Dean&Francis
ISSN 2959-409X

Structural Analysis of the SCN1A R1648H Mutation in Epilepsy Using Alpha Fold

Hanlin Hu[1, *], Jianwei Zhang[2]

[1]College of Engineering, The Ohio State University, Columbus, USA
[2]Southwest Jiaotong University, Chengdu, China

*Corresponding author: Hu.2305@osu.edu

Abstract:
Epilepsy is one of the most common chronic, noncommunicable neurological conditions in the world, affecting around 50 million people worldwide. The condition is characterized by repeated seizures caused by unusual electrical activity in the brain. The major causes of the condition include brain injury or infection, metabolic problems, neurotransmitter imbalance, and genetic factors. Besides the major causes listed above, genetic mutations also play an important role in certain forms. SCN1A is one of the key genes involved and is linked with Dravet syndrome—a severe developmental and epileptic condition—and a list of epileptic conditions. The SCN1A gene produces a specific protein, part of the sodium channel (NaV1.1), in nerve cells. This protein can regulate sodium ions across the cell membranes, which is critical for how the electrical signals move and can be controlled. However, the mutations in SCN1A can disrupt such electrical signal movement processes, leading to nerve transmission problems and possibly increasing the chance of seizures. This study employs AlphaFold, an artificial intelligence-based protein structure prediction tool, to explore the structural changes caused by the R1648H mutation. By comparing the wild-type and mutant SCN1A proteins, we aim to identify specific alterations contributing to sodium channel malfunctioning in epileptic conditions. Understanding these molecular consequences can bridge the gap between genetic mutations and clinical outcomes, potentially informing the development of targeted therapies and personalized medicine approaches for SCN1A-related epilepsies.

Keywords: Alpha Fold; epilepsy; the SCN1A R1648H mutation; structural analysis.

ISSN 2959-409X

1. Introduction

Epilepsy affects more than 50 million people globally [1], causing repeated seizures stemming from unusual electrical activity in the brain. Along with major factors like brain injuries or infections that can lead to epilepsy, genetic mutations are significant contributors to certain types [2,3]. Among the genes implicated, SCN1A stands out due to its association with a spectrum of epileptic disorders, including Dravet syndrome—a severe developmental and epileptic encephalopathy [4]. SCN1A is a gene that encodes the alpha subunit of the voltage-gated sodium channel NaV1.1, which is necessary to regulate neuronal electrical communication by directing the passage of sodium ions along the cell wall. Mutations in SCN1A can cause this channel to stop functioning properly, leading to malfunctioning nerve signaling and a heightened risk of seizures [5]. R1648H is the most well-known of the SCN1A mutations and is often associated with fatal epilepsy [6]. The mutation changes the geometry of the sodium channel, causing it to malfunction during normal functioning. Traditionally, investigating SCN1A mutations involved lab studies and clinical trials to discern their effects. However, computational software such as AlphaFold has reinvented the task with the capability to model three-dimensional protein structures and pinpoint the presence of individual mutations more precisely [7].

This study aims to explore how the R1648H mutation changes the protein's structure, called SCN1A, using AlphaFold. By contrasting the proteins' normal (wild-type) and mutated structures, we aim to discover the structural variants of the sodium channel that cause it to become defective in epilepsy. Such computational work allows for more information about the molecular impact of SCN1A mutations than can be reached with more conventional techniques.

This understanding of structural changes is important for linking the molecular changes to the clinical features of epilepsy [3]. By explaining how mutations such as R1648H induce more neuronal hyperactivity and seizures, we can inform the development of targeted therapeutic strategies. For example, drugs that normalize the mutated sodium channel or restore its function could be a solution. These insights also open up avenues for personalized medicine, where treatment depends on one's specific genetic makeup.

For the proteins of the wild-type SCN1A and R1648H mutant, we will use AlphaFold to predict the structure. After prediction, we'll compare these structures with molecular visualization software like PyMOL to find structural similarities. We'll focus on functionally important areas in the protein, such as the ion-conducting pore and voltage-sensor domains, to see how the mutation changes the channel's conformation. Such structural data will be integrated with functional data from other electrophysiological studies to provide a more detailed picture of the role of the mutation in sodium channel regulation in epilepsy. Finally, the work adds to a growing understanding of the molecular mechanisms of SCN1A-associated epilepsies. It paves the way for future research to find treatments that can counteract the impacts of these mutations.

2. Methods

2.1 Protein Sequence Acquisition

For the purpose of exploring the structural impact of the R1648H mutation of SCN1A, we acquired amino acid sequences of wild-type and mutant SCN1A. The wild-type SCN1A protein sequence was pulled from the UniProt database. The epileptic-related R1648H mutation replaces arginine (R) at position 1648 with histidine (H) and has been widely linked to very severe epilepsy conditions such as Dravet syndrome [2,4].

2.2 Structural Prediction with AlphaFold

The three-dimensional structures of the wild-type and R1648H mutant SCN1A proteins were predicted using AlphaFold2 implemented through the ColabFold platform. ColabFold provides an accessible interface to AlphaFold2, allowing high-accuracy protein structure predictions using Google Colaboratory notebooks with GPU acceleration.

Due to the large size of the SCN1A protein (2,009 amino acids), which poses computational challenges, a focused region encompassing the mutation site was selected for modeling. Specifically, residues 1600–1700 were chosen to cover the fourth voltage-sensing domain (VSD4) where the R1648H mutation resides.

The input sequences consisted of FASTA format files containing the wild-type and mutant amino acid sequences for residues 1600–1700. A monomer model without templates was used for the prediction. The default setting of three recycles was employed to optimize model accuracy. Five models were generated for each sequence, and the one with the highest predicted Local Distance Difference Test (pLDDT) score was selected for further analysis [7]. The predictions were run on Google Colaboratory using a Tesla T4 GPU.

2.3 Structural Analysis

All the structural analyses were conducted using PyMOL (version 2.5.5). The predicted structures of the wild-type and R1648H mutant proteins were imported into PyMOL.

Structures were aligned using the 'align' function to superimpose the models based on their alpha-carbon backbone atoms. The root mean square deviation (RMSD) was calculated to quantify structural differences between the wild-type and mutant structures [8,9].

2.4 Visualization and Assessment

We examined the immediate environment of residue 1648 to determine whether the R1648H substitution alters side-chain orientation, hydrogen bonding, and steric interactions [10]. The distance between the histidine side chain at residue 1648 and adjacent residues was calculated to identify any steric collisions or new interactions [8,9]. Additionally, we investigated whether the mutation might create new hydrogen bonds, given that a distance of approximately 2.9 Å is within typical hydrogen bond lengths [9].

The impact of the mutation on the S4 segment of VSD4 was analyzed, focusing on changes in helix alignment and orientation [10]. Potential distortions in transmembrane helices and their packing were investigated, which could affect the voltage-sensing mechanism of the channel [5].

2.5 Hydrogen Bond Analysis

Hydrogen bonds within the structures were identified using PyMOL, focusing on distances between 2.5 Å and 3.5 Å [8,9]. The hydrogen bond networks of the wild-type and mutant structures were compared to assess any disruptions or new interactions introduced by the mutation [10].

2.6 Functional Implication Prediction

To predict the functional impact of the R1648H mutation, we used the HOPE (Have (y)Our Protein Explained) module [11]. Both the wild-type and mutant sequences were input into HOPE to determine the effects on protein stability, hydrophobicity, and disruption of active sites. We cross-referenced our structural analysis against electrophysiological data from previous experiments, correlating these structural changes with observed alterations in sodium channel gating and conductance associated with the R1648H mutation [12, 13].

2.7 Software and Tools

The following software and tools were used in this study: UniProt (https://www.uniprot.org/) for protein sequence retrieval; SnapGene Viewer (version 5.3.2) for sequence editing; ColabFold (https://colabfold.com/) for structure prediction using AlphaFold2; PyMOL (version 2.5.5) for molecular visualization and structural analysis; and HOPE (https://www3.cmbi.umcn.nl/hope/) for mutation effect prediction.

2.8 Ethical Considerations

This study did not involve human participants or animal subjects. All data utilized were obtained from publicly accessible databases and tools. No ethical approval was required.

3. Results

3.1 Structural Analysis of the R1648H Mutation in SCN1A

Amino Acid Properties

The R1648H mutation involves the substitution of arginine (Arg, R) with histidine (His, H) at position 1648 in the SCN1A protein. Arginine is a large, positively charged amino acid essential for maintaining electrostatic interactions within the protein structure, particularly in voltage-sensing domains [9]. Histidine is smaller and can be positively charged or neutral depending on the local pH; at physiological pH, it is often neutral [13]. This substitution results in the loss of a positive charge and a decrease in side-chain size at this critical position, potentially affecting protein interactions and stability [11].

3.2 Localization and Structural Context: Voltage-Sensing Domain (VSD4)

Position 1648 is located within the S4 segment of the fourth voltage-sensing domain (VSD4) of the SCN1A protein [10]. The S4 segments are characterized by a series of positively charged residues, typically arginines or lysines at every third position [9]. These charges are crucial for the voltage-sensing mechanism of the sodium channel, allowing it to respond to changes in membrane potential by undergoing conformational shifts that open or close the channel (Figure1, Figure2) [12].

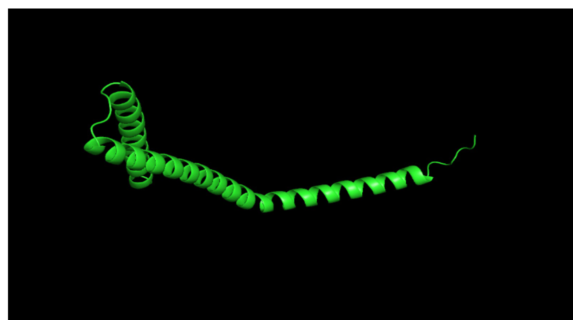

Figure 1. Rank1 model of Wildtype SCN1A residual 1600-1700 structure predicted by Alphafold.

ISSN 2959-409X

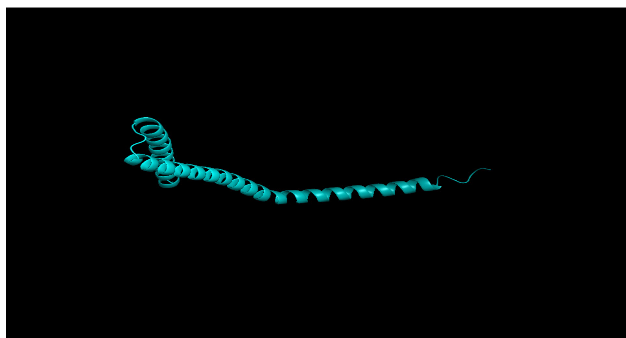

Figure 2. Rank1 model of Mutant R1648H residual 1600-1700 structure predicted by Alphafold.

3.3 Structural Alignment and RMSD Analysis

Figure 3. Aligned mutant and wildtype in PyMol.

To quantitatively assess the structural differences between the wild-type and R1648H mutant SCN1A proteins, structural alignments were performed using PyMOL [14]. The predicted structures for the selected region (residues 1600–1700) of both the wild-type and mutant proteins were superimposed to evaluate the extent of conformational changes induced by the mutation.

The alignment process began by reading the scoring matrix and assigning pairwise scores between residues, resulting in 101 × 101 assignments for the 101 residues in the selected region. All 101 residues of the wild-type protein were compared against those of the mutant, achieving an initial alignment score of 491.000. The 'align' function indicated that 810 atoms were aligned from both structures [14].

An iterative refinement was conducted to improve alignment accuracy. During the first cycle, 43 atoms were rejected due to significant deviations, resulting in a root mean square deviation (RMSD) of 1.86 Å. In the second cycle, 33 atoms were further rejected, lowering the RMSD to 1.01 Å. Subsequent cycles continued this refinement, with 10 atoms rejected in the third cycle (RMSD = 0.83 Å), 6 atoms in the fourth cycle (RMSD = 0.81 Å), and 2 atoms in the fifth and final cycle (RMSD = 0.80 Å). After the final cycle, the alignment yielded an RMSD of 0.795 Å across 716 atoms [14].

The decreasing number of rejected atoms and the lowering RMSD values across the cycles indicate that the majority of the protein structure aligns well between the wild-type and mutant forms [14]. An RMSD below 1 Å suggests that the global backbone conformation remains largely conserved despite the mutation. However, the initial higher RMSD and the atoms rejected in early cycles reflect localized structural deviations, particularly around the mutation site at residue 1648 [6].

These alignment results demonstrate that while the overall fold of the SCN1A protein is maintained, the R1648H mutation induces subtle but potentially significant conformational changes. The areas where atoms were rejected during alignment cycles likely correspond to regions affected by the mutation, including possible shifts in side-chain orientations and slight adjustments in secondary structure elements (Figure 3) [13].

3.4 Visualization and Assessment

3.4.1 Potential hydrogen bond formation

The substitution of arginine with histidine at position 1648 introduces the possibility of new hydrogen bond formations [15]. The measured distance of 2.9 Å between the side chain of histidine and neighboring atoms suggests that a hydrogen bond could form, as hydrogen bonds typically occur within distances of 2.5–3.5 Å [15, 16]. This new hydrogen bond may stabilize local interactions or alter existing ones, potentially affecting the conformation of the voltage-sensing domain (Figure 4, Figure 5) [13].

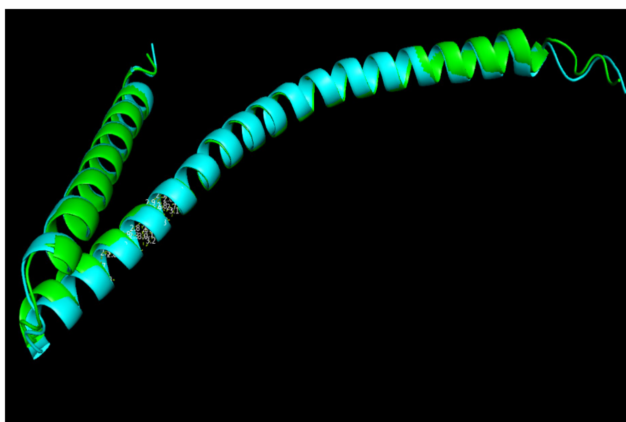

Figure 4. The measured distance of side chain of histidine and neighboring atoms at the mutation position.

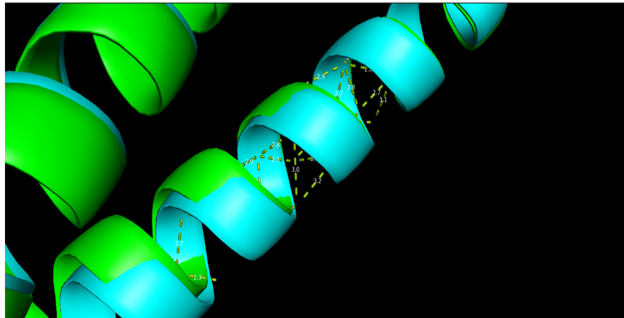

Figure 5. The zoom in view of the position of mutation.

3.4.2 Effect on structural stability

As shown in figure 6, histidine is smaller than arginine, which could lead to reduced van der Waals interactions with adjacent residues and create gaps within the protein structure [6]. The loss of a positive charge at position 1648 may disrupt electrostatic interactions critical for the stability and function of the voltage-sensing domain [3]. While the potential formation of a new hydrogen bond might partially compensate for the loss of previous interactions, it may not fully restore the original structural integrity [5].

Visualization of the mutant model revealed a significant overlap of van der Waals spheres between the histidine side chain at position 1648 and neighboring residues, indicating steric clashes [8,9]. These clashes suggest that the mutant residue is too close to adjacent atoms, causing steric hindrance. Such steric hindrance can lead to instability in the local protein conformation, disrupt the function of the voltage-sensing domain, and impede proper folding and maturation of the protein [6].

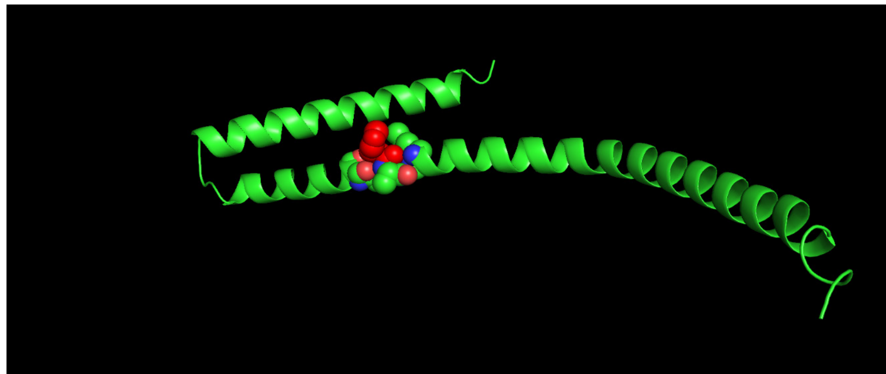

Figure 6. Mutant model with van der Waals spheres between the histidine side chain at position 1648 and nearby residues

To resolve these steric clashes, energy minimization using molecular modeling tools is recommended [11]. This process adjusts atomic positions to achieve a more energetically favorable conformation, potentially alleviating steric hindrance [13, 17].

3.5 Change in Chemical Properties

The loss of a positive charge at position 1648 may impair voltage sensing, as arginine contributes a positive charge essential for detecting changes in membrane potential [3]. Replacing it with histidine, which is neutral at physiological pH, reduces the net positive charge [6]. Additionally, the change in side-chain chemistry may affect interactions with surrounding residues, including hydrogen bonds and ionic interactions, disrupting the normal function of the voltage sensor [5].

3.6 Evaluation of the Voltage-Sensing Domain

3.6.1 Impact on helix alignment

The R1648H mutation may cause a shift or kink in the S4 helix due to changes in charge balance and steric environment. This alteration could affect the movement of the voltage sensor during membrane depolarization [8]. Furthermore, the mutation may impair the coupling between the voltage-sensing domain and the channel gating mechanism, leading to improper opening or closing of the channel [5].

3.6.2 Transmembrane helices

The mutation may induce slight distortions in the packing of transmembrane helices surrounding VSD4, affecting the overall architecture of the channel. Structural changes could alter the pathway for sodium ions, impacting ion selectivity and conductance. Changes in helix stability and orientation can disrupt the normal voltage-sensing mechanism, reducing channel sensitivity and affecting its role in action potential generation [5].

3.7 Overall Folding and Stability

3.7.1 RMSD analysis

The root mean square deviation (RMSD) between the

ISSN 2959-409X

mutant and wild-type structures for residues 1600–1700 is 1.903 Å, indicating moderate structural deviations due to the mutation. For the immediate region encompassing the mutation site (residues 1640–1650), the RMSD is 1.235 Å, suggesting minor local deviations and that the secondary structure remains largely conserved [8,9].

3.7.2 Hydrogen bond analysis

Hydrogen bond analysis of the selected region revealed a network of hydrogen bonds with distances ranging from 2.1 Å to 3.3 Å in the mutant structure. These distances fall within the typical range for stable hydrogen bonds, indicating that the stability of this region is largely preserved [6]. The presence and consistency of hydrogen bonds suggest that the mutation did not significantly disrupt the local hydrogen bond network, and the secondary structure elements remain intact [9]. The density of hydrogen bonds in the selected region is comparable to that in the wild-type, implying that the local stability and folding are maintained [15].

3.8 Conservation Analysis

Arginine at position 1648 is highly conserved across species, highlighting its critical role in channel function [5]. While some homologous proteins may have different residues at this position, the high conservation level implies that changes here could have significant functional consequences. The mutation is also located near highly conserved positions, which may amplify its impact on protein function [3].

3.9 MetaRNN Pathogenicity Score

The R1648H mutation has a MetaRNN score of 0.9407 on a scale from 0.0 to 1.0, where higher scores indicate a greater likelihood of pathogenicity [17]. This high score suggests that the R1648H mutation is likely deleterious and associated with disease [6].

3.9.1 Functional consequences

The loss of a positive charge reduces the effectiveness of the voltage-sensing mechanism, potentially leading to a failure in channel activation in response to membrane depolarization [3]. Structural changes may hinder the conformational transitions necessary for channel opening and closing, affecting sodium ion flow [6]. Disruption in sodium channel function can lead to increased neuronal firing, contributing to seizure activity observed in epilepsy [5].

3.9.2 Comparison with known mutations

The R1648H mutation is well-documented and associated with severe epilepsy syndromes, including Dravet syndrome and generalized epilepsy with febrile seizures plus (GEFS+) [2]. Previous electrophysiological studies have shown that R1648H results in altered channel kinetics, including impaired fast inactivation and persistent sodium currents, which contribute to neuronal hyperexcitability [6], [16].

3.9.3 Implications for protein function and epilepsy

The structural alterations predicted for the R1648H mutation provide a molecular basis for its pathogenicity [5,9]. Loss of charge and structural distortions impair the voltage-sensing ability, leading to dysfunctional gating [6]. The mutation exemplifies a channelopathy where altered ion channel function leads directly to disease [5]. Understanding the structural impact aids in developing targeted therapies that could stabilize the mutated channel or modulate its function [8,9].

4. Discussion

This study offers a computational model for the molecular basis of SCN1A-associated epilepsies by studying the structural impact of R1648H using protein structure prediction with AlphaFold [6, 7, 12]. Focusing on the central voltage-sensing domain (VSD4) of the SCN1A sodium channel, we provide a window into how structural changes could guide future therapies.

Importance of AlphaFold in Structural Prediction

A challenge for studies of membrane proteins such as SCN1A is obtaining high-resolution experimental structures because of their size, topology, and hydrophobicity [13]. AlphaFold has revolutionized structural biology by predicting protein structures using extremely high-confidence sequences of amino acids [6,7]. With AlphaFold, we were able to create stable structural models of the SCN1A protein VSD4 region that would have been time-consuming and difficult to obtain using traditional experimental approaches.

By inferring the 3-D structures of both wild-type and mutant R1648H forms, AlphaFold allowed us to conduct a comparative analysis of the mutation's structural consequences. The modeling of specific segments of the protein (such as residues 1600–1700 including the mutation site) was a convenient way to overcome the computational difficulties of such a large protein.

AlphaFold made it possible not only to visualize potential steric clashes and changes in the pattern of hydrogen bonding but also to evaluate fluctuations in electrostatic interactions crucial for protein function. That structural detail was vital for predicting the effect of the R1648H mutation, which interferes with the voltage-sensing system and leads to impaired channel gating and greater neuronal excitability [13].

Implications for Therapeutic Development

Understanding the structural consequences of the R1648H mutation provides valuable insights that can inform future therapeutic strategies. By identifying specific alterations in the voltage-sensing domain, targeted interventions can be designed to mitigate the functional deficits caused by the mutation [5].

A potential therapeutic strategy is to create small molecules or peptides that restore the positive charge or reposition VSD4's altered conformation. These compounds might improve the mutant channel's voltage-sensing capabilities and enhance its ability to sense changes in membrane potentials [15]. Furthermore, inhibitors that modulate the gating behavior of sodium channels can be targeted precisely to the mutant SCN1A version, reducing neuronal hyperexcitability [16].

Advantages of Computational Approaches

This article demonstrates the utility of computational methods—especially protein structure prediction with AlphaFold—in researching the molecular foundation of genetic diseases [6,12]. Computational analyses are accessible and efficient because they can be performed in a short time without special laboratory equipment, making them available to many researchers [13]. They are useful for obtaining structural information, which may be difficult to experimentally get, particularly for larger and more complex membrane proteins such as SCN1A [13]. Computational predictions can inform experimental setups by determining regions of interest and allocating resources to future laboratory experiments.

In this regard, AlphaFold was a breakthrough that helped us overcome the challenges of using structural biology as an overall method to understand the molecular effects of the mutation [6], [12].

Limitations and Future Directions

Computational models are useful for many things, but there are limitations. The calculations are based on static geometry and may not reflect the dynamics of proteins in a cellular environment [16]. It is crucial to establish through experiments such as electrophysiology, cryo-electron microscopy, or nuclear magnetic resonance spectroscopy that the mutation has a functional effect [13].

Future work will need to be both computational and experimental. Molecular dynamics simulations would allow more insight into how the mutation might affect the protein over time [17]. Research on other mutations in SCN1A might also shed light on broader structure-function connections in voltage-gated sodium channels [3, 5].

5. Conclusion

Overall, this paper adds new insights to the molecular mechanism of SCN1A-associated epilepsies by defining the structural role of the R1648H mutation through AlphaFold. In understanding how specific structural modifications in the voltage-sensing domain degrade channel activity, we lay the basis for designing targeted therapies. This success illustrates why AlphaFold will be critical for future biomedical studies in genetic disease contexts where structural information is of primary importance to design effective interventions.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] World Health Organization (WHO). Epilepsy. 2019. Retrieved from https://www.who.int/news-room/fact-sheets/detail/epilepsy

[2] Claes L., et al. De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. American Journal of Human Genetics, 2001, 68(6): 1327–1332.

[3] Escayg A., Goldin A. L. Sodium channel SCN1A and epilepsy: Mutations and mechanisms. Epilepsia, 2010, 51(9): 1650–1658.

[4] Dravet C. The core Dravet syndrome phenotype. Epilepsia, 2011, 52(Suppl 2): 3–9.

[5] Meng H., et al. Role of mutations in SCN1A in severe epilepsy syndromes: Dravet syndrome. Journal of Human Genetics, 2015, 60(7): 373–378.

[6] Lossin C., et al. Molecular basis of an inherited epilepsy. Neuron, 2002, 34(6): 877–884.

[7] Jumper J., et al. Highly accurate protein structure prediction with AlphaFold. Nature, 2021, 596(7873): 583–589.

[8] DeLano W. L. The PyMOL Molecular Graphics System. 2002.

[9] Catterall W. A. Sodium channels, inherited epilepsy, and antiepileptic drugs. Annual Review of Pharmacology and Toxicology, 2014, 54: 317–338.

[10] Spampanato J., et al. Functional effects of two voltage-gated sodium channel mutations that cause generalized epilepsy with febrile seizures plus type 2. Journal of Neuroscience, 2001, 21(19): 7481–7490.

[11] Yu F. H., et al. Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. Nature Neuroscience, 2006, 9(9): 1142–1149.

[12] Zuberi S. M., et al. Genotype-phenotype associations in SCN1A-related epilepsies. Neurology, 2011, 76(7): 594–600.

[13] Kuhlman B., Bradley P. Advances in protein structure prediction and design. Nature Reviews Molecular Cell Biology, 2019, 20(11): 681–697.

[14] Tunyasuvunakool K., et al. Highly accurate protein structure prediction for the human proteome. Nature, 2021, 596(7873):

ISSN 2959-409X

590–596.
[15] Callaway E. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. Nature, 2020, 588(7837): 203–204.
[16] UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research, 2021, 49(D1): D480–D489.
[17] Venselaar H., et al. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 2010, 11: 548.