UC Davis
UC Davis Previously Published Works

Title

Structural modeling of ion channels using AlphaFold2, RoseTTAFold2, and ESMFold

Permalink

https://escholarship.org/uc/item/4957x0w7

Journal

Channels, 18(1)

ISSN

1933-6950

Authors

Nguyen, Phuong Tran

Harris, Brandon John

Mateos, Diego Lopez

et al.

Publication Date

2024-12-31

DOI

10.1080/19336950.2024.2325032

Peer reviewed

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates

Structural modeling of ion channels using AlphaFold2, RoseTTAFold2, and ESMFold

Phuong Tran Nguyen[a], Brandon John Harris[a,b], Diego Lopez Mateos[a,b], Adriana Hernández González[a,b], Adam Michael Murray[c], and Vladimir Yarov-Yarovoy [a,d]

[a]Department of Physiology and Membrane Biology, University of California School of Medicine, Davis, CA, USA; [b]Biophysics Graduate Group, University of California School of Medicine, Davis, CA, USA; [c]Monterey Peninsula College, Monterey, CA, USA; [d]Department of Anesthesiology and Pain Medicine, University of California School of Medicine, Davis, CA, USA

ABSTRACT
Ion channels play key roles in human physiology and are important targets in drug discovery. The atomic-scale structures of ion channels provide invaluable insights into a fundamental understanding of the molecular mechanisms of channel gating and modulation. Recent breakthroughs in deep learning-based computational methods, such as AlphaFold, RoseTTAFold, and ESMFold have transformed research in protein structure prediction and design. We review the application of AlphaFold, RoseTTAFold, and ESMFold to structural modeling of ion channels using representative voltage-gated ion channels, including human voltage-gated sodium (Na$_V$) channel - Na$_V$1.8, human voltage-gated calcium (Ca$_V$) channel – Ca$_V$1.1, and human voltage-gated potassium (K$_V$) channel – K$_V$1.3. We compared AlphaFold, RoseTTAFold, and ESMFold structural models of Na$_V$1.8, Ca$_V$1.1, and K$_V$1.3 with corresponding cryo-EM structures to assess details of their similarities and differences. Our findings shed light on the strengths and limitations of the current state-of-the-art deep learning-based computational methods for modeling ion channel structures, offering valuable insights to guide their future applications for ion channel research.

Introduction

Ion channels play key roles in human physiology and have been established as important targets in drug discovery [1,2]. The atomic-scale structures of ion channels provide invaluable insights into a fundamental understanding of the molecular mechanisms of channel gating and modulation. The recent advancements in cryo-electron microscopy (cryo-EM) produced a remarkable increase in the number of high-resolution structures of ion channels [3–7]. Multiple ion channel structures have been resolved in various putative physiological states and in complex with auxiliary subunits, small molecules, and natural peptides, providing crucial insights into the molecular mechanisms underlying their modulation.

In parallel to advancements in cryo-EM, breakthroughs in deep learning-based computational methods, such as AlphaFold [8] from Google's DeepMind and RosetTTAFold [9,10] from David Baker's Institute for Protein Design at the University of Washington, have been transforming research in protein structure prediction. These methods utilize deep neural networks trained on co-evolution information from multiple sequence alignments derived from protein sequence database (UniProt) [11] and protein structural data derived from Protein Data Bank (PDB) [12] to predict protein structures. AlphaFold and RoseTTAFold based methods have been applied to protein design and modeling of protein complexes [13–19]. Additionally, large language models of protein sequences, such as Meta AI's ESMFold [20], trained on millions of protein sequences and using billions of parameters, provide rapid protein structure predictions, although with slightly lower accuracy compared to AlphaFold and RosettaFold. The ability to predict protein structures with high accuracy holds tremendous promise in transforming the field of drug

---

This article has been republished with minor changes. These changes do not impact the academic content of the article.

discovery. Notably, the AlphaFold Structural Database currently contains over 200 million protein models predicted by AlphaFold [21] and the ESM Metagenomic Atlas contains more than 700 million protein models predicted by ESMFold [20].

Despite the remarkable achievements in protein structure prediction using deep-learning-based methods, the performance of these methods on challenging targets like ion channels remains to be determined. This is particularly significant when considering the structural heterogeneity of ion channel physiologically relevant states. To examine the structural modeling of ion channels using the deep-learning-based methods, we applied AlphaFold2 [8], RosetTTAFold2 [9], and ESMFold [20] to predict structures of representative voltage-gated ion channels, including human voltage-gated sodium (Na_V) channel – $Na_V1.8$, human voltage-gated calcium (Ca_V) channel – $Ca_V1.1$, and human voltage-gated potassium (K_V) channel – $K_V1.3$. We compared AlphaFold2, RoseTTAFold2, and ESMFold structural models of $Na_V1.8$, $Ca_V1.1$, and $K_V1.3$ with corresponding cryo-EM structures of $Na_V1.8$ [22], $Ca_V1.1$ [23], and $K_V1.3$ [24] to assess details of their similarities and differences. Our findings shed light on the strengths and limitations of the current state-of-the-art deep learning-based computational methods for modeling ion channel structures, offering valuable insights to guide their future applications for ion channel research.

Structural modeling of voltage-gated sodium (Nav) channels

Voltage-gated sodium (Na_V) channels are responsible for initiating and propagating action potentials, the electrical signals facilitating communication between excitable cells [1,25–27]. There are nine Na_V channel subtypes, from $Na_V1.1$ to $Na_V1.9$. The $Na_V1.1$, $Na_V1.2$, and $Na_V1.6$ subtypes are predominantly expressed in the central nervous system [28]. The $Na_V1.4$ and $Na_V1.5$ subtypes are mainly expressed in skeletal and cardiac muscles, respectively [28]. The peripheral nervous system primarily expresses $Na_V1.7$, $Na_V1.8$, and $Na_V1.9$ subtypes [28]. Dysfunctions in these channels can lead to serious health issues, including epilepsy, cardiac arrhythmias, muscle weakness, and chronic pain.

The advancement of cryo-EM has facilitated the resolution of mammalian Na_V subtypes, ranging from $Na_V1.1$ to $Na_V1.8$, significantly enhancing our understanding of their structure, gating, and modulation [22,29–35].

The voltage-dependent gating, sodium conduction, and modulation by natural peptides and small molecule drugs are performed by the Na_V channel α subunit [1,25–27]. Auxiliary Na_V channel β subunits (β1-β4) are co-expressed with the α subunit and modulate the channel function [36]. We selected the $Na_V1.8$ channel α subunit, as an example Na_V channel, for structure prediction using AlphaFold2 [8] and ColabFold as the computational platform [37]. ColabFold's AlphaFold2 pipeline employs MMseqs2 multiple sequence alignment method [38,39], which is a more efficient alternative to Jackhmmer multiple sequence alignment method [40] used in DeepMind's original AlphaFold2 pipeline [8]. The MMseqs2 method has considerably accelerated the AlphaFold2 protein structure prediction pipeline performance while maintaining comparable accuracy [37]. The protein sequence of the SCN10A gene, which encodes the human $Na_V1.8$ (hNa_V 1.8) α subunit (UniProt ID: Q9Y5Y9), was used as input into ColabFold's AlphaFold_ mmseqs2 notebook for structure prediction. We assessed the quality of predicted AlphaFold2 models of $hNa_V1.8$ using predicted local distance difference test (pLDDT) confidence score. Generally, pLDDT values above 90 represent very high confidence, pLDDT values between 70 and 90 represent good confidence, pLDDT values between 50 and 70 represent low confidence, and pLDDT values below 50 represent very low confidence [8]. We also compared similarities and differences to the resolved $hNa_V1.8$ structures (PDB: 7WE4, 7WEL, 7WFR, and 7WFW) [22] using alpha carbon root mean square deviation (Cα RMSD). We assessed individual transmembrane voltage-sensing domains (VSD-I, VSD-II, VSD-III, and VSD-IV), the pore domain, the extracellular loop (ECL) regions, and the overall model topology.

Our results showed that AlphaFold2 could predict the majority of the $hNa_V1.8$ domains with very high confidence scores (pLDDT >90), ESMFold could predict VSDs, the pore domain,
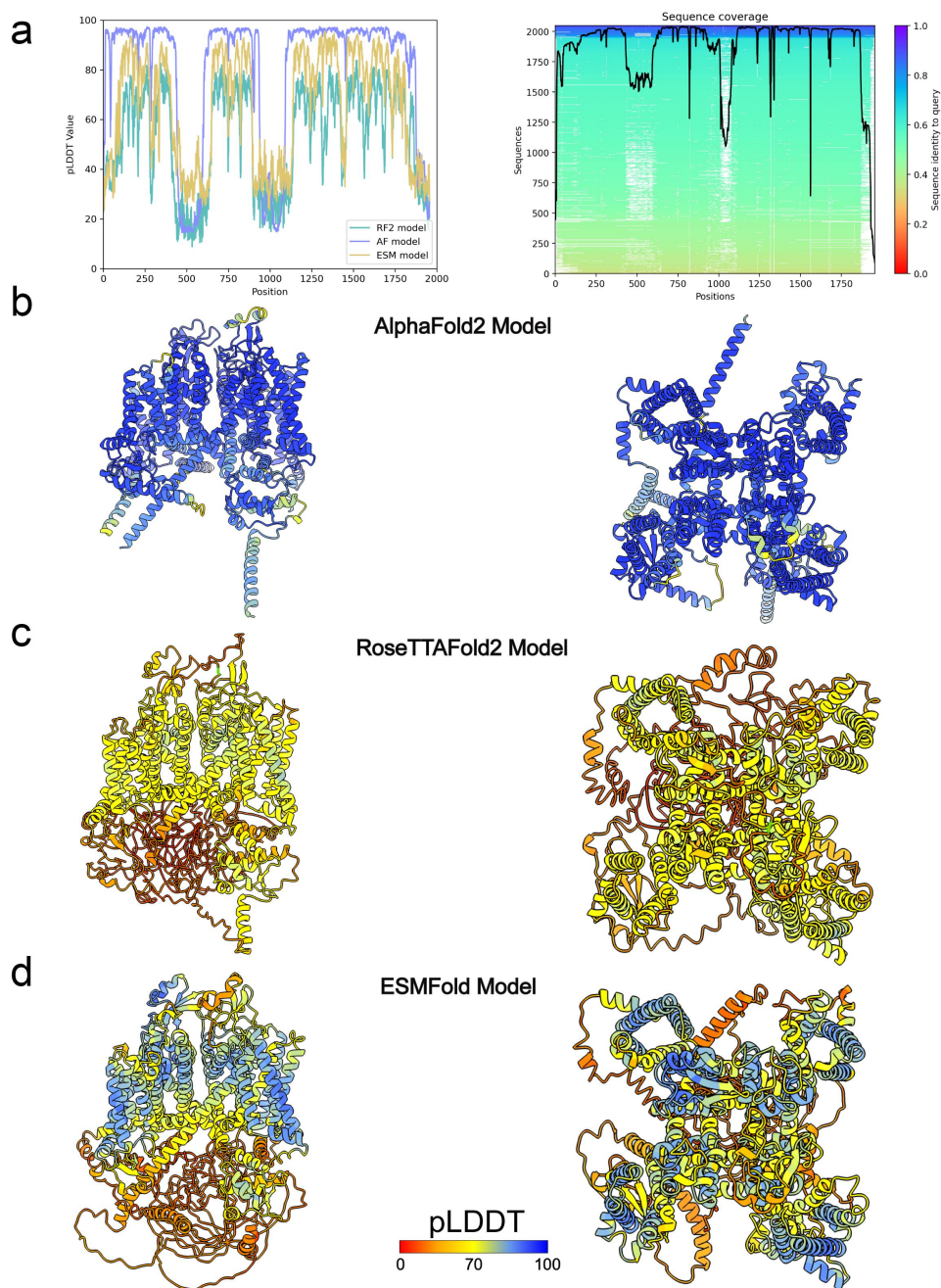
Figure 1. AlphaFold2, RoseTTAFold2, and ESMFold models of hNa$_V$1.8. a) plot of pLDDT confidence score versus hNa$_V$1.8 residue position for AlphaFold2 (AF), RoseTTAFold2 (RF2), and ESMFold (ESM) models. *right panel*, multiple sequence alignment of hNa$_V$1.8 sequence and its homologs identified by MMseqs2 method [38,39] and used for AlphaFold modeling of hNa$_V$1.8. a total number of homologous sequences identified per hNa$_V$1.8 residue position is shown by a black trace. b) transmembrane (*left panel*) and extracellular (*right panel*) views of AlphaFold model of hNa$_V$1.8. c) transmembrane (*left panel*) and extracellular (*right panel*) views of RoseTTAFold2 model of hNa$_V$1.8. d) transmembrane (*left panel*) and extracellular (*right panel*) views of ESMFold model of hNa$_V$1.8. AlphaFold2, RoseTTAFold2, and ESMFold models are colored by confidence score (pLDDT) from very low confidence (red) to good confidence (yellow) to high confidence (blue).

and ECL regions with good confidence (70 < pLDDT < 90), while RoseTTAFold2 predicted most transmembrane regions with low confidence (50 < pLDDT < 70) and predicted the pore domain with good confidence (Figure 1).

Relative to published structures, the overall topology of the models closest resembles the apo state (PDB: 7WFW), with AlphaFold2 having the lowest Cα RMSD at 2.0 Å (Figure 2(a)). This is exemplified by the AlphaFold2 model being able to predict VSD-
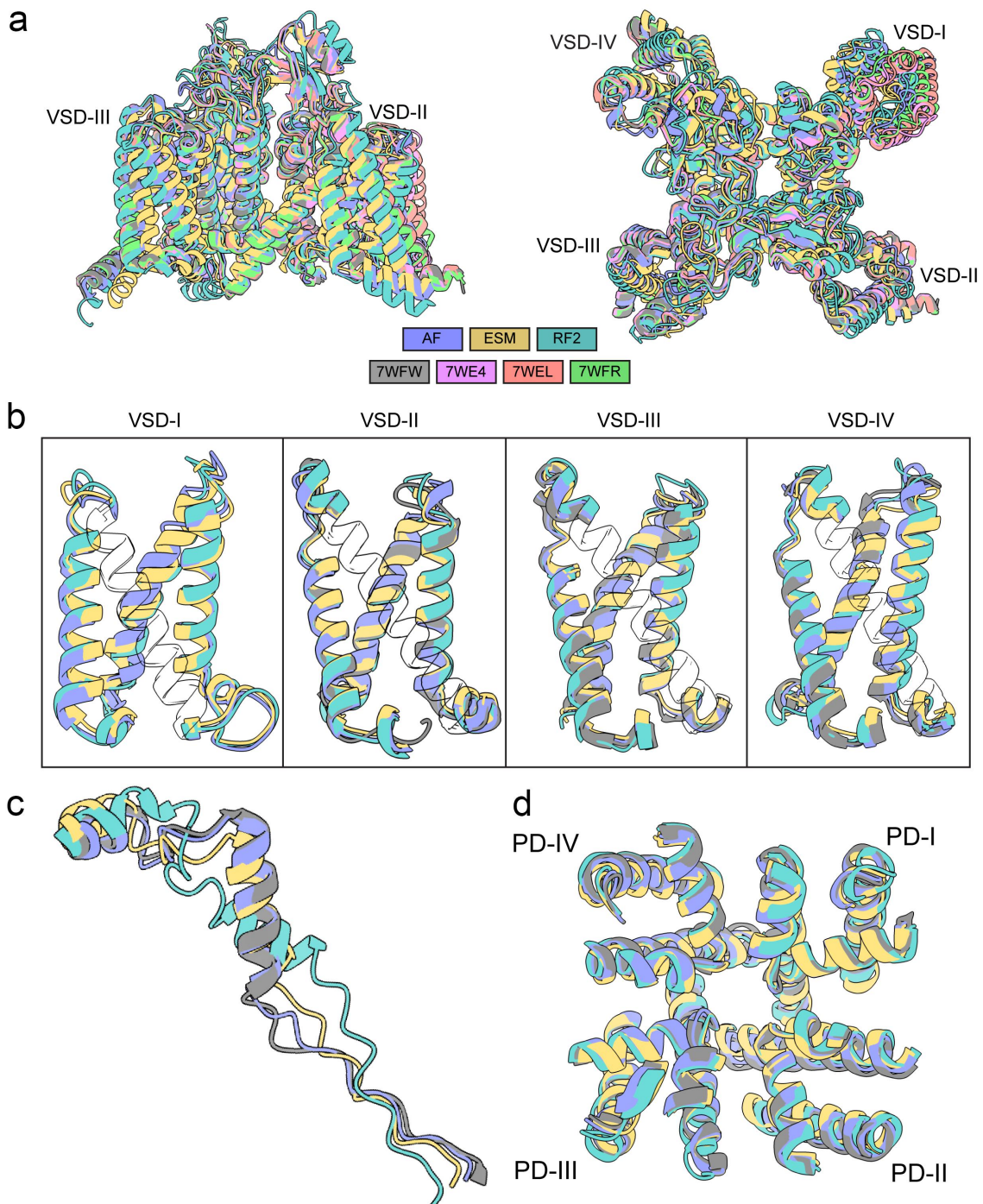
Figure 2. Comparison of AlphaFold2, RoseTTAFold2, and ESMFold models and cryoEM structures of hNa$_V$1.8. a) transmembrane (*left panel*) and extracellular (*right panel*) views of AlphaFold2, RoseTTAFold2, and ESMFold models superimposed to cryoEM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, 7WFR, and 7WFW) [22]. Intrinsically disordered *N*- and C-termini predicted by AlphaFold2, RoseTTAFold2, and ESMFold with low confidence were removed for clarity. b) transmembrane view of AlphaFold2, RoseTTAFold2, and ESMFold VSD-I, VSD-II, VSD-III, and VSD-IV models superimposed to apo-state hNa$_V$1.8 structure (PDB: 7WFW). c) extracellular view of AlphaFold2, RoseTTAFold2, and ESMFold domain III-IV intracellular linker models superimposed to apo-state hNa$_V$1.8 structure. d) extracellular view of AlphaFold2, RoseTTAFold2, and ESMFold pore domain (PD) models superimposed to apo-state hNa$_V$1.8 structure.

II, VSD-II, VSD-IV (Figure 2(b)), and the domain III-IV intracellular linker (Figure 2(c)) with less than 2.0 Å Cα RMSD to the apo state structure (VSD-I is not resolved in the apo state). Notably, all methods produced pore domain Cα RMSD 2.0 Å or less relative to the apo-state, with AlphaFold2 having the lowest Cα RMSD at 0.72 Å (Figure 2(d)).

For all three methods, the N-terminal domain (NTD) residues from M1 to N11, domain I-II intracellular loop region residues from L442 to D602, domain II-III intracellular loop region residues from S935 to L1092, and C-terminal domain (CTD) residues from N1875 to P1956 have low confidence scores (Figure 1). The low prediction confidence scores suggest that these regions are either highly flexible or inherently disordered in the absence of interacting intracellular partners. Notably, AlphaFold2 also predicted with good and high confidence scores parts of the NTD preceding the VSD-I (residues from N12 to S130), domain II-III intracellular loop region residues from N908 to R934, domain III-IV intracellular loop region from K1427 to N1463 that contains the intracellular gate important for Na_V channel fast inactivation, and segments of the CTD (residues from T1733 to S1874). The NTD, domain I-II intracellular loop region, domain I-III intracellular loop region, and CTD regions are absent in currently resolved structures of $hNa_V1.8$ (PDB IDs: 7WE4, 7WEL, 7WFR, and 7WFW) [22] but have been resolved in cryo-EM structures of other Na_V channel subtypes [22,29–35]. The domain III-IV intracellular linker residues from K1427 to N1463 consistently had the closest match between hNa_V 1.8 structures and AlphaFold2 models (Cα RMSD = 1.1 Å), compared to ESMFold models (Cα RMSD = 2.6–2.8 Å) and RoseTTAFold2 models (Cα RMSD = 5.5–5.6 Å) (Figure 2(c)).

Focusing on the pore region, the AlphaFold2, ESMFold, and RoseTTAFold2 models of $hNa_V1.8$ align closely with the experimentally resolved structures of $hNa_V1.8$ (Cα RMSD = 0.7–1.8 Å) Notably, the AlphaFold2 model exhibits only minor differences over the selectivity filter, P1-helix, P2-helix, S5, and S6 segments (Figure 2(d)). One noticeable difference is the conformation of domain I S6 segment (DI-S6). AlphaFold2's model shows a slight deviation in the helical turn near F386 residue in DI-S6,

causing its side chain to point toward the DI-IV fenestration and pack together with M1716 in DIV-S6 (Figure 3(a)). In contrast, the cryo-EM structures of $hNa_V1.8$ resolved in complex with a small molecule-based compound (A-803467) and in the apo state show F386 pointing downwards and M1716 oriented away from the fenestration. This difference between the AlphaFold2 model and the cryo-EM structure of $hNa_V1.8$ may arise due to a different conformation of DI-S6 captured in the AlphaFold2 model. Interestingly, AlphaFold2 predicted a conformation of the ECL in domain I (ECL-I) with high confidence, which was partially unresolved in the cryo-EM structures of $hNa_V1.8$. Specifically, residues 279 to 282 and 289 to 297 in ECL-I adopted a helical conformation with high confidence in the AlphaFold2 model of $hNa_V1.8$ but are unstructured in the cryo-EM structure of $hNa_V1.8$ (Figure 3(a)). For the same residues in ECL-I, ESMFold adopts a helical conformation between residues 283 to 291 with very low confidence (pLDDT <50), and RoseTTAFold2 adopts a loop conformation between residues 279 to 297 with very low confidence (Figure 3(a)).

The voltage sensor domains were predicted with high confidence in the AlphaFold2 model, good confidence in the ESMFold model, and low confidence in the RoseTTAFold2 model of hNa_V 1.8. However, the S3-S4 loop regions in VSD-I and VSD-II were predicted with good confidence in AlphaFold2, low confidence in ESMFold, and very low confidence in RoseTTAFold2, reflecting their flexibility (Figure 1). Despite the high confidence in the VSD predictions, comparison with the experimentally resolved $hNa_V1.8$ structures revealed several key differences. The cryo-EM structures of $hNa_V1.8$ in complex with A-803467 revealed various conformations of VSD-I, represented by class I, II, and III structures (PDB IDs: 7WE4, 7WEL, and 7WFR) [22]. VSD-I was not resolved in the cryo-EM structures of $hNa_V1.8$ in an apo state (PDB: 7WFW) [22]. However, conformations of VSD-I in the AlphaFold2, ESMFold, and RoseTTAFold2 models of hNa_V 1.8 don't align with the conformation of VSD-I in any of the cryo-EM structures of $hNa_V1.8$ (Figure 3(c)) (Cα RMSD = 2.5–3.8 Å). With a very high pLDDT confidence score, this
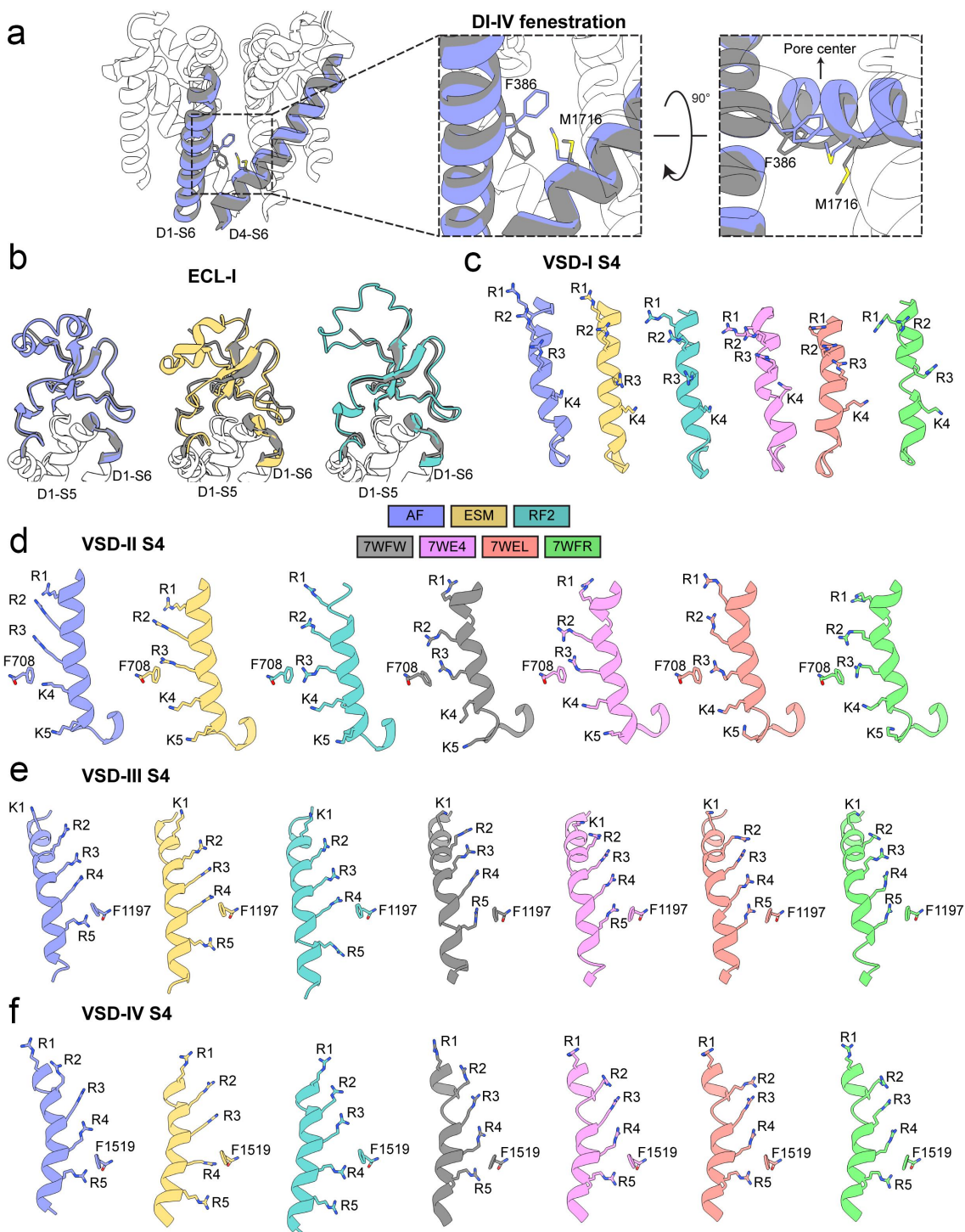
Figure 3. Comparison of specific regions in AlphaFold2, RoseTTAFold2, and ESMFold models to cryoEM structures of hNa$_V$1.8. a) transmembrane (*left and middle panel*), and extracellular (*right panel*) views of F386 relative to M1716 at the domains I-IV (DI-IV) interface fenestration region in AlphaFiold2 model superimposed to apo-state hNa$_V$1.8 structure (PDB: 7WFW). b) comparison of AlphaFold2, RoseTTAFold2, and ESMFold extracellular loop in domain I (ECL-I) model predictions relative to the partially resolved ECL-I in the apo-state hNa$_V$1.8 structure. c) comparison of VSD-I S4 gating charges in AlphaFold2, RoseTTAFold2, and ESMFold models relative to cryoEM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, and 7WFR) [22]. d) comparison of VSD-II S4 gating charges in AlphaFold2, RoseTTAFold2, and ESMFold models relative to cryoEM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, 7WFR, and 7WFW) [22]. e) comparison of VSD-III S4 gating charges in AlphaFold2, RoseTTAFold2, and ESMFold models relative to cryoEM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, 7WFR, and 7WFW) [22]. f) comparison of VSD-IV S4 gating charges in AlphaFold2, RoseTTAFold2, and ESMFold models relative to cryoEM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, 7WFR, and 7WFW) [22]. Side chains of gating charge-carrying residues in the S4 segments are shown in stick representation and labeled.

discrepancy raises a possibility that AlphaFold2's model of hNa$_V$1.8 represents another apo state of VSD-I. The AlphaFold2 model of hNa$_V$1.8 VSD-II appears to be in a fully "up" state, with the gating charges R1, R2, and R3 in the S4 segment of VSD-II positioned above the gating charge transfer residue F708 in the S2 segment of VSD-II. In contrast, the class I cryo-EM structure of hNa$_V$1.8 in complex with A-803467 (PDB: 7WE4) shows the gating charge R3 in the S4 segment of VSD-II at the gating charge transfer, considered to be a "half-click" down from the VSD-II state observed in the AlphaFold2 model (see Figure 2(b)). These observations may explain the difference in S3-S4 region conformations in VSD-II between the hNa$_V$1.8 structures and AlphaFold2 models (Figure 3(d)).

In contrast to the models of hNa$_V$1.8 VSD-I and VSD-II, AlphaFold2 conformations of VSD-III and VSD-IV appear to align closely with conformations of corresponding VSD-III and VSD-IV in the cryo-EM structures of hNa$_V$1.8 (VSD-III: Cα RMSD = 0.6–0.8 Å, VSD-IV: Cα RMSD = 1.2–1.3 Å) (Figure 2(b)). ESMFold and RoseTTAFold2 both performed worse in comparison to AlphaFold2 (ESMFold VSD-III: Cα RMSD = 1.5 Å, ESMFold VSD-IV: Cα RMSD = 1.8–1.9 Å, RoseTTAFold2 VSD-III: Cα RMSD = 1.8–1.9 Å, RoseTTAFold2 VSD-IV: Cα RMSD = 2.4 Å). The backbone conformations of VSD-III in the AlphaFold2 model almost identically match those in the resolved structure (PDB: 7WE4), with gating charges consistently at the same position relative to the gating transfer (Figure 3(e)). Similarly, the gating charges in the S4 segment of VSD-IV in the AlphaFold2 models and cryo-EM structures of hNa$_V$1.8 also occupy similar positions (Figure 3(f)). However, the S3-S4 conformations in VSD-IV are significantly different between the AlphaFold2 models and cryo-EM structures of hNa$_V$1.8. As with VSD-I, there is heterogeneity observed in the S3-S4 region of VSD-IV, with multiple conformations resolved in cryo-EM structures of hNa$_V$1.8 (PDB: 7WE4, 7WEL, and 7WFR, and 7WFW) [22]. Similar to the observations with VSD-II, the high confidence prediction of the VSD-IV S3-S4 region conformations by AlphaFold2 suggests a potentially different conformation of this region of hNa$_V$1.8.

Structural modeling of voltage-gated calcium (Cav) channels

Voltage-gated calcium (Ca$_V$) channels mediate Ca^{2+} influx upon depolarization of cell membrane potentials [1,2,25,41,42]. Ten subtypes of Ca$_V$ channels are divided into three main subfamilies, Ca$_V$1, Ca$_V$2, and Ca$_V$3, each serving distinct and crucial roles in physiological functions [28]. The Ca$_V$1 channels are responsible for muscle contraction, hormone secretion, and integrating synaptic inputs. The Ca$_V$2 channels play a key role in rapid communication in nerve cells. The Ca$_V$3 channels are crucial for the repetitive firing of action potentials in rhythmically firing cells, such as cardiac myocytes and thalamic neurons, contributing to regulating heart rhythm and synchronizing neural activities.

Cav1.1 structure is composed of α1, α2, β, γ, and γ subunits with α1 subunit responsible for voltage-dependent gating, calcium conduction, and modulation by small molecule drugs [23,43,44]. The protein sequence of CACNA1S gene encoded human Ca$_V$1.1 (hCa$_V$1.1) α subunit (UniProt ID: Q13698) was used as input into AlphaFold2, RoseTTAFold2, and ESMFold for structure prediction. We assessed the quality of predicted models using the pLDDT confidence score, compared how close they are to the resolved rabbit Ca$_V$1.1 structure (rCa$_V$1.1) (PDB: 5GJV) [43], and discussed the structural variations.

Our results showed that most of hCa$_V$1.1 structure was predicted with high confidence by AlphaFold2 (overall pLDDT = 71.4) and ESMFold (overall pLDDT = 70.6) and with low confidence by RoseTTAFold2 (overall pLDDT = 54.1) (Figures 4 and 5). Specifically, hCa$_V$1.1 VSD-I, VSD-II, VSD-III, and VSD-IV, pore domain, and ECL regions are predicted with high confidence (pLDDT >90) by AlphaFold2 and ESMFold and with good confidence by RoseTTAFold2 (60 < pLDDT < 90). However, most of the N-terminal domain formed by residues from M1 to K18, part of domain I-II intracellular loop region residues from L386 to A402, part of domain II-III intracellular loop region residues from L705 to E768, and most of the CTD regions formed by residues from Y1546 to S1780 and from G1835 to L1837 have been
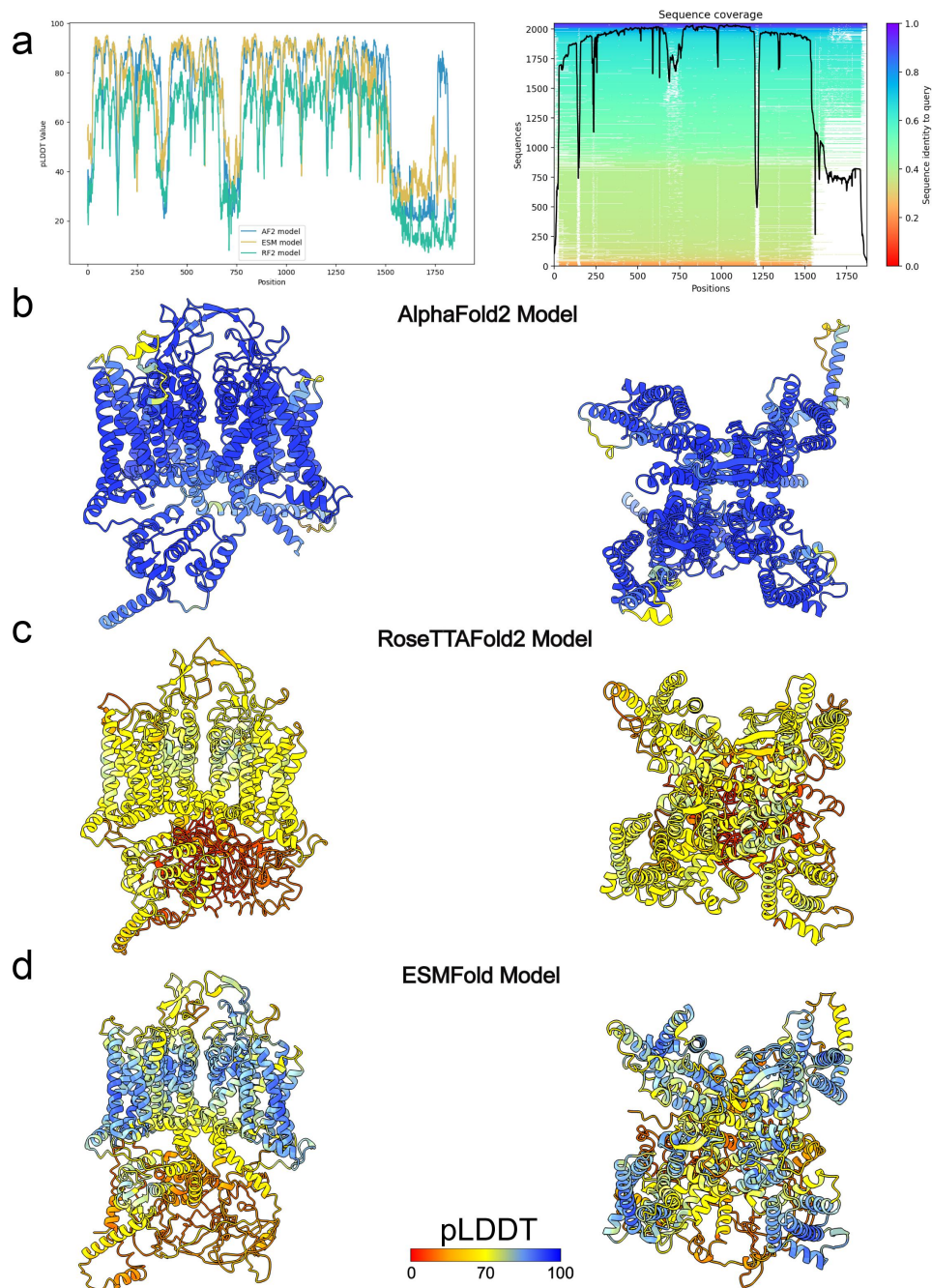
Figure 4. AlphaFold2, RoseTTAFold2, and ESMFold models of hCa$_V$1.1. a) plot of pLDDT confidence score versus hCa$_V$1.1 residue position for AlphaFold2 (AF), RoseTTAFold2 (RF2), and ESMFold (ESM) models. *right panel*, multiple sequence alignment of hCa$_V$1.1 sequence and its homologs identified by MMseqs2 method [38,39] and used for AlphaFold modeling of hCa$_V$1.1. a total number of homologous sequences identified per hCa$_V$1.1 residue position is shown by a black trace. b) transmembrane (*left panel*) and extracellular (*right panel*) views of AlphaFold model of hCa$_V$1.1. c) transmembrane (*left panel*) and extracellular (*right panel*) views of RoseTTAFold2 model of hCa$_V$1.1. d) transmembrane (*left panel*) and extracellular (*right panel*) views of ESMFold model of hCa$_V$1.1. AlphaFold2, RoseTTAFold2, and ESMFold models are colored by confidence score (pLDDT) from very low confidence (red) to good confidence (yellow) to high confidence (blue).

predicted with low confidence (pLDDT <50) by AlphaFold2, ESMFold, and RoseTTAFold2 (Figure 4(a)). This low confidence prediction suggests that these regions are either highly flexible or inherently disordered in the absence of

interacting intracellular partners (Figure 4(a)). Interestingly, AlphaFold2 also predicted parts of the NTD from P19 to K51, part of domain I-II intracellular loop region residues from R347 to K385, part of domain II-III intracellular loop
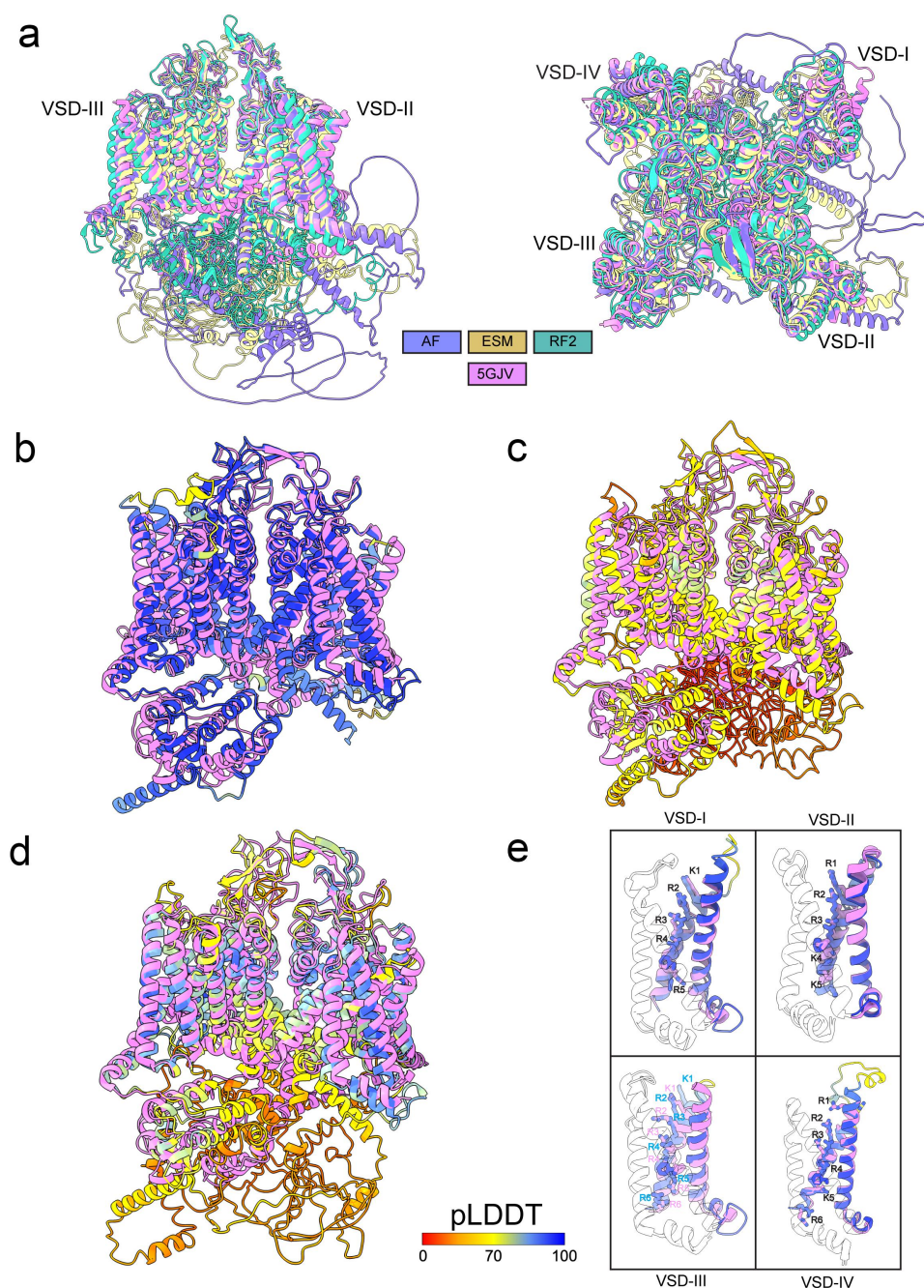
Figure 5. Comparison of AlphaFold2, RoseTTAFold2, and ESMFold models and cryoEM structures of hCa$_V$1.1. a) Transmembrane (*left panel*) and extracellular (*right panel*) views of AlphaFold2, RoseTTAFold2, and ESMFold models superimposed to cryoEM structure of hCa$_V$1.1 (PDB: 5GJV). b) transmembrane view of AlphaFold2 model of hCa$_V$1.1 superimposed to cryoEM structure of hCa$_V$1.1 (PDB: 5GJV). c) transmembrane view of RoseTTAFold2 model of hCa$_V$1.1 superimposed to cryoEM structure of hCa$_V$1.1 (PDB: 5GJV). d) transmembrane view of ESMFold model of hCa$_V$1.1 superimposed to cryoEM structure of hCa$_V$1.1 (PDB: 5GJV). AlphaFold2, RoseTTAFold2, and ESMFold models are colored by confidence score (pLDDT) from very low confidence (red) to good confidence (yellow) to high confidence (blue). e) transmembrane view of AlphaFold model of hCa$_V$1.1 VSD-I, VSD-II, VSD-III, and VSD-IV colored in blue superimposed with cryoEM structures of hCa$_V$1.1 (PDB: 5GJV) colored in purple. Side chains of gating charge-carrying residues in the S4 segments are shown in stick representation and labeled.

region residues from G689 to K704, domain III-IV intracellular loop region residues from G689 to K704 from E1073 to P1106, and CTD region residues from N1383 to G1545 with good

confidence (70 < pLDDT < 90) and high confidence (pLDDT >90) (Figure 4(a,b)).

In the AlphaFold2, ESMFold, and RoseTTAFold2 models of hCa$_V$1.1, the

extracellular regions of the pore relatively closely match the cryo-EM structure of rCa$_V$1.1 [43], exhibiting only minor differences over the selectivity filter, P1-helix, P2-helix, and ECL regions (Cα RMSD = 1.09 Å) (Figure 5(a-d)). However, conformations of the intracellular half of the S6 segments are captured in a different state in the AlphaFold2, ESMFold, and RoseTTAFold2 model of hCa$_V$1.1 compared to the cryo-EM structure of rCa$_V$1.1 (Cα RMSD = 4.96 Å) (Figure 5(a-d)).

The voltage sensor domains were predicted with high confidence (pLDDT >90) in the AlphaFold2 model of hCa$_V$1.1. However, a significant drop in confidence was observed in the S3-S4 regions, reflecting their higher flexibility (Figure 5(b,e)). Notably, VSD-I state in the AlphaFold2 model of hCa$_V$1.1 is matching closely VSD-I state in the cryo-EM structure of rCa$_V$1.1 (Figure 5(e)). Similarly, the gating charges in the S4 segment of VSD-I in the AlphaFold2 model of hCa$_V$1.1 and cryo-EM structures of rCa$_V$1.1 also occupy similar positions (Figure 5(e)). VSD-II state in the AlphaFold2 model of hCa$_V$1.1 has the S4 segment positioned a "half-click" down compared to VSD-II state in the cryo-EM structure of rCa$_V$1.1 (Figure 5(e)). VSD-III state in AlphaFold2 model of hCa$_V$1.1 has the S4 segment positioned a "half-click" up compared to VSD-III state in the cryo-EM structure of rCa$_V$1.1 (Figure 5(e)). VSD-IV state in AlphaFold2 model of hCa$_V$1.1 matches closely VSD-IV state in the cryo-EM structure of rCa$_V$1.1 (Figure 5(e)). Similarly, the gating charges in the S4 segment of VSD-IV in the AlphaFold2 model of hCa$_V$1.1 and cryo-EM structures of rCa$_V$1.1 also occupy similar positions. The S3-S4 loop regions in VSD-I, VSD-III, and VSD-IV have not been resolved in the cryo-EM structure of rCa$_V$1.1 (PDB: 5GJV) [43]. Notably, the S3-S4 loop regions in VSD-I, VSD-III, and VSD-IV AlphaFold2 models of hCa$_V$1.1 have good confidence (pLDDT >70) (Figure 5(e)).

Structural modeling of voltage-gated potassium (Kv) channels

Voltage-gated potassium (K$_V$) channels mediate K$^+$ efflux upon membrane depolarization and regulate membrane potential [1,2]. There are 40 subtypes of K$_V$ channels divided into 12 main subfamilies, from K$_V$1 to K$_V$12 [28]. Differently from Na$_V$ and Ca$_V$ channels, K$_V$ channels are tetramers where each VSD-PD pair forms a separate subunit. The K$_V$1.3 channels regulate membrane potential and calcium signaling in lymphocytes and oligodendrocytes [45]. These channels form homotetramers with the VSDs and PDs forming the membrane-spanning region and a cytosolic tetramerization (T1) domain [24]. The protein sequence of KCNA3 gene encoded human K$_V$1.3 (hK$_V$1.3) (UniProt ID: P22001) was used as input into AlphaFold2, RoseTTAFold2, and ESMFold for structure prediction. We assessed the quality of predicted models using the pLDDT confidence score, compared how close they are to the resolved hK$_V$1.3 structures (PDB: 7SSX and 7SSY) [24], and discussed the structural variations.

Our results showed that AlphaFold2 predicted the transmembrane region of hK$_V$1.3 formed by VSDs and pore domain with high confidence (pLDDT >90), as illustrated in Figures 6 and 7. The first part of the N-terminal region formed by residues from M1 to E104 was predicted with low confidence (pLDDT <50) (Figure 6(a)). The second part of the N-terminal region that comprises the T1 domain formed by residues from R105 to S230 was predicted with good confidence (70 < pLDDT < 90) or high confidence (pLDDT >90) and closely matched cryo-EM structures of hK$_V$1.3 [24] N-terminal region formed by the same region (Cα RMSD = 1.4 Å) (Figure 6(a)). All of the pore region formed by residues from M395 to T491 was predicted with good confidence (70 < pLDDT < 90) or high confidence (pLDDT >90) and closely matched cryo-EM structure of hK$_V$1.3 [24] pore region (PDB: 7SSX) (Cα RMSD = 1.4 Å) (Figure 6(a,b)). C-terminal region formed by residues from E492 to V575 has low confidence (pLDDT <50) (Figure 6(a)). This low prediction confidence suggests that this region is either highly flexible or inherently disordered in the absence of interacting partners. Most of the VSD region formed by residues from P232 to S381 was predicted with good confidence (70 < pLDDT < 90) or high confidence (pLDDT >90) and closely matched cryo-EM structure of hK$_V$1.3 [24] VSD region (Cα RMSD = 1.9 Å) (Figures 6(a) and 7(c)). However, the S1-S2
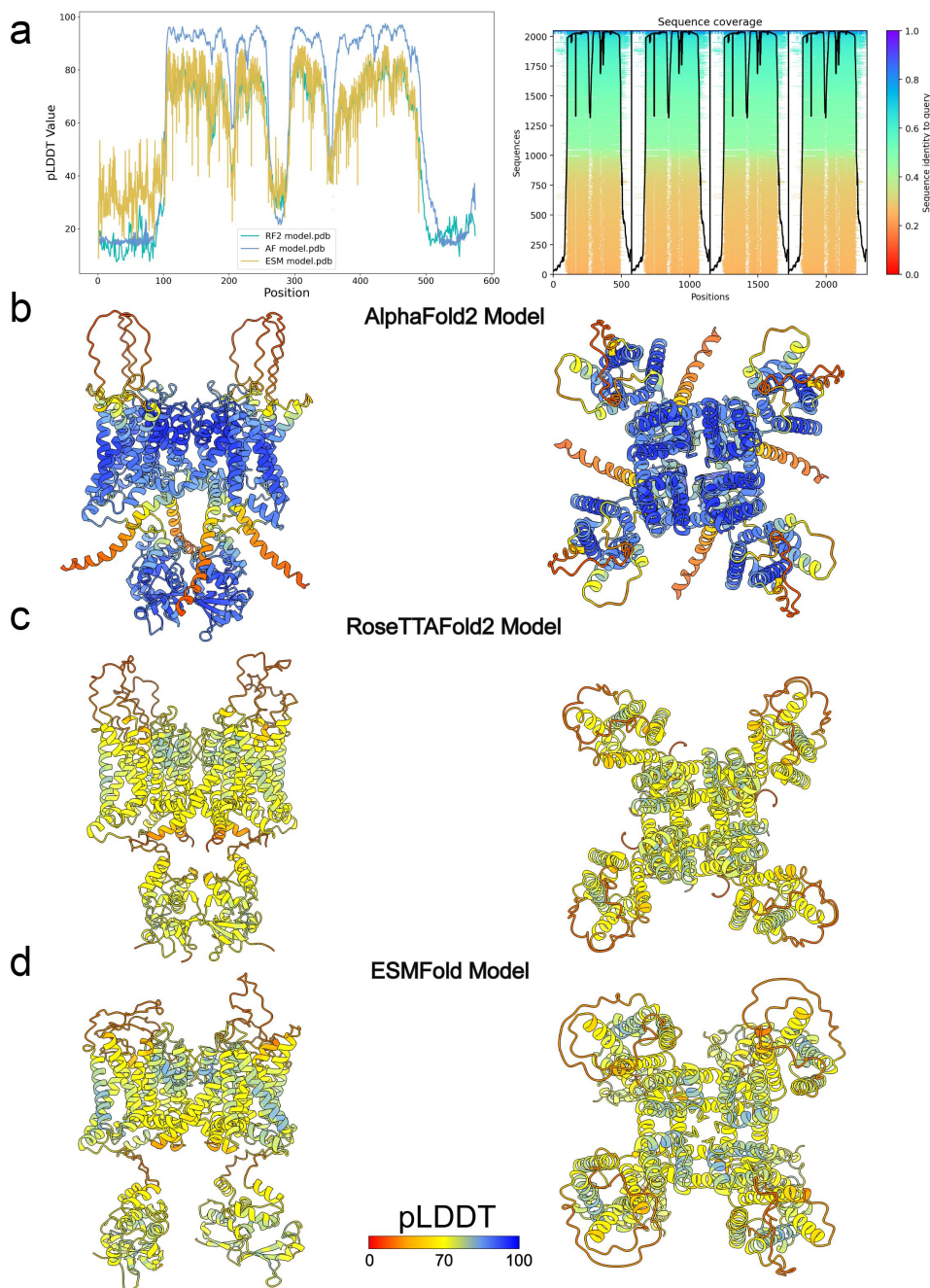
Figure 6. AlphaFold2, RoseTTAFold2, and ESMFold models of hK$_V$1.3. a) plot of pLDDT confidence score versus hK$_V$1.3 residue position for AlphaFold2 (AF), RoseTTAFold2 (RF2), and ESMFold (ESM) models. *right panel*, multiple sequence alignment of hK$_V$1.3 sequence and its homologs identified by MMseqs2 method [38,39] and used for AlphaFold modeling of hK$_V$1.3. a total number of homologous sequences identified per hK$_V$1.3 residue position is shown by a black trace. b) transmembrane (*left panel*) and extracellular (*right panel*) views of AlphaFold model of hK$_V$1.3. c) transmembrane (*left panel*) and extracellular (*right panel*) views of RoseTTAFold2 model of hK$_V$1.3. d) transmembrane (*left panel*) and extracellular (*right panel*) views of ESMFold model of hK$_V$1.3. Unstructured *N*- and C- terminals are not shown for clarity in AlphaFold, RoseTTAFold2 and ESMFold models. AlphaFold2, RoseTTAFold2, and ESMFold models are colored by confidence score (pLDDT) from very low confidence (red) to good confidence (yellow) to high confidence (blue).

and S3-S4 loop regions were predicted with low confidence (pLDDT <50), reflecting unstructured nature of these regions that have not been resolved in the cryo-EM structures of hK$_V$1.3 [24]

(Figure 7(c)). The relative position of the gating charges in the S4 segment revealed a similar state for the modeled VSD as the one observed in the cryo-EM experimental structures (Figure 7(c)).
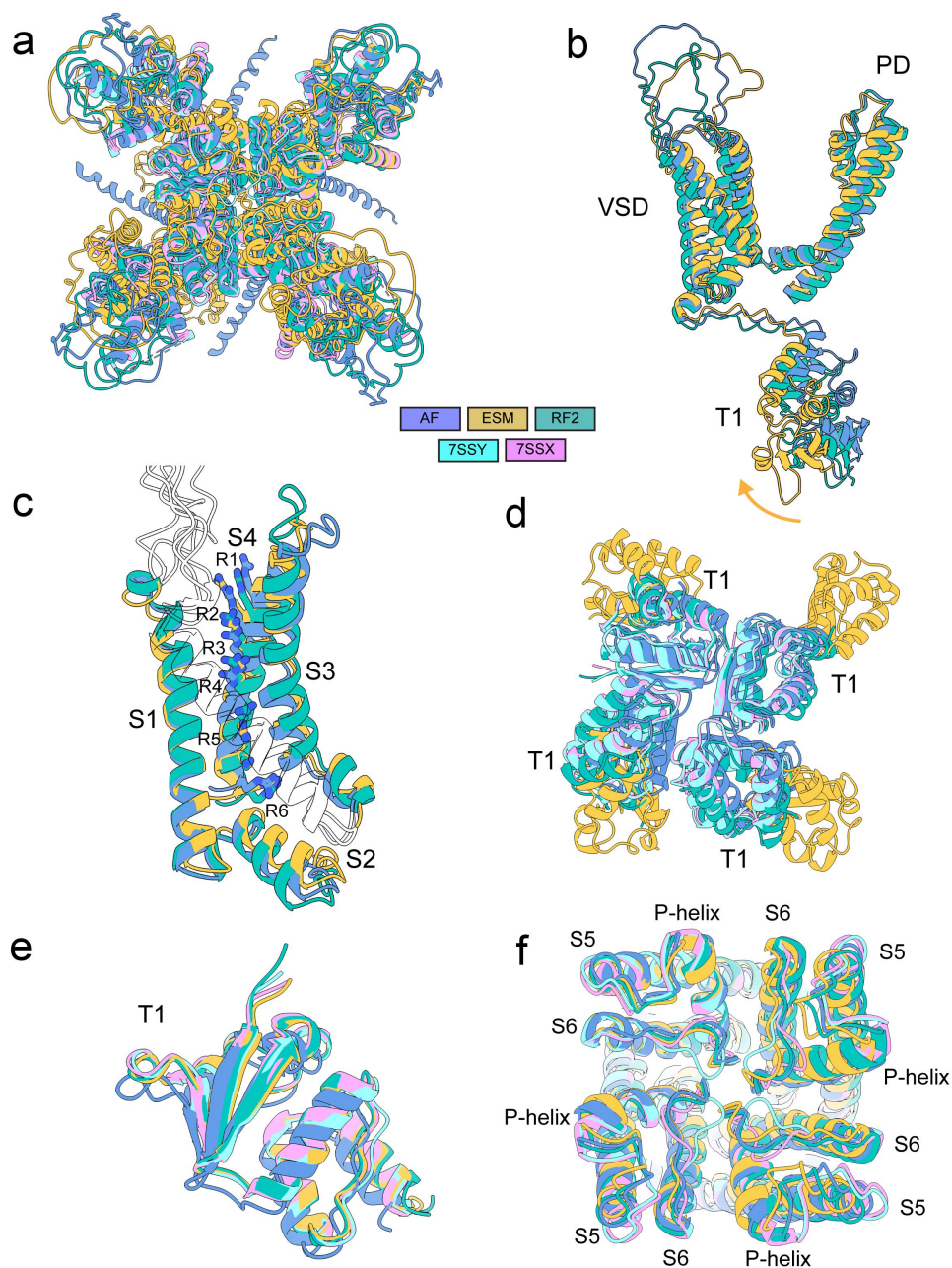
Figure 7. Comparison of specific regions in AlphaFold2, RoseTTAFold2, and ESMFold models to cryoEM structures of hK$_V$1.3. a) extracellular view of AlphaFold2, RoseTTAFold2, and ESMFold models of hK$_V$1.3 superimposed with cryoEM structures of hK$_V$1.3 (PDBs: 7SSX and 7SSY). b) superimposition of AlphaFold, RoseTTAFold2, and ESMFold models. The yellow arrow indicates the tilt observed in the ESMFold model of the T1 domain. c) transmembrane view of AlphaFold, RoseTTAFold2, and ESMFold models of hK$_V$1.3 VSD. The side chains of the gating charges located in the S4 segment are shown in stick representation and labeled. d) intracellular view of AlphaFold2, RoseTTAFold2, and ESMFold models of hK$_V$1.3 T1 domain after superimposition of the full models with the cryoEM structures of hK$_V$1.3 colored in purple (PDB: 7SSX) and cyan (PDB: 7SSY). e) superimposition of individual T1 domains from AlphaFold2, RoseTTAFold2, and ESMFold models with the T1 domains from cryoEM structures of hK$_V$1.3. f) extracellular view of AlphaFold2, RoseTTAFold2, and ESMFold models of hK$_V$1.3 pore domain superimposed with cryoEM structures of hK$_V$1.3 pore.

RoseTTAFold2 prediction for K$_V$1.3 showed the expected architecture with the transmembrane domains, building up the VSDs and PDs, and the intracellular T1 domain (Figure 6(a)). The confidence of the prediction was good at the well-structured regions (80 < pLDDT < 90), but lower compared to the AF models, while it presented low confidence (pLDDT <50) at the N and C terminal

regions, as well as at the S1-S2 and S3-S4 unstructured loops of the VSDs (Figure 6(a,c)). ESMFold prediction resulted in a similar model that also had good confidence prediction (80 < pLDDT < 90) at the well-structured transmembrane domains and T1 domain, and low (pLDDT <50) at the N and C terminal regions and VSD extracellular loops (Figures 6(a,d)). The predicted state of the VSDs was similar to the one in the AF models (Figure 7(c)).

Notably, the ESMFold model of $K_V1.3$ presented a different arrangement of the T1 domains, in which the domain independently has the expected fold when compared to the available cryo-EM structures (Figure 7(e)) (RMSD <1 Å), but the domain of each subunit shifts outwards (Figure 7(b), yellow arrow) breaking the interactions among the four subunit T1 domains that are observed in AlphaFold2 and RoseTTAFold2 models and solved structures (Figure 7(d)). The predicted structure for the $K_V1.3$ pore in AlphaFold2, RoseTTAFold2, and ESMFold models closely matched the one observed in experimental structures (RMSD = 1.5 Å) (Figure 7(f)).

Conclusions

Deep learning-based methods, such as AlphaFold [8], RosetTTAFold [9,10], and ESMFold [20] are useful for predicting structures of transmembrane regions of ion channels, including the voltage-sensing and pore domains, with high confidence. The extracellular and intracellular loop regions and intracellular *N*- and *C*-termini regions can be potentially predicted with high confidence if they are formed by α-helical or β-sheet secondary structure. Deep learning-based methods may predict alternative conformations of ion channels compared to known structures of identical or homologous ion channels. However, the accuracy of alternative ion channel conformations is only determined once confirmed by structural and experimental data. Modeling unstructured extracellular and intracellular loop regions and intracellular *N*- and *C*-termini regions remains challenging in the absence of potential protein partners to stabilize specific conformations of

these regions. Structure prediction of ion channels using deep learning-based methods might be useful for designing therapeutics and molecular probes targeting specific ion channel subtypes. Finally, structural modeling of ion channels in complex with other proteins deep learning-based methods might reveal molecular mechanisms of ion channel modulation by extracellular, transmembrane, and intracellular proteins.

Disclosure statement

V. Yarov-Yarovoy reported personal fees from Gerson Lehrman Group, Grunenthal, Novo Ventures, and Praxis Precision Medicines outside the submitted work; in addition, V. Yarov-Yarovoy had a patent to "Peptides targeting sodium channels to treat pain" pending. No other disclosures were reported.

Data availability statement

AlphaFold2, ESMFold, and RoseTTAFold2 models of hNaV1.8, hCaV1.1, and hKV1.3 are available for download through the DRYAD database (https://doi.org/10.5061/dryad.z08kprrn0).

ORCID

Vladimir Yarov-Yarovoy 🔴 http://orcid.org/0000-0002-2325-4834

References

[1] Hille B. Ion channels of excitable membranes. Sunderland, MA): Sinauer Associates; 2001.
[2] Zheng J, Trudeau MC. Textbook of ion channels. Boca Raton: CRC Press; 2023.

[3] Cao E, Liao M, Cheng Y, et al. TRPV1 structures in distinct conformations reveal activation mechanisms. Nature. 2013;504(7478):113–8. doi: 10.1038/nature12823

[4] Cheng Y. Single-particle cryo-EM at crystallographic resolution. Cell. 2015;161(3):450–7. doi: 10.1016/j.cell.2015.03.049

[5] Armache JP, Cheng Y. Single-particle cryo-EM: beyond the resolution. Natl Sci Rev. 2019;6(5):864–6. doi: 10.1093/nsr/nwz127

[6] Cheng Y. Single-particle cryo-EM—How did it get here and where will it go. Science. 2018;361(6405):876–880. doi: 10.1126/science.aat4346

[7] Yip KM, Fischer N, Paknia E, et al. Atomic-resolution protein structure determination by cryo-EM. Nature. 2020;587(7832):157–61. doi: 10.1038/s41586-020-2833-4

[8] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9. doi: 10.1038/s41586-021-03819-2

[9] Baek M, Anishchenko I, Humphreys IR, et al. Efficient and accurate prediction of protein structure using RoseTTAFold2. bioRxiv 2023:2023.05.24.542179.

[10] Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373 (6557):871–6. doi: 10.1126/science.abj8754

[11] UniProt C. UniProt: the universal protein knowledge-base in 2023. Nucleic Acids Res. 2023;51(D1):D523–D531. doi: 10.1093/nar/gkac1052

[12] Berman HM, Westbrook J, Feng Z, et al. The protein data bank. Nucleic Acids Res. 2000;28(1):235–242. doi: 10.1093/nar/28.1.235

[13] Humphreys IR, Pei J, Baek M, et al. Computed structures of core eukaryotic protein complexes. Science. 2021;374(6573):eabm4805. doi: 10.1126/science.abm4805

[14] Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2022:2021.10.04.463034.

[15] Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. Nature. 2021;596(7873):590–6. doi: 10.1038/s41586-021-03828-1

[16] Watson JL, Juergens D, Bennett NR, et al. De Novo design of protein structure and function with RFdiffusion. Nature. 2023;620(7976):1089–1100. doi: 10.1038/s41586-023-06415-8

[17] Watson JL, Juergens D, Bennett NR, et al. Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models. bioRxiv 2022:2022.12.09.519842.

[18] Wang J, Lisanza S, Juergens D, et al. Scaffolding protein functional sites using deep learning. Science. 2022;377(6604):387–94. doi: 10.1126/science.abn2100

[19] Rettie SA, Campbell KV, Bera AK, et al. Cyclic peptide structure prediction and design using AlphaFold. bioRxiv 2023:2023.02.25.529956.

[20] Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science. 2023;379(6637):1123–30. doi: 10.1126/science.ade2574

[21] Varadi M, Anyango S, Deshpande M, et al. AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50 (D1):D439–D44. doi: 10.1093/nar/gkab1061

[22] Huang X, Jin X, Huang G, et al. Structural basis for high-voltage activation and subtype-specific inhibition of human Na v 1.8. Proc Natl Acad Sci U S A. 2022;119 (30):e2208211119. doi: 10.1073/pnas.2208211119

[23] Wu J, Yan Z, Li Z, et al. Structure of the voltage-gated calcium channel Cav1.1 complex. Science. 2015;350 (6267):aad2395. doi: 10.1126/science.aad2395

[24] Selvakumar P, Fernandez-Marino AI, Khanra N, et al. Structures of the T cell potassium channel Kv1.3 with immunoglobulin modulators. Nat Commun. 2022;13 (1):3854. doi: 10.1038/s41467-022-31285-5

[25] Catterall WA, Lenaeus MJ, Gamal El-Din TM. Structure and pharmacology of voltage-gated sodium and calcium channels. Annu Rev Pharmacol Toxicol. 2020;60(1):133–54. doi: 10.1146/annurev-pharmtox-010818-021757

[26] Catterall WA. Forty years of sodium channels: structure, function, pharmacology, and epilepsy. Neurochem Res. 2017;42(9):2495–504. doi: 10.1007/s11064-017-2314-9

[27] Ahern CA, Payandeh J, Bosmans F, et al. The hitchhiker's guide to the voltage-gated sodium channel galaxy. J Gen Physiol. 2016;147(1):1–24. doi: 10.1085/jgp.201511492

[28] Alexander SP, Mathie A, Peters JA, et al. The concise guide to pharmacology 2021/22: ion channels. Br J Pharmacol. 2021;178 Suppl 1:S157–S245. doi: 10.1111/bph.15539

[29] Pan X, Li Z, Jin X, et al. Comparative structural analysis of human Na v 1.1 and Na v 1.5 reveals mutational hotspots for sodium channelopathies. Proc Natl Acad Sci U S A. 2021;118(11):118. doi: 10.1073/pnas.2100066118

[30] Pan X, Li Z, Huang X, et al. Molecular basis for pore blockade of human Na+ channel Na v 1.2 by the μ-conotoxin KIIIA. Science. 2019;363(6433):1309–1313. doi: 10.1126/science.aaw2999

[31] Li X, Xu F, Xu H, et al. Structural basis for modulation of human Na(V)1.3 by clinical drug and selective antagonist. Nat Commun. 2022;13(1):1286. doi: 10.1038/s41467-022-28808-5

[32] Pan X, Li Z, Zhou Q, et al. Structure of the human voltage-gated sodium channel Na v 1.4 in complex with β1. Science. 2018;362(6412):eaau2486. doi: 10.1126/science.aau2486

[33] Jiang D, Shi H, Tonggu L, et al. Structure of the cardiac sodium channel. Cell. 2020;180(1):122–34 e10. doi: 10.1016/j.cell.2019.11.041

[34] Fan X, Huang J, Jin X, et al. Cryo-EM structure of human voltage-gated sodium channel Na v 1.6. Proc Natl Acad Sci U S A. 2023;120(5):e2220578120. doi: 10.1073/pnas.2220578120

[35] Shen H, Liu D, Wu K, et al. Structures of human Nav1.7 channel in complex with auxiliary subunits and animal toxins. Science. 2019;363(6433):1303–8. doi: 10.1126/science.aaw2493

[36] Bouza AA, Isom LL. Voltage-Gated Sodium Channel beta Subunits and Their Related Diseases. Handb Exp Pharmacol. 2018;246:423–450.

[37] Mirdita M, Schutze K, Moriwaki Y, et al. ColabFold: making protein folding accessible to all. Nat Methods. 2022;19(6):679–82. doi: 10.1038/s41592-022-01488-1

[38] Mirdita M, Steinegger M, Soding J, et al. MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics. 2019;35(16):2856–8. doi: 10.1093/bioinformatics/bty1057

[39] Steinegger M, Soding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017;35(11):1026–8. doi: 10.1038/nbt.3988

[40] Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinf. 2010;11(1):431. doi: 10.1186/1471-2105-11-431

[41] Nanou E, Catterall WA. Calcium channels, synaptic plasticity, and neuropsychiatric disease. Neuron. 2018;98(3):466–81. doi: 10.1016/j.neuron.2018.03.017

[42] Catterall WA. Voltage-gated calcium channels. Cold Spring harbor perspectives in biology. Cold Spring Harbor Perspect Biol. 2011;3(8):a003947–a003947. doi: 10.1101/cshperspect.a003947

[43] Wu J, Yan Z, Li Z, et al. Structure of the voltage-gated calcium channel Cav1.1 at 3.6 Å resolution. Nature. 2016;537(7619):191–196. doi: 10.1038/nature19321

[44] Zhao Y, Huang G, Wu J, et al. Molecular basis for ligand modulation of a Mammalian Voltage-Gated Ca (2+) channel. Cell. 2019;177(6):1495–506 e12. doi: 10.1016/j.cell.2019.04.043

[45] Grissmer S, Nguyen AN, Aiyar J, et al. Pharmacological characterization of five cloned voltage-gated K[+] channels, types Kv1.1, 1.2, 1.3, 1.5, and 3.1, stably expressed in mammalian cell lines. Mol Pharmacol. 1994;45:1227–1234.