

Ao Professor Doutor José Miguel Ortega, pela orientação acerca do servidor Sagarana, a seus alunos que sempre estiveram a postos para ajudar e sanando dúvidas, de forma gentil, mesmo não sendo da responsabilidade deles.

Ao CNPq e à CAPES pelo financiamento do presente trabalho.

Ao técnico do Laboratório de Membranas Excitáveis José Narciso Fernandes, ótimo companheiro de altos papos.

Aos amigos e colegas que me acompanharam nestes últimos anos.

À minha família. Meus irmãos Marcel, Nicolas, Alessandra, e as minhas primas Jéssica e Lorena, que sabem o quanto sacrifiquei da convivência. Vô João, Vó Ana, Vó Zilá e Tio Márcio, *in memorian*, tendo certeza de estarem sempre me apoiando, onde quer que estejam.

A minha meia família, que me aceitou e ajudou em meu desenvolvimento pessoal, ao Sr. Rubens, Dona Maria, Rafael, Pedro, Ana Carolina, Fabiana e Maurício.

A Juliana (minha noiva, mas sempre no papel de esposa) lutando ao meu lado como ninguém, com paciência, compreensão e o mais importante, sabendo que, às vezes precisamos esperar, e acima de tudo, continua acreditando em mim.

Aos Deuses.

## EPÍGRAFE

“We choose to go to the moon...  
...not because they are easy, but  
because they are hard”. (John F.  
Kennedy)

## RESUMO

*Voltage-gated sodium channels* (VGSD) são proteínas responsáveis pela propagação da informação entre neurônios, fibras musculares esqueléticas, fibras musculares cardíacas e outros tipos de células excitáveis. Eles são rapidamente ativados quando a membrana celular é despolarizada, seguido de inativação, o que leva a um influxo de sódio transiente. Até recentemente não existia um modelo estrutural completo do canal de sódio humano.

Neste trabalho, realizamos a modelagem comparativa do canal hNav1.4, usando como *template* o canal de cálcio de coelho (Cav1.1), canal de sódio de barata (NavPaS) e o canal de sódio de enguia elétrica (EeNav1.4), com os programas Rosetta e SWISS-MODEL. Os modelos gerados foram designados hNav1.4/Cav, hNav1.4/PaS e hNav1.4/Ee, conforme o *template* usado. Estes modelos foram analisados de duas formas: na fidelidade do seu enovelamento em relação ao seu *template*, usando a ferramenta TM-Align; e pelos ângulos das ligações peptídicas entre os resíduos, utilizando o gráfico de Ramachandran. Todos os modelos foram considerados bem sucedidos pelos gráficos de Ramachandran e pelo TM-Align. Uma análise mais apurada das estruturas mostrou algumas incongruências no modelo hNav1.4/Cav ao compararmos com o funcionamento dos canais de sódio, tornando o modelo inadequado. O poro do canal se mostra fechado e seus domínios sensíveis a voltagem (VSD) se encontram ativados, o que torna este modelo falho em relação às propriedades funcionais do canal. Os modelos hNav1.4/PaS e hNav1.4/Ee, por outro lado, representam estados conformacionais fechado e aberto, respectivamente, e se mostraram consistentes com os dados biofísicos e farmacológicos conhecidos. Como critério para identificar em qual estado conformacional os canais se encontram, foi realizada a análise dos segmentos S4, e sua posição em relação à constrição hidrofóbica, bem como a verificação das interações dos resíduos carregados com resíduos de aminoácidos dos segmentos S1, S2 e S3. Considerando este critério, foi possível a contagem das cargas elementares necessárias para ativar o canal hNav1.4, chegando-se a um valor aproximado de 6,5.

Completando a análise das características dos canais, foi realizado o *docking* de alfa e beta-toxinas de escorpião no canal. A beta-toxina Ts1, se ligou ao sítio 4 do canal hNav1.4/PaS, mostrando alta afinidade com energia livre de ligação ( $\Delta G$ ) de -9.4 kcal/mol, confirmando que o segmento S4 do domínio DII se encontra na posição ativada, embora o canal esteja fechado, conforme análise feita pelo programa MOLEonline. A alfa-toxina AaHII

se ligou ao sítio 3 do modelo hNav1.4/PaS e teve sua energia livre de ligação em -10.7 kcal/mol. Um novo *docking* foi testado com o modelo estado aberto, hNav1.4/Ee, e não obteve sucesso. Este resultado confirma resultados experimentais que comprovam a incapacidade da toxina de se ligar ao canal quando o VSDIV está ativado. Uma análise mais profunda mostrou que o resíduo D1435 do canal, tido como de grande importância para a interação toxina-canal, se encontra em uma posição não favorável para ligação da toxina. Esta mudança é causada pelo movimento do segmento S4 que, ao se mover em direção ao meio externo durante sua ativação, provoca no segmento S3 um movimento para dentro seguido por um giro, colocando o resíduo D1435 na posição desfavorável para a ligação da toxina. Essa observação permite esclarecer uma característica das alfa-toxinas, que é o seu deslocamento do canal quando pulsos altamente despolarizantes são aplicados à membrana. Os resultados apresentados pelo *docking* das toxinas são compatíveis com os resultados obtidos experimentalmente.

Nossos resultados mostram que estes modelos teóricos são capazes de fornecer as bases estruturais para aspectos fisiológicos e farmacológicos desses canais, possibilitando assim uma nova gama de pesquisas teóricas envolvendo o canal hNav1.4 em diferentes estados conformacionais.



## ABSTRACT

Voltage-gated sodium channels (VGSD) are proteins responsible for the propagation of information between neurons, skeletal muscle fibers, cardiac muscle fibers and other types of excitable cells. They are rapidly activated when the cell membrane is depolarized, followed by inactivation, which leads to a transient sodium influx. Until recently there was no complete structural model of the human sodium channel.

In this work, we performed the comparative modeling of the hNav1.4 channel, using as template the rabbit calcium channel (Cav1.1), cockroach sodium channel (NavPaS) and the electric eel sodium channel (EeNav1.4), with the Rosetta and SWISS-MODEL programs. These models were analyzed in two ways: the fidelity of their folding in relation to their templates using the TM-Align tool, and by angles of the peptide bonds between the amino acid residues, using the Ramachandran plot. All models were considered successful based on these criteria. More accurate analysis of the structures showed that the hNav1.4/Cav model has inconsistencies when compared to the functioning of a sodium channel, making the model inappropriate. The pore of the channel is closed and its voltage sensitive domains (VSD) are activated, which makes this model faulty in relation to the functional properties of the channel. On the other hand, the hNav1.4/PaS and hNav1.4/Ee models were able to represent closed and open conformational states, respectively, and were consistent with known biophysical and pharmacological data. As a criterion to identify the conformational state the channels, the analysis of the segments S4 was performed checking its position in relation to the hydrophobic constriction and the interactions of the charged residues with segments S1, S2 and S3. Considering this criterion, it was possible to count the elementary charges necessary to activate the hNav1.4 channel, reaching an approximate value of 6.5.

Completing the analysis of the channel characteristics, the docking of alpha- and beta-scorpion toxins to the canal was performed. Beta-toxin Ts1 bound to channel site 4 with high affinity and binding free energy ( $\Delta G$ ) of -9.4 kcal/mol, thus confirming that the segment S4 of the DII domain is in the activated position, although the pore of the channel is closed state (as shown by the analysis by the MOLEonline program). The alpha toxin AaHII bound to site 3 of the closed-state model hNav1.4/PaS with a free energy of binding of -10.7 kcal/mol. A new docking was tested with the open-state model, hNav1.4/Ee, without success. This result is in agreement with the toxin's ability to bind to the VSDIV of the channel, but

not when it is activated. Further analysis showed that the D1435 residue of the channel, considered to be of great importance for the toxin-channel interaction, is at an unfavorable position for toxin binding. This change results from the outward movement of the S4 segment, which causes a rotation of S3, placing the D1435 in unfavorable position. This observation can account for one characteristic of the alpha-toxins, which is their displacement from the channel when strong depolarizing pulses are applied to the membrane. The results presented by the docking of the toxins can explain the results obtained experimentally.

Our results show that these theoretical models may provide structural basis to understand the physiological and pharmacological characteristics of these channels, thus enabling a new range of theoretical researches involving the hNav1.4 channel at different conformational states.

## LISTA DE ABREVIACOES

Å	Angstrom
C $\alpha$	Carbono alfa
Cav1.1	Canal de clcio de coelho
DEKA	Aspartato, Glutamato, Lisina, Alanina
EeNa $\nu$ 1.4	Canal de sodio de enguia eltrica
hNa $\nu$ 1.4	Canal de sodio humano de msculo esqueltico
ns	Nanossegundo
Na $\nu$ PaS	Canal de sodio de barata
VGSD	Canal de sodio sensvel a voltagem
VSD	Domnio do sensor de voltagem

## LISTA DE FIGURAS

Figura 1 - Arquitetura de canais de sódio dependentes de voltagem (VGSC)	16
Figura 2 - Aminoácidos conservados do domínio do sensor de voltagem (VSD)	18
Figura 3 - Toxinas AahII em Verde, Ts3 em Ciano, Ts1 em Roxo	21
Figura 4 - Modelo canal hNa <sub>v</sub> 1.4/Ca <sub>v</sub>	39
Figura 5 - Modelo estrutural hNa <sub>v</sub> 1.4/PaS e hNa <sub>v</sub> 1.4/Ee	42
Figura 6 - Sobreposição dos modelos destaque alça de inativação e filtro de seletividade	43
Figura 7 - Sobreposição entre os VSDs	44
Figura 8 - Constrição hidrofóbica	45
Figura 9 - Zoom do <i>docking</i> hNa <sub>v</sub> 1.4/PaS (Cinza), e AahII toxin (Roxo)	47
Figura 10 - Zoom no <i>docking</i> hNa <sub>v</sub> 1.4/PaS(Ciano) e a Ts1 toxina(Roxo)	48
Figura 11 - <i>Docking</i> Ts3 no canal hNa <sub>v</sub> 1.4/PaS	49
Figura 12 - RMSF	51
Figura 13 - Sobreposição pós <i>docking</i>	59

## LISTA DE TABELAS

Tabela 1 - Alinhamento: hNa <sub>v</sub> 1.4 x Cav1.1, hNa <sub>v</sub> 1.4 x EeNa <sub>v</sub> 1.4 e hNa <sub>v</sub> 1.4 x Na <sub>v</sub> Pas	63
Tabela 2 - Alinhamento para constrição hidrofóbica	67
Tabela 3a - Resíduos importantes para a ligação da toxina Ts1 ao canal	68
Tabela 3b - Resíduos importantes para a ligação da toxina AaHIII ao canal	68
Tabela 4 - Interações iônicas entre resíduos do S4 e resíduos do S1, S2 e S3	69

## SUPLEMENTAR

<i>Script</i> Rosetta	70
Protocolo para inserção do canal na membrana e <i>script</i> para dinâmica	72
<i>Script</i> python MDAnalysis	77

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>15</b>
1.1 Canais de sódio dependentes de voltagem	15
1.2 Alfa e Beta toxinas	19
<b>2 OBJETIVOS</b>	<b>22</b>
2.1 Objetivos específicos	22
<b>3 MÉTODOS</b>	<b>23</b>
3.1 Modelagem com Rosetta	23
3.1.1 Seleção do template	23
3.1.2 Modelagem Comparativa	25
3.1.3 Cluster	27
3.1.4 TM-align	27
3.2 Modelagem Swiss-Model	28
3.3 Docking HADDOCK	30
3.3.1 Ts1	31
3.3.2 AaHII	32
3.3.3 Ts3	32
<b>4 DINÂMICA MOLECULAR</b>	<b>33</b>
4.1 SWISS-MODEL e HADDOCK	33
<b>5 RESULTADOS</b>	<b>37</b>
5.1 Modelagem Rosetta	37
5.2 Modelagem SWISS-MODEL	40
5.3 Docking	46
5.3.1 Docking da AaHII	46
5.3.2 Docking da Ts1	47
5.3.3 Docking da TS3	48
5.4 Dinâmica molecular hNav1.4/PaS com e sem a toxina AaHII	49
<b>6 DISCUSSÃO</b>	<b>51</b>
6.1 Modelagem	51
6.2 Fisiologia molecular do canal	53
6.3 Interações toxina-canal	57
<b>7 CONCLUSÃO</b>	<b>61</b>
<b>8 PERSPECTIVAS</b>	<b>62</b>

# 1 INTRODUÇÃO

## 1.1 Canais de sódio dependentes de voltagem

Os canais de sódio dependentes de voltagem (*Voltage-gated sodium channels* - VGSD) são proteínas transmembrana heteroméricas formadas por uma subunidade  $\alpha$  com ~260 kDa e subunidade  $\beta$  com 30-40 kDa (CATTERALL, W A, 1986, 2000).

Em mamíferos, existem nove isoformas distintas da subunidade  $\alpha$  do canal de sódio, identificadas como Na<sub>v</sub>1.1 a Na<sub>v</sub>1.9. Essas isoformas do canal de sódio são mais expressas em determinados tecidos: Na<sub>v</sub>1.1, Na<sub>v</sub>1.2, Na<sub>v</sub>1.3 e Na<sub>v</sub>1.6 são expressos em grande concentração no sistema nervoso central; Na<sub>v</sub>1.4 no músculo esquelético; Na<sub>v</sub>1.5 no músculo cardíaco; Na<sub>v</sub>1.7, Na<sub>v</sub>1.8 e Na<sub>v</sub>1.9 no sistema nervoso periférico (CATTERALL, W A, 2000). Elas são, pelo menos 50% idênticas na sua sequência de aminoácidos nas porções transmembranas e no domínio extracelular (GOLDIN *et al.*, 2000). A estrutura do canal de sódio é composta por 4 domínios homólogos, com aproximadamente 2.000 resíduos (AHMED *et al.*, 1992). Esses domínios são identificados como DI a DIV, sendo que cada domínio contém 6 segmentos transmembrana identificados como S1-S6 (AHMED *et al.*, 1992). O domínio do sensor de voltagem (VSD) é formado pelos segmentos S1-S4, sendo que os segmentos S4 contêm uma sequência conservada de Arginina ou Lisina a cada três resíduos (RxxRxxRxxR), tornando o segmento S4 carregado positivamente e agindo como um sensor de voltagem (figura 1 e 2) (PAYANDEH *et al.*, 2011). A parte da estrutura composta pelos segmentos S5-S6 forma o poro condutor e o filtro de seletividade (figura 1) (PAYANDEH *et al.*, 2012). O filtro de seletividade é formado a partir do arranjo criado pelos *loops* P1-P2. O filtro de seletividade é formado por um anel de resíduos de aminoácidos conhecido como DEKA, abreviações para os respectivos aminoácidos que formam esse anel; Aspartato (domínio DI DIP1-DIP2 *loop*), Glutamato (domínio DII DIIP1-DIIP2 *loop*), Lisina (domínio DIII DIIP1-DIIP2 *loop*) e Alanina (domínio DIV DIVP1-DIVP2 *loop*), conjunto esse que permite a seletividade ao íon de sódio. Em publicação LIPKIND e FOZZARD (2008) concluíram que o sódio parcialmente hidratado passa através do filtro de seletividade, interagindo principalmente com os 2 grupos de oxigênio do glutamato, à medida que passa



através do filtro de seletividade. Durante este processo os resíduos da DEKA, se deslocam para formar pontes de hidrogênio e interações de van der Waals. No caso do potássio, que é um íon significativamente maior, ele forma uma fraca interação com o glutamato e é repelido pela lisina (FAVRE; MOCZYDLOWSKI; SCHILD, 1996; HEINEMANN *et al.*, 1992).

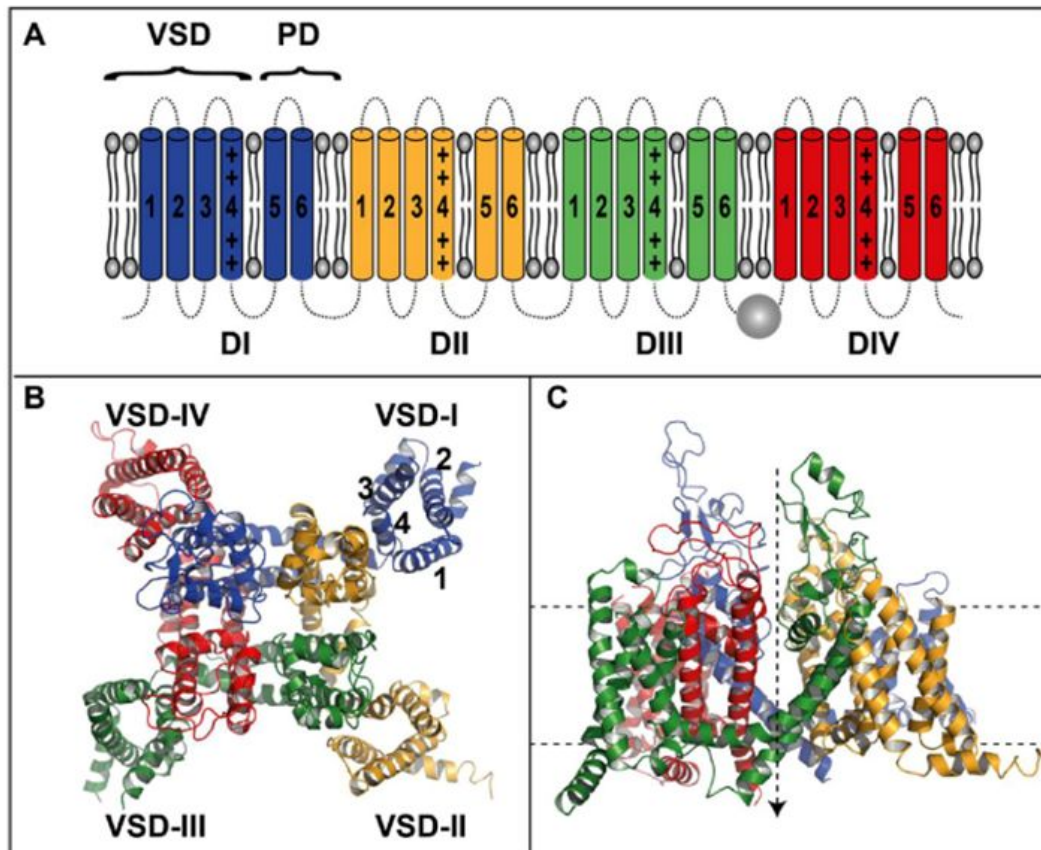


FIGURA 1 - Arquitetura de canais de sódio dependentes de voltagem (VGSC). A) Imagem esquemática das seis hélices transmembranares (S1-S6) de cada um dos quatro domínios (DI-DIV) de um VGSC. As alças entre os segmentos estão pontilhadas, uma vez que variam significativamente em comprimento, dependendo do tipo de VGSC. O domínio do sensor de voltagem (VSD) e o domínio do poro (PD) do primeiro domínio são anotados. A partícula de inativação (IFM) do VGSC no *linker* DIII-DIV é representada como uma bola. Os seis segmentos de cada um dos quatro domínios se organizam radialmente para formar o canal funcional. B) Vista superior do modelo Na<sub>v</sub>1.4Ee. Os quatro domínios estão anotados e as quatro hélices que compõem o VSD do domínio I também estão marcadas. O poro condutor de íons está no centro oco do canal. C) Vista lateral do canal mostrado em B, onde as grandes alças extracelulares do canal são evidentes. A posição aproximada da bicamada lipídica é indicada pelas linhas pontilhadas. O caminho de condução iônica é indicado pela seta pontilhada. Adaptado de (ZHANG. *et al.*, 2018).

O VSD está ligado ao domínio do poro por um *linker* intracelular entre os segmentos transmembrana S4 e S5 (CAPES *et al.*, 2013; GOLDIN, ALAN L, 2003). A atividade de condução iônica dos canais de sódio funciona em uma escala muito curta de tempo entre o processo de ativação e inativação. Em resposta à despolarização da membrana, os segmentos transmembrana S4, dos domínios DI, DII e DIII, carregados positivamente, se movem ao longo do campo elétrico transmembrana. Este movimento é transferido ao domínio do poro via *linker* S4-S5, provocando uma mudança conformacional resultando, assim, na abertura do poro (CAPES *et al.*, 2013). Com a repolarização os segmentos S4 retornam à posição de repouso e o poro é fechado. No entanto, mesmo com a despolarização mantida o canal se fecha por um mecanismo chamado inativação rápida (figura 1) (CAPES *et al.*, 2013; GOLDIN, ALAN L, 2003).

Em cerca de 1 (um) milissegundo após a ativação, o canal de sódio é inativado (HODGKIN; HUXLEY, 1952). Em células excitáveis, rápida e completa inativação dos canais de sódio é essencial. Mutações que prejudicam a inativação rápida causam paralisia periódica hereditária músculo esquelético, arritmias no coração e epilepsia (BOOKER; WHYTE; LADUSANS, 2003).

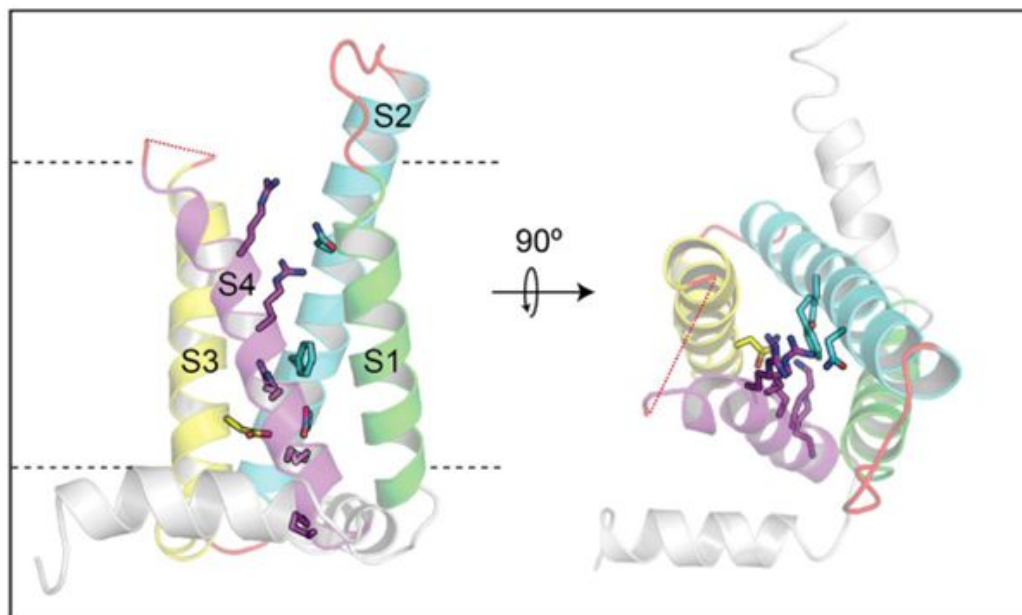


FIGURA 2 - Modelo representativo do VSD do canal  $\text{Na}_v1.4\text{Ee}$ . Em destaque os aminoácidos conservados nas hélices S1-S4. A alça S3-S4 está faltando na estrutura usada - o linker ausente é indicado por uma linha pontilhada. As cadeias laterais dos resíduos (argininas e lisinas) carregadas positivamente do S4 são mostradas. As cadeias laterais de resíduos conservados em S2 e S3 também são mostradas, as quais incluem os resíduos com carga oposta no S2-S3 e a constrição hidrofóbica (fenilalanina) em S2. Adaptado de (ZHANG, A. H. *et al.*, 2018).

A inativação rápida é realizada por quatro resíduos no linker entre os domínios III e IV do canal de sódio: I-isoleucina, F-fenilalanina, M-metionina e T-treonina (IFMT) figura 1. Esta região muito conservada é necessária para a rápida inativação, funcionando como uma tampa, obstruindo pelo lado intracelular o poro do canal, impedindo o sódio de mover-se através do canal (WEST *et al.*, 1992; GOLDIN, A L, 2003). Com a repolarização da célula, a região do IFMT se desliga do poro do canal (CAPES *et al.*, 2013). Outra forma de cessar a condutância do canal é a inativação lenta (RICHMOND *et al.*, 1998). Enquanto a inativação rápida ocorre em poucos milissegundos, a inativação lenta pode chegar a uma escala de tempo de segundos. Esta inativação lenta ocorre quando há uma repetitiva ou prolongada despolarização do canal de sódio limitando sua disponibilidade por um longo período de tempo (RICHMOND *et al.*, 1998).

As subunidades  $\beta$  são estruturas proteicas que estão aderidas ao canal de sódio, com importante papel na modulação da cinética do canal e na regulação da densidade do canal, ou seja, o número de canais em uma determinada área. Elas são classificadas em subunidades  $\beta 1$ ,

$\beta 2$ ,  $\beta 3$ ,  $\beta 4$  e  $\beta 1B$  e são componentes importantes para a fisiologia dos canais de sódio (ISOM *et al.*, 1992, 1995).

## 1.2 Alfa e Beta toxinas

Os canais iônicos são alvo de um grande número de compostos químicos e proteínas. Recentemente foi reportado que todos os peptídeos moduladores do canal iônico descobertos até o momento, perfazem cerca de 954 estruturas (MOBLI; UNDHEIM; RASH, 2017). Deste total, 407 são moduladores de canais de sódio, 278 são moduladores do canal de potássio e 98 são moduladores de canais de cálcio. Toxinas de escorpião parecem ser particularmente uma fonte rica de moduladores de canais dependentes de voltagem com ~350 peptídeos reportados, em segundo lugar vêm ~200 peptídeos de aranhas.

Dentro desta gama de modificadores de canais de sódio, aqueles derivados de veneno de escorpião têm um papel importante. Venenos de escorpião são misturas altamente complexas de enzimas, peptídeos, nucleotídeos, lipídios e outras substâncias ainda desconhecidas. Toxinas formadas por peptídeos são o grupo mais abundante no veneno de escorpião, e os mais estudados. As toxinas de escorpião que agem no canal de sódio podem ser divididas em duas classes, de acordo com seu modo de ação e suas propriedades distintas de ligação:  $\alpha$ -toxinas e  $\beta$ -toxinas (CESTÈLE *et al.*, 1998; KOPEYAN *et al.*, 1974; ROCHAT; BERNARD; COURAUD, 1979).  $\alpha$ -toxinas e  $\beta$ -toxinas se ligam ao receptor sítio 3 e sítio 4, para inibir a rápida inativação e afetando a ativação do canal (CATTERALL *et al.*, 2007).

Estas toxinas são formadas por polipeptídeos de 61-76, resíduos de aminoácidos interligados por quatro pontes dissulfeto, e o núcleo estrutural da toxina é formado por uma  $\alpha$ -hélice, com 2 a 3 motivos folhas-beta. As  $\alpha$ -toxinas de escorpião são divididas em três grupos de acordo com sua toxicidade em insetos e mamíferos:  $\alpha$ -toxinas clássicas, são altamente ativas apenas em canais de sódio sensíveis a voltagem, com alta afinidade, em cérebros de ratos, exemplos: AaHII (*Androctonus australis hector*), LqhII (*Leiurus quinquestriatus hebraeus*);  $\alpha$ -toxinas anti-insetos sendo altamente ativas apenas em insetos, exemplo: Lqq3.  $\alpha$ -like toxinas que são ativas em mamíferos e insetos, exemplo: Lqh-3 (GORDON *et al.*, 2007; GORDON; GUREVITZ, 2003).

A estrutura cristalográfica da toxina AaHII (*Androctonus australis hector* - 1PTX) tornou-se disponível para pesquisadores com resolução de 1.3Å (HOUSSET *et al.*, 1994) (figura 3). Sua estrutura é formada por uma  $\alpha$ -hélice e três folhas  $\beta$ , composta por 64 resíduos de aminoácido. Sua superfície funcional é composta por dois domínios distintos: um conservado “*core domain*”, formado por resíduos das alças, entre a  $\alpha$ -hélice e uma folhas  $\beta$ , e entre duas folhas  $\beta$ , conectando os elementos da estrutura secundária do núcleo da molécula, e um variável “*NC-domain*”, formada por uma volta de cinco resíduos (resíduos 8-12) e um segmento C-terminal (resíduos 56-64) (KAHN *et al.*, 2009; KARBAT *et al.*, 2004).

A estrutura da toxina Ts3 (*Tityus serrulatus* - 5CY0), obtida com resolução de 1.93Å, é formada por 64 resíduos, como ocorre com a toxina AaHII, e é dividida em “*core domain*”, “*NC-domain*” e C-terminal. Toda sua estrutura é formada em uma  $\alpha$ -hélice e três folhas  $\beta$  (figura 3) (DANG *et al.*, 2016).

As  $\beta$ -toxinas são toxinas formadas por uma cadeia polipeptídica com 61-76 resíduos de aminoácidos, estabilizada por quatro pontes dissulfeto (POSSANI *et al.*, 1999), ligando-se ao sítio 4 (entre o segmento S3-S4 do domínio DII). A interação desta toxina com o seu local de ligação específico, é dependente do estado conformacional ativado do VSD do DII. O mecanismo proposto para as  $\beta$ -toxinas envolve a ligação na alça S3-S4, do domínio DII quando seu segmento S4 está em posição ativada. Foi apontado que a toxina aprisiona e estabiliza esse segmento em sua posição ativada (para fora), mantendo-o assim mesmo no repouso e facilitando assim a ativação do canal em resposta à despolarização subsequente. O resultado é o deslocamento negativo na dependência de voltagem de ativação (CESTÈLE *et al.*, 1998; ZHANG, J. Z. *et al.*, 2012).

A Ts1 (do *Tityus serrulatus* - 1NPI) é uma toxina da classe das  $\beta$ -toxinas de escorpião (figura 3) (PINHEIRO *et al.*, 2003). Sua estrutura foi elucidada com 1.16 Å, e é formada por 61 resíduos de aminoácido, com alta afinidade para mamífero e inseto. Sua estrutura é formada por uma  $\alpha$ -hélice e três folhas  $\beta$  com quatro pontes dissulfeto estabilizando a toxina (PINHEIRO *et al.*, 2003).

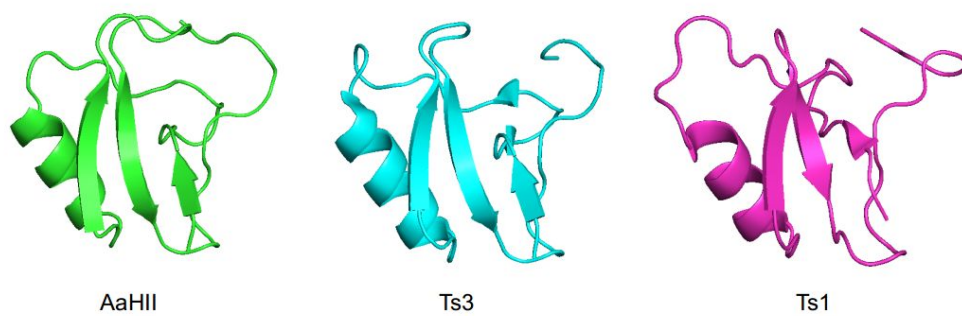


Figura 3 - Toxinas AaHII em Verde, Ts3 em Ciano, Ts1 em Roxo.

## 2 OBJETIVOS

- O principal objetivo do presente trabalho foi obter modelos do canal hNa<sub>v</sub>1.4 para melhor entender o seu funcionamento, bem como as ligações de toxinas peptídicas de escorpião e as consequências funcionais dessas ligações.

### 2.1 Objetivos específicos

- Gerar modelos de canais Na<sub>v</sub>1.4 capazes de fornecer as bases estruturais para funções conhecidas deste canal.
- Análise das estruturas modeladas.
- Realizar o docking de  $\alpha$  e  $\beta$  toxinas de escorpião nessas estruturas modeladas.
- Entender as bases estruturais para os efeitos das toxinas  $\alpha$  e  $\beta$  de escorpião.

## 3 MÉTODOS

### 3.1 Modelagem com Rosetta

#### 3.1.1 Seleção do *template*

Embora canais de sódio sejam encontrados desde organismos filogeneticamente primitivos, sua estrutura básica, principalmente nos seus segmentos transmembrana, é bastante conservada. Por esse motivo buscou-se selecionar como *template*, para esta modelagem, aquelas estruturas conhecidas que estivessem mais próximas filogeneticamente do *Homo sapiens*. Ao longo do desenvolvimento dessa tese surgiram novas estruturas que forçaram a elaboração de novos modelos baseados nas mesmas, e que serão comparados.

Inicialmente dispunha-se apenas estruturas de canais de sódio originadas de bactérias, que foram o ponto de partida. Como novos e mais apropriados *templates* surgiram nos últimos anos, estes modelos gerados a partir de canais bacterianos não foram incluídos neste documento.

Nosso alvo para modelagem é o canal de sódio humano músculo esquelético hNav1.4, sua sequência é registrada com o código UniProtKB/Swiss-Prot P35499 com 1836 aminoácidos.

A estrutura do canal de cálcio ( $Ca_v$ ), publicada por Wu e colaboradores (WU *et al.*, 2016), foi a primeira estrutura de canal iônico em mamífero mais próxima estruturalmente ao canal de sódio. O  $Ca_v$  em mamíferos, é classificado em 3 famílias,  $Ca_v1$ ,  $Ca_v2$  e  $Ca_v3$ . Os canais  $Ca_v1$  ( $Ca_v1.1$ - $Ca_v1.4$ ) são canais do tipo L e, como VGSD, são altamente dependentes de voltage, mas não têm mecanismo equivalente de inativação (CATTERALL, W A 2011). A subunidade responsável pela condução iônica,  $\alpha1$ , tem aproximadamente 1873 resíduos de aminoácidos e é formada por 4 domínios (DI-DIV), e cada domínio contém 6 segmentos transmembrana (S1-S6). O domínio do sensor de voltage (VSD) é formado pelos segmentos S1-S4, no segmento S4 estão presentes cargas positivas, argininas ou lisinas, a cada terceira ou quarta posição no segmento, e o poro do canal é formado pelos segmentos S5-S6, tendo



seu filtro de seletividade formado por uma sequência de glutamato (E292/E614/E1014/E1323). A subunidade transmembranar  $\alpha 1$  do  $\text{Ca}_v1.1$ , responsável pela condução iônica dependente de voltagem, contém 1873 resíduos de aminoácidos e é estruturalmente semelhante aos canais de sódio. Por esse motivo, consideramos a subunidade  $\alpha 1$ , como nosso candidato a *template* para a modelagem comparativa.

O  $\text{Ca}_v1.1$  tem sua sequência de resíduos de aminoácidos registrada no UniProtKB/Swiss-Prot com código P07293. Foi realizado o alinhamento da sequência com a sequência de  $\text{hNa}_v1.4$ , código UniProtKB/Swiss-Prot: P35499 usando o ClustalOmega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (tabela 1). Com esse alinhamento, foi possível identificar regiões de referência, como segmentos transmembrana e cargas positivas dos segmentos S4, que ajudaram a determinar como seria a divisão dos domínios e dos segmentos. Para analisar o alinhamento e identificação, usamos o código P35499 para acessar os dados da família de canais iônicos no banco de dados Pfam (<https://pfam.xfam.org/>). Este possibilitou identificar as regiões que correspondiam ao início e fim para cada um dos domínios do  $\text{hNa}_v1.4$ . A fim de identificar quais sequências de resíduos correspondiam ao segmentos transmembranas, consideramos os segmentos transmembranas do  $\text{Ca}_v1.1$  como referência. O banco de dados RCSB-PDB (<https://www.rcsb.org/>) armazena estruturas e respectivas informações destas proteínas. Uma destas informações é a conformação secundária formada da sua sequência de aminoácidos. Como são estruturas conservadas pelo alinhamento, identificamos a região compatível com os segmentos transmembrana e os segmentos intermediários foram consideradas como sendo alças intra ou extracelulares.

O protocolo do RosettaCM (SONG *et al.*, 2013), mostra que o nível de confiabilidade nas cadeias laterais e no *backbone* são maiores quando a identidade é maior do que 15%. Usando o programa Needle (<http://www.ebi.ac.uk/Tools/psa/>) o resultado ficou em 23.3% de identidade e 38.9% de similaridade, ou seja, mesmo sendo estruturas de canais iônicos diferentes, a identidade ainda ficou acima do requerido para o protocolo Rosetta.

### 3.1.2 Modelagem Comparativa

O programa ROSETTA (ROHL *et al.*, 2004) é uma *suite* com muitos pacotes integrados, que permite desenvolver várias linhas de pesquisas utilizando o mesmo pacote de programas, desde o enovelamento *ab initio* de proteínas (que envolve a modelagem sem um modelo como referência), modelagem de proteínas (comparativa ou por homologia), *docking* proteína-proteína, *docking* proteína-ligante ou mesmo alteração no *design* de uma estrutura a procura de uma nova função. A maioria dos protocolos de amostragem conformacional do ROSETTA usa o algoritmo baseado em Monte Carlo - Motrópolis para guiá-lo (KAUFMANN *et al.*, 2010).

A estratégia usada no ROSETTA, como descrito por Rohl e colaboradores (ROHL *et al.*, 2004) faz com que seu algoritmo trabalhe a procura de estruturas locais nas sequências que têm semelhança, ou relação a estruturas já conhecidas. Desta forma, ele acessa bibliotecas de fragmentos que representam a variação local para todos os segmentos curtos da cadeia de proteínas selecionados a partir de dados estruturais conhecidos. As estruturas curtas são montadas de forma aleatória seguindo o algoritmo de Monte Carlo (KAUFMANN *et al.*, 2010). Os resultados desta montagem são avaliados de acordo com uma função de pontuação, baseada em estatísticas de conformação estrutural de proteínas.

A representação que o ROSETTA utiliza para a proteína no espaço é resumida em uma lista de ângulos de torção da cadeia principal. Modificações estruturais devem ocorrer no espaço de torção, com intuito de avaliar a energia de conformação correspondente à representação da proteína em um espaço cartesiano. Desta forma, são geradas coordenadas atômicas que correspondem aos átomos pesados na cadeia principal da proteína, assumindo ângulos e comprimentos ideais para os resíduos de forma individual. Duas representações são utilizadas, a depender da necessidade da função de energia. Nos resíduos base, os termos potências devem ser descritos em uma forma reduzida, onde cada cadeia lateral assuma uma representação em forma de centróide, localizado no seu centro de massa, usando como referência o C $\beta$ . Para a glicina, o centróide está localizado no C $\alpha$ . De tal modo, a posição dos centróides de cada tipo de resíduo vai ser determinada pela média observada nas cadeias laterais de proteínas conhecidas de forma a aumentar a sensibilidade. As coordenadas atômicas dos átomos das cadeias laterais são acrescentadas com os hidrogênios. Essas cadeias

laterais são montadas de forma restrita, de acordo com uma biblioteca de rotâmeros. As posições que as cadeias laterais podem ser montadas são escolhidas usando o critério de Monte Carlo.

A função de pontuação do ROSETTA é descrita em duas formas: a) *low-resolution*, em que os termos correspondentes a solvatação e efeitos eletrostáticos estão baseados na distribuição dos resíduos, já observados em estruturas de proteínas, e as pontes de hidrogênio não são descritas de forma explícita. A sobreposição estérica de átomos estruturais e centróides são tratados como penalidades na função de pontuação e, no caso das forças de van der Waals, são modeladas apenas para estruturas compactas. As interações locais não são avaliadas de forma direta, já que são consideradas inclusas na biblioteca de fragmentos; b) *high-resolution*, as interações de van der Waals são simplificadas com o modelo 6-12 do potencial Lennard-Jones, que compensa atenuando a função linear no regime repulsivo, de forma a compensar a representação discreta dos rotâmeros nas cadeias laterais. Com a solvatação, segue-se o modelo de LAZARIDIS, & KARPLUS (1999), onde as pontes de hidrogênio estão inclusas, usando uma orientação dependente de estrutura secundária baseada em análises geométricas. As energias eletrostáticas são calculadas de acordo com um potencial de resíduo base em pares, o qual é similar ao utilizado no *low-resolution*. Com isso todas as interações energéticas são incluídas na energia total da estrutura (ROHL *et al.*, 2004). Para maiores informações sobre equações e outras aplicações consideradas na função de pontuação do Rosetta, consultar a literatura referenciada (ROHL *et al.*, 2004).

O protocolo empregado na modelagem comparativa, foi baseado no artigo de SONG *et al.* (2013). Para a modelagem do canal, foi utilizada a sequência de aminoácidos completa do hNav1.4. Os arquivos necessários para a modelagem foram: arquivo no modelo FASTA referente ao alvo a ser modelado, arquivo tipo *fragment files 3mers* e *9mers* (ALTSCHUL *et al.*, 1997; GRONT *et al.*, 2011; JONES, 1999), *spanfile*, um arquivo usado no Rosetta para modelagem de alfa hélice em proteína de membrana, que informa o número, tamanho e a posição da sequência. Este arquivo foi gerado usando a sequência FASTA com a ferramenta de predição OCTOPUS (VIKLUND; ELOFSSON, 2008) ( <http://octopus.cbr.su.se/> ), e este, convertido para o formato *spanfile* com o *script octopus2span.pl* no *suite* do Rosetta, e um protocolo no formato XML. Todos os protocolos envolvendo o Rosetta foram rodados nos servidores do Altix-xe do LNCC - Laboratório Nacional de Computação Científica (<http://www.lncc.br/estrutura/default.php>) e GAUSS do Cesup - Centro Nacional de

Supercomputação - Universidade Federal do Rio Grande do Sul (<http://www.cesup.ufrgs.br/>). O script utilizado para a modelagem comparativa se encontra na seção complementar.

### 3.1.3 Cluster

Como houve um grande número de candidatos gerados (*poses*), para encontrar o melhor candidato, foi utilizado o programa Calibur (LI, S. C.; NG, 2010) para gerar *clusters*. O programa Calibur é uma ferramenta de agrupamento muito eficiente, tanto para métodos *ab initio* (com grande número de candidatos), quanto para modelagem comparativa (com número menor de modelos gerados porém grandes estruturas proteicas). O programa rodou seguindo os parâmetros: `(./calibur -t r z.txt)` (*-t* especifica a estratégia de *threshold* a ser utilizada, *r* usa um *threshold* baseado nos parâmetros do Rosetta, *z.txt* é um arquivo com os endereços dos arquivos *pdb*s a serem usados).

Neste processo houve perda de dados todos os dados foram apagados do servidor. Sendo assim, não temos os números e os resultados gerados pelo programa. O candidato, considerado pelo programa como mais adequado, foi utilizado para um protocolo de relaxamento do Rosetta, o MPRelax.

O protocolo do Rosetta, MPRelax (ALFORD *et al.*, 2015; BARTH; SCHONBRUN; BAKER, 2007; FLEISHMAN *et al.*, 2011; TYKA *et al.*, 2011), faz uma série de pequenos movimentos da cadeia principal da estrutura, com novas tentativas de reposicionamento das cadeias laterais (biblioteca de rotâmetros) e minimização energética. Foram gerados cerca de 1.000 novos candidatos.

### 3.1.4 TM-align

O resultado final do protocolo MPRelax, foi de aproximadamente 1.000 candidatos. Alguns critérios foram considerados na escolha do modelo adequado entre os candidatos gerados: o modelo com menor *score* dos 1.000 e utilizando a ferramenta TM-align (ZHANG, Y.; SKOLNICK, 2005). A inexistência de ferramentas para aferir a qualidade de grandes estruturas proteicas nos levou a optar pelo uso do programa TM-align. Seu algoritmo é baseado na ferramenta TM-score, que se tornou muito conhecido por usar de ferramentas

matemáticas que compensam as diferenças encontradas no tamanho das sequências (XU, JINRUI; ZHANG, 2010). O programa funciona medindo a similaridade entre duas estruturas de proteínas. O TM-align gera uma nota usando os mesmos critérios que o TM-score (XU, JINRUI; ZHANG, 2010) que indica a similaridade entre o *template* e o modelo, gerando uma nota entre 0 e 1.

$$\text{TM-score} = \text{Max} \left[ \frac{1}{L_N} \sum_{i=1}^{L_T} \frac{1}{1 + \left(\frac{d_i}{d_0}\right)^2} \right]$$

A fórmula acima descreve como o TM-score é calculado para avaliar a semelhança topológica de duas estruturas de proteínas, onde  $L_N$  é o comprimento da proteína alvo a que outra estrutura está alinhada;  $L_T$  é o número de resíduos equivalentes nas 2 proteínas;  $d_i$  é a distância do  $i$ -ésimo par do resíduo equivalente que depende da matriz de sobreposição; Max representa o valor máximo após a superposição espacial ideal;  $d_0 = \sqrt[3]{L - 15 - 1.8}$  é definido para normalizar o TM-score de uma forma que a magnitude do TM-score médio para os pares aleatórios de proteína sejam independentes da dimensão das proteínas. A nota gerada fica entre 0 e 1, sendo que 1 indica que as duas estruturas são perfeitamente equivalentes. Quando o TM-score fica abaixo de 0.20, corresponde a estruturas proteicas não relacionadas, enquanto uma pontuação acima de 0.5 entende-se que as estruturas são equivalentes (XU, JINRUI; ZHANG, 2010; ZHANG, Y.; SKOLNICK, 2005). Com base nestes critérios, o melhor candidato foi selecionado para a dinâmica molecular.

### 3.2 Modelagem Swiss-Model

O SWISS-MODEL (<https://swissmodel.expasy.org/>) (ARNOLD *et al.*, 2006; KOPP; SCHWEDE, 2006; WATERHOUSE *et al.*, 2018) é uma ferramenta *online* para modelagem de proteínas totalmente automatizada. Sua modelagem comparativa funciona seguindo as etapas:

1. Informação do alvo a ser modelado é fornecida ao sistema, sendo uma sequência de aminoácidos, ou um código de acesso UniProtKB.
2. O programa realiza uma busca por *template* compatível com a sequência. Para otimizar a procura utiliza o BLAST (ALTSCHUL *et al.*, 1997; CAMACHO *et al.*, 2009) e o HHblits (REMMERT *et al.*, 2011) adicionando precisão ao método.
3. Seleção do *template* acontece baseada em um resultado de qualidade dos modelos, usando Global Model Quality Estimate (GMQE) (BIASINI *et al.*, 2013) e Quaternary Structure Quality Estimate (QSQE) (BIASINI *et al.*, 2013). Estes critérios são utilizados para a criação de um *ranking* em que as melhores estruturas são ordenadas.
4. Construção do modelo é realizada gerando um modelo da proteína 3D para cada *template* selecionado. Um alinhamento *template*/alvo é realizado transformando as coordenadas dos átomos. A modelagem é feita usando OpenStructure (BIASINI *et al.*, 2013) e o ProMod3 (BIASINI *et al.*, 2013). O ProMod3 é um programa de modelagem baseado no OpenStructure. Este funciona extraindo informações estruturais iniciais da base do *template*, inserções e deleções são definidas pela sequência do alinhamento, e são resolvidas primeiro procurando em um banco de dados estrutural por candidatos viáveis. O candidato final é então selecionado usando critérios estatísticos. Se nenhum candidato for encontrado, a técnica de Monte Carlo é aplicada para este espaço conformacional. A melhor conformação para cadeias laterais é estimada usando o algoritmo TreePack (XU, JINBO; JIAO; BERGER, 2005) e minimizado usando função de energia SCWRL4 (KRIVOV; SHAPOVALOV; DUNBRACK, 2009). O passo final pequenas distorções estruturais, iterações desfavoráveis ou choques são introduzidos durante o processo de modelagem e são resolvidos por minimização energética.
5. A qualidade do modelo é testada usando o QMEAN (BENKERT; BIASINI; SCHWEDE, 2011), que usa uma função de pontuação para testar de forma global e por resíduo a qualidade do modelo. O QMEAN se baseia na combinação linear de descritores estruturais, que depende do potencial de

interação envolvendo a distância baseado no C $\beta$  e todos os tipos de átomos usados para interação de longa distância, ângulo de torção aplicado para três aminoácidos consecutivos usado para analisar a geometria local da cadeia principal da estrutura e um potencial de solvatação para descrever o estado dos resíduos, o cálculo usado para prever a estrutura secundária e acessibilidade do solvente envolvem dois termos, sendo para estrutura secundária PSIPRED (JONES, 1999) e para acessibilidade do solvente ACCpro (CHENG *et al.*, 2005), e o cálculo da estrutura secundária é feita pelo DSSP (KABSCH; SANDER, 1983).

A sequência dos aminoácidos do canal de sódio hNa $_v$ 1.4 foi obtida no NCBI (<https://www.ncbi.nlm.nih.gov/protein/292495096>) com código UniProtKB/Swiss-Prot P35499 (<http://www.uniprot.org/uniprot/P35499>). Usando o programa ClustalOmega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), realizou-se o alinhamento entre as sequências do hNa $_v$ 1.4 e Na $_v$ PaS e EeNa $_v$ 1.4, estruturas que tiveram seu modelo 3D recém publicados utilizando a tecnologia de *cryo-EM* (SHEN *et al.*, 2017; YAN *et al.*, 2017). Este alinhamento serviu para verificar a posição dos segmentos transmembranas, junto com as informações fornecidas pelo NCBS-PDB (<https://www.rcsb.org/>), que é um banco de dados de estruturas de proteínas, e tem a informação da estrutura secundária dos canais Na $_v$ PaS (código 5X0M) e EeNa $_v$ 1.4 (código 5XSY). A sequência do hNa $_v$ 1.4 foi usada no servidor do SWISS-MODEL (<https://swissmodel.expasy.org/>), que identificou as estruturas do Na $_v$ PaS e EeNa $_v$ 1.4 como as com maior identidade em relação a estrutura do hNa $_v$ 1.4.

### **3.3 Docking HADDOCK**

O servidor HADDOCK2.2 (<http://haddock.science.uu.nl/>) (DOMINGUEZ; BOELEN; BONVIN, 2003; VAN ZUNDERT *et al.*, 2016) é um sistema *online* usado para o *docking* entre estruturas.

O HADDOCK usa informações bioquímicas e biofísicas, como perturbações químicas obtidas de experimentos por NMR (*nuclear magnetic resonance*) ou informações de mutagenesis. A informação da interação dos resíduos é introduzida como *ambiguous*

*interaction restraints* (AIR) para direcionar o *docking*. O resultado final é ranqueado de acordo com sua energia intermolecular, que é a soma das energia eletrostáticas, van der Waals e AIR.

O protocolo do *docking* funciona em três estágios:

1. Sorteio das posições e minimização das energias para *rigid body*.
2. Simulação do anelamento onde a estrutura se encontra em um estado semi-rígido para seu ângulo de torsão.
3. Refinamento final no plano cartesiano com solvente explícito.

Todos os processos envolvendo o *docking*, foram realizados utilizando o sistema Guru Interface do HADDOCK web-server (<http://haddock.science.uu.nl/>). Com esse tipo de *interface*, é possível realizar algumas alterações para o *docking*. Como a estrutura do canal é grande e complexa, algumas alterações foram realizadas. O número de estruturas foi alterado para 10.000 (o padrão do programa é gerar 1000 candidatos) e o número de estruturas para refinamento semi-flexível e refinamento com solvente explícito fixado em 400, sendo o padrão 200 estruturas. Para o método de *cluster* foi utilizado o RMSD com linha de corte de 2 Å, onde o programa sugere o RMSD abaixo de 7.5 Å e número de estruturas para análise 1.000 (a configuração original é marcada em 200 estruturas).

### 3.3.1 Ts1

Os resíduos K12, W39, W54 da toxina Ts1 (código de acesso 1NPI) (PINHEIRO *et al.*, 2003) foram considerados como necessários para a ligação, de acordo com resultados experimentais, e foram considerados pelo programa como resíduos ativos (MARTIN-EAUCLAIRE; BOUGIS; DE LIMA, 2018). Os resíduos passivos, são aqueles que de alguma forma colabora para a ligação da proteína, normalmente estando posicionado ao redor dos resíduos ativos, estes foram definidos automaticamente pelo programa. O modelo hNa<sub>v</sub>1.4/PaS, teve o alvo definido no sítio 4, localizado entre o DIIS1-DIIS4, com o resíduos correspondentes de números 596-668, conforme demonstrado na literatura (tabela 3a) (ZHANG *et al.*, 2011).



### 3.3.2 AaHII

Os resíduos F15, R18, W38 e H64 da toxina AaHII (código de acesso 1PTX) (HOUSSET *et al.*, 1994) foram considerados como necessários para a ligação, de acordo com resultados teóricos e experimentais, e os resíduos passivos foram definidos automaticamente pelo programa (CHEN, R.; CHUNG, 2012; HOUSSET *et al.*, 1994; KAHN *et al.*, 2009; WANG *et al.*, 2011). O modelo hNav1.4/PaS, teve o alvo definido no sítio 3, localizado entre o DIVS3 e DIVS4, com os resíduos correspondentes de números 1373-1442, conforme demonstrado na literatura (tabela 3b) (ROGERS *et al.*, 1996; WANG *et al.*, 2011).

### 3.3.3 Ts3

Os resíduos informados ao programa como necessários para a ligação, foram baseados em uma sobreposição entre a toxina AaHII e a Ts3. Os resíduos I15, N18, C38 e K63 da toxina Ts3 (código de acesso 5CY0) (DANG *et al.*, 2016) foram considerados para a ligação. Como não foram encontradas informações específicas, apenas deduções, sobre quais resíduos da toxina e do canal fazem parte da ligação, tomamos como base uma toxina da mesma classe (DANG *et al.*, 2016). O modelo hNav1.4/PaS, teve o alvo definido todo o VSDDIV, incluindo o sítio 3, localizado entre o DIVS3 e DIVS4, com os resíduos correspondentes de números 1373-1442, conforme demonstrado na literatura (ROGERS *et al.*, 1996).

## 4 DINÂMICA MOLECULAR

### 4.1 SWISS-MODEL e HADDOCK

A dinâmica molecular (DM) é uma das ferramentas computacionais mais versáteis para estudo de sistemas moleculares. A metodologia aplicada é baseada na mecânica clássica (dinâmica molecular clássica) onde as forças entre as partículas são tratadas de forma clássica e não quântica. A DM fornece informações sobre o comportamento dinâmico microscópico, dependente do tempo e dos átomos individuais que compõem o sistema. Na mecânica clássica, as moléculas são tratadas como um conjunto de átomos que podem ser descritos por forças newtonianas, sendo assim, são tratados como uma coleção de partículas mantidas unidas por forças harmônicas e elásticas. As simulações de dinâmica, podem trabalhar com sistemas tão simples quanto uma molécula de metano a uma estrutura mais complexa como uma proteína ou membrana.

O algoritmo que governa o NAMD (KALÉ *et al.*, 1999; PHILLIPS *et al.*, 2005) considera o movimento dos átomos de acordo com as equações de movimento newtonianas:

$$m_{\alpha} \vec{r}_{\alpha} = - \frac{\partial}{\partial \vec{r}_{\alpha}} U_{\text{total}} (\vec{r}_1, \vec{r}_2, \vec{r}_3 \dots \vec{r}_N), \alpha = 1, 2 \dots N.$$

onde  $m_{\alpha} \vec{r}_{\alpha}$  é a massa do átomo e sua posição,  $U_{\text{total}}$  é a energia potencial total que depende da posição de todos os átomos (PHILLIPS *et al.*, 2005). O potencial de energia na dinâmica molecular é tratado como campo de força, que é a forma de representar com máxima fidelidade as interações entre os átomos. Este campo de força segue uma fórmula geral a ser aplicada:

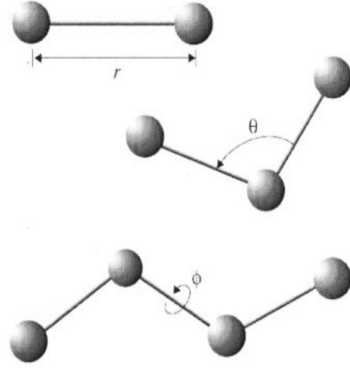
$$U_{\text{total}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{sdW}} + U_{\text{Coulomb}}$$

Na construção desta fórmula, os 3 primeiros são responsáveis por descrever as forças regentes das funções de energia potencial provenientes de interações entre átomos ligados covalentemente:

$$U_{\text{bond}} = \sum_{\text{bonds } i} k_i^{\text{bond}} (r_i - r_{0i})^2,$$

$$U_{\text{angle}} = \sum_{\text{angles } i} k_i^{\text{angle}} (\theta_i - \theta_{0i})^2,$$

$$U_{\text{dihedral}} = \sum_{\text{dihedral } i} \begin{cases} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i - \gamma_i)], & n_i \neq 0 \\ k_i^{\text{dihe}} (\phi_i - \gamma_i)^2 & n_i = 0, \end{cases}$$



*Bonds* representam cada ligação covalente no sistema, *angles* são os ângulos entre cada par de ligação covalente compartilhada com um simples átomo, no centro, e o *dihedral* descreve pares de átomos que separados, por exatamente 3 ligações covalentes em que a ligação central tem um Ângulo de torção  $\phi$  (PHILLIPS *et al.*, 2005). Os dois termos seguintes são responsáveis pelas interações não ligantes,

$$U_{\text{vdW}} = \sum_i \sum_{j>i} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right],$$

$$U_{\text{Coulomb}} = \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}},$$

$U_{\text{vdWaals}}$  correspondem, às forças de van der Waals (com uma aproximação de Lennard-Jones) que é descrita como a energia potencial gerada pela atração fraca entre partículas distantes e pela repulsão forte entre partículas excessivamente próximas,  $U_{\text{Coulomb}}$  é a energia potencial da interação de Coulomb entre duas partículas carregadas eletricamente.

Todos os processos envolvendo Dinâmica Molecular, seguiram exatamente os mesmos passos. Nos modelos do canal de humano, hNav1.4/PaS com e sem o *docking*, foi utilizado

para simular a membrana o lipídio POPC (1-palmitoyl-2-oleoyol-sn-gluco-3-phosphocholine) com 30 ns para o total da simulação. Os arquivos para a dinâmica molecular foram preparados utilizando os programa VMD (HUMPHREY; DALKE; SCHULTEN, 1996) e NAMD (PHILLIPS *et al.*, 2005). O protocolo de montagem da membrana com o canal iônico, e dinâmica molecular, foram feitos baseados no protocolo da *University of Illinois at Urbana-Champaign Beckman Institute for Advanced Science and Technology Theoretical and Computational Biophysics Group Computational Biophysics Workshop*, denominado *Membrane Proteins Tutorial* (<http://www.ks.uiuc.edu/Training/Tutorials/>) (Alek Aksimentiev, Marcos Sotomayor, David Wells February 2012).

A membrana lipídica foi gerada a partir do programa Charmm-gui (JO *et al.*, 2008; JO; KIM; IM, 2007; LEE *et al.*, 2016) no site (<http://www.charmm-gui.org/>) em *input generator > membrane builder > bilayer builder*, empregando os seguintes parâmetros: 200Å no eixo X e Y, 586 lipídios do tipo POPC, e uma camada de 30Å no eixo Z e no -Z para as moléculas de H<sub>2</sub>O TIP3P (JORGENSEN *et al.*, 1983). Lipídios próximos à 1.0Å do canal foram removidos. O sistema foi solvatado com moléculas de H<sub>2</sub>O a 1.5 Å da proteína e uma camada com 30 Å acima e abaixo da membrana lipídica. Todas as moléculas de água dentro da membrana foram removidas. Íons foram adicionados ao sistema em concentração final de 0.15 M de NaCl. O campo de força utilizado para o canal iônico foi CHARMM36 (HUANG, J.; MACKERELL, 2013), para H<sub>2</sub>O TIP3P, e para os lipídeos foi utilizado o NBFIX (DARRÉ; DOMENE, 2015) .

O protocolo específico para proteína em membrana é dividido em quatro etapas, para diminuir os choques estéricos iniciais, seguiu-se três séries de minimização com 10.000 passos e 0.25 ns de simulação e restrições específicas ao sistema: i) curta dinâmica para equilíbrio das caldas da membrana lipídica, mantendo fixas as moléculas de água, íons e proteína com temperatura de 280K; ii) restrição ao movimento da proteína e equilíbrio entre membrana lipídica, íons e água realizado com temperatura de 280K; iii) sem restrições e com um *script* em TCL que é aplicado para evitar que o H<sub>2</sub>O atravessasse a membrana lipídica, uma curta dinâmica é realizada para o equilíbrio de todo sistema a temperatura de 300K, iv) dinâmica realizada com sistema por 30ns com temperatura de 310K. A pressão de todo o sistema foi mantida em 1 atm em todas as etapas, em acordo com método Langevin Piston (MILLER *et al.*, 2012). Para tratamento das interações eletrostáticas de longa distância foi

usado a PME, e para forças de Van der Waals, foram consideradas as distâncias de 10-12 Å. Toda a dinâmica molecular foi rodada no *cluster* Sagarana do ICB - UFMG com 224 núcleos.

O RMSD  $\rho(t)$  em função do tempo é calculado como

$$\rho(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N w_i (\mathbf{x}_i(t) - \mathbf{x}_i^{\text{ref}})^2}$$

É a distância Euclidiana no espaço geral da posição atual (possivelmente após a translação e rotação ideais) subtraído de uma configuração inicial e dividido por  $\sqrt{1/N}$  onde  $N$  é o número das coordenadas.

Os pesos  $w_i$  são calculados a partir dos pesos de entrada  $w'_i$  em relação à média

$$w_i = \frac{w'_i}{\langle w' \rangle}$$

O RMSF de um átomo  $i$  é calculado como:

$$\rho_i = \sqrt{\langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)^2 \rangle}$$

As simulações foram analisadas usando o MDAnalysis (GOWERS *et al.*, [S.d.]), e as imagens geradas usando VMD (HUMPHREY; DALKE; SCHULTEN, 1996).

## 5 RESULTADOS

A princípio questionamo-nos se era possível modelar um canal de sódio parcialmente ou completo e, o mais importante, de forma que ele preservasse suas características funcionais. A ausência de modelos estruturais nos trouxe o primeiro problema: no início da pesquisa, havia apenas o modelo estrutural de um canal de potássio Kv1.2 e um canal de sódio de bactéria, o Na<sub>v</sub>Ab (de *Arcobacter butzleri*). Durante o desenvolvimento deste projeto, foi publicada a primeira estrutura para o canal de cálcio de mamífero, Cav1.1 de coelho (WU *et al.*, 2016). Outra dificuldade, era trabalhar com uma sequência de resíduos muito longa, ao todo 1.836 aminoácidos para o hNa<sub>v</sub>1.4, para modelagem molecular. Apesar da disponibilidade de modelos baseados em bactéria e canais estruturalmente similares como potássio, lacunas teóricas sobre o mecanismo, e mesmo sobre conceitos importantes da estrutura, permaneciam sem resposta. Vários trabalhos experimentais descrevem o canal de sódio em diversos estados e condições, mas ainda havia brechas em relação a informações envolvendo o canal como um todo. Aqui apresentamos modelos gerados por programas diferentes e com *templates* de diferentes organismos. Tais modelos gerados, em princípio, nos trouxeram informações importantes. Também evidenciam as vantagens e limitações oferecidas pelas duas técnicas utilizadas.

### 5.1 Modelagem Rosetta

O protocolo do Rosetta gerou aproximadamente 20.000 candidatos na modelagem a partir do modelo Ca<sub>v</sub>1.1. Para encontrar o melhor candidato a ser usado no protocolo de relaxamento do Rosetta, foi utilizado o programa Calibur, o qual gera o *cluster* das estruturas, e as divide em grupos. Segundo SHORTLE; SIMONS; BAKER, (1998) constatou-se que estruturas com as menores energias conformacionais estão cercadas pelo maior número de estruturas com baixa energia conformacional. O programa foi configurado para obedecer aos critérios do programa Rosetta, que considera como referência principal o *scoring* da estrutura. O candidato mais representativo do maior grupo foi utilizado no protocolo MPRelax. O

*scoring* do modelo antes do protocolo era de -2310.28. Após este protocolo, o candidato com menor *score* (-3258.12) foi o escolhido. Este recebeu nome de hNa<sub>v</sub>1.4/Ca<sub>v</sub>.

Como análise inicial da qualidade da estrutura do modelo hNa<sub>v</sub>1.4/Ca<sub>v</sub>, o gráfico Ramachandran foi gerado (este não adicionado a este documento) com o programa MolProbity (DAVIS *et al.*, 2007). O diagrama de Ramachandran foi desenvolvido em 1963 por Gopalasamudram Ramachandran, e é uma técnica ou forma de visualizar ângulos diedrais  $\psi$  e  $\phi$  dos resíduos de aminoácidos na estrutura de proteínas. Estes ângulos de torção que ocorrem em uma cadeia polipeptídica são chamados de ângulos de Ramachandran, que descrevem a rotação do polipeptídeo entre N-C $\alpha$  (chamado Phi,  $\phi$ ) e C $\alpha$ -C (chamado Psi,  $\psi$ ). Esta é uma forma de verificar os ângulos  $\phi$  e  $\psi$  na estrutura de uma proteína e, também, para determinar se os valores desses ângulos de torção se encontram em regiões consideradas como favoráveis ou desfavoráveis. Os ângulos de torção estão entre os parâmetros mais importantes para o enovelamento de uma proteína, já que estão diretamente envolvidos com a flexibilidade necessária ao enovelamento da estrutura central do polipeptídeo (DAVIS *et al.*, 2007).

O modelo hNa<sub>v</sub>1.4/Ca<sub>v</sub> teve 87.74% dos resíduos com ângulos favoráveis, e 2.61% em ângulos considerados desfavoráveis. A estrutura do Ca<sub>v</sub>1.1 (código pdb 5GJW), que foi usada como referência para a modelagem, tem 86.87% dos resíduos em ângulos favoráveis e 1.74% em ângulos desfavoráveis, como mostrado na tabela 1.

Como ocorre com o canal de sódio em mamíferos, o filtro de seletividade é composto por resíduos conservados presentes em posições específicas em cada domínio do canal. No caso do Ca<sub>v</sub>1.1 é formado por resíduos de glutamato (EEEE) (WU *et al.*, 2016). As posições dos respectivos resíduos DEKA, do modelo hNa1.4/Ca<sub>v</sub>, apresentam uma diferença considerável em relação ao *template*, com distâncias de 1.71 Å para o E761 a 2.29 Å para o A1536 (figura 4).

Outra ferramenta utilizada para aferir a qualidade do modelo foi o TM-align (ZHANG, Y.; SKOLNICK, 2005). O TM-align é um algoritmo para comparação de estruturas proteicas, que usa uma métrica para medir a similaridade estrutural entre dois modelos de proteína. Sua nota varia entre 0 e 1, onde 1 indica que existe uma similaridade perfeita entre os modelos (XU, JINRUI; ZHANG, 2010). O programa prioriza aferir a similaridade do enovelamento global, e é menos sensível a variação estrutural local. TM-align foi utilizado para comparar as

estruturas entre o modelo  $\text{Ca}_v1.1$  e o  $\text{hNa}_v1.4/\text{Ca}_v$ , a nota gerada foi de 0.87762 e o RMSD entre as estruturas foi 4.28 Å.

Para determinar em qual estado conformacional o modelo  $\text{hNa}_v1.4/\text{Ca}_v$  se encontra, analisamos as posições dos resíduos para cada um dos 4 domínios nos seus respectivos segmentos S4. De acordo com a posição das cargas positivas (arginina e lisina), o canal, aparentemente, se encontra no estado aberto, considerando que nos domínios DI e DII 3 resíduos acima da contração hidrofóbica e no DIII 4 resíduos se encontram acima.

Para compreender melhor a conformação do poro do canal, foi utilizado o programa MOLEonline 2.0 (BERKA *et al.*, 2012). MOLEonline é um programa que permite a análise e visualização das dimensões do poro em canais iônicos, que podem incluir estudos de acessibilidade a fármacos e anestésicos, ou análises mais específicas comparando canais iônicos em diferentes estados (JIANG *et al.*, 2018). O resultado gerado pelo programa MOLEonline mostra o filtro de seletividade com a passagem de 0.3 Å de diâmetro, e a parte inferior do poro do canal com 1 Å, concluindo assim que temos o poro fechado neste modelo (figura 4).

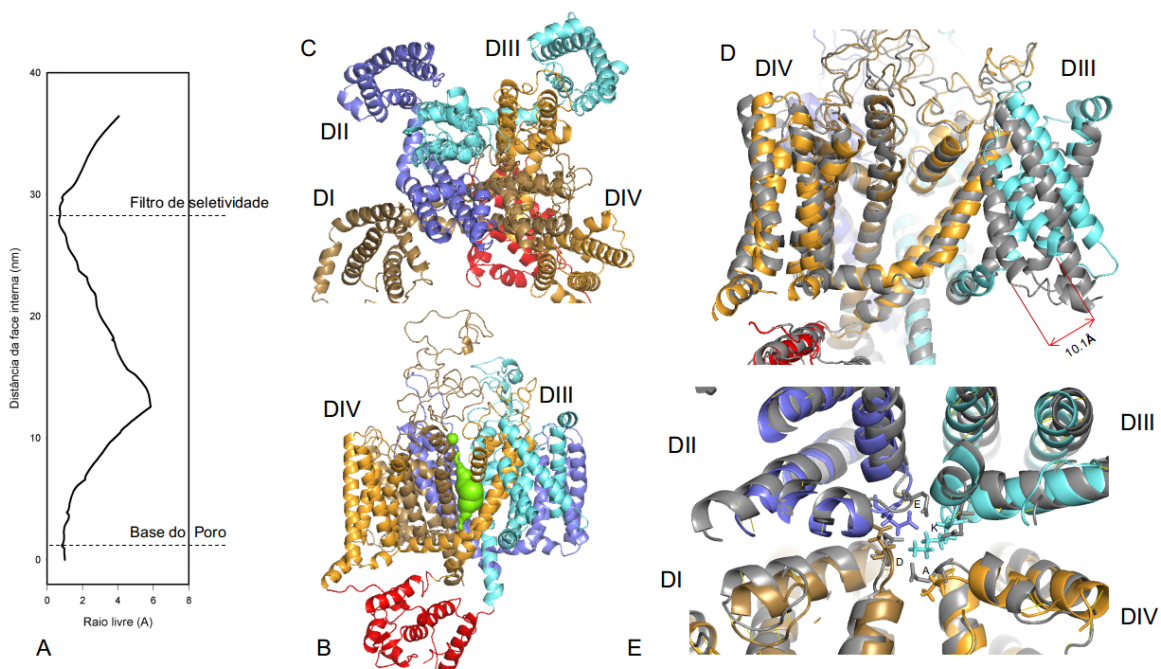


Figura 4 - Modelo canal  $\text{hNa}_v1.4/\text{Ca}_v$ . A - Gráfico feito com o programa MOLEonline representa toda a passagem do poro do canal. B - visão lateral do canal, com vista para o domínio DIV na cor Laranja e DIII na cor Ciano, em destaque no centro em Verde Limão o poro do canal. C - Visão de cima do canal com os quatro domínios. D - Visão lateral do canal, sobreposto com  $\text{Ca}_v1.1$  na cor Cinza, em destaque a diferença no posicionamento do segmento S1 do domínio DIII. E - Vista de cima aproximada, com destaque para o filtro de seletividade, modelo  $\text{hNa}_v1.4/\text{Ca}_v$  e  $\text{Ca}_v1.1$  sobrepostos.



## 5.2 Modelagem SWISS-MODEL

Com o desenvolvimento deste projeto, e por limitações computacionais, novos programas de modelagem foram pesquisados, e dada preferência que seu sistema fosse todo *online*. O candidato encontrado foi o SWISS-MODEL (<https://swissmodel.expasy.org/>), por ser o único a demonstrar capacidade para modelar grandes estruturas proteicas.

Neste meio tempo, novas estruturas de canais iônicos foram surgindo, incluindo as estruturas EeNa<sub>v</sub>1.4 e Na<sub>v</sub>PaS obtidas por *cryo-EM*.

O SWISS-MODEL gerou um modelo para cada *template* inicialmente selecionado - EeNa<sub>v</sub>1.4 e Na<sub>v</sub>PaS - e esses dois modelos tiveram suas estruturas analisadas. Para facilitar, os modelos foram denominados, hNa<sub>v</sub>1.4/Ee, para o modelo gerado a partir do *template* EeNa<sub>v</sub>1.4, e hNa<sub>v</sub>1.4/PaS para o modelo gerado a partir do Na<sub>v</sub>PaS. A comparação das sequências de aminoácidos entre o hNa<sub>v</sub>1.4 e o Na<sub>v</sub>PaS, e o hNa<sub>v</sub>1.4 e EeNa<sub>v</sub>1.4 mostrou a identidade de 42.12% e 66.17%, respectivamente, e com uma cobertura respectivamente de 0.84 e 0.92. O parâmetro cobertura corresponde a quanto da sequência é complementar entre o *template* e o alvo. No caso do resultado ser igual a 1, toda a sequência de aminoácidos do canal hNa<sub>v</sub>1.4 está representada na estrutura.

Como realizado no modelo hNa<sub>v</sub>1.4/Ca<sub>v</sub>, o gráfico de Ramachandran foi gerado para os dois modelos. O modelo hNa<sub>v</sub>1.4/PaS teve 83,91% dos resíduos de aminoácidos com ângulo considerado favorável, e 6,48% em região desfavorável. Se compararmos ao modelo Na<sub>v</sub>PaS, ele teve a relação de 81,85% dos resíduos em região favorável e 5,24% em região desfavorável. Como mostrado no alinhamento da tabela 1, todos os resíduos que foram considerados com ângulos desfavoráveis pelo Ramachandran se encontram nos *loops* das estruturas, no início ou ao final do segmento transmembrana e, em alguns casos, a posição destes resíduos desfavoráveis é a mesma no canal usado como *template*.

Para o modelo hNa<sub>v</sub>1.4/Ee, a mesma análise foi realizada, com 92,25% dos resíduos com ângulo favorável e apenas 1,07%, para ângulo desfavorável. Quando comparado ao seu *template*, o EeNa<sub>v</sub>1.4 teve como resultado 88,52% dos resíduos com ângulo favorável e 3,79% para ângulo desfavorável, resultado que pode ser visualizado no alinhamento suplementar 2.

Os canais usados para modelagem, Na<sub>v</sub>PaS e EeNa<sub>v</sub>1.4, além dos domínios estruturais bem definidos, têm como outras características de destaque suas alças extracelulares. Estas alças foram bem correspondidas nos modelos hNa<sub>v</sub>1.4/PaS e hNa<sub>v</sub>1.4/Ee. Por outro lado, com relação às alças intracelulares, que correspondem às sequências entre os resíduos 464-560, entre os segmentos DI-DII, e 806-1008, entre DII-DIII, não foi possível obter sucesso na sua modelagem, já que nenhum dos dois *templates* oferece uma solução para elas. Buscou-se o alinhamento dessas sequências no banco de dados do RCSB-PDB, à procura de estruturas que pudessem servir de *template*, mas nenhuma estrutura foi encontrada.

O resultado obtido com a ferramenta TM-align, para o modelo hNa<sub>v</sub>1.4/PaS com seu *template* Na<sub>v</sub>PaS foi 0.99149 e o RMSD de 0.51 Å. Para o modelo hNa<sub>v</sub>1.4/Ee e o *template* EeNa<sub>v</sub>1.4 a pontuação foi de 0.99889, com RMSD de 0.37 Å.

Para uma melhor compreensão da estrutura dos poros dos dois modelos, foi utilizado o programa MOLEonline (BERKA *et al.*, 2012). É possível identificar que os dois modelos estão em dois estados bem distintos (figura 5).

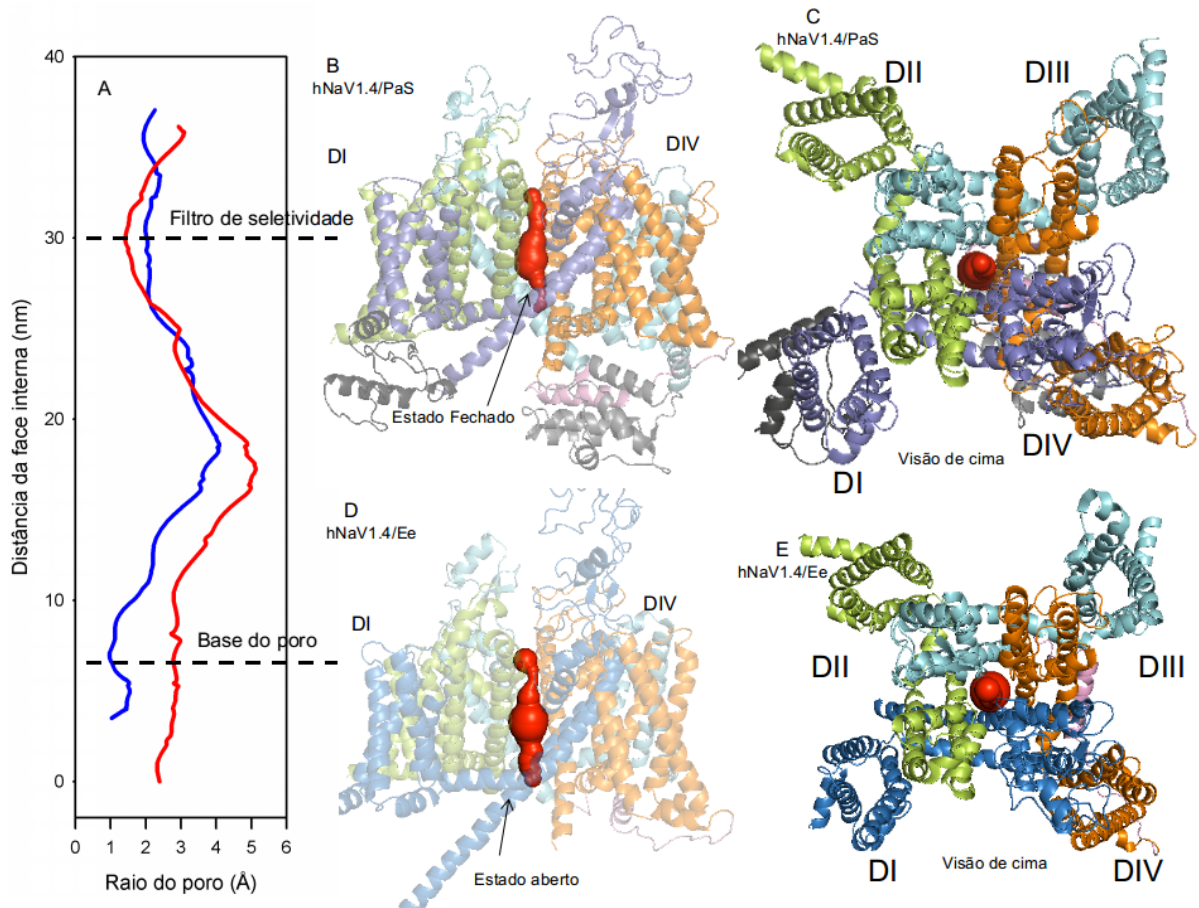


Figura 5 - A) O gráfico da passagem do poro do canal, em Vermelho o modelo  $Na_v1.4/Ee$  e em Azul o modelo  $hNa_v1.4/PaS$ . B) Visão lateral do canal  $hNa_v1.4/PaS$ , vista do domínio DI e DIV, em Vermelho o poro do canal. C) Visão superior do canal  $hNa_v1.4/PaS$ . D) Visão lateral do canal  $hNa_v1.4/Ee$ , vista do domínio DI e DIV, em Vermelho o poro do canal. E) Visão superior do canal  $hNa_v1.4/Ee$ . Os domínios dos canais estão coloridos da seguinte forma: DI Azul, DII Verde, DIII Ciano, DIV Laranja.

Realizou-se uma sobreposição, entre o  $hNa_v1.4/PaS$  e o  $hNa_v1.4/Ee$ , para analisar a posição dos resíduos responsáveis pela seletividade do canal, Aspartato (D406), Glutamato (E761), Lisina (K1244) e Alanina (A1536) (DEKA), entre os modelos, empregando como referência a posição do  $C\alpha$  de cada um destes resíduos. As diferenças observadas de  $0.6 \text{ \AA}$  para D406 e de  $1.3 \text{ \AA}$  para E1536, devem ser consideradas com cautela, levando em consideração que a resolução dos *templates* gira perto dos  $4 \text{ \AA}$ . Como a distância entre os resíduos é menor do que a resolução conseguida por *cryo-EM* do *template*, a confiabilidade nas posições deles diminui muito para pequenas variações. Quando consideramos as cadeias laterais, tendo uma liberdade de movimento, essa distância fica ainda maior  $0.9 \text{ \AA}$  para D406 e  $2.4 \text{ \AA}$  para E761 (figura 6).

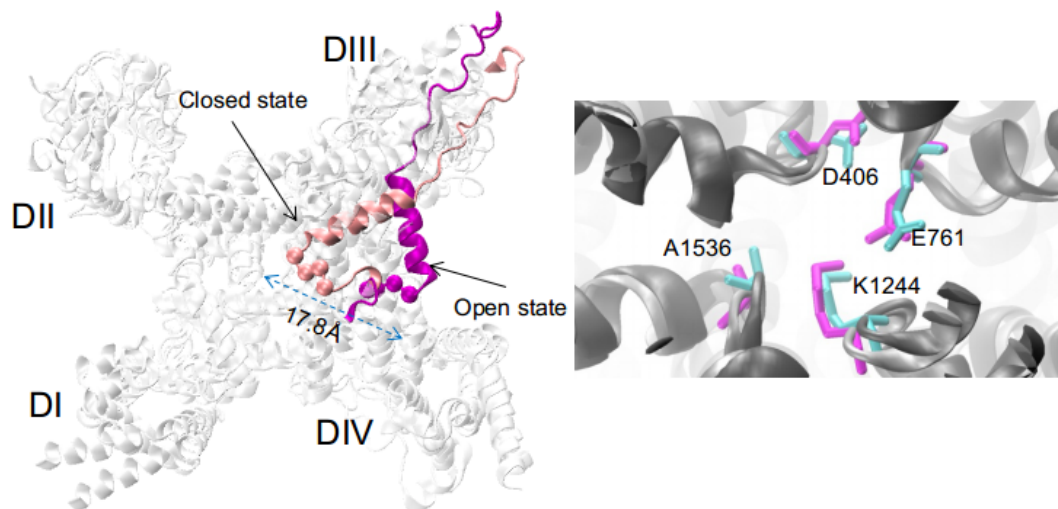


Figura 6 - A) Sobreposição dos canais hNav<sub>v</sub>1.4/PaS e hNav<sub>v</sub>1.4/Ee, com ênfase nas alças de inativação, quando se observa um deslocamento de 17.8Å. Na cor Rosa o estado fechado, e em Roxo o estado aberto. B) Sobreposição dos canais com destaque para o filtro de seletividade e os resíduos que formam o DEKA.

Com essa sobreposição de modelos, foi possível notar que existe um posicionamento diferente entre os resíduos dos segmentos S4. Este deslocamento, medido para as argininas do segmento S4 do DIV, varia entre 8.69 Å para o R1(R1448) a 10.68 Å no caso do R6(R1463). No DIII a distância varia um pouco menos, se comparado ao DIV, mas na ordem de 4.28 Å no R3(R1129) a 5.77 Å no K5(K1138) (figura 7). Os segmentos S6 de todos os domínios também apresentaram uma nova posição, quando comparando os dois modelos. A mudança de posição do segmento S6 no modelo hNav<sub>v</sub>1.4/PaS leva a uma constrição do poro, o que é evidenciado com o gráfico gerado pelo programa MOLEonline (figura 5).

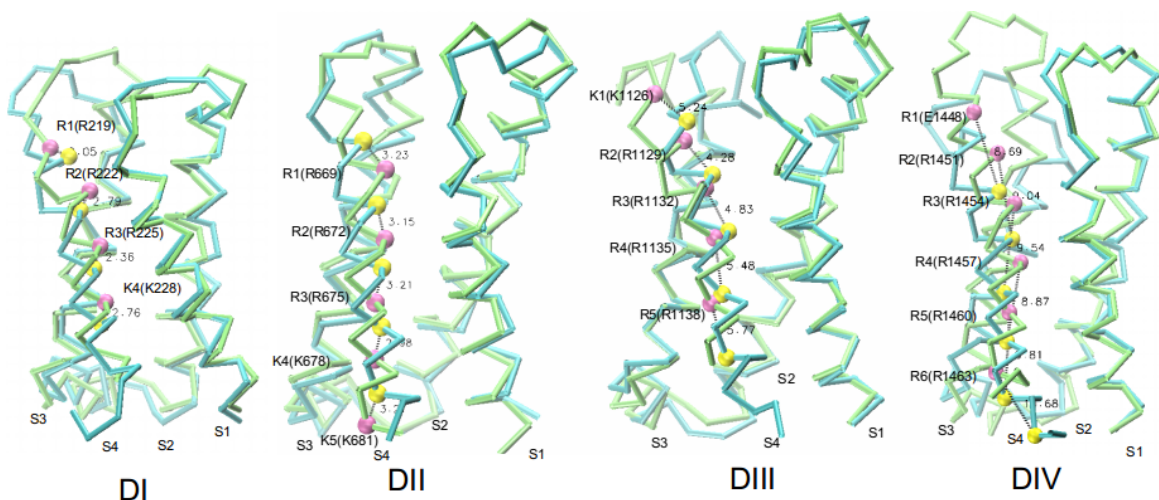


Figura 7 - Sobreposição entre os VSDs hNa<sub>v</sub>1.4/Ee Limão, hNa<sub>v</sub>1.4/PaS Ciano. Os Cα dos resíduos carregados do S4 de cada um dos domínios estão destacados em cores diferentes, sendo Amarelo para o modelo hNa<sub>v</sub>1.4/PaS e Rosa para o modelo hNa<sub>v</sub>1.4/Ee.

Com relação à alça de inativação, ocorre uma mudança no seu posicionamento. No modelo hNa<sub>v</sub>1.4/PaS ela se encontrava próximo a abertura do poro do canal, enquanto no modelo hNa<sub>v</sub>1.4/Ee ela foi deslocada para a lateral, uma mudança de posição de 17.8 Å em relação ao estado anterior (figura 7).

Observou-se uma região conservada em cada VSD, conhecida como constrição hidrofóbica, pela qual as cargas positivas dos segmentos S4 passam durante a despolarização da membrana (LACROIX *et al.*, 2013, 2014). A maneira utilizada para identificar os resíduos que compõem a constrição hidrofóbica no hNa<sub>v</sub>1.4, foi comparando as sequências dos VSDs do canal de sódio humano com o canal Shaker Kv (RSCB cod. 3LUT). Em estudo realizado por Lacroix e colaboradores (LACROIX *et al.*, 2013, 2014) utilizando mutação sítio dirigida, observou-se que os resíduos Isoleucina (I241), Fenilalanina (F290) e Isoleucina (I320) no canal Shaker Kv são de grande importância para a formação desta constrição hidrofóbica. O alinhamento realizado usando o ClustalOmega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) possibilitou identificar os resíduos equivalentes para todos os domínios do hNa<sub>v</sub>1.4 (tabela 2). Os resíduos identificados foram I141, I588, I1041 e I1363 para o segmento S1 de cada domínio, respectivamente. No caso do DIS2 foi identificado o Y168, e para os domínios DII, DIII e DIV foram identificados F621, F1076 e F1396, respectivamente. Já para o DIS3 foi identificado o I201, e para os domínios DII, DIII e DIV os resíduos que compõem a constrição foram identificados como V650, V1105 e V1424, respectivamente (figura 8).

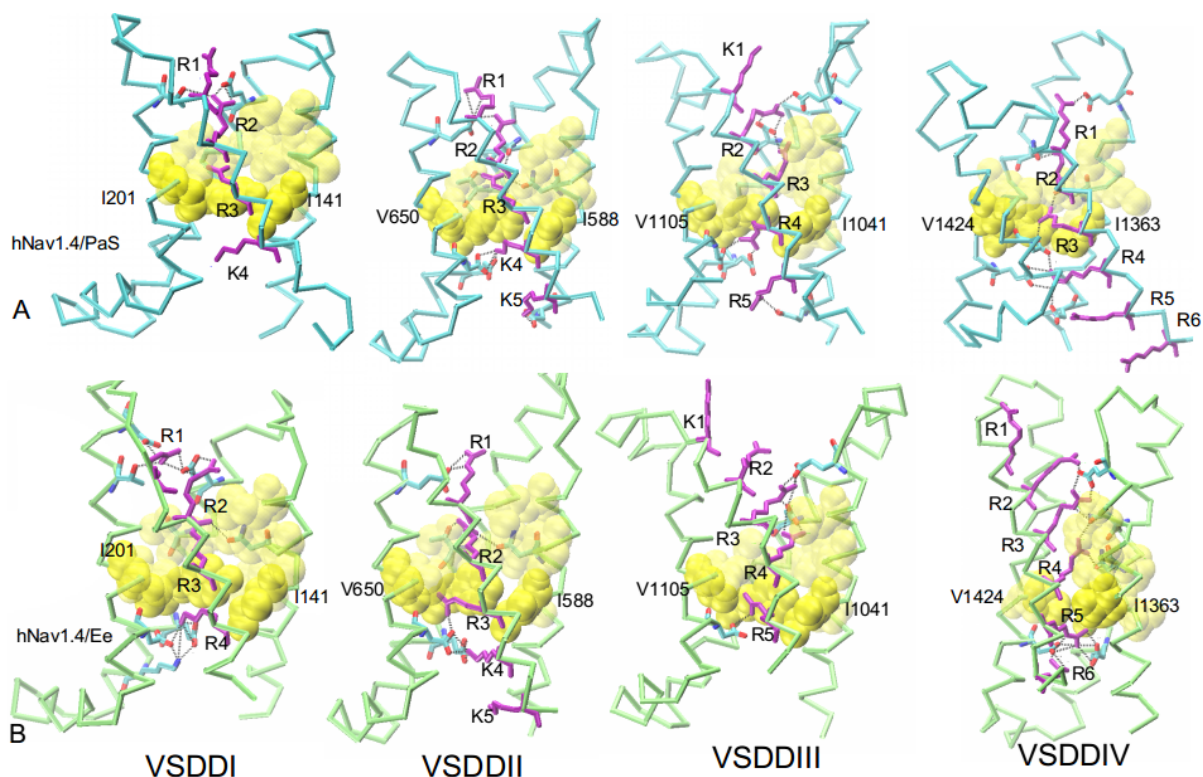


Figura 8 - hNav<sub>v</sub>1.4/PaS VSD de DI-DIV: Ciano para Arginina e Lisina em Roxo, e constrição hidrofóbica em Amarelo. B) hNav<sub>v</sub>1.4/Ee VSD para DI-DIV: em Limão para Arginina e Lisina em Roxo, e resíduos da constrição hidrofóbica em Malva. Os resíduos dos segmentos que ajudam a estabilizar a conformação foram desenhados em *Licore* e na cor azul para os carbonos e vermelho para o oxigênio.

Os resíduos que compõem a constrição hidrofóbica possibilitaram identificar uma transição das cargas positivas do segmento S4 entre o modelo aberto e o modelo fechado. No modelo hNav<sub>v</sub>1.4/PaS foi possível notar que os resíduos R1(R219), R2(R222) e R3(R225) do domínio DI, R1(R669), R2(R672) e R3(R675) do domínio DII, K1(K1126), R2(R1129) e R3(R1132) do DIII, e R1(R1148), R2(R1451) e R3(R1454) do domínio DIV se encontram acima da constrição hidrofóbica. Por outro lado, os resíduos K4(K228) do DI, K4(K678) e K5(K681) do DII, R4(R1135) e R5(R1138) do DIII, e R4(R1457), R5(R1460), e R6(R1463) do domínio DIV encontram-se na parte inferior da constrição hidrofóbica. Já no modelo de estado aberto hNav<sub>v</sub>1.4/Ee, uma diferente situação pôde ser observada: os resíduos K1(K1126), R2(R1129), R3(R1132), R4(R1135) do DIII, e R1(R1448), R2(R1451), R3(R1454), R4(R1457) do domínio DIV se encontram acima da constrição hidrofóbica. As posições dos resíduos nos domínios DI e DII são as mesmas nos dois modelos, R1(R219), R2(R222), e R3(R225) do DI e R1(R669), R2(R672) e R3(R675) para o domínio DII (figura 8).

Os arquivos pdb, referentes às estruturas hNa<sub>v</sub>1.4/PaS e hNa<sub>v</sub>1.4/Ee estão depositadas no Protein Model Database sob o ID PM0081855 e PM0081856 respectivamente.

### 5.3 Docking

#### 5.3.1 Docking da AaHII

O HADDOCK agrupou 347 candidatos em 10 grupos, que correspondem a 86,75% dos candidatos gerados, disponibilizando os 5 melhores grupos. A pontuação que o melhor modelo recebeu de acordo com HADDOCK *score* foi -291.7 e a média do grupo a que pertencia o melhor resultado foi de  $-149.7 \pm 5.7$ . Esse grupo é composto por 113 modelos e o RMSD em relação à estrutura de menor energia foi de 1.1Å.

O melhor modelo foi analisado pelo programa PRODIGY (PROtein binDing enerGY prediction) (VANGONE; BONVIN, 2017; XUE *et al.*, 2016), que é uma ferramenta que usa a correlação entre o número de interfaces de contato do complexo proteína-proteína e a força de ligação. Seu cálculo também é influenciado pela superfície de não interface, que entra no cálculo de regressão linear, resultado no cálculo da energia livre de ligação ( $\Delta G$ ) (VANGONE; BONVIN, 2015; XUE *et al.*, 2016). O valor resultante da energia livre de ligação é usado no cálculo da constante de dissociação ( $K_d$ ), que é realizado de acordo com a seguinte fórmula  $\Delta G = RT \ln K_d$ , onde R é a constante ideal dos gases (em kcal K<sup>-1</sup>mol<sup>-1</sup>), T é a temperatura absoluta e  $\Delta G$  é a energia livre calculada. Para todos os resultados foi estabelecida a temperatura em 310.15 K (37.0 °C). Os valores resultantes para a energia livre de ligação e constante de dissociação são -10.7 kcal/mol e 31 nM, respectivamente.

Como o PRODIGY realiza apenas os cálculos de  $\Delta G$  e  $K_d$ , foi utilizado outro programa para identificar as ligações entre a toxina AaHII e o modelo hNa<sub>v</sub>1.4/PaS. O PISA server (PDBePISA) ([http://www.ebi.ac.uk/msd-srv/prot\\_int/](http://www.ebi.ac.uk/msd-srv/prot_int/)) é uma ferramenta para exploração da interface macromolecular (KRISINEL; HENRICK, 2007). Usando o PyMol (Molecular Graphics System, Version 2.0 Schrodinger, LLC) e o PISA *server*, foram identificadas as interações toxina-canal: D3-N325/S329, F15-L1436, G17-D1435, R18-G1430/L1431/L1433/D1435, Y21-L1436, P41-K377, Y47-S329, R62-N362 (figura 9).

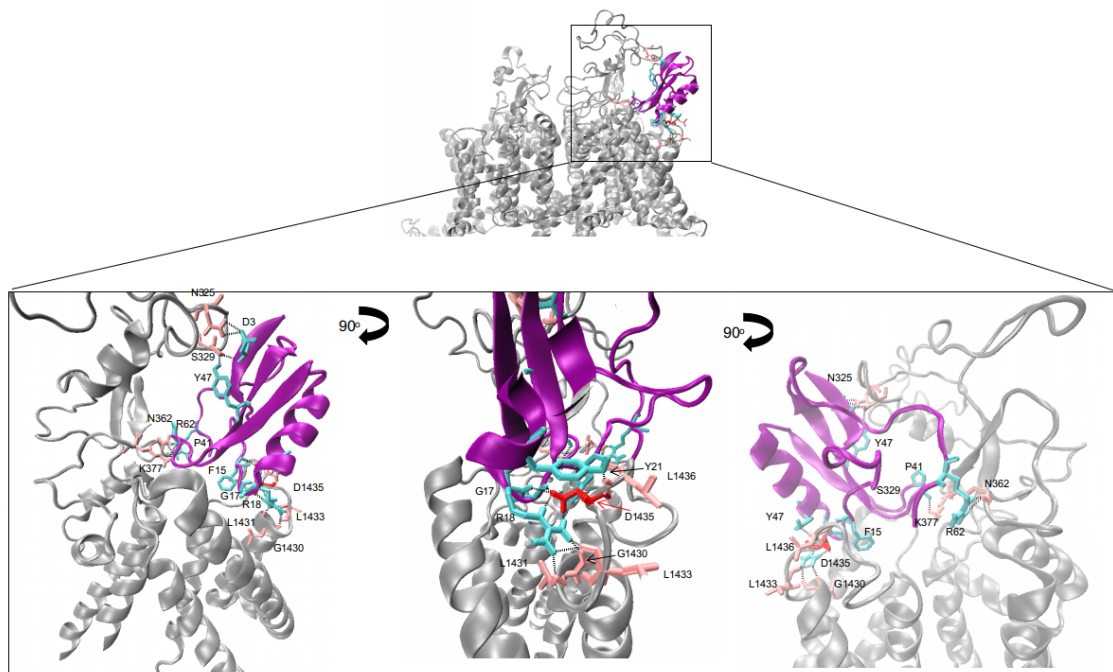


Figura 9 - Zoom do docking hNav<sub>v</sub>1.4/PaS (Cinza), e AahII toxin (Roxo). Resíduos do modelo que se ligam a toxina: os resíduos S329, N362 e K377 se encontram na alça entre o segmento S5 e S6 do domínio DI, e os resíduos G1430, L1431, L1433, D1435 e L1436 estão localizados entre os segmentos S3 e S4 do domínio DIV. Resíduos da toxina AahII que se ligam ao canal: F15, G17, R18, Y21, P41, Y47 e R62.

### 5.3.2 Docking da Ts1

O HADDOCK agrupou 153 candidatos em 18 grupos correspondentes a 76.5% dos candidatos gerados, e disponibilizou os 8 melhores grupos que se formaram. A pontuação que o melhor candidato recebeu, de acordo com HADDOCK *score*, foi -216.5 e a média do grupo o qual pertencia o melhor resultado foi de  $-123.8 \pm 5.6$ . Neste grupo estão alocados 91 modelos, com RMSD de 3.5Å. O PRODIGY server calculou a energia livre de ligação ( $\Delta G$ ) em -9.4 kcal/mol e a constante de dissociação ( $K_d$ ) em 240 nM. Usando o PyMol e o PISA *server*, foram identificadas as interações toxina-canal Y4-Q663/L665, K12-D607/T611, R18-E604/K1206, S20-N1205/S1207, C37-K1259, W39-Y600, W54-Q663, R56-N661/S653, A57-N661, R56-S653/E656/N661, e C61-N608 (Figura 10).





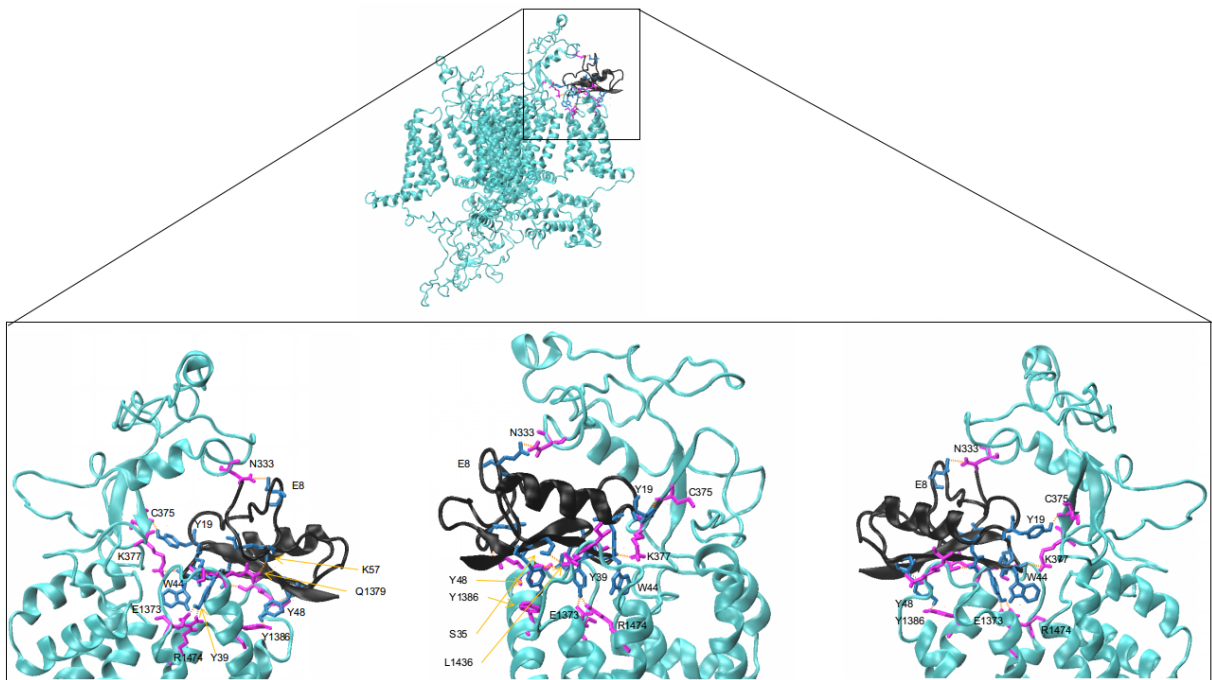


Figura 11 - *Docking* Ts3 no canal hNa<sub>v</sub>.1.4/PaS. Toxina em colorido em Preto. VSDDIV do canal hNa<sub>v</sub>.1.4/PaS em Ciano. Resíduos da Toxina em destaque na cor Roxa. Resíduos do canal na cor Azul.

#### 5.4 Dinâmica molecular hNa<sub>v</sub>.1.4/PaS com e sem a toxina AaHII

Os dados da dinâmica molecular foram analisados de duas formas: (i) como a toxina afeta o equilíbrio da estrutura do canal, e (ii) se a toxina afeta a flexibilidade do DIV-S4 com o *docking*. Para isso foram utilizadas dinâmicas do canal hNa<sub>v</sub>.1.4/PaS e dinâmica do canal hNa<sub>v</sub>.1.4/PaS com *docking* da AaHII, estando o canal nas duas dinâmicas inserido em membrana lipídica (protocolo de inserção na membrana descrito no *script* suplementar).

Para as duas análises, foi utilizado o programa MDAnalysis (GOWERS *et al.*, [S.d.]; MICHAUD-AGRAWAL *et al.*, 2011). MDAnalysis consiste em uma biblioteca na linguagem Python orientada a objeto, para análise de trajetória em simulação de dinâmica molecular. Foi utilizado um pequeno *script* em Python (ver material suplementar).

O RMSF (root mean square fluctuation) é calculado baseado na fórmula logo abaixo e foi realizado para analisar a flexibilidade estrutural da proteína. Seu resultado é capaz de identificar flutuações atômicas locais ou ao longo da cadeia (*script* suplementar). A redução da flexibilidade pela ligação da toxina pode indicar uma alteração local que pode reduzir

drasticamente seu movimento. A análise do gráfico gerado para o RMSF do DIV-S4 foi realizada com o *script* específico, sendo marcado o *backbone* da estrutura de todos os resíduos do segmento S4(1444-1465). No gráfico pode-se observar que o *docking* afeta a os resíduos a partir do número 1455, que chegam no caso do resíduo 1462 e 1464, a um RMSF próximo a 2.4Å, e com a toxina ligada ao canal, este movimento limitou-se a aproximadamente 1.8Å para o resíduo 1462 e 1.7Å para o resíduo 1464 (figura 12).

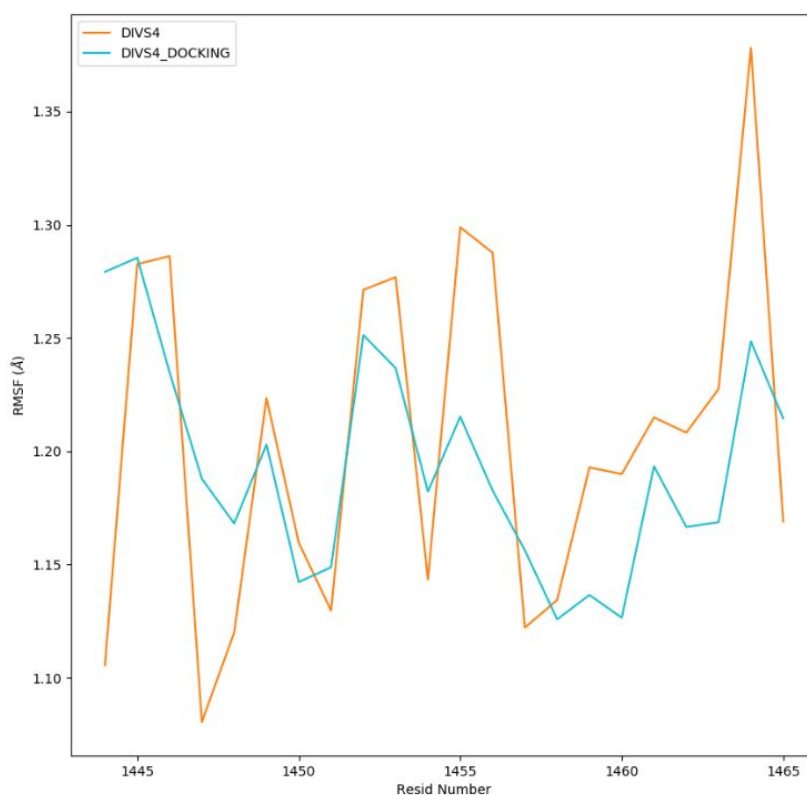


Figura 12 - RMSF da dinâmica molecular realizada para o *docking* da AaHII com canal hNa<sub>v</sub>1.4/PaS. Gráfico referente ao domínio DIV segmento S4. Na cor Laranja a linha referente à dinâmica sem o *docking* e a linha na cor Ciano dinâmica com *docking*.

## 6 DISCUSSÃO

### 6.1 Modelagem

Inicialmente devemos ter cautela com as medidas feitas a partir dos modelos. Como a resolução dos *templates* é limitada, embora as medidas representem valores mais prováveis, elas poderiam levar a valores enganosos, especialmente quando abaixo do poder de resolução dos modelos. Por exemplo, os valores apresentados referentes das distâncias entre os resíduos do filtro de seletividade são abaixo da resolução conseguida nos *templates*. Por esse motivo, essas distâncias devem ser consideradas imprecisas.

Os resultados gerados por modelagem molecular pelo programa Rosetta e pelo servidor SWISS-MODEL preenchem os critérios usados como padrão para análise estrutural de modelos teóricos de estrutura protéica. Também se encontram, muito próximos ao modelo recém publicado do hNa<sub>v</sub>1.4 obtido por *cryo-EM* (PAN *et al.*, 2018).

O modelo hNa<sub>v</sub>1.4/Cav, gerado pelo Rosetta, demonstrou um bom resultado no critério enovelamento, mostrando semelhança considerável ao *template*, com a nota gerada pelo TM-align de 0.84449 (todos os resultados gerados pelo TM-align foram gerados a partir das estruturas sem as alças internas, as quais não foram modeladas). Alguns pontos, que talvez tenham prejudicado um resultado mais próximo, estão relacionados à formação da estrutura secundária, que no segmento entre DIIS2 e DIIS3 seria formada por uma alfa hélice. No entanto, formou-se uma alça no modelo e apareceram pequenas discrepâncias no posicionamento do segmento transmembrana no domínio DIII. O RMSD alto de 4.28 Å está relacionado a algumas porções dos segmentos transmembrana, onde o posicionamento não é o mesmo do *template*. No domínio DIII, o segmento pré-S1 e o S1, quando comparados aos mesmos segmentos do Ca<sub>v</sub>1.1, nota-se uma diferença no posicionamento, que chegou a 10.1 Å (medida feita no início do segmento S1) (figura 4d). Neste caso, a diferença se deve à grande alça que liga o DII ao DIII. Essa diferença no posicionamento provavelmente leva a um efeito cascata no restante do VSD, afetando os segmentos S2 (7.6 Å), S3 (6.6 Å) e S4 (4.1 Å), e apenas voltando ao melhor alinhamento no *linker* S4-S5.

Ao comparar o modelo hNa<sub>v</sub>1.4/Ca<sub>v</sub> com os modelos gerados pelo SWISS-MODEL, hNa<sub>v</sub>1.4/PaS e hNa<sub>v</sub>1.4/PaS, observou-se que há semelhança entre eles, mesmo sendo originados de programas que usam técnicas e *templates* diferentes. São especialmente parecidos nas posições dos segmentos transmembranares, salvo pequenas discordâncias. Os modelos hNa<sub>v</sub>1.4/PaS e hNa<sub>v</sub>1.4/Ee tiveram as notas 0.98631 e 0.99541, respectivamente, quando comparados aos respectivos *templates* (os RMSD foram abaixo de 0.5 Å). Reproduzindo esta comparação entre hNa<sub>v</sub>1.4/PaS e o hNa<sub>v</sub>1.4/Ca<sub>v</sub>, temos o resultado de 0.79317 e o RMSD 5.37 Å. Levando em consideração que os dois canais se encontram em estados diferentes (o hNa<sub>v</sub>1.4/Ca<sub>v</sub> aparentemente se encontra com seus VSDs ativados) pode-se considerar grande a semelhança. No caso do modelo hNa<sub>v</sub>1.4/Ee comparado ao hNa<sub>v</sub>1.4/Ca<sub>v</sub>, o resultado foi 0.75703 com o RMSD de 4.49 Å. Nessa comparação pode-se notar na sobreposição em determinados domínios em que as posições dos segmentos transmembrana quase que se sobrepõem completamente, como nos segmentos que formam o filtro de seletividade, e com destaque para o DIVS1 e DIVS2, onde a sobreposição é quase completa. No segmento DIVS3, ocorre uma pequena discrepância, denotando que o seu final se encontra um pouco afastado. Para o DIVS4 esse deslocamento acontece logo no início do segmento, fazendo com que os finais desses segmentos nos modelos se sobreponham. No DIII, como tem o problema já relatado causado talvez pela grande alça DII-DIII, a sobreposição não ocorre em boa parte do domínio. É algo interessante a se observar, já que entre as duas estruturas, a região com a sobreposição quase completa entre os segmentos do filtro de seletividade P1 e P2 ocorre entre os segmentos S5 e S6.

O resultado gerado pelo programa MOLEonline para o hNa<sub>v</sub>1.4/Ca<sub>v</sub>, junto às informações geradas pelas posições dos segmentos S4, faz com que não tenhamos segurança em afirmar em qual estado se encontra o canal. Segundo Wu e colaboradores (WU *et al.*, 2016) “O poro do canal da  $\alpha$ 1-subunidade está fechado e todos os quatro domínios do sensor de voltagem se encontram na conformação para cima, sugerindo um estado potencialmente inativado”. Esse fato chama a atenção para um aspecto importante e discrepante quanto aos mecanismos de ativação e inativação do canal Ca<sub>v</sub>1.1, que são diferentes do funcionamento dos canais de sódio. Sendo assim, o modelo hNa<sub>v</sub>1.4/Ca<sub>v</sub>, embora construído a partir de *template* mais próximo do que os canais bacterianos, falha em relação às propriedades funcionais do canal.

Quando analisados os aspectos estruturais dos canais hNa<sub>v</sub>1.4/PaS e o hNa<sub>v</sub>1.4/Ee, pode-se afirmar que eles se encontram em estados diferentes, como discutido abaixo.

Ao comparar os resíduos que compõem o filtro de seletividade (DEKA), foi possível notar que, em relação aos C $\alpha$ , o resíduo D406 tem o deslocamento de 0.6 Å e sua cadeia lateral de 1.3 Å. No E761, a posição dos C $\alpha$  é similar, com 0.8 Å de diferença, mas sua cadeia lateral está deslocada de 2.9 Å. No resíduo K1244, o C $\alpha$  tem uma diferença de 1.6 Å e a cadeia lateral está deslocada de 1.2 Å. No caso do resíduo A1536, ele está deslocado de 1.9 Å, considerando como referência seu C $\alpha$ . O programa MOLEonline também foi capaz de medir as dimensões do poro do canal para o modelo hNa<sub>v</sub>1.4/PaS quanto no modelo hNa<sub>v</sub>1.4/Ee. No modelo estado fechado, a medida foi de 1.4 Å de raio no filtro de seletividade e 0.7 Å de raio na base do poro do canal entre os resíduos A444 e L1598. Já o modelo hNa<sub>v</sub>1.4/Ee, estado aberto, a distância no filtro de seletividade é de 2.0 Å de diâmetro entre E761 e K1244, é de 2.8 Å de diâmetro entre os resíduos V446, I1295 e I1597 na base do poro (figura 5).

Baseado nas informações do poro do canal fornecidas pelo programa MOLEonline, e as medidas obtidas pelo programa, podemos concluir que o poro do canal do modelo hNa<sub>v</sub>1.4/PaS se encontra fechado tendo seu raio em 0.7 Å impossibilitando a passagem de um íon de sódio, e o poro do canal do model hNa<sub>v</sub>1.4/Ee com a medida de 2.8 Å se encontra no estado aberto.

## 6.2 Fisiologia molecular do canal

Entre os modelos aberto e fechado, é possível identificar que novas ligações são formadas envolvendo os resíduos carregados do segmento S4, arginina e lisina, e resíduos dos segmentos S1, S2 e S3 (tabela 4). Este rearranjo de ligações, ajudam a estabilizar os resíduos carregados e a manter estável a conformação do VSD nesses diferentes estados. As ligações referentes ao domínio DI e DII não apresentam mudanças significativas, diferentemente do que ocorre com os domínios DIII e DIV. No domínio DIV, no modelo fechado, os resíduos R1451(R2) e R1454(R3) formam ligações com o resíduo D1420 do segmento S3, já no modelo estado aberto os resíduos que interagem com o D1420 são R1460(R5) e R1463(R6),

mostrando desta forma um reposicionamento para fora do DIVS4 na transição de estados. Novas interações são vistas em todos os domínios.

SULA *et al.* (2017) compararam três modelos do canal de sódio, sendo dois originados de bactéria, Na<sub>v</sub>Ms (*Magnetococcus marinus*) no estado ativado, o Na<sub>v</sub>Ab (*Arcobacter butzleri*) no estado pré-ativado e um de planta no estado repouso AtTPC1 (*Arabidopsis thaliana*), chegando a resultado similar. Os TPC (two-pore channel) são canais de cátions simultaneamente expressados em organelas de animais e plantas, e contêm duas subunidades, cada uma com dois domínios com seis segmentos transmembranas, *Shaker-like*, em cada subunidade. Na análise feita entre os modelos, os autores observaram a transição das ligações das cargas positivas do S4, evidenciando novas interações iônicas que ocorrem entre o S4 e resíduos negativos no S1, S2 e S3, no estado aberto e de repouso. O resíduo R2 do TPC1 (repouso), se liga ao E478 na parte inferior do segmento S2 no canal, já no modelo Na<sub>v</sub>Ms (ativado), o R2 se liga ao resíduo E46 na parte superior do segmento S2, quase exposto a superfície extracelular. Estes dois modelos mostram a transição do resíduo R2, ultrapassando a constrição hidrofóbica no estado ativado.

Ao analisarmos as posições dos resíduos positivos do segmento S4, dos domínios DIII e DIV, a sobreposição desses domínios dos modelos hNa<sub>v</sub>1.4/Ee e hNa<sub>v</sub>1.4/PaS mostra que ocorre um giro e um novo posicionamento do segmento para cima. Esta mudança de posição entre os dois modelos varia para cada resíduo. No domínio DIII, a distância ficou em 4.28 Å para o resíduo R2 (R1129) e 5.77 Å para K5 (K1138). No domínio DIV este deslocamento foi ainda maior: para o resíduo R1 (R1448) essa distância foi de 8.69 Å e chegando a 10.68 Å, para o resíduo R6 (R1463). O trabalho realizado por Souza e colaboradores (SOUZA; AMARAL; TREPTOW, 2014) constatou que essa distância variou entre 2 Å a 10 Å, na posição das cargas positivas, resultados próximos aos que obtivemos. Em estudo Li e colaboradores (LI, Q. *et al.*, 2014) determinaram as estruturas cristalográficas da *Ciona intestinalis* VSD (Ci-VSD) nas conformações ativada e de repouso. No seu trabalho, a posição dos resíduos carregados do S4 sofre um deslocamento de ~5Å. Ambos os trabalhos de SOUZA; AMARAL; TREPTOW (2014) quanto o de LI, Q. *et al.* (2014) informam um ângulo de giro do S4, cálculo não realizado neste estudo. Este movimento realizado pelo segmento S4 é denominado como *sliding-helix* ou *helix-screw*, definindo o movimento realizado pelo segmento S4 durante sua ativação. Outros trabalhos relatam o mesmo movimento, originalmente descrito por GUY e SEETHARAMULU (1986) e CATTERALL, W A (1986)

como sendo um movimento de  $\sim 5 \text{ \AA}$  de deslocamento e  $\sim 60^\circ$  graus de giro. No trabalho realizado por Vargas e colaboradores (VARGAS *et al.*, 2012) com canal de potássio, além do movimento realizado pelo S4, que neste caso foi de  $7\text{-}10 \text{ \AA}$ , notou-se também que ocorreu uma diferença na inclinação do segmento, situação que não foi possível constatar em nossos modelos.

Houve uma dúvida sobre o posicionamento das cargas positivas nos domínios DI e DII do modelo hNa<sub>v</sub>1.4/Ee, causado pela pobre resolução das cadeias laterais do *template* EeNa<sub>v</sub>1.4, conforme relato do autor. Isso ficou evidente quando aferimos as distâncias das cargas positivas do S4 no DI e DII em comparação com as mesmas no modelo hNa<sub>v</sub>1.4/PaS. Causou-nos surpresa ao perceber que no domínio DI a distância entre elas era pequena e no DII deu-se a entender que o segmento S4, que deveria estar deslocado para fora, parecia estar deslocado para dentro cerca de  $\sim 3 \text{ \AA}$ , nos tirando-nos a confiabilidade em usar este critério.

Um novo critério foi então aplicado para identificar as posições dos resíduos carregados. Para isso considerou-se a constrição hidrofóbica como limiar de transição entre o lado interno e o externo em relação ao campo elétrico sobre o canal. A constrição hidrofóbica, ou barreira hidrofóbica, é formada por resíduos nos segmentos S1, S2 e S3 em cada um dos 4 VSDs, funcionando como uma barreira dielétrica à permeabilidade da água e íons, criando uma separação, e ainda, definindo os lados extra e intracelular para o campo elétrico transmembrana, como foi descrito por LI, Q. *et al.*, (2014). Sua importância não se restringe a apenas uma separação física entre meio interno e externo. Em pesquisa realizada por LACROIX e colaboradores (LACROIX *et al.*, 2013, 2014) ficou comprovado que essa constrição opera concentrando o campo elétrico junto ao VSD, colaborando também com a estabilidade do estado aberto, e a hidrofobicidade da barreira hidrofóbica que constitui na principal barreira energética para o movimento das cargas de *gating*. Essa conformação nos permitiu identificar a transição das cargas positivas, auxiliando na compreensão sobre o posicionamento interno ou externo das cargas dos segmentos S4, e sua relação com o estado aberto e fechado nos modelos obtidos.

Os resíduos aromáticos Fenilalanina no domínio DII-DIV, e Tirosina no domínio DI, têm papel crucial em estabilizar as cargas positivas do segmento S4 em um ambiente hidrofóbico (LACROIX *et al.*, 2013). Considerando estes resíduos como barreira para as cargas positivas, é possível notar que 3 resíduos carregados positivamente se encontram externos à esta barreira em todos os domínios do hNa<sub>v</sub>1.4/PaS, e no modelo hNa<sub>v</sub>1.4/Ee o



número de 3 resíduos é encontrado no domínio DI e DII, e 4 resíduos carregados para o domínio DIII. Assim sendo, o número de aminoácidos carregados positivamente em posição externa no VSDI e VSDII dos modelos hNa<sub>v</sub>1.4/PaS e Na<sub>v</sub>1.4/Ee é o mesmo, levando a entender que o VSDI e o VSDII, encontram-se na posição ativada. Essas informações levam-nos a crer que o movimento incompleto do VSDIII, como visto aqui, determina o estado fechado do canal hNa<sub>v</sub>1.4/PaS.

Cargas de *gating* são o número de cargas elementares (*e*) que percorrem o campo elétrico da membrana no processo de ativação e/ou inativação do canal. Estas cargas variam conforme o tipo de canal. HIRSCHBERG *et al.* (1995) utilizaram uma técnica que mediu a probabilidade de abertura como uma função do potencial de membrana, assim produziu um indicador quantitativo do número de cargas de *gating* que se moviam através do campo elétrico quando o canal passava do estado fechado para o estado aberto. HIRSCHBERG *et al.* (1995) concluíram, ainda, que ao menos 12 cargas elementares estavam envolvidas neste processo.

No trabalho desenvolvido por GOSSELIN-BADAROUNDINE *et al.* (2012), foi realizada a substituição por histidina da primeira arginina do segmento S4 dos domínios DI-DIII de canais Na<sub>v</sub>1.4. Desta forma foi possível determinar a posição do segmento S4 em relação a constrição hidrofóbica no estado de repouso, onde os autores concluíram que os resíduos R219, R669 e R1129 (dos domínios DI, DII e DII, respectivamente) se encontram no nível da constrição hidrofóbica no estado de repouso. Usando a mesma metodologia foi verificado que no domínio DIV a constrição hidrofóbica é mais longa (cerca de 10Å), sendo por isso mais imprecisa a localização dos resíduos carregados em relação a ela. Tendo isso em mente, nos nossos modelos calculamos as cargas elementares voltadas para o lado externo levando em consideração dois fatores: i) considerar apenas as cargas acima da constrição hidrofóbica, e ii) caso o resíduo apresente interação com resíduos negativos abaixo da constrição hidrofóbica, este seria desconsiderado para o cálculo das cargas.

Considerando que no estado em repouso do canal o resíduo DI/R1, DII/R1 e DIII/R2 se encontram no nível da constrição hidrofóbica, sua diferença em relação ao número das cargas do hNa<sub>v</sub>1.4/Ee são de: no domínio DI 2.5*e*, DII 1.5*e* e DIII 2.5*e* cargas externas à constrição hidrofóbica. Podemos concluir que cerca de 6.5 cargas estão associadas a ativação do hNa<sub>v</sub>1.4. As cargas do DIV não foram acrescentadas à contagem de cargas elementares necessária à ativação, porque esse segmento não participa desse processo. Como no estudo do

GOSSELIN-BADAROUDINE *et al.* (2012) ficou constatado que a constrição hidrofóbica no domínio DIV é maior se comparado ao DI, DII e DIII, que é de  $\sim 2\text{\AA}$ , ajudando a explicar o motivo pelo qual o DIV sofre um atraso na sua ativação, já que ele precisa de mais energia para deslocar o S4 e ultrapassar uma barreira energética maior (CHANDA; BEZANILLA, 2002; SHEETS; HANCK, 1995).

As análises do número de cargas acima da constrição hidrofóbica nos domínios DI e DII do  $hNa_v1.4/PaS$  sugere que seus sensores de voltagem estejam no estado ativado. A fim de verificar esta possibilidade, decidiu-se testar o *docking* da toxina de escorpião Ts1, que é conhecida por se ligar ao VSDII ativado e, assim, mantê-lo na posição (ZHANG, J. Z. *et al.*, 2011).

### 6.3 Interações toxina-canal

O *docking* realizado com a toxina Ts1, demonstrou novas e importantes ligações entre a toxina e o canal. Estudos empregando modificações químicas identificaram que substituições dos resíduos K12, W39 e W54 apresentam diminuição da toxicidade em ratos e diminuição da afinidade de ligação em cérebros de ratos, revelando que são importantes componentes da sua superfície bioativa. Entretanto, outros resíduos como Y4, R18, P20, C37, R56, A57 e W54 se mostraram presentes nesta interação. A participação de resíduos do domínio DIII trouxe-nos novos dados, acrescentando a informação de que as alças externas entre o S5-S6 participam da ligação, auxiliando na estabilização e ligação da toxina. Estudos de mutação identificaram resíduos importantes para ligação e *voltage-dependent trapping* da toxina C<sub>ss</sub>IV no canal  $rNa_v1.2$  (ZHANG, J. Z. *et al.*, 2011). O *voltage-dependent trapping* ocorre quando a toxina do tipo  $\beta$  de escorpião, como a C<sub>ss</sub>IV, estabiliza e retém na posição para fora o segmento S4 do domínio (posição ativada), aumentando a probabilidade de abertura do canal e, portanto, produzindo o deslocamento para potenciais mais negativos da dependência de voltagem da ativação. O alinhamento realizado entre o domínio DII do canal  $rNa_v1.2$  e o canal  $hNa_v1.4$  sugere que os resíduos que mais contribuem para a ligação da toxina são E598, Y600, P601, E656, L659, A660, G664, L665 e S666. Tanto a C<sub>ss</sub>IV, quanto a Ts1 ligam-se ao sítio 4 dos canais de sódio e são assim classificadas como  $\beta$ -toxina. Como não são toxinas idênticas, é possível que outros resíduos estejam envolvidos na ligação da

Ts1. No caso da toxina Ts1, é importante evidenciar que a toxina não precisa de um prepulso de despolarização para se ligar (YATANI *et al.*, 1988). Verificamos que os resíduos Y600, E656 e L665 estão envolvidos com o docking da Ts1 e são parte importante da sequência do canal. De forma interessante, o resíduo N661 também participa do *docking*, onde mostrou-se importante, ligando-se tanto ao R56 quanto ao A57 da toxina. No estudo realizado com CssIV, o resíduo N842 para o canal rNa<sub>v</sub>1.2, quando substituído por arginina, permite que a toxina CssIV atue sem o prepulso (ZHANG, J. Z. *et al.*, 2011). A mutação S788D aumenta a afinidade do canal rNa<sub>v</sub>1.2 à toxina CssIV (ZHANG, J. Z. *et al.*, 2011). No canal de humano, o resíduo equivalente ao S788 no hNa<sub>v</sub>1.4 é D607, exatamente um resíduo aspartato, que mostrou participação no *docking* da Ts1 ao canal humano.

Para verificar se os modelos hNa<sub>v</sub>1.4/PaS e hNa<sub>v</sub>1.4/Ee podem responder pelas propriedades de ligação de uma outra classe de toxina de escorpião, as toxinas do tipo alfa, realizamos os mesmos procedimentos de *docking* para a toxina AaHII.

As interações que ocorreram entre a toxina AaHII e o canal hNa<sub>v</sub>1.4/PaS, reproduzem muito bem os resultados experimentais e teóricos vistos até o momento, principalmente na participação do resíduo R18 da toxina e D1435 do canal (CHEN, R.; CHUNG, 2012; GUR *et al.*, 2011).

Observamos a formação de interações entre a toxina e a alça P1 do domínio DI (Y47-S329, R62-N362), resultado que ainda não havia sido predito em estudos teóricos, já que os modelos propostos, usaram apenas o VSDDIV. Tejedor e Catterall, em 1988, apresentaram pela primeira vez que existe uma participação da alça entre o S5 e S6 do domínio DI, no *docking* da alfa-toxina (TEJEDOR; CATTERALL, 1988). Mais tarde ficou determinado que não apenas a alça S3-S4 faz parte da ligação, mas também a alça S1-S2. Juntas, essa região foi nomeada como receptor primário e a região do loop S5-S6 do domínio DI como receptor secundário (ROGERS *et al.*, 1996; WANG *et al.*, 2011).

Um estudo realizado por KSOURI *et al.* (2018) gerou-se um modelo de um anticorpo baseado em uma tecnologia chamada *Nanobody system*. Este *Nanobody* chamado NbAaHII10, se liga à toxina AaHII na sua face bioativa, neutralizando a toxina. Neste sistema a ligação entre o NbAaHII10 e AaHII é feita principalmente pelo resíduo R18 da toxina, mais uma vez mostrando a importância deste resíduo na bioatividade da toxina.

Para uma análise mais aprofundada, buscou-se realizar o *docking* da mesma toxina usando o canal hNa<sub>v</sub>1.4/Ee, que se encontra no estado aberto, sem o mesmo sucesso.

Comparando os dois modelos de canal, pode-se perceber que os resíduos importantes para o *docking* estão em posições diferentes no estado aberto, em relação ao estado fechado dos modelos do hNa<sub>v</sub>1.4. De forma interessante, foi observado que, no modelo de estado fechado, o arranjo dos resíduos importantes para a ligação da toxina AaHII estão em uma disposição favorável para essa ligação, promovendo uma alta energia de ligação. Por outro lado, as posições dos mesmos resíduos, no modelo estado aberto, são desfavoráveis para a ligação. A comparação dos modelos fechado e aberto traz a conclusão de que o VSDIV, no processo de ativação, tem o seu segmento S4 movendo-se para fora, o que causa uma mudança de posição no segmento S3. Este, por sua vez, faz um giro e um movimento para dentro. Este movimento do S3, faz com que o resíduo D1435 gire para o lado de fora e realize um pequeno movimento descendente, tornando o resíduo indisponível para ligação da toxina (figura 13).

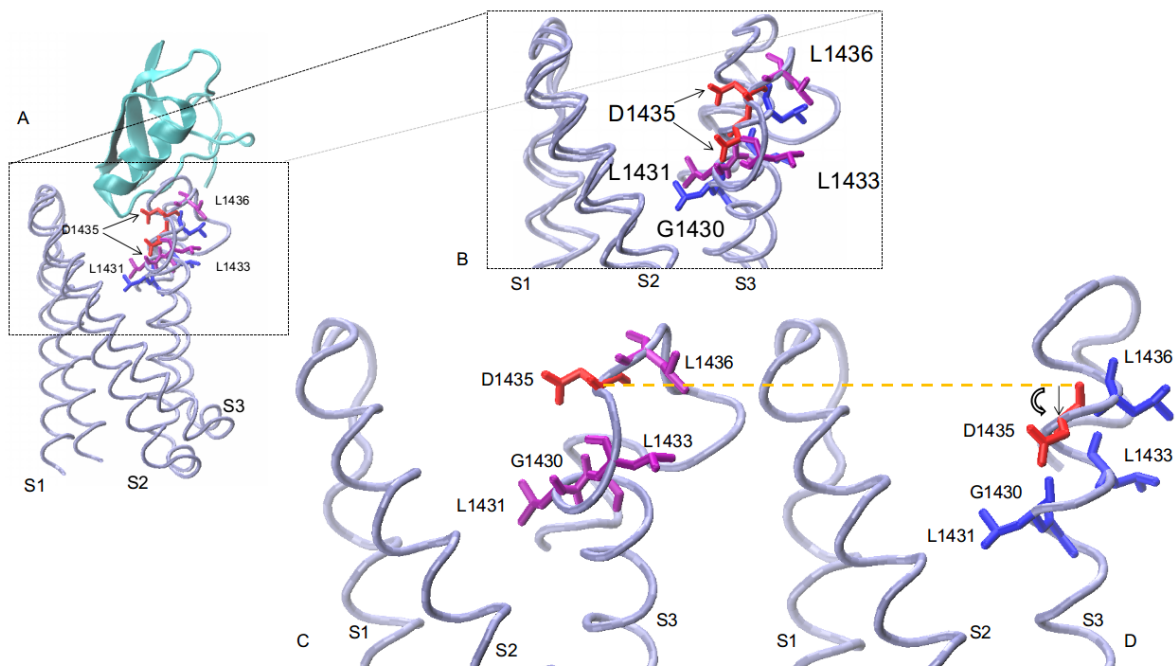


Figura 13 - A) Sobreposição canal hNa<sub>v</sub>1.4/PaS (estado repouso) com *docking* e canal hNa<sub>v</sub>1.4/Ee (estado aberto) sem *docking*. B) Zoom na região onde acontece o *docking*, a toxina foi retirada para melhor visualização. Na figura é mostrado o VSDDIV, com o segmento S4 retirado. Em destaque em Vermelho resíduo D1435, considerado mais importante para o *docking* da AaHII. Outros resíduos que participam do *docking* estão em destaque. A) Sobreposição com a toxina. B) Sobreposição sem a toxina. C e D) Comparando os a posição dos resíduos nos dois canais, é possível notar que o resíduo D1435 no canal aberto é movido para baixo, e gira para o lado de fora.

Estas observações fornecem a base estrutural para uma propriedade bem conhecida das alfa-toxinas de escorpião, a sua remoção do canal quando são aplicados pulsos

despolarizantes fortes à membrana. Os modelos desenvolvidos mostram claramente que os resíduos de aminoácidos importantes para a ligação da toxina se movem para arranjos desfavoráveis à ligação. Para ocorrer o desligamento da toxina, faz-se necessário o fornecimento de energia suficiente para superar a energia de ligação. Essa energia é fornecida pelo pulso elétrico, que força o movimento para fora do S4 do domínio IV. Esta questão foi abordada por CAMPOS e BEIRÃO (2006), que calcularam a energia necessária para a remoção de Ts3 dos canais de sódio GH3. A energia estimada, fornecida pelo pulso elétrico, foi de 27 kJ/mol. Mesmo não tendo informações experimentais, para comprovar o *docking* da toxina Ts3, o resultado da energia de ligação calculada neste estudo, no *docking* foi de 41 kJ/mol. Considerando as diferenças nos procedimentos e preparações, podemos considerar equivalentes.

A dinâmica do *docking* nos trouxe um resultado interessante para o segmento S4 do domínio DIV. O seu RMSF mostra que o final do segmento S4 perde um pouco da sua flexibilidade na parte final do segmento, a partir do resíduo 1452. Levando em consideração que a toxina causa um efeito na rápida inativação, tornando-a mais lenta, a menor flexibilidade do S4 pode estar relacionada ao efeito que a toxina causa no canal (figura 12).

## 7 CONCLUSÃO

Os modelos do canal hNa<sub>v</sub>1.4 gerados a partir das estruturas dos canais Na<sub>v</sub>PaS e EeNa<sub>v</sub>1.4 forneceram bases estruturais para o entendimento de dois estados do canal hNa<sub>v</sub>1.4, um estado aberto e um estado fechado.

O *docking* da toxina Ts1 ao modelo hNa<sub>v</sub>1.4/PaS mostrou que, embora esse modelo esteja no estado fechado, ele não está completamente no repouso, mas sim em um dos vários estados conformacionais em que estão ativados (em posição mais externa) apenas parte dos segmentos S4 necessários para a abertura do canal. Esse resultado validou a nossa abordagem de analisar a posição dos segmentos S4 em relação à constrição hidrofóbica. A partir dessa abordagem concluímos que cerca de 6,5 cargas elementares estão associadas ao processo de ativação. Além disso, foi possível sugerir as ligações toxina-canal que estabilizam essa ligação.

O *docking* das toxinas AaHII e Ts3, por outro lado, além de sugerir as interações toxina-canal, forneceram as bases estruturais para o mecanismo de ação dessas toxinas, confirmando estudos do nosso laboratório.

## 8 PERSPECTIVAS

A criação de modelos de canais de sódio em dois estados funcionais diferentes traz novos horizontes para a compreensão do seu funcionamento.

Com as estruturas que foram modeladas, o desenvolvimento da pesquisa fazendo uso mais intenso da técnica de dinâmica molecular com dinâmicas mais longas em torno de 700 ns, como realizado por Kasimova e colaboradores (KASIMOVA *et al.*, 2018), podem revelar as transições entre estados funcionais.

Esses modelos podem propiciar o desenvolvimento de estudos como o realizado por Huang e colaboradores (HUANG, W. *et al.*, 2017), voltado para uma análise específica de doenças genéticas, realizando as análises diretamente no modelo estrutural.

Novos estudos de *docking* podem ser realizados, abrangendo novos grupos de toxinas, como as toxinas que interagem com filtro de seletividade, como por exemplo a  $\mu$ -conotoxina KIIIA (MCARTHUR *et al.*, 2011), ou mesmo testes de competição por sítios de ligação entre toxinas diferentes.

Finalmente, os modelos podem proporcionar estudos mais abrangentes: o desenvolvimento de drogas específicas para o canal (LEELANANDA; LINDERT, 2016).

TABELA 1

hNa<sub>v</sub>1.4 X Ca<sub>v</sub>1.1

121	RRGAIKVLH	ALFSMFIMITILTNCVFMT	MSDPPPWS-----	KNVEYFTFTGIYTFESLIKIL	177		
41	RKACISIVE	KPFETIILLTIFANCVAVL	VYLPMPEDDN	SLNLGL	EKLEYFFLTVFSIEAAMKII	106	
	*..*..:	* :***:***:*** :	* * .	::** *	:***: **		
178	ARGFCVDDFTFLRD	EWNWLDVSVIMMAYL	TEFVDLG-----	N	214		
107	AYGFLFHQDAYLRS	GWNVLDVFIIVFLGVF	TAILEQVNVIQSN	TAPMSSKAGLD	160		
	:* ** ..:	::** . ** ***	:***: **	:			
215	ISAL	RTFRVLRALKTITV	IPGL	LKTIVGALIQSVK	KLSDVMILTVFCLSVFALVGLQLFMG	274	
161	VKAL	RAFRVLRPLRLVSG	VPS	LQVVLNSIFKAMPLPHI	ALLVLFMVIIYAIIGLELFGK	220	
	:.***:***** *	: : :*.***:***:*** :	* .:	:*.* :	:***:***:*** *		
394	WAFALFRMLTQ	YWENLFQLTLRAAGKTY-MI	FFVVIIFLGSFYLINLILAVVAMAYAE	452			
280	FSMLTVYQCITM	EGWTDVLYWVND	DAIGNEWPVI	YFVTLILLGSFFILNLVGLVLSGE	FTK	339	
	::***:***: **	: * * : : . * * : :	* ** : :	* ** : : * ** : : * ** : : * ** : : * ** : : :			
572	HLIVMDPF	VDLGITICIVLNTLFMA	MEHYPM	EHFDNVLTVGNLVFTGIFTAEMVLKLI	IA	631	
434	-----	FYWLIVLIVALNTLSIA	SEHHP	ELWLT	HLQDIANRVLLSLFTIEMLLKMYG	485	
	.	: : :.*** * * ** :	: : :	:.* * :	:** ** :***: .		
632	MD	PYEFYQQGWNIFDSIIVT-----	LSLVELGLANVQGL	SVLRSFRLLRVFKLA	KSWPTL	686	
486	LGLRQYFMS	IFNRFDCFVVCSGILELL	LVESGAMTPLGI	SVLRCIRLLRLFKIT	KYWTSL	545	
	:. :** . :*	** :***:*** *	* *** *	. * :***:***:***:***:*** *	:*		
687	NMLIKIIGNSV	GALGNLTLVLAII	VFI	FAVGMQLF	KSYK	CVCKIALDCNLPRWHMHD	746
546	SNLVASLLNS	SIRSIASLLLLLFLFII	IIFALLGMQLFGGRYDFEDT-----	EVRRSNFDN	599		
	. * :	: ** : :*** * * * :***:***:*** * .		: : * :***:			
747	FHFSFLIVFRILCG-	EWIETMWDCMEVAG-----	QAMCLTVFLMVMVIGNLVVNLNLFALLLS	803			
600	FPQALISVFQVLTC	EDWNSVMYNGIMAYGG	SYPG	VLVCI-YFII	LFVCGNYILLNVFLAIAVD	662	
	* :***: ** :*	* * ..***: . *		:* :***:*** ** :***:***: . :			
1031	NWFETFIVFMILLSSGAL	AFEDIYIEQRR	VIRTILEYADKVFTYIFIMEMLLK	WVAYGFKVYF	1093		
799	TWFTNFILLFILLSSAALAE	EDPI-RAES	VRNQILGYFDIAFTSVFTVEIVLKM	TTYGAFL-H	859		
	. ** ** :***:*** .*** **	. . *	. ** * * . ** :	:***:*** :*** :	. :		
1094	TNAWC-----	WLDFLIVDVSIIISLVAN	WLGYSSELGP	KSLRTLRLALRPLRALS	RFEGRV	1148	
860	KGSFCRNYFNILDLLVAVSLIS---	MGLESSTISVV	KILRVLRLRPLRAINRAKGLKH	916			
	..:***	** :***:*** ** :***	* * .:	:* ** .***.***:***:*** * :***:			
1149	VVNALLGAIP	SIMNVLLVCLIFWLIFS	IMGVNLFAGKFYICINTT	SERFDISE--	VNNK	1206	
917	VVQCVFAIRTIGNIVLVTLLQ	FMFACIGVQLFKGKFF	SCNDLSKMTEEE	CRGYYVYK	976		
	** :***: ** * * * :***:***	: : * : :***:*** ** : *	: . . :	*			
1207	SECESLM	GTQVRWLVNVKVN	YDNVGLGYLSLLQVATF	KGWMDIMYAAVDSREKEEQPYE	1266		
977	GDPTQ	MELRPRQWIHN	DFHFDNVL	SAMMSLFTVSTF	EGWPQLLYRAIDSNEEDMGPVYN	1036	
	. : *	:* : :***:*** . :***: * ** :***:*** :***:***:*** *		:*			
1267	VN	LYMYLYFVIFII	FGSFFTLNLF	IGVIIDNFNQKKK	LGGKDF	EMTEEQKKYINAMKKL	1326
1037	NRV	EMAIFFIYIILIAFF	MMNIFVGFVIVTF	QEQGETEYKN-CELDKNQRQCV----	QY	1091	
	. : * :***:***:*** :	** :***:***:*** *	:***:*** :	: : :***:*** :			









## TABELA 2

### Alinhamento contração hidrofóbica

	S1 I241	S2 F290	S3 I320
Kv1.2	AIVSVMVILISIVSF	VETLCIIWFSFEFLVREFFA	NIIDIVAIIIPYFI
DI	SMFIMITILTNCVFM	VEYTFGTGIFTFESLIKILA	NWLDFSVIMMAYL
DII	DLGITICIVLNTLFM	GNLVFTGIFTAEMVLKLIA	NIFDSIIVTSLV
DIII	ETFIVFMILLSSGAL	ADKVFTYIFIMEMLLKWVA	CWLDFLIVDVSII
DIV	DITIMILICLNMVTM	INMIFIIIFTGECVLKMLA	NIFDFVVVILSIV

Tabela 2: Em destaque colorido, resíduos que fazem parte da contração hidrofóbica no Kv1.2 alinhados com os domínios do hNa<sub>v</sub>1.4. Magenta para Isoleucina, Ciano para Fenilalanina ou Tirosina, Laranja para Isoleucina ou Valina. Em azul são outros resíduos considerado na literatura como fazendo parte da contração hidrofóbica (CHEN, X. *et al.*, 2010).

### TABELA 3

Alinhamento rNa<sub>v</sub>1.2 x hNa<sub>v</sub>1.4 sítios de afinidade com toxina Ts1.

hNav1.4S1-S2	597	MEHYPMTEHFDNVL	610
rNav1.2S1-S2	778	MEHYPMTEQFSSVL	791
		*****:*.**. **	
hNav1.4S3-S4	656	ELGLANVQGLS	666
rNav1.2S3-S4	837	ELGLANVEGLS	847
		*****:***	

Tabela 3a: Em laranja resíduos importantes para a ligação da toxina ao canal. E598, Y600, P601, E656, L659, A660, G664, L665 E S666.

Alinhamento hNa<sub>v</sub>1.4 x rNa<sub>v</sub>1.2 resíduos de afinidade com toxina AaHII

hNav1.4	1353	QAFDITIMILICLNMTMMVETDNQS	QLKVDILYNINMIFIIIFTGECVLKMLALRQYF	1412
rNav1.2	1532	-VFDISIMILICLNMTMMVETDDQS	QEMTNILYWINLVFIVLFTGECVLKLSLRHYF	59
		.***:*****:***	.:*** **:*:**:*****:::***:***	
hNav1.4	1413	TVGWNIFDFVIVLSIVGLALS	DLIQYFVSPTLFRVIRLARIGRVLRLI	1463
rNav1.2	1590	TIGWNIFDFVIVLSIVGMFLA	LIEKYFVSPTLFRVIRLARIGRILRLI	1640
		*:*****:*. **:***:*****:***		

Tabela 3b: Resíduos importante para o docking destacado em laranja, E1613 no rNa<sub>v</sub>1.2 e D1435 no hNa<sub>v</sub>1.4. Na cor ciano, segmentos S3 e S4.

**TABELA 4**

**Tabela de interações iônicas entre os resíduos carregados do S4 e resíduos do S1, S2 e S3.**

Domínio	Na <sub>v</sub> 1.4/PaS		Na <sub>v</sub> 1.4/Ee	
	Aminoácido carregado	Aminoácidos interagindo	Aminoácido carregado	Aminoácidos interagindo
DI			R219(R1)	E161/T207/D211
	R222(R2)	E161/T207	R222(R2)	E161
			R225(R3)	N144 T165
			K228(K4)	E171/K175/D197
DII	R669(R1)	E656	R669(R1)	E656
	R672(R2)	E656	R672(R2)	N591/T618
	R675(R3)	N591/N614/T618/E624	R675(R3)	D646
	K678(K4)	E624/D646	K678(K4)	E624/ D646
	K681(K5)	D581		
DIII	R1129(R2)	E1051/D1069		
	R1132(R3)	D1069	R1132(R3)	E1051/D1069
	R1135(R4)	E1079/D1101	R1135(R4)	D1069
	R1138(R5)	E1034	R1138(R5)	D1101
DIV	R1448 (R1)	E1374		
	R1451 (R2)	D1420/N1389	R1451 (R2)	E1373
	R1454 (R3)	D1420 / N1366	R1454 (R3)	E1373/N1389
	R1457 (R4)	D1356 / E1399/D1420	R1457 (R4)	N1366/N1389
			R1460 (R5)	E1399 / D1420
			R1463 (R6)	D1420

Tabela 4 - interações iônicas entre os resíduos carregados do S4 e resíduos do S1, S2 e S3

## SUPLEMENTAR

### Script Rosetta

#### Rosetta *fragment file*

```
/home/jrcf/Rosetta/main/source/bin/fragment_picker.linuxgccrelease \  
-database /home/jrcf/Rosetta/main/database \  
-in::file::vall /home/jrcf/Rosetta/tools/fragment_tools/vall.apr24.2008.extended.gz \  
-in::file::fasta 3rvyCH1.fasta \  
-frags::ss_pred t001_.psipred_ss2 predA \  
-frags::bounded_protocol \  
-frags::scoring::config simple.wgths \  
-frags::frag_sizes 9 3 \  
-frags::n_candidates 1000 \  
-frags::n_fragments 200 \  
-out::file::frag_prefix frags \  
-overwrite \  
-frags::describe_fragments frags.fsc
```

#### RosettaCM

```
mpiexec  
/home/u/jrcfer/Rosetta_src_2016.11.58563_bundle/main/source/bin/Rosetta_scripts.mpi.linuxgcc  
release \  
-jd2:mpi_work_partition_job_distributor \  
-database /home/u/jrcfer/Rosetta_src_2016.11.58563_bundle/main/database \  
-in:file:fasta scn4a_human.fasta \  
-parser:protocol Rosetta_cm.xml \  
-relax:minimize_bond_angles \  
-relax:minimize_bond_lengths \  
-relax:jump_move true \  
-default_max_cycles 200 \  
-relax:min_type lbfgs_armijo_nonmonotone \  
-relax:jump_move true \  
-score:weights stage3_rlx_membrane.wts \  
-use_bicubic_interpolation \  
-hybridize:stage1_probability 1.0 \  
-chemical:exclude_patches LowerDNA UpperDNA Cterm_amidation SpecialRotamer VirtualBB  
ShoveBB VirtualDNAPhosphate VirtualNTerm CTermConnect sc_orbitals pro_hydroxylated_case1  
pro_hydroxylated_case2 ser_phosphorylated thr_phosphorylated tyr_phosphorylated  
tyr_sulfated lys_dimethylated lys_monomethylated lys_trimethylated lys_acetylated  
glu_carboxylated cys_acetylated tyr_diiodinated N_acetylated C_methylamidated  
MethylatedProteinCTerm \  
-membrane \  
-in:file:spanfile scn4a.span \  
-membrane:no_interpolate_Mpair \  
-membrane:Menv_penalties \  
-out:levels all:warning core.init:info protocols.jd2.JobDistributor:info \  
-out:file:silent_struct_type binary \  
-out:file:silent_h_coelho_silent.out \  
-out:file:fullatom \  
-scorefile h_coelho.sc \  
-ignore_unrecognized_res \  
-nstruct 20000 \  
-overwrite
```

## Rosetta\_cm.xml

```
<dock_design>
  <TASKOPERATIONS>
</TASKOPERATIONS>
  <SCOREFXNS>
    <stage1 weights="stage1_membrane.wts" symmetric=0>
      <Reweight scoretype=atom_pair_constraint weight=1/>
    </stage1>
    <stage2 weights="stage2_membrane.wts" symmetric=0>
      <Reweight scoretype=atom_pair_constraint weight=0.5/>
    </stage2>
    <fullatom weights="stage3_rlx_membrane.wts" symmetric=0>
      <Reweight scoretype=atom_pair_constraint weight=0.5/>
    </fullatom>
  </SCOREFXNS>
  <FILTERS>
</FILTERS>
  <MOVERS>
    <Hybridize name=hybridize stage1_scorefxn=stage1 stage2_scorefxn=stage2
fa_scorefxn=fullatom batch=1 stage1_increase_cycles=1.0 stage2_increase_cycles=1.0
linmin_only=1>
      <Fragments 3mers="frags.200.3mers" 9mers="frags.200.9mers"/>
      <Template pdb="5gjb_A.pdb" cst_file="AUTO" weight= 1.000 />
    </Hybridize>
  </MOVERS>
  <APPLY_TO_POSE>
</APPLY_TO_POSE>
  <PROTOCOLS>
    <Add mover=hybridize/>
  </PROTOCOLS>
</dock_design>
```

## MPRelax

```
/hpc/Rosetta-2015/Rosetta_bin_linux_2015.20.57849_bundle/main/source/bin/relax.mpi.linuxgcc
release \
-jd2:mpi_work_partition_job_distributor \
-database /hpc/Rosetta-2015/Rosetta_bin_linux_2015.20.57849_bundle/main/database \
-relax:fast \
-relax:constrain_relax_to_start_coords \
-ignore_unrecognized_res \
-in:file:fullatom \
-in:file:s model.pdb \
-in:file:spanfile scn4.span \
-score:weights membrane_highres_Menv_smooth.wts \
-out:prefix mrelax- \
-out:file:scorefile mrelax-.fsc \
-nstruct 2000 \
-out:output \
-overwrite
```



## Protocolo para inserção do canal na membrana

### Alinhamento da membrana e proteína

```
set popc [atomselect top all]
set kcsamol [mol new kcsa_solv.psf]
mol addfile kcsa_solv.pdb
set kcsa [atomselect $kcsamol all]
$popc moveby [vecinvert [measure center $popc weight mass]]
$popc writepdb popc_TEMP.pdb
set vest [atomselect $kcsamol "protein and resid 814 to 824"]
$kcsa moveby [vecinvert [measure center $vest weight mass]]
display resetview
$kcsa move [transaxis y 10]
$kcsa move [transaxis z 15]
$kcsa writepdb kcsa_TEMP.pdb
mol delete all
package require psfgen
resetpsf
readpsf popc.psf
coordpdb popc_TEMP.pdb
readpsf kcsa_solv.psf
coordpdb kcsa_TEMP.pdb
writepsf kcsa_popc_raw.psf
writepdb kcsa_popc_raw.pdb
file delete kcsa_TEMP.pdb
file delete popc_TEMP.pdb
```

### Combinando membrana com proteína

```
mol delete all
mol new kcsa popc raw.psf
mol addfile kcsa popc raw.pdb
set POPC "resname POPC"
set all [atomselect top all]
$all set beta 0
set seltext1 "$POPC and same residue as (name P and z>0 and abs(x)<15 and abs(y)<15)"
set seltext2 "$POPC and same residue as (name P and z<0 and abs(x)<10 and abs(y)<10)"
set seltext3 "$POPC and same residue as (within 0.6 of protein)"
set sel1 [atomselect top $seltext1]
set sel2 [atomselect top $seltext2]
set sel3 [atomselect top $seltext3]
$sel1 set beta 1
$sel2 set beta 1
$sel3 set beta 1
set badlipid [atomselect top "name P and beta > 0"]
set seglistlipid [$badlipid get segid]
set reslistlipid [$badlipid get resid]
set seltext4 "(water and not segname WCA WCB WCC WCD WF SOLV) and same residue as within 3 of ((same residue as (name P and beta>0)) or protein)"
set seltext5 "segname SOLV and same residue as within 3 of lipids"
set sel4 [atomselect top $seltext4]
set sel5 [atomselect top $seltext5]
$sel4 set beta 1
$sel5 set beta 1
set badwater [atomselect top "name OH2 and beta >0"]
set seglistwater [$badwater get segid]
set reslistwater [$badwater get resid]
mol delete all
resetpsf
```

```

readpsf kcsa popc raw.psf
coordpdb kcsa popc raw.pdb
foreach segid $seglistlipid resid $reslistlipid {delatom $segid $resid}
foreach segid $seglistwater resid $reslistwater {delatom $segid $resid}
writepsf kcsa popc.psf
writepdb kcsa popc.pdb

```

## Solvatar e ionizar

```

mol delete all
mol new kcsa popc.psf
mol addfile kcsa popc.pdb
set water [atomselect top water]
measure minmax $water
package require solvate
solvate kcsa_popc.psf kcsa_popc.pdb -o kcsa_popc_water_TEMP -b 1.5 -minmax {{-105 -106 -60}
{107 103 52}}
set all [atomselect top all]
$all set beta 0
set seltext "segid WT1 to WT99 and same residue as abs(z) <25"
set sel [atomselect top $seltext]
$sel set beta 1
set badwater [atomselect top "name OH2 and beta > 0"]
set seglist [$badwater get segid]
set reslist [$badwater get resid]
mol delete all
package require psfgen
resetpsf
readpsf kcsa_popc_water_TEMP.psf
coordpdb kcsa_popc_water_TEMP.pdb
foreach segid $seglist resid $reslist {delatom $segid $resid}
writepdb kcsa_popcw.pdb
writepsf kcsa_popcw.psf
file delete kcsa_popc_water_TEMP.psf
file delete kcsa_popc_water_TEMP.pdb

```

A ionização foi realizada diretamente pelo programa VMD, seguindo o processo *Extensions* → *Modeling* → *Add Ions*, de acordo com os parâmetros abaixo:

Opção	Configuração
Entrada PSF	kcsa_popcw.psf
Entrada PDB	kcsa_popcw.pdb
Saída prefix	kcsa_popcwi
Concentração	0.15
Neutralizar	Yes
Distância mínima da molécula	5
Distância mínima entre ions	5
Seguimento ID	ION

A ionização foi feita procurando o equilíbrio, com 0.15 M de NaCl

Para rodar a simulação, o processo foi dividido em quatro partes para otimizar a dinâmica:

1. Equilíbrio das caldas lipídicas a 280K de temperatura

2. Minimização e equilíbrio com restrições a proteína a 290K de temperatura
3. Equilíbrio com a proteína livre a 300K de temperatura
4. Dinâmica a 310K de temperatura

Os arquivos de configuração são padronizados, com apenas algumas diferenças entre cada etapa, sendo assim, o primeiro arquivo serve como base para todos os outros, acrescentando ou retirando alguns parâmetros. Todos os passos da simulação foram rodados usando o NAMD2.11, pressão de 1 atm. Parâmetros de campo de força utilizados foram para o CHARMM36 ([http://mackerell.umaryland.edu/charmm\\_ff.shtml](http://mackerell.umaryland.edu/charmm_ff.shtml)) arquivo toppar\_c36\_feb16.tgz.

### Arquivos de configuração

```
#####
## ADJUSTABLE PARAMETERS ##
#####

structure      myfile.psf
coordinates    myfile.pdb
outputName     0030ch1POPC-

set temperature 300

# Continuing a job from the restart files
if {0} {
set inputname  kcsa-popcwi
binCoordinates $inputname.restart.coor
binVelocities  $inputname.restart.vel
extendedSystem $inputname.restart.xsc
}
firsttimestep 0
#####
## SIMULATION PARAMETERS ##
#####
# Input
paraTypeCharmm on
parameters    par_all27_prot_lipidNBFIX.prm
parameters    par_all36_lipid.prm
parameters    par_all36_prot.prm
temperature   $temperature

if {1} {
cellBasisVector1 212. 0. 0.
cellBasisVector2 0. 209. 0.
cellBasisVector3 0. 0. 112.
cellOrigin        0.5857862830162048 -0.8354904651641846 -4.566994667053223
}
wrapWater      on
wrapAll        on

# Force-Field Parameters
exclude        scaled1-4
1-4scaling     1.0
cutoff         12.
switching      on
switchdist     10.
pairlistdist   13.5
# Integrator Parameters
timestep       2.0 ;# 2fs/step
```

```

rigidBonds          all
nonbondedFreq      1
fullElectFrequency  2
stepspercycle      20
#PME (for full-system periodic electrostatics)
if {1} {
PME                 yes
PMEGridSizeX       212
PMEGridSizeY       209
PMEGridSizeZ       112
}
# Constant Temperature Control
langevin            on
langevinDamping     1
langevinTemp        $temperature
# Constant Pressure Control (variable volume)
if {0} {
useGroupPressure    yes
useFlexibleCell     yes
useConstantArea     no
langevinPiston      on
langevinPistonTarget 1.01325
langevinPistonPeriod 200.
langevinPistonDecay 50.
langevinPistonTemp  $temperature
}
restartfreq         1000
dcdfreq            1000
xstFreq            1000
outputEnergies     50
outputPressure      50
# Fixed Atoms Constraint (set PDB beta-column to 1)
if {1} {
fixedAtoms          on
fixedAtomsFile      kcsa_popcwi.fix
fixedAtomsCol       B
fixedAtomsForces    on
}
#####
## EXTRA PARAMETERS                                     ##
#####
# Put here any custom parameters that are specific to
# this job (e.g., SMD, TclForces, etc...)
#####
## EXECUTION SCRIPT                                     ##
#####
# Minimization
if {1} {
minimize            10000
reinitvels          $temperature
}
run 250000 ;# 0.25 ns

```

O arquivo com os parâmetros específicos para fixar a bicamada lipídica (kcsa\_popcwi.fix) foi gerado via *tk console* no próprio VMD;

```

mol new kcsa_popcwi.psf
mol addfile kcsa_popcwi.pdb
set all [atomselect top "all"]
$all set beta 0
set fixed [atomselect top "water or name CLA POT or protein or
(name N C12 H12A H12B C13 H13A H13B H13C C14 H14A H14B H14C C15 H15A H15B H15C C11 H11A
H11B P O13 O14 O12 O11)"]
$fixed set beta 1
$all writepdb kcsa_popcwi.fix
exit

```

Minimização e equilíbrio com restrições a proteína:

Neste ponto, o processo procura chegar ao mínimo de energia entre as estruturas em uma configuração de espaço. Este equilíbrio vai ser alcançado restringindo os movimentos da proteína. A configuração do arquivo seguiu com as seguintes alterações:

```
# Continuing a job from the restart files
if {1} {
  set inputname      0030ch1POPC-
  binCoordinates     $inputname.restart.coor
  binVelocities      $inputname.restart.vel
  extendedSystem     $inputname.restart.xsc
}
firsttimestep       251000
#####
## EXTRA PARAMETERS                                ##
#####
constraints on
consexp 2
consref myfile.pdb
conskfile kcsa_popcwi.cnst
conskcol B
margin 4
tclforces on
set waterCheckFreq      100
set lipidCheckFreq      100
set allatompdb          myfile.pdb
tclForcesScript         keep_water_out.tcl
```

Novamente foi criado um arquivo de restrições, mas neste momento, restrições harmônicas para a proteína. Da mesma forma que o arquivo anterior, este também foi gerado diretamente pelo programa VMD utilizando o *tk console*.

```
mol new kcsa_popcwi.psf
mol addfile kcsa_popcwi.pdb
set all [atomselect top "all"]
$all set beta 0
set prot [atomselect top "protein"]
$prot set beta 1
$all writepdb kcsa_popcwi.cnst
exit
```

## Script python mdanalysis

### RMSD

```
import MDAnalysis as mda
import MDAnalysis.analysis.rms
import matplotlib.pyplot as plt

u = mda.Universe("kcsa_popcwi.psf", "kcsa_popcwieq-04_all.dcd")
ref = mda.Universe("kcsa_popcwi.psf", "kcsa_popcwieq-04_all.dcd") #referencia em que começa
a contar o frame = 0

R = mda.analysis.rms.RMSD(u, ref,
                          select="name CA", # sobrepor todo o backbone de toda a proteina
                          groupselections=["name CA and resid 134-232", # DI
                                           "name CA and resid 578-679", # DII
                                           "name CA and resid 1031-1142", # DIII
                                           "name CA and resid 1353-1465" ], # DIV
                          filename="rmsd_all_grupos.dat")

R.run()
R.save()

rmsd = R.rmsd.T # transposição que facilita a plotagem
time = rmsd[0]*2/1000

u2 = mda.Universe("../HADDOCKAah2/kcsa_popcwi.psf",
                  "../HADDOCKAah2/kcsa_popcwieq-04_Wout_all.dcd")
ref2 = mda.Universe("../HADDOCKAah2/kcsa_popcwi.psf",
                    "../HADDOCKAah2/kcsa_popcwieq-04_Wout_all.dcd")

R2 = mda.analysis.rms.RMSD(u2, ref2,
                           select="name CA",
                           groupselections=["name CA and resid 134-232", # DI
                                             "name CA and resid 578-679", # DII
                                             "name CA and resid 1031-1142", # DIII
                                             "name CA and resid 1353-1465" ], # DIV
                           filename="rmsd_all_grupos_dock.dat")

R2.run()
R2.save()

rmsd2 = R2.rmsd.T # transpose makes it easier for plotting
time2 = rmsd2[0]*2/1000
fig = plt.figure(figsize=(10,10))
ax = fig.add_subplot(111)
ax.plot(time, rmsd[2], 'C0', label="DI")
ax.plot(time2, rmsd2[2], 'C1', label="DI_Docking")
ax.legend(loc="best")
ax.set_xlabel("time (ns)")
ax.set_ylabel(r"RMSD ($\AA$)")
fig.savefig("rmsd_DI_DiDocking.png")
```

### RMSF

```
import MDAnalysis as mda
import matplotlib.pyplot as plt
from MDAnalysis.analysis.rms import RMSF
from MDAnalysis.analysis.align import AlignTraj

u = mda.Universe("kcsa_popcwi.psf", "kcsa_popcwieq-04_all.dcd")
```

```
protein = u.select_atoms("resid 1444-1465")

reference_coordinates = u.trajectory.timeseries(asel=protein).mean(axis=1)
reference = mda.Merge(protein).load_new(
    reference_coordinates[:, None, :], order="afc")

aligner = AlignTraj(u, reference, select="resid 1444-1465 and backbone").run()

calphas = protein.select_atoms("backbone")
rmsfer = RMSF(calphas, verbose=True).run()

fig = plt.figure(figsize=(10,10))
ax = fig.add_subplot(111)
ax.plot(calphas.resnums, rmsfer.rmsf, 'k-', label="DIVS4")
ax.legend(loc="best")
ax.set_xlabel("resid number")
ax.set_ylabel("RMSF (Å)")
fig.savefig("rmsfBackBone.png")
```

## 9 Bibliografia

- AHMED, C. M. *et al.* Primary structure, chromosomal localization, and functional expression of a voltage-gated sodium channel from human brain. *Proceedings of the National Academy of Sciences of the United States of America*, v. 89, n. 17, p. 8220–8224, 1 Sep. 1992.
- ALFORD, R. F. *et al.* An integrated framework advancing membrane protein modeling and design. *PLoS Computational Biology*, v. 11, n. 9, p. e1004398, 1 Sep. 2015.
- ALTSCHUL, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, v. 25, n. 17, p. 3389–3402, 1 Sep. 1997.
- ARNOLD, K. *et al.* The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, v. 22, n. 2, p. 195–201, 15 Jan. 2006.
- BARTH, P.; SCHONBRUN, J.; BAKER, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proceedings of the National Academy of Sciences of the United States of America*, v. 104, n. 40, p. 15682–15687, 2 Oct. 2007.
- BENKERT, P.; BIASINI, M.; SCHWEDE, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, v. 27, n. 3, p. 343–350, 1 Feb. 2011.
- BERKA, K. *et al.* MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. *Nucleic Acids Research*, v. 40, n. Web Server issue, p. W222-7, Jul. 2012.
- BIASINI, M. *et al.* OpenStructure: an integrated software framework for computational structural biology. *Acta Crystallographica. Sect. D, Biological Crystallography*, v. 69, n. Pt 5, p. 701–709, May 2013.
- BOOKER, P. D.; WHYTE, S. D.; LADUSANS, E. J. Long QT syndrome and anaesthesia. *British Journal of Anaesthesia*, v. 90, n. 3, p. 349–366, Mar. 2003.
- CAMACHO, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics*, v. 10, p. 421, 15 Dec. 2009.
- CAMPOS, F. V.; BEIRÃO, P. S. L. Effects of bound ts3 on voltage dependence of sodium channel transitions to and from inactivation and energetics of its unbinding. *Cell Biochemistry and Biophysics*, v. 44, n. 3, p. 424–430, 2006.



- CAPES, D. L. *et al.* Domain IV voltage-sensor movement is both sufficient and rate limiting for fast inactivation in sodium channels. *The Journal of General Physiology*, v. 142, n. 2, p. 101–112, Aug. 2013.
- CATTERALL, W. A. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron*, v. 26, n. 1, p. 13–25, Apr. 2000.
- CATTERALL, W. A. Molecular properties of voltage-sensitive sodium channels. *Annual Review of Biochemistry*, v. 55, p. 953–985, 1986.
- CATTERALL, W. A. Voltage-gated calcium channels. *Cold Spring Harbor Perspectives in Biology*, v. 3, n. 8, p. a003947, 1 Aug. 2011.
- CATTERALL, W. A. *et al.* Voltage-gated ion channels and gating modifier toxins. *Toxicon*, v. 49, n. 2, p. 124–141, Feb. 2007.
- CESTÈLE, S. *et al.* Voltage sensor-trapping: enhanced activation of sodium channels by beta-scorpion toxin bound to the S3-S4 loop in domain II. *Neuron*, v. 21, n. 4, p. 919–931, Oct. 1998.
- CHANDA, B.; BEZANILLA, F. Tracking voltage-dependent conformational changes in skeletal muscle sodium channel during activation. *The Journal of General Physiology*, v. 120, n. 5, p. 629–645, Nov. 2002.
- CHEN, R.; CHUNG, S.-H. Binding modes and functional surface of anti-mammalian scorpion  $\alpha$ -toxins to sodium channels. *Biochemistry*, v. 51, n. 39, p. 7775–7782, 2 Oct. 2012.
- CHEN, X. *et al.* Structure of the full-length Shaker potassium channel Kv1.2 by normal-mode-based X-ray crystallographic refinement. *Proceedings of the National Academy of Sciences of the United States of America*, v. 107, n. 25, p. 11352–11357, 22 Jun. 2010.
- CHENG, J. *et al.* SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Research*, v. 33, n. Web Server issue, p. W72-6, 1 Jul. 2005.
- DANG, B. *et al.* Elucidation of the covalent and tertiary structures of biologically active ts3 toxin. *Angewandte Chemie*, v. 55, n. 30, p. 8639–8642, 18 Jul. 2016.
- DARRÉ, L.; DOMENE, C. Binding of capsaicin to the TRPV1 ion channel. *Molecular Pharmaceutics*, v. 12, n. 12, p. 4454–4465, 7 Dec. 2015.

- DAVIS, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, v. 35, n. Web Server issue, p. W375-83, Jul. 2007.
- DOMINGUEZ, C.; BOELEN, R.; BONVIN, A. M. J. J. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, v. 125, n. 7, p. 1731–1737, 19 Feb. 2003.
- FAVRE, I.; MOCZYDLOWSKI, E.; SCHILD, L. On the structural basis for ionic selectivity among Na<sup>+</sup>, K<sup>+</sup>, and Ca<sup>2+</sup> in the voltage-gated sodium channel. *Biophysical Journal*, v. 71, n. 6, p. 3110–3125, Dec. 1996.
- FLEISHMAN, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *Plos One*, v. 6, n. 6, p. e20161, 24 Jun. 2011.
- GOLDIN, A. L. Mechanisms of sodium channel inactivation. *Current Opinion in Neurobiology*, v. 13, n. 3, p. 284–290, Jun. 2003.
- GOLDIN, A. L. *et al.* Nomenclature of voltage-gated sodium channels. *Neuron*, v. 28, n. 2, p. 365–368, Nov. 2000.
- GORDON, D. *et al.* The differential preference of scorpion alpha-toxins for insect or mammalian sodium channels: implications for improved insect control. *Toxicon*, v. 49, n. 4, p. 452–472, 15 Mar. 2007.
- GORDON, D.; GUREVITZ, M. The selectivity of scorpion alpha-toxins for sodium channel subtypes is determined by subtle variations at the interacting surface. *Toxicon*, v. 41, n. 2, p. 125–128, Feb. 2003.
- GOSSELIN-BADAROUNDINE, P. *et al.* Gating pore currents and the resting state of Nav1.4 voltage sensor domains. *Proceedings of the National Academy of Sciences of the United States of America*, v. 109, n. 47, p. 19250–19255, 20 Nov. 2012.
- GOWERS, R. *et al.* MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. [S.d.].
- GRONT, D. *et al.* Generalized fragment picking in Rosetta: design, protocols and applications. *Plos One*, v. 6, n. 8, p. e23294, 24 Aug. 2011.

GUR, M. *et al.* Elucidation of the molecular basis of selective recognition uncovers the interaction site for the core domain of scorpion alpha-toxins on sodium channels. *The Journal of Biological Chemistry*, v. 286, n. 40, p. 35209–35217, 7 Oct. 2011.

GUY, H. R.; SEETHARAMULU, P. Molecular model of the action potential sodium channel. *Proceedings of the National Academy of Sciences of the United States of America*, v. 83, n. 2, p. 508–512, Jan. 1986.

HEINEMANN, S. H. *et al.* Calcium channel characteristics conferred on the sodium channel by single mutations. *Nature*, v. 356, n. 6368, p. 441–443, 2 Apr. 1992.

HIRSCHBERG, B. *et al.* Transfer of twelve charges is needed to open skeletal muscle Na<sup>+</sup> channels. *The Journal of General Physiology*, v. 106, n. 6, p. 1053–1068, Dec. 1995.

HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, v. 117, n. 4, p. 500–544, Aug. 1952.

HOUSSET, D. *et al.* Crystal structure of toxin II from the scorpion *Androctonus australis* Hector refined at 1.3 Å resolution. *Journal of Molecular Biology*, v. 238, n. 1, p. 88–103, 22 Apr. 1994.

HUANG, J.; MACKERELL, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry*, v. 34, n. 25, p. 2135–2145, 30 Sep. 2013.

HUANG, W. *et al.* Structure-based assessment of disease-related mutations in human voltage-gated sodium channels. *Protein & cell*, v. 8, n. 6, p. 401–438, Jun. 2017.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. *Journal of molecular graphics*, v. 14, n. 1, p. 33–8, 27, Feb. 1996.

ISOM, L. L. *et al.* Primary structure and functional expression of the beta 1 subunit of the rat brain sodium channel. *Science*, v. 256, n. 5058, p. 839–842, 8 May 1992.

ISOM, L. L. *et al.* Structure and function of the beta 2 subunit of brain sodium channels, a transmembrane glycoprotein with a CAM motif. *Cell*, v. 83, n. 3, p. 433–442, 3 Nov. 1995.

- JIANG, D. *et al.* Structural basis for gating pore current in periodic paralysis. *Nature*, v. 557, n. 7706, p. 590–594, 16 May 2018.
- JO, S. *et al.* CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, v. 29, n. 11, p. 1859–1865, Aug. 2008.
- JO, S.; KIM, T.; IM, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *Plos One*, v. 2, n. 9, p. e880, 12 Sep. 2007.
- JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, v. 292, n. 2, p. 195–202, 17 Sep. 1999.
- JORGENSEN, W. L. *et al.* Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, v. 79, n. 2, p. 926, 1983.
- KABSCH, W.; SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, v. 22, n. 12, p. 2577–2637, Dec. 1983.
- KAHN, R. *et al.* Molecular requirements for recognition of brain voltage-gated sodium channels by scorpion alpha-toxins. *The Journal of Biological Chemistry*, v. 284, n. 31, p. 20684–20691, 31 Jul. 2009.
- KALÉ, L. *et al.* NAMD2: greater scalability for parallel molecular dynamics. *Journal of Computational Physics*, v. 151, n. 1, p. 283–312, May 1999.
- KARBAT, I. *et al.* Molecular basis of the high insecticidal potency of scorpion alpha-toxins. *The Journal of Biological Chemistry*, v. 279, n. 30, p. 31679–31686, 23 Jul. 2004.
- KASIMOVA, M. A. *et al.* A hypothetical molecular mechanism for TRPV1 activation that invokes rotation of an S6 asparagine. *The Journal of General Physiology*, v. 150, n. 11, p. 1554–1566, 5 Nov. 2018.
- KAUFMANN, K. W. *et al.* Practically useful: what the Rosetta protein modeling suite can do for you. *Biochemistry*, v. 49, n. 14, p. 2987–2998, 13 Apr. 2010.
- KOPEYAN, C. *et al.* Disulfide bonds of toxin II of the scorpion *Androctonus australis* Hector. *European Journal of Biochemistry / FEBS*, v. 47, n. 3, p. 483–489, 16 Sep. 1974.

KOPP, J.; SCHWEDE, T. The SWISS-MODEL Repository: new features and functionalities. *Nucleic Acids Research*, v. 34, n. Database issue, p. D315-8, 1 Jan. 2006.

KRISSINEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, v. 372, n. 3, p. 774–797, 21 Sep. 2007.

KRIVOV, G. G.; SHAPOVALOV, M. V.; DUNBRACK, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, v. 77, n. 4, p. 778–795, Dec. 2009.

KSOURI, A. *et al.* Homology modeling and docking of AahII-Nanobody complexes reveal the epitope binding site on AahII scorpion toxin. *Biochemical and Biophysical Research Communications*, v. 496, n. 4, p. 1025–1032, 19 Feb. 2018.

LACROIX, J. J. *et al.* Molecular bases for the asynchronous activation of sodium and potassium channels required for nerve impulse generation. *Neuron*, v. 79, n. 4, p. 651–657, 21 Aug. 2013.

LACROIX, J. J. *et al.* Moving gating charges through the gating pore in a Kv channel voltage sensor. *Proceedings of the National Academy of Sciences of the United States of America*, v. 111, n. 19, p. E1950-9, 13 May 2014.

LAZARIDIS, T.; KARPLUS, M. Effective energy function for proteins in solution. *Proteins*, v. 35, n. 2, p. 133–152, 1 May 1999.

LEE, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation*, v. 12, n. 1, p. 405–413, 12 Jan. 2016.

LEELANANDA, S. P.; LINDERT, S. Computational methods in drug discovery. *Beilstein journal of organic chemistry*, v. 12, p. 2694–2718, 12 Dec. 2016.

LI, Q. *et al.* Structural mechanism of voltage-dependent gating in an isolated voltage-sensing domain. *Nature Structural & Molecular Biology*, v. 21, n. 3, p. 244–252, Mar. 2014.

LI, S. C.; NG, Y. K. Calibur: a tool for clustering large numbers of protein decoys. *BMC Bioinformatics*, v. 11, p. 25, 13 Jan. 2010.

LIPKIND, G. M.; FOZZARD, H. A. Voltage-gated Na channel selectivity: the role of the conserved domain III lysine residue. *The Journal of General Physiology*, v. 131, n. 6, p.

523–529, Jun. 2008.

MARTIN-EAUCLAIRE, M.-F.; BOUGIS, P. E.; DE LIMA, M. E. Ts1 from the Brazilian scorpion *Tityus serrulatus*: A half-century of studies on a multifunctional beta like-toxin. *Toxicon*, v. 152, p. 106–120, 15 Sep. 2018.

MCARTHUR, J. R. *et al.* Interactions of key charged residues contributing to selective block of neuronal sodium channels by  $\mu$ -conotoxin KIIIA. *Molecular Pharmacology*, v. 80, n. 4, p. 573–584, Oct. 2011.

MICHAUD-AGRAWAL, N. *et al.* MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, v. 32, n. 10, p. 2319–2327, 30 Jul. 2011.

MILLER, A. *et al.* La Crosse viral infection in hospitalized pediatric patients in Western North Carolina. *Hospital pediatrics*, v. 2, n. 4, p. 235–242, Oct. 2012.

MOBLI, M.; UNDHEIM, E. A. B.; RASH, L. D. Modulation of Ion Channels by Cysteine-Rich Peptides: From Sequence to Structure. *Advances in Pharmacology*, v. 79, p. 199–223, 24 Apr. 2017.

PAN, X. *et al.* Structure of the human voltage-gated sodium channel Nav1.4 in complex with  $\beta$ 1. *Science*, v. 362, n. 6412, 19 Oct. 2018.

PAYANDEH, J. *et al.* Crystal structure of a voltage-gated sodium channel in two potentially inactivated states. *Nature*, v. 486, n. 7401, p. 135–139, 20 May 2012.

PAYANDEH, J. *et al.* The crystal structure of a voltage-gated sodium channel. *Nature*, v. 475, n. 7356, p. 353–358, 10 Jul. 2011.

PHILLIPS, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1781–1802, Dec. 2005.

PINHEIRO, C. B. *et al.* Structural analysis of *Tityus serrulatus* Ts1 neurotoxin at atomic resolution: insights into interactions with Na<sup>+</sup> channels. *Acta Crystallographica. Sect. D, Biological Crystallography*, v. 59, n. Pt 3, p. 405–415, Mar. 2003.

POSSANI, L. D. *et al.* Scorpion toxins specific for Na<sup>+</sup>-channels. *European Journal of Biochemistry / FEBS*, v. 264, n. 2, p. 287–300, Sep. 1999.

- REMMERT, M. *et al.* HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods*, v. 9, n. 2, p. 173–175, 25 Dec. 2011.
- RICHMOND, J. E. *et al.* Slow inactivation in human cardiac sodium channels. *Biophysical Journal*, v. 74, n. 6, p. 2945–2952, Jun. 1998.
- ROCHAT, H.; BERNARD, P.; COURAUD, F. Scorpion toxins: chemistry and mode of action. *Advances in cytopharmacology*, v. 3, p. 325–334, 1979.
- ROGERS, J. C. *et al.* Molecular determinants of high affinity binding of alpha-scorpion toxin and sea anemone toxin in the S3-S4 extracellular loop in domain IV of the Na<sup>+</sup> channel alpha subunit. *The Journal of Biological Chemistry*, v. 271, n. 27, p. 15950–15962, 5 Jul. 1996.
- ROHL, C. A. *et al.* Protein structure prediction using Rosetta. [S.l: s.n.], 2004. v. 383. p. 66–93
- SHEETS, M. F.; HANCK, D. A. Voltage-dependent open-state inactivation of cardiac sodium channels: gating current studies with Anthopleurin-A toxin. *The Journal of General Physiology*, v. 106, n. 4, p. 617–640, Oct. 1995.
- SHEN, H. *et al.* Structure of a eukaryotic voltage-gated sodium channel at near-atomic resolution. *Science*, v. 355, n. 6328, 3 Mar. 2017.
- SHORTLE, D.; SIMONS, K. T.; BAKER, D. Clustering of low-energy conformations near the native structures of small proteins. *Proceedings of the National Academy of Sciences of the United States of America*, v. 95, n. 19, p. 11158–11162, 15 Sep. 1998.
- SONG, Y. *et al.* High-resolution comparative modeling with RosettaCM. *Structure*, v. 21, n. 10, p. 1735–1742, 8 Oct. 2013.
- SOUZA, C. S.; AMARAL, C.; TREPTOW, W. Electric fingerprint of voltage sensor domains. *Proceedings of the National Academy of Sciences of the United States of America*, v. 111, n. 49, p. 17510–17515, 9 Dec. 2014.
- SULA, A. *et al.* The complete structure of an activated open sodium channel. *Nature Communications*, v. 8, p. 14205, 16 Feb. 2017.
- TEJEDOR, F. J.; CATTERALL, W. A. Site of covalent attachment of alpha-scorpion toxin derivatives in domain I of the sodium channel alpha subunit. *Proceedings of the National*

*Academy of Sciences of the United States of America*, v. 85, n. 22, p. 8742–8746, Nov. 1988.

TYKA, M. D. *et al.* Alternate states of proteins revealed by detailed energy landscape mapping. *Journal of Molecular Biology*, v. 405, n. 2, p. 607–618, 14 Jan. 2011.

VAN ZUNDERT, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of Molecular Biology*, v. 428, n. 4, p. 720–725, 22 Feb. 2016.

VANGONE, A.; BONVIN, A. PRODIGY: A Contact-based Predictor of Binding Affinity in Protein-protein Complexes. *Bio-protocol*, v. 7, n. 3, 2017.

VANGONE, A.; BONVIN, A. M. Contacts-based prediction of binding affinity in protein-protein complexes. *eLife*, v. 4, p. e07454, 20 Jul. 2015.

VARGAS, E. *et al.* An emerging consensus on voltage-dependent gating from computational modeling and molecular dynamics simulations. *The Journal of General Physiology*, v. 140, n. 6, p. 587–594, 1 Dec. 2012.

VIKLUND, H.; ELOFSSON, A. OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics*, v. 24, n. 15, p. 1662–1668, 1 Aug. 2008.

WANG, J. *et al.* Mapping the receptor site for alpha-scorpion toxins on a Na<sup>+</sup> channel voltage sensor. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 37, p. 15426–15431, 13 Sep. 2011.

WATERHOUSE, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, v. 46, n. W1, p. W296–W303, 2 Jul. 2018.

WEST, J. W. *et al.* A cluster of hydrophobic amino acid residues required for fast Na(+)-channel inactivation. *Proceedings of the National Academy of Sciences of the United States of America*, v. 89, n. 22, p. 10910–10914, 15 Nov. 1992.

WU, J. *et al.* Structure of the voltage-gated calcium channel Ca(v)1.1 at 3.6 Å resolution. *Nature*, v. 537, n. 7619, p. 191–196, 8 Sep. 2016.



- XU, J.; JIAO, F.; BERGER, B. A tree-decomposition approach to protein structure prediction. *Proceedings / IEEE Computational Systems Bioinformatics Conference, CSB. IEEE Computational Systems Bioinformatics Conference*, p. 247–256, 2005.
- XU, J.; ZHANG, Y. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics*, v. 26, n. 7, p. 889–895, 1 Apr. 2010.
- XUE, L. C. *et al.* PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*, v. 32, n. 23, p. 3676–3678, 1 Dec. 2016.
- YAN, Z. *et al.* Structure of the Nav1.4- $\beta$ 1 Complex from Electric Eel. *Cell*, v. 170, n. 3, p. 470–482.e11, 27 Jul. 2017.
- YATANI, A. *et al.* Effects of New World scorpion toxins on single-channel and whole cell cardiac sodium currents. *The American Journal of Physiology*, v. 254, n. 3 Pt 2, p. H443-51, Mar. 1988.
- ZHANG, A. H. *et al.* A complicated complex: Ion channels, voltage sensing, cell membranes and peptide inhibitors. *Neuroscience Letters*, v. 679, p. 35–47, 13 Jul. 2018.
- ZHANG, J. Z. *et al.* Mapping the interaction site for a  $\beta$ -scorpion toxin in the pore module of domain III of voltage-gated Na(+) channels. *The Journal of Biological Chemistry*, v. 287, n. 36, p. 30719–30728, 31 Aug. 2012.
- ZHANG, J. Z. *et al.* Structure-function map of the receptor site for  $\beta$ -scorpion toxins in domain II of voltage-gated sodium channels. *The Journal of Biological Chemistry*, v. 286, n. 38, p. 33641–33651, 23 Sep. 2011.
- ZHANG, Y.; SKOLNICK, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, v. 33, n. 7, p. 2302–2309, 22 Apr. 2005.



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)

## Closed- and open-state models of human skeletal muscle sodium channel

João R.C. Fernandes, Lucas Bleicher, Paulo S.L. Beirão\*

Department of Biochemistry and Immunology – ICB, Universidade Federal de Minas Gerais - UFMG, Av. Antonio Carlos, 6627, 31270-901, Belo Horizonte, MG, Brazil

### ARTICLE INFO

#### Article history:

Received 10 October 2018

Accepted 17 October 2018

Available online xxx

#### Keywords:

Human muscle sodium channel  
Molecular model  
Voltage sensor  
Charge movement  
Scorpion toxin

### ABSTRACT

Voltage-gated sodium channels play important roles in human physiology. However, their complexity hinders the understanding of their physiology and pathology at atomic level. We took advantage of the structural reports of similar channels obtained by cryo-EM (EeNav1.4, and NavPaS), and constructed models of human Nav1.4 channels at closed and open states. The open-state model is very similar to the recently published cryo-EM structure of hNav1.4. The comparison of both models shows shifts of the voltage sensors (VS) of DIII and DIV. The activated position of VS-DIII in the closed model was demonstrated by Ts1 docking, thereby confirming the requirement that VS-DI, VS-DII and VS-DIII must be activated for the channel to open. The interactions observed with VS-DIII suggest a stepwise, yet fast, transition from resting to activated state. These models provide structural insights on the closed-open transition of the channel.

© 2018 Published by Elsevier Inc.

### 1. Introduction

Voltage-gated sodium channels play a crucial role in the generation and transmission of action potentials of neurons and vertebrate skeletal and cardiac muscles. This fact makes sodium channels important targets for toxins and drugs. There are nine types of human sodium channels, named hNav<sub>v</sub>1.1 through hNav<sub>v</sub>1.9, that can be identified by their amino acid sequence, biophysical properties, pharmacology and tissue where they are expressed [1–3], being hNav1.4 characteristic of skeletal muscle. The basic structure of voltage-dependent ion channels is well preserved in evolution, especially when the positions of the transmembrane segments are concerned.

The major functional structure of eukaryotic sodium channels is the alpha subunit (mw. ~260 kDa), which contains the Na<sup>+</sup> conducting pore and the voltage-sensing mechanism, and may be associated with beta-subunits. The alpha subunit contains four homologous domains, named DI through DIV, each one composed of six transmembrane segments, denominated S1 to S6. Within each domain, we can identify a voltage-sensing domain VSD

(formed by S1 through S4) and an ion-conduction pore domain formed by the pseudo-tetrameric arrangement of the S5 and S6 segments of the four domains [3]. The selectivity filter is formed by loops between the S5 and S6 of each domain, radially organized, with the narrowest part formed by a ring of four amino acids (aspartate, glutamate, lysine, and alanine, from DI, DII, DIII and DIV, respectively), therefore known as DEKA. The voltage sensitivity is conferred by the S4 segments, which contain a repeated sequence of basic amino acids (mostly arginine) located at every third position of the segment [3], referred to as gating charge (GC). The “Hydrophobic constriction” or “gating pore” is a region formed by hydrophobic residues that prevents the passage of water and forms an electric energy barrier that is crossed by the positively charged residues of S4 during the activation of the channel [4,5].

Despite all important contributions provided by the biophysical, pharmacological and molecular biology methods, one important obstacle to the understanding of the human sodium channel function is the near absence of experimentally determined structures. This shortcoming is due to their complex structure that so far prevented their crystallization and X ray diffraction analysis or NMR studies. This lacuna hinders the rational prediction of drugs that could alter their function and the understanding of the consequences of mutations in their primary structures. On the other hand, relevant works were published with the experimental determination at atomic details on the structure of the voltage-

*Abbreviations:* hNav1.4/PaS, model of closed human skeletal muscle sodium channel; hNav1.4/Ee, model of open human skeletal muscle sodium channel.

\* Corresponding author.

E-mail address: [pslb@ufmg.br](mailto:pslb@ufmg.br) (P.S.L. Beirão).

<https://doi.org/10.1016/j.bbrc.2018.10.110>

0006-291X/© 2018 Published by Elsevier Inc.

gated sodium channel of bacteria [6] and, using electron cryo-microscopy (cryo-EM), the structures of eukaryotic sodium channels (cockroach NavPaS and electric eel EeNav1.4) [7,8]. Very recently, the structure of hNav1.4 obtained by cryo-EM was published [9].

The present work aims to generate reliable models of human hNav1.4 channel at different conformational states using these known structures as templates. We observed that the model originated from NavPaS reflects a closed state of the channel, whereas the one produced from EeNav1.4 is open, and is very similar to the recently published structure of hNav1.4 [9]. We have extensively compared these models and concluded that they provide important insights to the closed-open voltage-dependent transition.

## 2. Material and methods

### 2.1. Molecular modeling

The primary sequence of the human sodium channel 1.4 was recovered from the NCBI (<https://www.ncbi.nlm.nih.gov/protein/292495096>) under the code SCN4A\_HUMAN, and UniProt code P35499 (<http://www.uniprot.org/uniprot/P35499>). The cryo-EM structures of NavPaS [7] and EeNav1.4 [8] served as structural templates for the modeling of hNav1.4 channel. The pdb file of hNav1.4 structure was kindly provided by Dr. Yang.

By comparing the amino acid sequence of hNav1.4 to these structures using SWISS-MODEL [10], an alignment was generated and models of the human channel were created using EeNav1.4 and NavPaS as templates. The transmembrane regions of these models coincided with the previously reported sequence alignment, and matched the structure of hNav1.4.

As the length and sequence of the internal loops are less conserved, the absence of more detailed structural information and lack of homology with known structures prevented us from obtaining reliable models of these loops, and therefore they are omitted in most images shown in the present paper. The models passed by structural comparisons with TM-align [11] and were validated with Ramachandran plot [12].

In order to identify which residues correspond to the region of the hydrophobic constriction [4], we compared the sequences of the VSDs of the hNav1.4 with the VSD of the Shaker Kv channels (RSCB cod. 3LUT). Directed mutation studies showed that the residues Isoleucine (I241), Phenylalanine (F290) and Isoleucine (I320) in the Shaker Kv channel have a great importance in producing this hydrophobic barrier. We could identify the corresponding hydrophobic amino acids on our alignment, considering their positions in the segments S1, S2 and S3, in agreement with the results reported by Lacroix et al. [4,13] (see Supplementary material).

### 2.2. Toxin docking

The model generated for hNav1.4/PaS was used in conjunction with the Ts1 toxin (access code 1NPI) [14] for protein-protein docking using HADDOCK web server v2.2 [15], a robust global and local docking program based on molecular mechanics and rigid body energy minimization. Residues K12, W39, and W54 of the toxin were considered important for binding, as previously reported in the literature [16]. The segment of the human Nav1.4 channel expected to interact with Ts1 is between A596-L668 [17].

HADDOCK generated 153 structures divided into 8 clusters, which represents 76.5% of the water-refined models. The PRODIGY [18,19] webserver was used to predict the binding affinity in protein-protein complexes. PISA server was used to identify the likely interactions involved in the docking [20].

## 3. Results

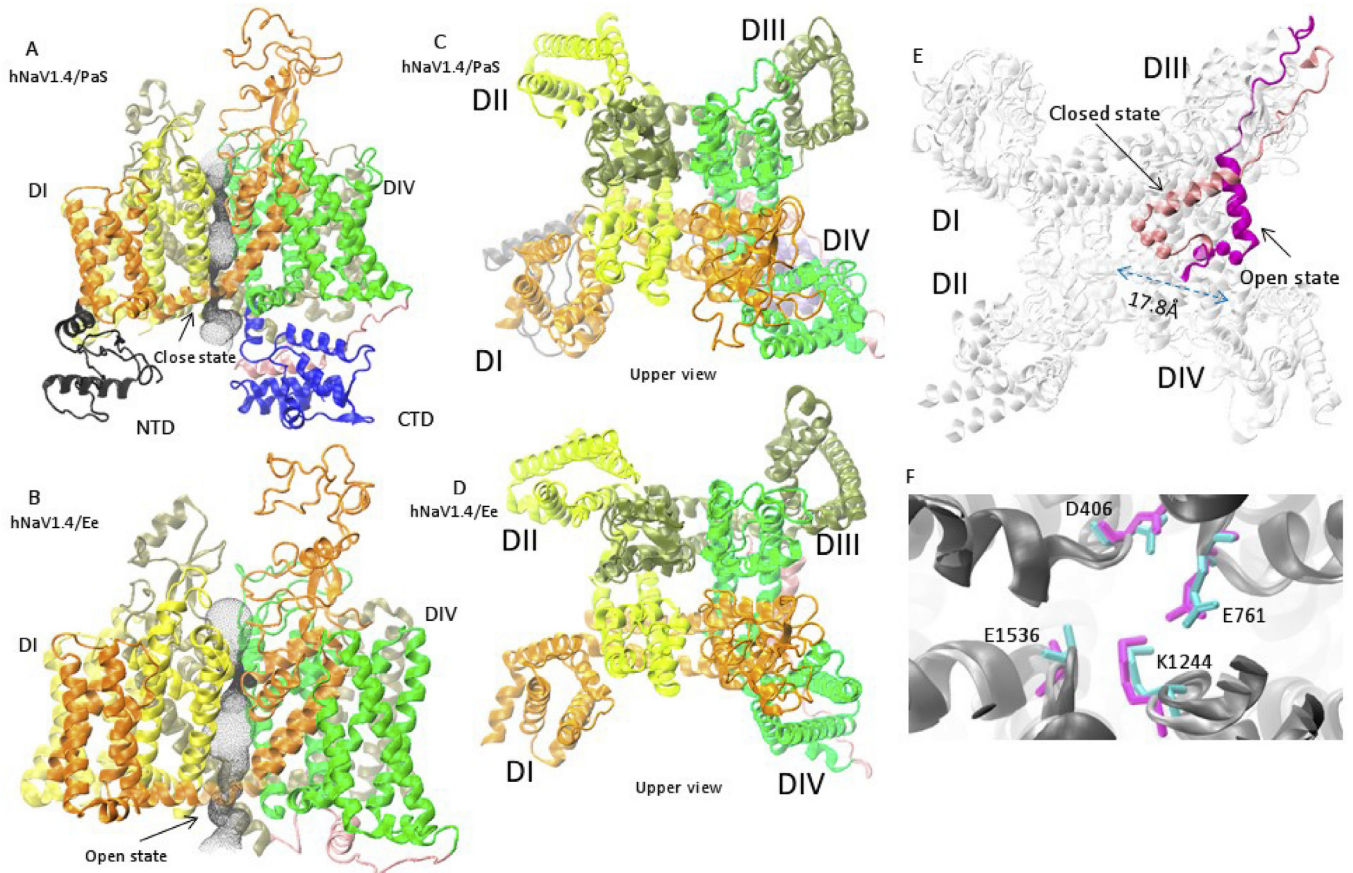
Two models of hNav1.4 were produced in the present work: one model – hNav1.4/PaS (see Supplementary material)– was generated using as template the reported structure NavPaS (PDB code: 5X0M), a closed-state eukaryotic sodium channel, and the other – hNav1.4/Ee (see Supplementary material)– originated from the reported structure of EeNav1.4 (PDB code: 5XSY), an open-state eukaryotic sodium channel. The amino acid sequence of NavPaS has an identity of 42.12% with hNav1.4, with a coverage of 0.84, whereas the sequence of EeNav1.4 has an identity of 66.17%, with a coverage of 0.92. The models are depicted in Fig. 1. The pdb files are deposited in the Protein Model Database under the ID PM0081855 (hNav1.4/PaS) and PM0081856 (hNav1.4/Ee).

The quality of the hNav1.4/PaS structure was analyzed using the Ramachandran plot: 83.91% of the amino acid residues were in favorable regions and only 6.48% in disallowed region. As comparison, we confronted this result with the original structure of NavPaS, analyzed by the same procedure. Similar results were obtained: 81.85% of its amino acid residues were in favorable regions, 5.24% in disallowed region. In both cases, most of them are located in loop regions or at extremities of transmembrane segments (see Supplementary material). The same procedure was performed with the hNav1.4/Ee structure. Ramachandran plot showed 92.25% of the amino acids with favorable angles and 1.07% considered unfavorable angles. EeNav1.4 showed 88.52% of the residues at favorable angles and 3.79% at unfavorable angles. This model was further validated by the structure of hNav1.4 obtained by cryo-EM [9], as examined in the Discussion. The sodium channels used for modeling (NavPaS, EeNav1.4) had their extracellular loops better defined than the intracellular loops, which favored modeling the external face of the channel.

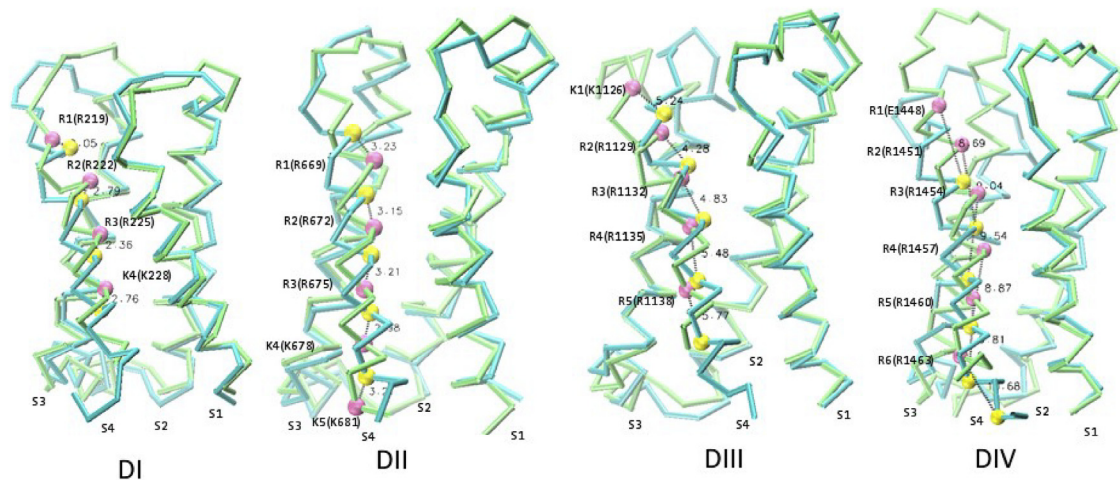
Using the TM-align tool, we overlapped our model hNav1.4/PaS with NavPaS and the similarity score (without the internal loops) was 0.98631. Likewise, hNav1.4/Ee was compared with EeNav1.4, receiving a score of 0.99541, and the comparison of hNav1.4/Ee with hNav1.4 received a score of 0.97846. This result shows that the model generated by molecular modeling based on EeNav1.4 is very similar to the experimental model, thereby supporting this modeling approach. In fact, only small differences were found between the hNav1.4/Ee and hNav1.4 models. The position of the transmembrane segments provides a near perfect match of the C $\alpha$  positions, including the charged S4 residues and the selectivity filter (DEKA), all within the resolution of the Cryo-EM structures. The inactivation loops are in the same position in both models.

Fig. 1 shows the well-defined domains. The ion-conducting pore calculated by HOLE [21] is shown with black dots. It is possible to identify two well-defined states: a closed-state (hNav1.4/PaS), and an open-state (hNav1.4/Ee). By superimposing the models, we may identify regions that are important for the closed-open transition. The DEKA region remains essentially unchanged (Fig. 1F). Fig. 2 shows outward shifts of positions of the arginine residues in the S4 of domain IV of Nav1.4/Ee, and these shifts ranged from 8.69 Å in the case of R1(R1448) to 10.68 Å of R6(R1463). In DIII, the distance variation between charged residues comparing the two models was 4.28 Å in R2(R1129) and 5.77 Å in K5(K1138). It is also possible to observe that the S6 segments of all domains are in different positions in the open state and in the closed state of the channel, leading to a constriction of the pore as seen in hNav1.4/PaS (Fig. 1A). Fig. 1E shows that the inactivation loop undergoes a shift of 17.8 Å between the closed-state and the open-state models.

To better determine this shift of the GC, we identified the hydrophobic constriction through which the positive charges of the S4 segment pass upon depolarization. This was achieved as described in Methods, and the amino acids that form these



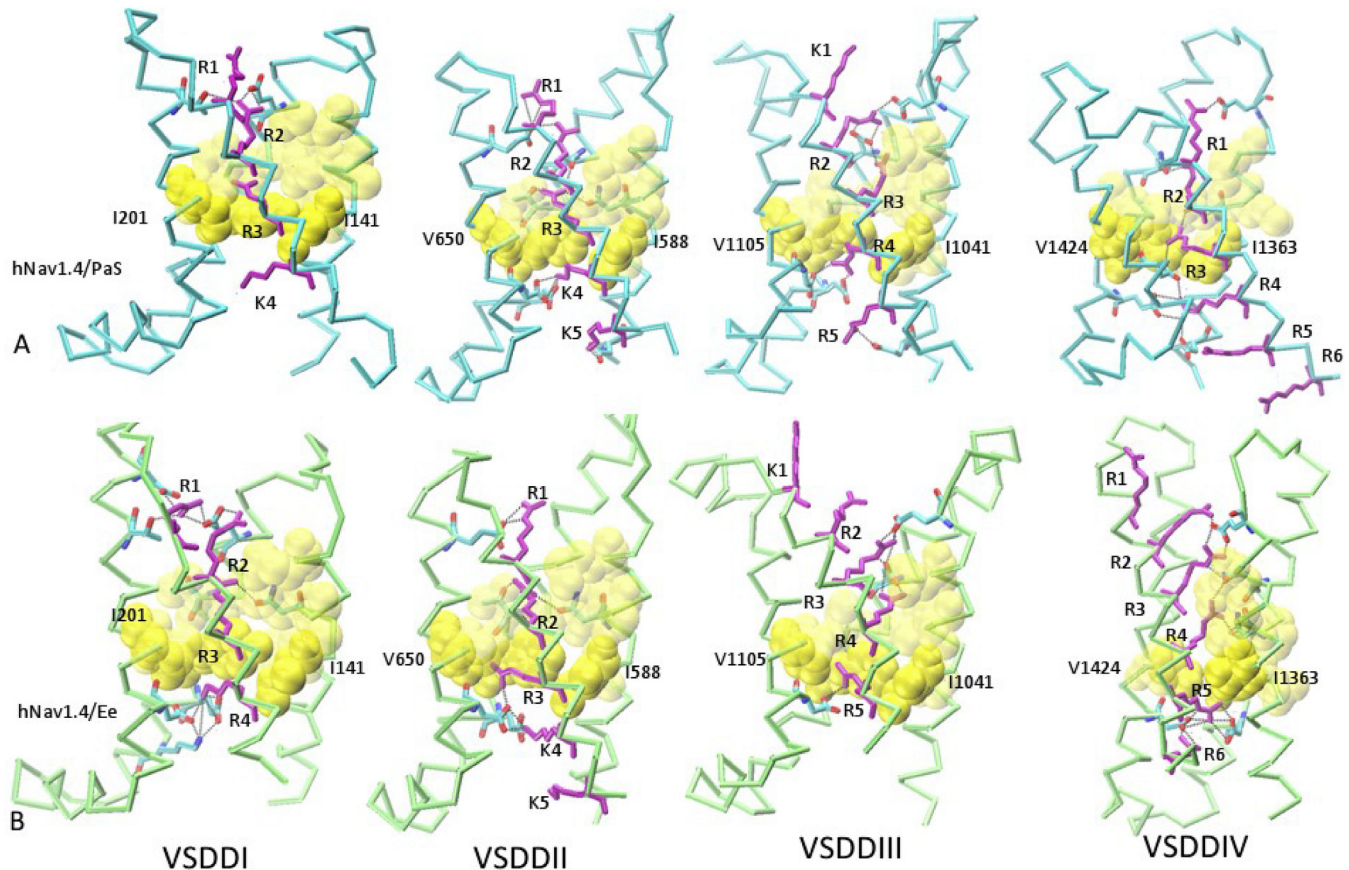
**Fig. 1.** A – hNav1.4/PaS side view, the conduction pore is shown with black dots; B – hNav1.4/Ee side view, the conduction pore is shown with black dots; C – hNav1.4/PaS upper view; D – hNav1.4/Ee upper view. In A–C: Colored structures as follows: DI orange, DII yellow, DIII mauve, DIV green, IFM pink, NTD black and CTD blue. E – hNav1.4/PaS and hNav1.4/Ee overlapped, showing the inactivation loop shifted by 17.8 Å: pink closed state, purple open state. F – hNav1.4/PaS and hNav1.4/Ee overlapped, showing the selectivity filter DEKA. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 2.** Overlap of VSDs hNav1.4/Ee lime, hNav1.4/PaS cyan. The C $\alpha$  of the charged residues of S4 of each of domain are highlighted in different color, being yellow for hNav1.4-NavPaS and mauve for hNav1.4-EeNav1.4. The numbers mark the distance of the corresponding C $\alpha$  of these residues. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

structures are shown in Fig. 3. Forming the hydrophobic constriction, we could identify I141, I588, I1041 and I1363 in segments S1 of each domain. In S2-DI we identified Y168, whereas F621, F1076 and F1396 are present in DII, DIII and DIV, respectively. In S3-DI we identified I201, and in DII, DIII and DIV the hydrophobic

constriction is composed by V650, V1105 and V1424, respectively. The positions of these amino acids are shown in Fig. 3. This figure also shows the positions of the charged residues in the closed and open models. One can observe the relative positions of the positive charges in respect to the hydrophobic constriction formed by these



**Fig. 3.** A - hNav1.4/PaS VSD for DI-DIV in cyan, Arginine and Lysine in purple, and hydrophobic constriction residues in yellow. B - hNav1.4/Ee VSD for DI-DIV in lime, Arginine and Lysine in purple, and hydrophobic constriction residue in mauve. In S1, S2 and S3 the residues that help stabilize the positive charges are shown in licorice. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

amino acids, using as reference the aromatic amino acids Phenylalanine and Tyrosine. It is visible in the closed model (hNav1.4/PaS) the outer position of residues R1(R219), R2(R222) and R3(R225) of DI, R1(R669), R2(R672) and R3(R675) of DII, K1(K1126), R2(R1129) and R3(R1132) of DIII, and R1(R1148), R2(R1451) and R3(R1454) of DIV. In contrast, residues at inner positions relative to the constriction are K4(K228) of DI, K4(K678) and K5(K681) of DII, R4(R1135) and R5(R1138) of DIII, and R4(R1457), R5(R1460), and R6(R1463) of DIV. On the other hand, one can notice in the open model (hNav1.4/Ee) and in hNav1.4 a different situation: the residues K1(K1126), R2(R1129), R3(R1132), R4(R1135) of DIII, and R1(R1448), R2(R1451), R3(R1454), R4(R1457) of DIV are at outer positions. The positions of the residues in DI and DII are the same in both cases: R1(219), R2(R222), and R3(R225) of DI and R1(R669), R2(R672) and R3(R675) of DII appear external to the hydrophobic constriction.

The analysis of the number of charges at outer positions on domains DI and DII of hNav1.4/PaS suggests that their voltage-sensors may be at activated state. In order to verify this possibility we decided to try the docking of the scorpion toxin Ts1, which is known to bind to the activated VS-DII and thereby holding it at the activated position [17].

### 3.1. Docking of Ts1

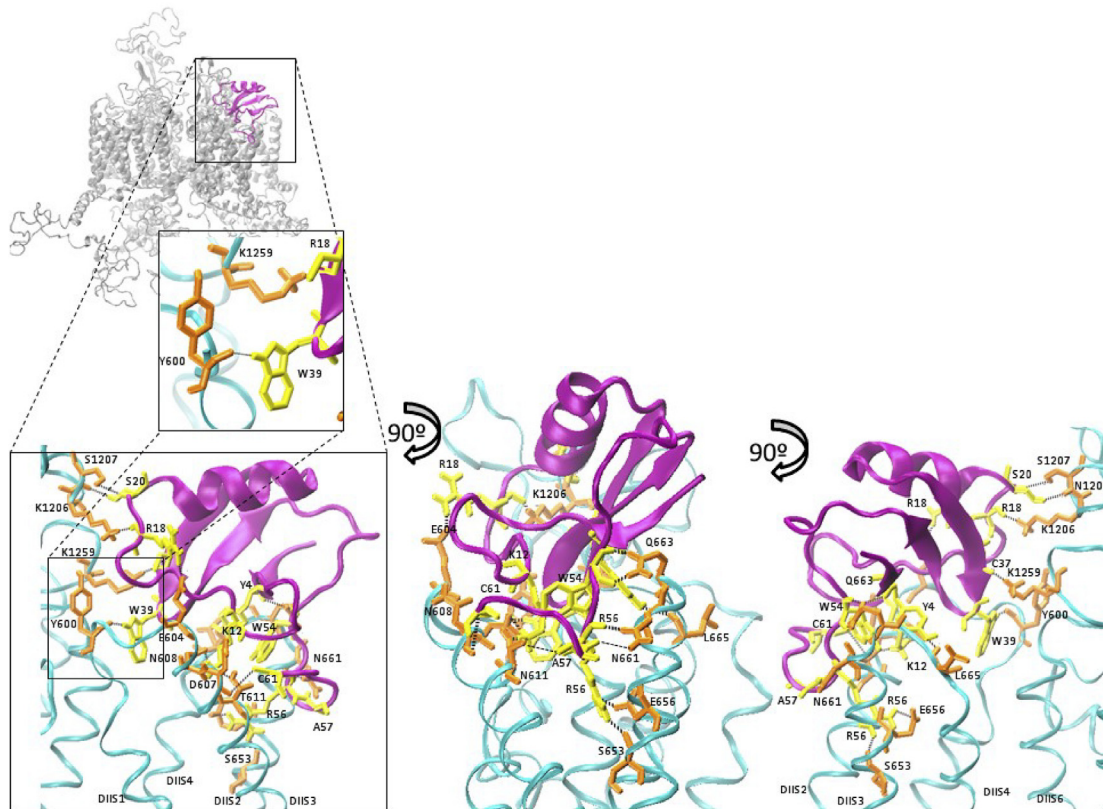
We successfully obtained a high affinity docking of Ts1 to DII of hNav1.4/PaS. The best result among the 153 of the cluster generated by HADDOCK was analyzed using the PRODIGY program [18,19],

and the calculated binding energy ( $\Delta G$ ) was  $-9.4$  kcal/mol, corresponding to  $2.4 \cdot 10^{-7}$  M for the dissociation constant. This result is compatible with the experimental dissociation constant of  $5.0 \cdot 10^{-8}$  M measured with a chimera formed with hNav1.4 external loop [22], especially if one considers that the docking was carried out in the absence of  $H_2O$ , thus reducing the effect of hydrophobic interactions.

It has been experimentally demonstrated that Ts1 residues K12, W39 and W54 are fundamental for the bioactivity of the toxin [23]. Fig. 4 shows that these amino acids interact with the VS-DII. Using PyMol (Molecular Graphics System, Version 2.0 Schrödinger, LLC) and PISA server we could identify the interaction of K12 with D607 and T611 of S2-DII, W39 with Y600 of the S1-S2 loop and W54 with Q663 of S3-DII. Additional interacting residues were identified (toxin-channel): Y4-Q663, R56-N661/S653, A57-N661, R56-E656, R18-E604/K1206, S20-N1205/S1207, C37-K1259 and C61-N608. Note that R18, S20 and C37 interact with residues of the loop between S5-DIII and S6-DIII.

## 4. Discussion

Despite the great importance of sodium channels for human physiology and pharmacology, their complexity has hindered the advance of the knowledge on their molecular structures. By using cryo-EM, important advances were achieved with the resolution at atomic level of sodium channel structure of cockroach (NavPaS), eel (EeNav1.4), and more recently the human Nav1.4. We took advantage of these structures to generate models of human Nav1.4



**Fig. 4.** A – Zoom of the docking of the scorpion toxin Ts1 (purple) to hNav1.4/PaS (cyan). Residues important for the binding are highlighted: In yellow residues from toxin Y4, K12, R18, S20, C37, W54, R56 and A57. In orange residues from hNav1.4/PaS E604, N608, T611, S653, N661, Q663, K1206, S1207, K1259. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

channel at open and closed states. These models can be considered trustworthy by bioinformatics criteria, and the near coincidence of hNav1.4/Ee with the recently published structure of hNav1.4 validated our open-state model and our modeling procedure. The results generated by the TM-align program shows a very good match of the models with their templates. The narrowing of the pore of the channel, together with the number of positive charges external

and internal to the hydrophobic constriction, leads us to conclude that the two models represent different functional states: closed, Nav1.4/PaS derived from the closed channel PaSNav, and open (activated), derived from the open channel Nav1.4Ee. By comparing both models, one can identify putative movements associated with the transition of closed to open state on human Nav1.4 channel. The outward shift of charges in the S4 segments of domains III and IV

**Table 1**  
Residues found to interact with the GC. In bold are the residues found in both models.

Domain	Charged amino acid	Nav1.4/PaS Interacting amino acids	Nav1.4/Ee Interacting amino acids
DI	R219(R1) R222(R2) R225(R3) K228(K4)	<b>E161/T207</b>	<b>E161/T207/D211</b> <b>E161</b> N144/T165 E171/K175/D197
DII	R669(R1) R672(R2) R675(R3) K678(K4) K681(K5)	<b>E656</b> <b>E656</b> <b>N591/N614/T618/E624</b> <b>E624/D646</b> D581	<b>E656</b> <b>N591/T618</b> <b>D646</b> <b>E624/D646</b>
DIII	R1129(R2) R1132(R3) R1135(R4) R1138(R5)	<b>E1051/D1069</b> <b>D1069</b> E1079/ <b>D1101</b> E1034	<b>E1051/D1069</b> <b>D1069</b> <b>D1101</b>
DIV	R1448 (R1) R1451 (R2) R1454 (R3) R1457 (R4) R1460 (R5) R1463 (R6)	E1374 <b>D1420/N1389</b> <b>D1420/N1366</b> D1356/ <b>E1399/D1420</b>	E1373 E1373/ <b>N1389</b> <b>N1366/N1389</b> <b>E1399/D1420</b> <b>D1420</b>

are clearly identified. The comparison of the closed- and open-state models suggests that during the transition between the two states new interactions are formed between the charged residues of S4 and amino acids of other segments of the VSD, which happen to transiently interact with the GC. This is shown in Fig. 3 and Table 1. One can see that for DI and DII there is not a significant change in the amino acids that interact with specific charged residues of S4. In contrast, one can notice that for DIII and DIV the charged amino acids interact with different groups in each model. For example, while in the closed-state model R1451(R2) and R1454(R3) interact with D1420, in the open-state model this amino acid interacts with R1460(R5) and R1463(R6), which reflects an outward movement of S4-DIV in the transition between these states. In all domains, it is possible to verify consistent interactions between the charged residue with other amino acids of the VSD.

The observed outward shift of the S4-DIII and S4-DVI ranged from ~4 Å to ~10 Å, along with a small rotation, which is consistent with data reported by other authors [24,25], and the positions of the alpha carbons suggest movements such as the 'sliding-helix' [26].

With the structural information provided by models in different states, it is possible to verify a charge translocation crossing the important hydrophobic barrier in the VSD, formed by residues in the segments S1, S2, and S3, as previously reported [4,13]. The aromatic amino acids Phenylalanine in DII-DIV and Tyrosine in DI play a crucial role in stabilizing positively charged GC in a hydrophobic environment (Fig. 3). Relative to these residues, in hNav1.4/PaS the number of positively charged amino acids at outer positions is 3 in all domains, and in hNav1.4/Ee the number is 3 for DI and DII, and 4 for DIII and DIV. Therefore, despite being closed, the number of positively charged amino acids at outer positions in VS-DI and VS-DII of hNav1.4/PaS is the same as in Nav1.4/Ee (and hNav1.4), which suggests that they are in an "activated" position. This is not surprising if one considers that cryo-EM uses channels in depolarized condition. This interpretation is further supported by the docking of Ts1 to hNav1.4/Pas, because it is known that this toxin binds and holds the VS-DII in an "activated" position.

We conclude that incomplete outward movement of VS-DIII holds the channel at the closed state. Interestingly, we could identify interactions of the GC with S1, S2 and S3 at this intermediate state, thus suggesting that, although very rapid, the outward movement is stepwise.

The substitution of specific arginine residues of S4 for histidine generates H<sup>+</sup> current at resting potentials. This evidences that R215, R669 and R1132 are at the level of the hydrophobic constrictions at rest. Considering that in the open state we can observe that 3 charges of DI and DII and 4 of DIII are "out", we conclude that about 6.5 charges are associated with hNav1.4 activation. DIV was excluded from this calculation because experimental evidence indicate that its activation is much slower, and its movement is not necessary for activation of the channel.

The residues previously reported as important for the Ts1 activity (K12, W39, and W54), along with residues Y4, R18, P20, C37, R56, A57, and C61, are shown to interact with the channel (Fig. 4). New information was obtained regarding the participation of DIII loops (K1206, S1207 and K1259) in the docking [27].

Mutation studies have identified residues important for the binding and voltage-dependent trapping of CssiV toxin on rNav1.2 [17]. The alignment between rNav1.2 and hNav1.4 shows that the most important residues for the binding are present in hNav1.4: E598, Y600, P601, E656, L659, A660, G664, L665, and S666. Although both toxins bind to site 4 of sodium channels and are classified as  $\beta$ -toxin, they are not identical and different residues are expected to be involved in Ts1 binding. One important difference is that Ts1 does not need a strong depolarizing pulse to exert

its effect [28]. We could identify the involvement of Y600, E656 and L665 with the docking of Ts1. Interestingly, the docking of Ts1 predicts an important interaction of N661 with its residues R56 and A57. The equivalent residue N842 in rNav1.2, when replaced by arginine, allows CssiV (which have the sequence PLP at the equivalent position) to act without the prepulse. Interestingly, the mutation S788D increases the affinity of rNav1.2 to CssiV, and D607 is the equivalent residue in hNav1.4 and is important for the docking of Ts1.

The present models may be used to test functional alterations produced by single mutations and possible specific drug activities. The similarity of the model hNav1.4/Ee with the recently reported cryo-EM structure of hNav1.4 validates the procedure used in the present paper and opens wide perspectives for modeling other human sodium channels. By producing a solution to the structure at atomic resolution, the present procedures open new possibilities to the development of theoretical research with ion channels and to inspire the generation of new structural models.

### Conflicts of interest

The authors declare no conflict of interest.

### Acknowledgements

This work was supported by Fundação CAPES - Programa de Toxinologia and Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, Brazil. JRCF was recipient of a fellowship from CAPES.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbrc.2018.10.110>.

### References

- [1] W.A. Catterall, G. Wisedchaisri, N. Zheng, The chemical basis for electrical signaling, *Nat. Chem. Biol.* 13 (2017) 455–463, <https://doi.org/10.1038/nchembio.2353>.
- [2] W.A. Catterall, S. Cestèle, V. Yarov-Yarovoy, F.H. Yu, K. Konoki, T. Scheuer, Voltage-gated ion channels and gating modifier toxins, *Toxicol.* 49 (2007) 124–141, <https://doi.org/10.1016/j.toxicol.2006.09.022>.
- [3] W.A. Catterall, Voltage-gated sodium channels at 60: structure, function and pathophysiology, *J. Physiol. (London)* 590 (2012) 2577–2589, <https://doi.org/10.1113/jphysiol.2011.224204>.
- [4] J.J. Lacroix, H.C. Hyde, F.V. Campos, F. Bezanilla, Moving gating charges through the gating pore in a Kv channel voltage sensor, *Proc. Natl. Acad. Sci. U.S.A.* 111 (2014) E1950–E1959, <https://doi.org/10.1073/pnas.1406161111>.
- [5] X. Chen, Q. Wang, F. Ni, J. Ma, Structure of the full-length Shaker potassium channel Kv1.2 by normal-mode-based X-ray crystallographic refinement, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 11352–11357, <https://doi.org/10.1073/pnas.1000142107>.
- [6] S. Ahuja, S. Mukund, L. Deng, K. Khakh, E. Chang, H. Ho, et al., Structural basis of Nav1.7 inhibition by an isoform-selective small-molecule antagonist, *Science* 350 (2015), <https://doi.org/10.1126/science.aac5464> aac5464.
- [7] H. Shen, Q. Zhou, X. Pan, Z. Li, J. Wu, N. Yan, Structure of a eukaryotic voltage-gated sodium channel at near-atomic resolution, *Science* 355 (2017), <https://doi.org/10.1126/science.aal4326>.
- [8] Z. Yan, Q. Zhou, L. Wang, J. Wu, Y. Zhao, G. Huang, et al., Structure of the Nav1.4- $\beta$ 1 complex from electric eel, *Cell* 170 (2017) 470–482, <https://doi.org/10.1016/j.cell.2017.06.039>, e11.
- [9] X. Pan, Z. Li, Q. Zhou, H. Shen, K. Wu, X. Huang, et al., Structure of the human voltage-gated sodium channel Nav1.4 in complex with  $\beta$ 1, *Science* (2018), <https://doi.org/10.1126/science.aau2486>.
- [10] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, et al., SWISS-MODEL: homology modelling of protein structures and complexes, *Nucleic Acids Res.* 46 (2018) W296–W303, <https://doi.org/10.1093/nar/gky427>.
- [11] Y. Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, *Nucleic Acids Res.* 33 (2005) 2302–2309, <https://doi.org/10.1093/nar/gki524>.
- [12] I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, et al.,

- MolProbity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res.* 35 (2007) W375–W383, <https://doi.org/10.1093/nar/gkm216>.
- [13] J.J. Lacroix, F.V. Campos, L. Frezza, F. Bezanilla, Molecular bases for the asynchronous activation of sodium and potassium channels required for nerve impulse generation, *Neuron* 79 (2013) 651–657, <https://doi.org/10.1016/j.neuron.2013.05.036>.
- [14] C.B. Pinheiro, S. Marangoni, M.H. Toyama, I. Polikarpov, Structural analysis of Tityus serrulatus Ts1 neurotoxin at atomic resolution: insights into interactions with Na<sup>+</sup> channels, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 59 (2003) 405–415.
- [15] G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastiris, E. Karaca, et al., The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes, *J. Mol. Biol.* 428 (2016) 720–725, <https://doi.org/10.1016/j.jmb.2015.09.014>.
- [16] M.-F. Martin-Eauclaire, P.E. Bougis, M.E. de Lima, Ts1 from the Brazilian scorpion Tityus serrulatus: a half-century of studies on a multifunctional beta like-toxin, *Toxicon* 152 (2018) 106–120, <https://doi.org/10.1016/j.toxicon.2018.07.024>.
- [17] J.Z. Zhang, V. Yarov-Yarovoy, T. Scheuer, I. Karbat, L. Cohen, D. Gordon, et al., Structure-function map of the receptor site for  $\beta$ -scorpion toxins in domain II of voltage-gated sodium channels, *J. Biol. Chem.* 286 (2011) 33641–33651, <https://doi.org/10.1074/jbc.M111.282509>.
- [18] L.C. Xue, J.P. Rodrigues, P.L. Kastiris, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein-protein complexes, *Bioinformatics* 32 (2016) 3676–3678, <https://doi.org/10.1093/bioinformatics/btw514>.
- [19] A. Vangone, A.M. Bonvin, Contacts-based prediction of binding affinity in protein-protein complexes, *Elife* 4 (2015), e07454, <https://doi.org/10.7554/eLife.07454>.
- [20] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.* 372 (2007) 774–797, <https://doi.org/10.1016/j.jmb.2007.05.022>.
- [21] O.S. Smart, J.G. Neduelil, X. Wang, B.A. Wallace, M.S. Sansom, HOLE: a program for the analysis of the pore dimensions of ion channel structural models, *J. Mol. Graph.* 14 (1996) 354–360, [https://doi.org/10.1016/S0263-7855\(97\)00009-X](https://doi.org/10.1016/S0263-7855(97)00009-X), 376.
- [22] F. Bosmans, M.-F. Martin-Eauclaire, K.J. Swartz, Deconstructing voltage sensor function and pharmacology in sodium channels, *Nature* 456 (2008) 202–208, <https://doi.org/10.1038/nature07473>.
- [23] O. Hassani, P. Mansuelle, S. Cestèle, M. Bourdeaux, H. Rochat, F. Sampieri, Role of lysine and tryptophan residues in the biological activity of toxin VII (Ts gamma) from the scorpion Tityus serrulatus, *Eur. J. Biochem.* 260 (1999) 76–86.
- [24] F. Bezanilla, Gating currents, *J. Gen. Physiol.* 150 (2018) 911–932, <https://doi.org/10.1085/jgp.201812090>.
- [25] C.S. Souza, C. Amaral, W. Treptow, Electric fingerprint of voltage sensor domains, *Proc. Natl. Acad. Sci. U.S.A.* 111 (2014) 17510–17515, <https://doi.org/10.1073/pnas.1413971111>.
- [26] H.R. Guy, P. Seetharamulu, Molecular model of the action potential sodium channel, *Proc. Natl. Acad. Sci. U.S.A.* 83 (1986) 508–512, <https://doi.org/10.1073/pnas.83.2.508>.
- [27] J.Z. Zhang, V. Yarov-Yarovoy, T. Scheuer, I. Karbat, L. Cohen, D. Gordon, et al., Mapping the interaction site for a  $\beta$ -scorpion toxin in the pore module of domain III of voltage-gated Na<sup>(+)</sup> channels, *J. Biol. Chem.* 287 (2012) 30719–30728, <https://doi.org/10.1074/jbc.M112.370742>.
- [28] F.V. Campos, B. Chanda, P.S.L. Beirão, F. Bezanilla, beta-Scorpion toxin modifies gating transitions in all four voltage sensors of the sodium channel, *J. Gen. Physiol.* 130 (2007) 257–268, <https://doi.org/10.1085/jgp.200609719>.