Molecular complexity of voltage-gated sodium channels

Citation for published version (APA):

Xenakis, M. (2021). *Molecular complexity of voltage-gated sodium channels: theory and applications in mutation-response prediction*. [Doctoral Thesis, Maastricht University]. ProefschriftMaken. <https://doi.org/10.26481/dis.20210526mx>

Document status and date:

Published: 01/01/2021

DOI:

[10.26481/dis.20210526mx](https://doi.org/10.26481/dis.20210526mx)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Molecular complexity of voltage-gated sodium channels; theory and applications in mutation-response prediction

Markos Xenakis

Molecular complexity of voltage-gated sodium channels; theory and applications in mutation-response prediction

a Dissertation conducted at the intersection of
computational biology, medicine, and physics.

Authored by Markos Xenakis

Methods and experiments curated in this Dissertation were designed and performed *in silico*. Research was carried out within the framework of PROPANE (Probing the Role of Sodium Channels in Painful Neuropathies) study. Largest research part was performed at Faculty of Health, Medicine and Life Sciences of Maastricht University, within the MH&NS (Mental Health and Neuroscience) School, the Clinical Genomics Section of the Department of Toxicogenomics, and the Department of Data Science and Knowledge Engineering. Minor research part was carried out at the School of Medicine of Yale University, within the Department of Neurology, and Center for Neuroscience and Regeneration Research.

© Markos Xenakis, 2021

All rights reserved

ISBN: 978-94-6423-260-8

Cover design: ProefschriftMaken || www.proefschriftmaken.nl

Cover photo: Meri Linna

Lay-out: Custom-made L^AT_EX design by Markos Xenakis

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Published by: ProefschriftMaken || www.proefschriftmaken.nl

Molecular complexity of voltage-gated sodium channels; theory and applications in mutation-response prediction

Dissertation

to obtain the degree of Doctor at Maastricht University, on the authority
of the Rector Magnificus Prof. Dr. Rianne M. Letschert in accordance
with the decision of the Board of Deans, to be defended in public on
Wednesday, the 26th of May 2021, at 04:00 PM

by

Markos Xenakis
born in Viersen NRW, DE, the 5th of May 1986

Promotor

Prof. Dr. Hubert J.M. Smeets

Co-promotors

Dr. Ronald L. Westra

Dr. Lindsey J. Patrick

Assessment committee

Prof. Dr. Peter J. Peters (chair)

Prof. Dr. Angelika Lampert (Aachen University, DE)

Prof. Dr. Yoram Rudy (Washington University in St. Louis, USA)

Prof. Dr. David E. Linden

Dr. Pieter J. Collins

.. to my dear friend, and father, Nikos

Contents

1	Introduction & Dissertation Outline	1
2	Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution	17
3	Non-extensivity and criticality of atomic hydropathicity around a voltage-gated sodium channel's pore; a modeling study	59
4	Hydropathicity-based prediction of pain-causing NaV1.7 variants	95
5	General Discussion	149
	Appendices	163
	Dissertation Summary	169
	Impact Paragraph	177
	Acknowledgments	179
	Curriculum Vitae	181

Chapter 1

Introduction & Dissertation Outline

Voltage-gated sodium channels: a brief outlook

Structure and function

Voltage-gated sodium channels (NaVChs) are pore-forming proteins embedded in cell membranes. They belong to the ion channels superfamily and their main biological function is to facilitate transport of sodium ions across the membrane. To achieve that, NaVChs have evolved into highly-efficient molecular sieves catalyzing transport of sodium ions with high specificity. This sets the molecular basis for cell excitation and transmission of electrical signals along neuronal tissue.

NaVChs are expressed in a great variety of excitable cells including axons, neuron cell bodies, vertebrate skeletal and cardiac muscles, and endocrine glands [1]. NaVChs are also present in nonexcitable cells where their physiological role remains unclear [2]. The first high-resolution crystal structure of a NaVCh was published in 2011 capturing a pre-open conformation state of the bacterial *Arcobacter butzleri* (NaVAb) channel at 2.7 Å resolution [3]. Due to its place in evolutionary lineage, the homomeric

NaVAb channel represents a prototype structure providing structural biologists with a starting point for understanding larger and more complex NaVCh structures such as those of mammalian species. Although nine NaVChs are expressed in mammalian species (NaV1.1-NaV1.9) their crystallographies remain still unsolved. Nevertheless, a structural model of the transmembrane helices of the NaV1.7 channel was recently constructed via homology modeling procedures based on the NaVAb template [4].

Figure 1: **Structure of a NaVCh.** (a), Linear schematic of the α -subunit of the human NaV1.7 channel consisting of four homologous domains (DI-DIV). Each domain comprises the S1-S4 and S5-S6 helices forming the voltage-sensors and the pore module, respectively. Domain interconnection is achieved via three loops (L1-L3). (b), Cartoon illustration of the atomic structure of a NaV1.7 channel (view from the intracellular side). Atomic structure is constructed according to computational procedures described in [4].

NaVCh functional architectures typically emerge from assembly of four homologous domains (DI-DIV). Each of the four homologous domains comprises six transmembrane helices (S1-S6) with the S1-S4 helices forming the voltage sensor (VS), and with the S5-S6 helices forming the pore module (PM) (see Figure 1). This conserved architectural motif allows for efficient "sensing" of changes in ionic concentrations across the membrane. If the concentration of negative charges outside the cell becomes large enough,

the outward movement of positive charges in the S4 segments induces a sliding-helix [5] effect that pulls the S4-S5 linker and, eventually, forces the pore to open. Transmission of gate-open signals crucially depends on allosteric communication between the PM and the VSs established via their electromechanical coupling. The open pore can facilitate the flow of sodium ions at a rate nearing free diffusion [3].

NaVCh functional architectures comprise tens of hundreds of interacting amino acids capable of generating a rich dynamics content across different temporal and spatial scales. The coarse-grained dynamics picture for a NaVCh is that of a dynamical system visiting more-or-less the same locations of its phase space. These locations, that can be understood as meta-stable attractors coexisting within NaVCh's phase space [6], are structurally attributed to distinct NaVCh conformational states. At least two conformational states are necessary in order to achieve on-off control of ionic flow, namely, a closed (i.e., non-conducting) state and an open (i.e., conducting) state. At fine-grained temporal scales, molecular dynamics simulations have shed light upon ion transport mechanisms through the pore but also gating currents generated by closed-to-open conformational changes [7]. Zooming-in the pore's nanoenvironment revealed that gating phenomena can occur also at a nanometer scale; a switching of the pore from an open to a closed state occurs due to hydrophobic interactions inducing stochastic liquid-vapor transitions, a phenomenon termed hydrophobic gating [8,9].

Is it possible to investigate spatiotemporal dynamics generated by NaVChs biological entities in a unified manner? And, if yes, what kind of phenomenological approaches are needed? In order to answer this question, we adopt throughout this Dissertation a complex system's viewpoint upon the problem of decoding the structure of a NaVCh. Accordingly, the notion of molecular complexity is introduced and used as an umbrella term referring to macroscopic structural features emerging from microscopic interactions underlying a NaVCh's conformational state. In complex system's glossary, macroscopic structural features are termed "emergent properties" [10] and describe how individual, spatially-distributed system components co-aggregate into functional architectures.

Health and disease

Defects in NaVCh genes can have profound implications in electrophysiological activity leading to inherited human disorders such as neuropathic pain syndromes, cardiac arrhythmias and epilepsy (for a review see [11]). Throughout this Dissertation, we focus on neuropathic pain which has been established as one of the most well-studied clinical, as well as, biophysical channelopathy cases. The NaV1.7 channel plays a central role in pain research as it is known to act as a crucial regulator of pain signals [12]. In humans, the NaV1.7 channel is encoded by the *SCN9A* gene and is preferentially expressed in peripheral sensory neurons such as nociceptive dorsal root ganglion (DRG) neurons. Several functional studies have investigated *in vitro* how missense *SCN9A* mutations affect cellular electrophysiological output based on expression of mutant NaV1.7 channel in DRG neurons. Findings of these studies have established a mechanistic link between defects in the *SCN9A* gene and pathophysiology of pain; gain-of-function *SCN9A* mutations can disrupt physiological firing of affected DRGs by depolarizing their membrane potential thus triggering abnormal pain sensation [12]. The clinically-observed spectrum of inherited pain disorders is broad including diseases such as inherited erythromelalgia (IEM) [13–21], paroxysmal extreme pain disorder (PEPD) [22–25] and small fiber neuropathy (SFN) [26, 27].

IEM is clinically characterized by a burning pain sensation usually localized at feet and hands. Electrophysiology experiments have shown that IEM-related mutations impair NaV1.7 activation by inducing a hyperpolarizing shift rendering channel opening more likely. This is realized via enhancing channel response to slow and small depolarizations or by slowing down closing of the channel after cessation of depolarizing stimuli [12]. Expression of IEM NaV1.7 mutants in DRG neurons renders them hyperexcitable leading to abnormal pain sensation.

Clinical characteristics of PEPD are associated with severe perirectal pain typically starting in infancy. Molecular mechanisms underlying PEPD differ from IEM as PEPD-related mutations impede channel's fast inactivation so that the number of channels available for activation increases which,

in turn, leads to hyperexcitability of affected neurons [22].

SFN can damage small somatic nerve fibers thus triggering burning pain sensation and tingling typically felt at patient's limbs. Functional analysis of SFN-related mutations revealed that they can lead to hyperexcitability of affected DRG neurons via a variety of mechanistic paths such as impaired slow inactivation, depolarized slow and fast inactivation, and enhanced resurgent currents [26].

Although the NaV1.7 has emerged as a potential molecular target for pain treatment, drug therapies are still at a trial-and-error stage. In particular, lidocaine and carbamazepine pharmacological interventions in patients carrying specific *SCN9A* mutations were shown to have a pain relieve effect suggesting that some NaV1.7 mutants are well-responsive to these compounds [12]. Specifically, patients with the PEPD-related I1461T and T1464I mutations demonstrated favorable response to carbamazepine treatment which is believed to counteract fast inactivation impairment of NaV1.7 mutants [22]. However, patients with IEM-related mutations are generally not responsive to pharmacological treatment [12]. Nevertheless, a patient carrying the IEM-related V872G mutation was responsive to lidocaine or mexiletine treatment that led to a substantial decrease in patient's pain symptoms [28]. Moreover, control of pain symptoms was achieved by treating three family members carrying the IEM-related V400M mutation with carbamazepine [29].

Translational computational modeling

Computational pipelines for modeling molecular mechanisms underlying channelopathies are still at their infancy. Despite the long tradition of phenomenological models attempting to explain ion channel gating (for a review see [30]), algorithmic pipelines that can bridge the gap between channel's functional architecture, mutation response and cellular electrophysiology toward applications with clinical translational value are not available. The main reason for this is that ion channels translational research is genuinely multidisciplinary and inevitably addressing molecular complexity challenges thus demanding significant amount of resources.

A key-property for understanding structural stability and mutation response of proteins is hydrophobicity [31] (from now on, we will use the term *hydropathicity* in order to indicate that a molecule can be either hydrophobic or hydrophilic). A major obstacle toward accurate modeling of hydropathic interactions in protein structures is the spatial inhomogeneity of underlying interaction fields. In particular, a hydropathic field [32] provides a summary of fundamental molecular interactions (such as coulomb and van der Waals interactions) that cannot be reduced into a single force-distance equation due to non-scalability of the hydropathic-interaction laws [33]. Along an ion channel’s pore, the interplay of hydrophobic-vs-hydrophilic components not only shapes microscopic pore gating dynamics, e.g., via formation of hydrophilic selectivity filters [3] or hydrophobic gates [34], but also contributes to pore’s stability via peripheral expansion of hydrogen bonds networks [3, 35]. The role of hydropathic interactions is also dominant in the VSs where local mutations of hydrophobic residues can induce potentially-large destabilizing effects [36].

Atomic-level computational modeling in combination with functional channel characterization can elucidate connections between hydropathic NaV1.7 characteristics and pain disease. Guidelines for targeted pharmacological interventions, as well as, a phenomenological basis for investigating disease-causing molecular mechanisms triggered by mutations can then be introduced. Toward this direction, authors of [4] showed via thermodynamic analysis that the IEM-related V400M and S241T mutations, located only 2.4 Å apart within the NaV1.7 structure, are energetically coupled thus both exhibiting pharmacoresponsiveness to carbamazepine. Moreover, the authors of [37, 38] proposed that a hydrophobic-ring disruption at the putative activation gate (AG) of NaV1.7 can act as a disease-causing mechanism leading to IEM-related symptoms. Interestingly, a long-range, mutation-induced perturbation effect was also considered as a plausible scenario for explaining how AG destabilizations can affect distant channel components [39]. Taken together, these observations suggest that mutations might have a non-local destabilizing effect on the NaVCh functional architecture which is initiated by local intrachannel disruptions of hydropathic characteristics.

Hydropathic dipole moments [32] provide with a simple, but yet efficient, tool for modeling hydropathic interactions at play. Utilization of cumulative hydrophobic moments allows for investigation of slowly-varying, hydropathic-field characteristics with respect to function-relevant intraprotein locations [40,41]. Importantly, hydropathic characteristics of the transition from a globular protein's interior to its exterior were shown to exhibit universality as a consequence of the formation of a hydrophobic core in a protein's interior [42]. When we focus on pore-forming proteins such as ion channels, hydropathic-field spatial profiling can be performed with respect to the pore axis corresponding to the reaction coordinate [43], i.e., ion conduction pathway, that, in the case of homomeric architectures, accounts also for structural symmetries.

Functional validation of a missense *SCN9A*-gene mutation via electrophysiology is laborious and time-consuming thus making it practically impossible to routinely investigate every candidate variant's pathophysiological profile. *In silico* algorithms for variants classification aim at providing predictions for a variant's pathophysiological profile by utilizing information about the location, frequency and nature of the mutation within the NaVCh structure, as well as, about the degree of conservation of the substituted or deleted amino acid. Current classification guidelines for NaVCh variants of unknown pathophysiological profile are based on [44] and aim at placing a variant into one of the following categories; pathogenic, probably pathogenic, possible pathogenic, variants with uncertain clinical significance, and non-pathogenic variants [44,45]. Despite the plethora of available bioinformatic tools for modeling protein response to mutation-induced perturbations (e.g., see [46]), predicting *in silico* the clinical phenotype of a NaVCh mutation can only be done with moderate success [45]. This is mainly because 3D structures of NaVChs were made available only recently to the scientific community (e.g., the crystal structure of the NaVA β was published in 2011) but also because of scarce genetic information regarding homologues to NaVCh-encoding genes.

Dissertation outline

The aim of this Dissertation is threefold;

(1), to implement novel computational tools for cumulative hydrophobic analysis of NaVCh structures.

(2), to demonstrate how cumulative hydrophobic analysis provides with an opportunity to study molecular complexity of NaVCh structures.

(3), to utilize molecular complexity metrics in order to predict the pathophysiological profile of a NaV1.7 variant.

In order to accomplish this aim, this Dissertation builds systematically upon previous concepts and observations regarding cumulative hydrophobic analysis and incorporates them into a complexity theory for atomic-level modeling of hydrophobic interactions underlying a NaVCh's functional architecture. As a first step, we develop a computational tool for mapping cumulative hydrophobic NaVCh pore characteristics and translating them into stability- and function-relevant topological descriptors. This leads to the novel concept of cumulative hydrophobic topology (CHT) presented in **Chapter 2**. Methodological basis of CHT is a spherical sampling procedure that allows for studying cumulative hydrophobic characteristics of the atomic environment around a pore point (see Figure 2). Atomic-level resolution is achieved by carefully choosing a binary hydrophobic atomic scale so that each atom is assigned with a hydrophobic value.

Next, we focus on how to identify critical points [47, 48] of a NaVCh structure and embed them into a framework for analyzing the functional architecture. Stated differently, we test the self-organized criticality (SOC) [49] hypothesis for the cumulative hydrophobic environment around the pore of a NaVCh. The theoretical vehicle underlying this approach is scale-invariance, i.e., power-law behavior [50], as it can establish a connection between packing of atoms around the pore, hydrophobic field topological configurations and conservation. Accordingly, a mean-field spherical hydrophobic model is constructed that cumulatively incorporates the atomic environment around the pore. This provides with computational evidence for the validity of the SOC hypothesis in NaVCh structures which is presented in **Chapter 3**. In order to demonstrate how cumulative hydrophobic

analysis methodologies of **Chapters 2** and **3** can open a window to the realm of a NaVCh’s molecular complexity, computations are performed upon the pre-open NaVAb channel structure [3].

Figure 2: **Cumulative hydropathic analysis of a NaVCh.** Illustration of the structure of the human NaV1.7 channel (view from the intracellular side). Each atom is visualized as a colored sphere (domain colors are the same as in Figure 1). The atomic environment around a pore point is step-wise sampled by increasing spatial scale, i.e., sphere radius. The smallest sphere corresponds to a microscopic spatial scale incorporating pore-lining atomic components. The largest sphere corresponds to a macroscopic spatial scale incorporating the NaV1.7 structure as a whole. At each sampling step hydropathic density and dipole-field characteristics of the atomic environment are calculated in accordance to [40, 41, 43].

In **Chapter 4** we employ the NaV1.7 structural model from [4] and scrutinize its atomic porous environment by applying methods of **Chapters 2** and **3**. Our goal is to assess whether structural locations of pain-related missense *SCN9A*-gene mutations occupy intra-channel regions that exhibit higher sensitivity to structural perturbations. We demonstrate that pain-related mutation sites are found in proximity to a hydropathic patch lining NaV1.7’s pore while maintaining a critical hydropathic-interactions

distance from NaV1.7's selectivity filter. This observation is used as a basis for predicting the pathophysiological profile of a NaV1.7 variant by classifying it as either pathogenic or non-pathogenic.

References

- [1] Hille, B. *Ion Channels of Excitable Membranes*, 3rd ed. Sinauer Associates Inc., Sunderland, MA (2001).
- [2] Black, J.A., Waxman, S.G. Noncanonical roles of voltage-gated sodium channel. *Neuron* **80**, 280-91 (2013).
- [3] Payandeh, J., Scheuer, T., Zheng, N., Catterall, W.A. The crystal structure of a voltage-gated sodium channel. *Nat* **475**, 353-358 (2011).
- [4] Yang, Y., Dib-Hajj, S.D., Zhang, J. et al. Structural modelling and mutant cycle analysis predict pharmacoresponsiveness of a Na(V)1.7 mutant channel. *Nat Commun* **3**, 1186 (2012).
- [5] Catterall, W.A. Ion channel voltage sensors: structure, function, and pathophysiology. *Neuron* **67**, 915-28 (2010).
- [6] Liebovitch, L.S., Toth, T.I. A model of ion channel kinetics using deterministic chaotic rather than stochastic processes. *J Theor Biol* **148**, 243-267 (1991).
- [7] Amaral, C., Carnevale, V., Klein, M.L., Treptow, W. Exploring conformational states of the bacterial voltage-gated sodium channel NavAb via molecular dynamics simulations. *Proc Natl Acad Sci USA* **109**, 21336-21341 (2012).
- [8] Beckstein, O., Sansom, M.S.P. Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc Natl Acad Sci USA* **100**, 7063-7068 (2003).
- [9] Beckstein, O., Sansom, M.S.P. The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Phys Biol* **1**, 42 (2004).
- [10] Damper, R.I. Editorial for the Special Issue on 'Emergent Properties of Complex Systems': Emergence and levels of abstraction. *Int J Syst Sci* **31**, 811-818 (2000).

- [11] Catterall, W.A. Voltage-gated sodium channels at 60: structure, function and pathophysiology. *J Physiol* **590**, 2577-2589 (2012).
- [12] Dib-Hajj, S.D., Yang, Y., Black, J.A., Waxman, S.G. The Nav1.7 sodium channel: from molecule to man. *Nat Rev Neurosci* **14**, 49-62 (2013).
- [13] Black, J.A., Liu, S., Tanaka, M., Cummins, T.R., Waxman, S.G. Changes in the expression of tetrodotoxin-sensitive sodium channels within dorsal root ganglia neurons in inflammatory pain. *Pain* **108**, 237-247 (2004).
- [14] Han, C., Rush, A.M., Dib-Hajj, S.D. et al. Sporadic onset of erythromalgia: a gain-of-function mutation in Nav1.7. *Ann Neurol* **59**, 553-558 (2006).
- [15] Harty, T.P., Dib-Hajj, S.D., Tyrrell, L., Blackman, R., Hisama, F.M., Rose, J.B., Waxman, S.G. Nav1.7 mutant A863P in erythromalgia: effects of altered activation and steady-state inactivation on excitability of nociceptive dorsal root ganglion neurons. *J Neurosci* **26**, 12566-12575 (2006).
- [16] Lampert, A., Dib-Hajj, S.D., Eastman, E.M., Tyrrell, L., Lin, Z., Yang, Y., Waxman, S.G. Erythromalgia mutation L823R shifts activation and inactivation of threshold sodium channel Nav1.7 to hyperpolarized potentials. *Biochem Biophys Res Commun* **390**, 319-324 (2006).
- [17] Yang, Y., Wang, Y., Li, S. et al. Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythromalgia. *J Med Genet* **41**, 171-174 (2004).
- [18] Dib-Hajj, S.D., Rush, A.M., Cummins, T.R. et al. Gain-of-function mutation in Nav1.7 in familial erythromalgia induces bursting of sensory neurons. *Brain* **128**, 1847-1854 (2005).
- [19] Drenth, J.P., te Morsche, R.H.M., Guillet, G., Taieb, A., Lee Kirby, R., Jansen, J.B.M. SCN9A mutations define primary erythromalgia as a neuropathic disorder of voltage gated sodium channels. *J Invest Dermatol* **124**, 1333-1338 (2005).
- [20] Lee, M.J., Yu, H.S., Hsieh, S.T. et al. Characterization of a familial case with primary erythromalgia from Taiwan. *J Neurol* **254**, 210-214 (2007).
- [21] Drenth, J.P., Waxman, S.G. Mutations in sodium-channel gene SCN9a cause a spectrum of human genetic pain disorders. *J Clin Invest* **117**, 3603-3609 (2007).

- [22] Fertleman, C.R., Baker, M.D., Parker, K.A. et al. SCN9A mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52**, 767-774 (2006).
- [23] Jarecki, B.W., Sheets, P.L., Jackson, J.O. 2nd, Cummins, T.R. Paroxysmal extreme pain disorder mutations within the D3/S4-S5 linker of Nav1.7 cause moderate destabilization of fast inactivation. *J Physiol* **586**, 4137-4153 (2008).
- [24] Dib-Hajj, S.D., Estacion, M., Jarecki, B.W. et al. Paroxysmal extreme pain disorder M1627K mutation in human Nav1.7 renders DRG neurons hyperexcitable. *Mol Pain* **4**, 37 (2008).
- [25] Theile, J.W., Jarecki, B.W., Piekarz, A.D., Cummins, T.R. Nav1.7 mutations associated with paroxysmal extreme pain disorder, but not erythromelalgia, enhance Nav β 4 peptide-mediated resurgent sodium currents. *J Physiol* **589**, 597-608 (2011).
- [26] Faber, C.G., Hoeijmakers, J.G., Ahn, H.S. et al. Gain of function Nav1.7 mutations in idiopathic small fiber neuropathy. *Ann Neurol* **71**, 26-39 (2012).
- [27] Hoeijmakers, J.G., Han, C., Merkies, I.S. et al. Small nerve fibres, small hands and small feet: a new syndrome of pain, dysautonomia and acromesomelia in a kindred with a novel NaV1.7 mutation. *Brain* **135**, 345-358 (2012).
- [28] Choi, J.S., Zhang, L., Dib-Hajj, S.D. et al. Mexiletine-responsive erythromelalgia due to a new NaV1.7 mutation showing use-dependent current fall-off. *Exp Neurol* **216**, 383-389 (2009).
- [29] Fischer, T.Z., Gilmore, E.S., Estacion, M., et al. A novel NaV1.7 mutation producing carbamazepine-responsive erythromelalgia. *Ann Neurol* **65**, 733-741 (2009).
- [30] Sigg, D. Modeling ion channels: Past, present, and future. *J Gen Physiol* **144**, 7-26 (2014).
- [31] Chiti, F., Stefani, M., Taddei, N. et al. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nat* **424**, 805-808 (2003).
- [32] Eisenberg, D., Weiss, R.M., Terwilliger, T.C., Wilcox, W. Hydrophobic moments and protein structure. *Faraday Symp Chem Soc* **17**, 109-120 (1982).
- [33] Hammer, M.U., Anderson, T.H., Chaimovich, A., Shell, M.S., Israelachvili, J. The search for the hydrophobic force law. *Faraday Discuss* **146**, 299-308 (2010).

- [34] Kitaguchi, T., Sukhareva, M., Swartz, K.J. Stabilizing the closed S6 gate in the Shaker K_v channel through modification of a hydrophobic seal. *J Gen Physiol* **124**, 319-332 (2004).
- [35] Yonkunas, M., Kurnikova, M. The hydrophobic effect contributes to the closed state of a simplified ion channel through a conserved hydrophobic patch at the pore-helix crossing. *Front Pharmacol* **6**, 284 (2015).
- [36] Bendahhou, S., O'Reilly, A.O., Duclohier, H. Role of hydrophobic residues in the voltage sensors of the voltage-gated sodium channel. *Biochim Biophys Acta* **1768**, 1440-1447 (2007).
- [37] Lampert, A., O'Reilly, A.O., Dib-Hajj, S.D., Tyrrell, L., Wallace, B.A., Waxman, S.G. A Pore-blocking Hydrophobic Motif at the Cytoplasmic Aperture of the Closed-state Nav1.7 Channel Is Disrupted by the Erythromelalgia-associated F1449V Mutation. *J Biol Chem* **283**, 24118-24127 (2008).
- [38] Yang, Y., Estacion, M., Dib-Hajj, S.D., Waxman, S.G. Molecular architecture of a sodium channel S6 Helix: radial tuning of the voltage-gated sodium channel 1.7 activation gate. *J Biol Chem* **288**, 13741-13747 (2013).
- [39] Gurkiewicz, M., Korngreen, A., Waxman, S.G., Lampert, A. Kinetic modeling of Nav1.7 provides insight into erythromelalgia-associated F1449V mutation. *J Neurophysiol* **105**, 1546-1557 (2011).
- [40] Meirovitch, H., Scheraga, H.A. Empirical studies of hydrophobicity. 3. Radial distribution of clusters of hydrophobic and hydrophilic amino acids. *Macromolecules* **14**, 340-345 (1981).
- [41] Silverman, B.D. Hydrophobic moments of protein structures: Spatially profiling the distribution. *Proc Natl Acad Sci USA* **98**, 4996-5001 (2001).
- [42] Zhou R., Silverman, B.D., Royyuru, A.K., Athma, P. Spatial profiling of protein hydrophobicity: Native vs. decoy structures. *Proteins* **52**, 561-572 (2003).
- [43] Silverman, B.D. Hydrophobicity of transmembrane proteins: Spatially profiling the distribution. *Protein Sci* **12**, 586-599 (2003).
- [44] Wallis, Y., Payne, S., McAnulty, C. et al. Practice guidelines for the evaluation of pathogenicity and reporting of sequence variants in clinical molecular genetics. *ACGS* (2013). http://www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf (accessed May 25, 2013).

- [45] Waxman, S.G., Merkies, I.S.J., Gerrits, M.M. et al. Sodium channel genes in pain-related disorders: phenotype-genotype associations and recommendations for clinical use. *Lancet Neurol* **13**, 1152-1160 (2014).
 - [46] Feyfant, E., Sali, A., Fiser, A. Modeling mutations in protein structures. *Protein Sci* **16**, 2030-2041 (2007).
 - [47] Phillips, J.C. Scaling and self-organized criticality in proteins I. *Proc Natl Acad Sci USA* **106**, 3107-3112 (2009).
 - [48] Phillips, J.C. Scaling and self-organized criticality in proteins II. *Proc Natl Acad Sci USA* **106**, 3113-3118 (2009).
 - [49] Bak, P., Tang, C., Wiesenfeld, K. Self-organized criticality: an explanation of 1/f noise. *Phys Rev Lett* **59**, 381-384 (1987).
 - [50] Marković, D., Gros, C. Power laws and self-organized criticality in theory and nature. *Phys Rep* **536**, 41-74 (2014).
-

Chapter 2

Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution

Markos N. Xenakis, Dimos Kapetis, Yang Yang, Jordi Heijman, Stephen G. Waxman, Giuseppe Lauria, Catharina G. Faber, Hubert J. Smeets, Ronald L. Westra, Patrick J. Lindsey

Proteins, Volume 88, Issue 10, pages 1319-1328, October 2020

Abstract

Voltage-gated sodium channels (NaVChs) are biological pores that control the flow of sodium ions through the cell membrane. In humans, mutations in genes encoding NaVChs can disrupt physiological cellular activity thus leading to a wide spectrum of diseases. Here, we present a topological connection between the functional architecture of a NaVA bacterial channel and accumulation of atomic hydropathicity around its pore. This connection is established via a scaling analysis methodology that elucidates how intra-channel hydropathic density variations translate into hydropathic dipole field configurations along the pore. Our findings suggest the existence of a non-random cumulative hydropathic topology that is organized parallel to the membrane surface so that pore's stability, as well as, gating behavior are guaranteed. Given the biophysical significance of the hydropathic effect, our study seeks to provide a computational framework for studying cumulative hydropathic topological properties of NaVChs and pore-forming proteins in general.

Introduction

Voltage-gated sodium channels (NaVChs) are fundamental components of electrically excitable cells such as neurons and muscle cells [1]. They belong to the superfamily of ion channels and their primary function is to facilitate the transport of sodium ions across a cell's membrane in response to changes in the membrane potential. Genetic dysfunction of NaVChs has profound implications for physiological neuronal or muscle activity, leading to inherited human disorders [2]. For example, genetic and functional studies have shown that missense mutations in the *textitSCN9A* gene encoding the Nav1.7 channel are causally related to neuropathic pain syndromes such as inherited erythromelalgia [3–12], paroxysmal extreme pain disorder [13–16] and small fibre neuropathy [17, 18]. Although molecular mechanisms underlying missense *textitSCN9A* mutations remain largely unexplored, structural modeling of the Nav1.7 revealed that disrupting hy-

dropathic lining properties of the Nav1.7's pore can act as a disease-causing molecular trigger [19, 20].

The crystal structure of the pre-open I217C NaVAb bacterial channel (PDB code: 3RVY) [21] provides a prototype for understanding the functional architecture of the NaVCh family. The NaVAb channel consists of four homologous subunits, each of which comprises a pore domain (PD) interlinked with a voltage-sensing domain (VSD). This conserved architectural motif results in an ion-conducting central pore lined by hydropathically-diverse and, hence, functionally-distinct structural channel components (see Figure 2d in [21]). Hydropathic effects not only contribute to the stability of an ion channel's native structure [22, 23] but also regulate ion conduction through the pore via "hydropathic gates" [24–29]. However, the spatial complexity underlying hydropathic interactions [30–32] poses a significant obstacle in the development of computational tools for hydropathic analysis of protein structures.

Hydropathic moments (usually referred to as "hydrophobic moments" [33]) provide a simple, yet efficient tool for modeling of molecular mechanisms (e.g., see [34]). The method of cumulative hydropathic moments [35] has been used for analyzing the spatial profile of cumulative hydropathicity relative to a function-relevant geometry (e.g., the geometric center of the residues distribution) [36–39]. Interestingly, it was shown that data extracted from the vanishing, i.e., zero-crossing, behavior of the cumulative zero- and second-order residue hydropathic moments exhibited invariance over a large number of non-redundant globular proteins [37]. This finding highlighted the universality of the spatial transition from a protein's hydrophobic core to its hydrophilic exterior implying the existence of hydropathic scaling laws spanning all different molecular scales.

In principle, hydropathic characteristics along a NaVCh's pore are studied at a microscopic scale, i.e., by focusing on the hydropathic profile of pore-lining atomic structures (e.g., see [21, 40]). In this work, we extend this methodology to incorporate hydropathic characteristics of the spatial transition from the pore's microenvironment toward a macroscopic molecular regime that includes the VSDs. This methodological extension allows for a continuous mapping of hydropathic, as well as, structural properties

of a NaVCh on two dimensions. Specifically, instead of performing a three-dimensional (3D) mapping of hydrophathic characteristics on the NaVCh structure at residue-level resolution (as already implemented by available tools, e.g., pymol [41] and chimera [42]), we present a two-dimensional hydrophathic mapping procedure of atomic resolution by exploiting pore-forming geometrical principles. The main advantage of this approach is that it elucidates how hydrophathic dipole accumulation occurs with respect to pore’s geometry thus providing information about the pore’s gating behavior [24]. In order to achieve that we utilized the tool of cumulative hydrophathic moments and developed an atomic sampling algorithm that adapts to geometrical characteristics of the pore so that hydrophathic density variations, as well as, their corresponding hydrophathic dipole field configurations are visualized and, consequently, analyzed across different molecular scales. Accordingly, the relevant dimensions of the presented methodology are the pore axis coordinate and the molecular scale. We illustrate our methods for the pre-open I217C NaVAb channel where we report implications of our observations on molecular stability, as well as, on the formation of hydrophathic gates along the pore.

Methods

3D structure preparation

The I217C NaVAb model (PDB code: 3RVY) was selected for analysis as it provides a crystallography of the NaVAb with the highest resolution (2.7 Å resolution). The structure was protonated using the WHAT IF software [43, 44] and its principal axes were estimated using the VMD software [45]. A coordinate system $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ with origin O was introduced and the protonated structure was placed within it, so that the principal pore axis, i.e., the axis approximating the direction of the channel’s pore, is aligned with the z -axis. Orientation of the structure was set from the extracellular side (ES) to the intracellular side (IS) with respect to $\hat{\mathbf{z}}$ and the channel’s molecular mass center $\mathbf{e} = \frac{1}{M} \sum_{i=1}^{N_c} m_i \cdot \mathbf{c}_i$ was set to coincide with O where $\mathbf{c}_i = (c_{x,i}, c_{y,i}, c_{z,i})$ is the atomic center of the i -th atom, m_i

is the mass of the i -th atom, $N_c = 14776$ is the total number of atoms and $M = \sum_{i=1}^{N_c} m_i$ is the total molecular mass (see Dissertation Appendices, Table A1 for atomic masses values).

Geometrical characteristics of the pore

We considered P to represent the principal pore axis and $\mathbf{p} = (p_x, p_y, p_z) \in P$ to be a pore point. The radius of the smallest sphere that can be squeezed through the pore at $\mathbf{p}$ i.e., the pore radius at $\mathbf{p}$, is given by [46]

$$R(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| - vdW_i \} \quad (\text{m1})$$

where $\|\cdot\|$ is the euclidean norm and vdW_i is the Van der Waals radius of the i -th atom (see Dissertation Appendices, Table A1). Consequently, the distance between $\mathbf{p}$ and its nearest neighbor atom is given by

$$D(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| \} \quad (\text{m2})$$

In analogy to equation m1, we introduced the outer surface radius at $\mathbf{p}$ with

$$L(\mathbf{p}) = \max_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| + vdW_i \} \quad (\text{m3})$$

The unit of measurement for $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ is expressed in [Å].

Atomic sampling around the pore

In order to investigate cumulative atomic properties with respect to P , we introduced the $\mathbf{p}$ -dependent atomic sampling radius

$$l_\alpha(\mathbf{p}) = D(\mathbf{p}) + \alpha \cdot \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha} \text{ for } \begin{cases} K_\alpha \in \mathbb{Z}^+ \\ 0 < \alpha \leq K_\alpha \end{cases} \quad (\text{m4})$$

where K_α is the total number of sampling spheres centered at $\mathbf{p}$, α denotes the index of the sampling sphere. Note that $l_\alpha(\mathbf{p})$ plays the role of the molecular scale, i.e., it provides a $\mathbf{p}$ -dependent measurement of the size of

the channel in [$\text{\AA}$]. The number of sampled atoms within a sphere of radius $l_\alpha(\mathbf{p})$ centered at $\mathbf{p}$ is given by

$$N(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \quad (\text{m5})$$

where $\theta(\cdot)$ is the heaviside function.

Cumulative hydrophathic pore moments

We employed the zero-order hydrophathic pore moment [33]

$$h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \cdot HI_i^\chi \quad (\text{m6})$$

and the first-order (or dipole) hydrophathic pore moment [33]

$$\begin{aligned} \vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) &= \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \cdot HI_i^\chi \cdot \vec{r}_{\mathbf{p},i} = \\ &\underbrace{h_x(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{x}} + h_y(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{y}} + h_z(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{z}}}_{\vec{h}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))} \end{aligned} \quad (\text{m7})$$

where $HI_i^\chi = HI_i + \chi_i$ is the hydrophathic value of the i -th atom set according to the Kapcha-Rosky atomic hydrophathic indices [47] (see Appendices, Table A2) with additive Gaussian noise $\chi_i \in \mathcal{N}(\mu = 0, \sigma = 0.001)$, $\vec{r}_{\mathbf{p},i}$ is the vector from $\mathbf{p}$ to $\mathbf{c}_i$ and the measurement units of $h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ and of $\|\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))\|$ are roughly given by [kcal/mol] and [kcal· $\text{\AA}$ /mol], respectively. Division of equations m6 and m7 with m5 provided with an estimation of the hydrophathic density [37]

$$m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))}{N(\mathbf{p}, l_\alpha(\mathbf{p}))} \quad (\text{m8})$$

and of the hydrophathic imbalance [37]

$$\vec{m}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))}{N(\mathbf{p}, l_\alpha(\mathbf{p}))} \quad (\text{m9})$$

pore functions for increasing molecular scale $l_\alpha(\mathbf{p})$, respectively. Due to the spherical sampling procedure all scalar and vector pore functions remain invariant to rotations of the channel around P .

Discretization of the principal pore axis

We discretized P by introducing the equidistant grid $Q = \{\mathbf{p}_1, \mathbf{p}_1 + \Delta\mathbf{p}, \dots, \mathbf{p}_{N_p} - \Delta\mathbf{p}, \mathbf{p}_{N_p}\} \subset P$ where N_p is the total number of grid pore points, $\|\Delta\mathbf{p}\|$ is the sampling distance between two consecutive pore points and $\mathbf{p}_1 = (0, 0, p_{z,1})$, $\mathbf{p}_{N_p} = (0, 0, p_{z,N_p})$ are boundary pore points. Consequently, Q was constructed by setting $p_{z,1} = \text{round}(\min_{i=1,2,\dots,N_c}(c_{z,i}), 1) = -27.1$ and $p_{z,N_p} = \text{round}(\max_{i=1,2,\dots,N_c}(c_{z,i}), 1) = 26.8$ with $\text{round}(x \in \mathbb{R}, n \in \mathbb{Z}^+)$ returning the value of x rounded up to the n -th decimal digit and by setting $N_p = 540$ so that $\|\Delta\mathbf{p}\| = 0.1 \text{ \AA}$.

Detection of zero-crossing points of a scalar pore function along the principal pore axis

We consider $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ to represent a scalar pore function. If for a given scaling index α there is a pair $\{\mathbf{p}' = \mathbf{p} - \Delta\mathbf{p}, \mathbf{p}\} \in Q$ for which the sign-change condition $f(\mathbf{p}', l_\alpha(\mathbf{p}')) \cdot f(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ is satisfied, then the four-dimensional point

$$(\mathbf{s} = \mathbf{p}' + \frac{|f(\mathbf{p}', l_\alpha(\mathbf{p}'))|}{|f(\mathbf{p}', l_\alpha(\mathbf{p}'))| + |f(\mathbf{p}, l_\alpha(\mathbf{p}))|} \cdot \Delta\mathbf{p}, l_\alpha(\mathbf{s})) \quad (\text{m10})$$

represents a zero-crossing point of $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction where $|\cdot|$ returns the absolute value of f and $\mathbf{s}$ is obtained by linear interpolation. The set of all detected zero-crossing points of $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction for a given scaling index α is represented as $\Gamma(\alpha)$. Note that due to the addition of noise to the hydropathic indices, the probability of $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ being zero was practically eliminated so that $f(\mathbf{p}', l_\alpha(\mathbf{p}')) \cdot f(\mathbf{p}, l_\alpha(\mathbf{p})) \neq 0$ is always satisfied.

Results

Hydrophathic density variations around pre-open I217C NaVAb's pore

The first step in this study was to investigate how atomic hydrophathicity of the pore's microenvironment changes for increasing molecular scale, $l_\alpha(\mathbf{p})$ (see eq. m4). To do so, we employed the hydrophathic density pore function, $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ (see eq. m8), and illustrated it as a contour map (see Figure 1(a)). "Reading" of the contour map was performed from left to right, namely, for a given pore point, $\mathbf{p}$, we investigated the zero-crossing behavior of the hydrophathic density pore function for increasing $l_\alpha(\mathbf{p})$ based on a color gradient that illustrates negative and positive values of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ as blue as red, respectively. Accordingly, blue and red contour map domains represent hydrophobic and hydrophilic atomic structures, respectively. Observations were interpreted with respect to the conserved architectural motif dictating the structural separation of the PDs from the VSDs [48]. For that, a geometrical representation of the structural position of the PDs relatively to the VSDs (and vice-versa) was introduced by approximating the scales $\nu(\mathbf{p})$ for which an equilibrium of the corresponding radial distribution functions is achieved (see SI Section S2). In that way, the line $\nu(\mathbf{p})$ appearing in Figure 1(a) (and in Figure 2(a)) indicates that the number of PD atoms within a sampling radius $l_\alpha(\mathbf{p}) \leq \nu(\mathbf{p})$ is larger than that of VSD atoms thus roughly approximating the structural boundary between PDs and the VSDs. On the other hand, for $l_\alpha(\mathbf{p}) > \nu(\mathbf{p})$ the number of VSD atoms gradually increases so that the spatial transition from the PDs to the VSDs is realized.

At first glimpse, the contour map of Figure 1(a) reveals that the spatial organization of PD atoms around P results in the formation of a hydrophobic centrally-located cavity (CC). This becomes evident if we focus on the contour area enclosed within $-4.4 \leq p_z \leq 7.3$ and $l_\alpha(\mathbf{p}) < 10.0$ Å (Figure 1(a)). Deviating from the center of the pore toward the IS, we observe the narrowing and, eventually, occlusion of the pore by a hydrophilic microstructure as illustrated in Figure 1(a) for $p_z > 19.0$ and $l_\alpha(\mathbf{p}) < 10.0$ Å.

Figure 1: Spatial profile of the hydropathic density pore function. (a), Contour map of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in Q$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. Blue and red color contour domains represent hydrophobic and hydrophilic atomic structures around P , respectively. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict structural pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ indicates the structural boundary between the PDs and the VSDs. Black arrows indicate the clustering behavior of zero-crossing points of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ found in $\Omega^{(0)}$ resulting in the formation of the contour domains $T_1^{(0)}$, $T_2^{(0)}$, $T_3^{(0)}$ and $T_4^{(0)}$. (b), Histogram of $\Omega^{(0)}$ along $l_\alpha(\mathbf{p})$ -direction, $\psi^{(0)}$. (c), Histogram of $\Omega^{(0)}$ along $\mathbf{p}$ -direction, $\phi^{(0)}$ where $sm(\phi^{(0)})$ represents a smoothed version of $\phi^{(0)}$.

Macroscopically, occlusion of the pore translates into a qualitative change in the profile of the outer pore radius, $L(\mathbf{p})$ (see eq. m3), as $L(\mathbf{p})$ monotonically decreases and increases for $p_z \leq 19.0$ and for $p_z > 19.0$, respectively. In that way, a funnel-like outer pore surface with an opening at the IS is formed by 2π -rotation of $L(\mathbf{p})$ around P (Figure 1(a)). Deviating from the center of the pore toward the ES, we observe a widening of the pore so that an extracellular opening (or "mouth") is formed (see Figure 1(a)).

for $p_z < -12.0$). For $l_\alpha(\mathbf{p}) > 30 \text{ \AA}$ hydrophathic density variations cannot anymore be deduced via visual investigation of Figure 1(a), indicating that the structural combination of the PDs with the VSDs tends to increase hydrophathic uniformity of the atomic environment around the pore. This tendency results macroscopically in a down-regulation of hydrophobicity as the channel appears to be weakly-hydrophilic with

$$\begin{aligned} m_{total}^{(0)} &= m^{(0)}(\mathbf{p}, L(\mathbf{p})) \\ &= N_c^{-1} \sum_i^{N_c} HI_i^X = 0.00487 \sim \text{kcal/mol} \end{aligned} \tag{r1}$$

that accounts for the total, i.e., whole-channel, hydrophathic density.

A detailed examination of hydrophathic density variations around the pore was performed by analyzing the zero-crossing behavior of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing molecular scale $l_\alpha(\mathbf{p})$. To do so, we detected zero-crossing points of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing scaling index α and constructed the sets $\Gamma^{(0)}(\alpha)$ (see Methods). A zero-crossing point of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction identifies a pore point around which the hydrophathicity of the atomic environment becomes vanishingly small (see SI Section S3) and changes polarity, and is illustrated as a blue-to-red contour map transition (see Figure 1(a)). The arrangement of zero-crossing points on the contour map of Figure 1(a) revealed the boundaries among visually distinguishable, but also among visually indistinguishable contour domains as indicated by black arrows for $l_\alpha(\mathbf{p}) \leq \nu(\mathbf{p})$ and $l_\alpha(\mathbf{p}) > \nu(\mathbf{p})$, respectively. Accordingly, the contour map was partitioned into four distinct domains, namely, the hydrophobic contour domain $T_3^{(0)}$, the hydrophilic contour domains $T_1^{(0)}$ and $T_2^{(0)}$, and the weakly hydrophilic contour domain $T_4^{(0)}$. This partitioning indicates that the CC is a smaller part of a larger domain that covers the largest area of the contour map, namely of $T_3^{(0)}$ (Figure 1(a)). Around $T_3^{(0)}$ the hydrophilic $T_1^{(0)}$, $T_2^{(0)}$ and $T_4^{(0)}$ are formed in accordance with the overall structural organization of the PDs relatively to the VSDs as minimization of $\nu(\mathbf{p})$ roughly coincides with the formation of the CC (Figure 1(a)). On the other hand, deviating toward the ES or toward the

IS we observe an expansion of the PDs, as indicated by the monotonic increase of $\nu(\mathbf{p})$ for deviating from the CC, in accordance with an increase in atomic hydrophilicity that is expressed by the formation of $T_1^{(0)}$ with $T_2^{(0)}$ (Figure 1(a)). Taken together, these observations reflect the channel’s strategy to maintain a hydrophobic core, i.e., a hydrophobic CC, by increasing the density of hydrophobic atoms around its mass center, which in turn places hydrophilic atoms elsewhere. The advantage behind this configuration might be an increase in NaVAb’s stability due to reduction of water-accessible hydrophobic surface achieved by ”burying” hydrophobic atoms around the mass center.

In order to obtain a statistical summary of our observations, we introduced the distributions $\psi^{(0)}$ and $\phi^{(0)}$, respectively (see SI Section S4 for construction). This allowed for identifying molecular scales and locations along P where hydropathic density variations tend to increase or decrease as implied by extremization of $\psi^{(0)}$ and $\phi^{(0)}$, respectively. As we can see in Figure 1(b), $\psi^{(0)}$ can be decomposed into two sub-distributions, namely, $\psi_A^{(0)}$ and $\psi_B^{(0)}$ (Figure 1(b)), revealing that hydropathic density variations within the PDs occur more frequently as the majority of zero-crossing points (approximately 65%) are found in the contour map for $l_\alpha(\mathbf{p}) \leq \nu(\mathbf{p})$. Thus, separation of $\psi_A^{(0)}$ from $\psi_B^{(0)}$ is another indication of the structural and functional separation of the PDs from the VSDs.

The minimum structural information required in order to navigate through the pore’s microenvironment is to identify pore-lining residues. To do that, we partitioned P into distinct pore regions based on a minimal-distance geometrical criterion that takes into account only direct-neighboring residues (see SI Section S5). This approach revealed that the global minimum of $sm(\phi^{(0)})$ at $p_z = -12.0$ identifies a hydrophilic and narrow pore region that co-localizes with the hydrophilic-to-hydrophilic S178-E177 side chain transition that is part of the selectivity filter (SF) (Figure 1(a),(c) and SI Table S3). The strongly-hydrophilic and narrow environment that is locally formed creates favorable conditions for dehydration of incoming ions (Figure 1(a)). On the other hand, the global maximum of $sm(\phi^{(0)})$ at $p_z = 19.0$ indicates that closing of the channel at the IS requires a

hydropathically-diverse atomic environment as indicated by the hydrophilic-to-hydrophobic C217-M221 side chain transition (Figure 1(a),(c) and SI Table S3). The local minimum of $sm(\phi^{(0)})$ at $p_z = 7.3$ identifies the CC and co-localizes with the hydrophobic-to-hydrophobic I210-V213 side chain transition (Figure 1(a),(c) and SI Table S3). The local maximum of $sm(\phi^{(0)})$ at $p_z = -4.4$ co-localizes with the hydrophilic-to-hydrophobic E177-L176 side chain transition and identifies a transient pore region located between the narrow, hydrophilic pore region and the CC (Figure 1(a),(c) and SI Table S3). Finally, the local maximum of $sm(\phi^{(0)})$ at $p_z = -20.3$ found at the ES mouth of the pore co-localizes with the hydrophobic-to-hydrophilic M181-S178 side chain transition (Figure 1(a),(c) and SI Table S3). Notably, the hydrophobic-to-hydrophilic M181-S178 side chain transition becomes evident in Figure 1(a) as a color change from blue-to-red occurring in the vicinity of the $D(\mathbf{p})$ line with respect to $\hat{\mathbf{z}}$ (Figure 1(a)). Accordingly, the location $p_z = -20.3$ captures the transition from the wide ES mouth shaped by the hydrophobic M181 side chain toward a narrow and hydrophilic pore region surrounded by the hydrophilic S178 and E177 side chains.

Cumulative hydropathic topology of pre-open I217C NaVAb's pore

In this section we investigated the dipole field topology underlying hydropathic density variations illustrated in the contour map of Figure 1(a). In order to do so, we utilized the hydropathic imbalance pore function, $\vec{m}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ (see eq. m7) that can be equivalently understood as the mean (or average) cumulative hydropathic dipole moment (see connection of eq. m7 with eq. m9) thus providing information about accumulation of hydropathic dipoles around the pore. Given that the gating behavior of a nanopore is largely determined by its geometry and dipole accumulation parallel to its lining [24], we adopted here a similar approach to the previous section and analyzed how microscopic, pore-lining dipoles scale up and form larger polarized structures that incorporate, both, the PDs and the VSDs. By exploiting radial structural symmetries underlying NaVAb's tetrameric conformation it was straightforward to reduce $\vec{m}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ to

its z -component, $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) = m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{z}}$, as the amplitude of the radial component of dipole hydropathic moment, $\|\vec{h}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))\|$ (see eq. m7), is shown to be vanishingly small (see SI Section S6). Accordingly, our analysis focused on the spatial profile of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ that is illustrated on the contour map of Figure 2(a). A zero-crossing point of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction accounts for a change in dipole field's orientation, as well as, for a vanishingly small hydropathic imbalance effect (see SI Section S7), i.e., for the formation of a hydropathic dipole center on P for a given scaling index α . From a topological viewpoint, this vanishing event corresponds to an equilibrium point of the hydropathic imbalance field acting along P for a given α . Depending on the orientation of the local field, an equilibrium point acts either as molecular attractor (i.e., "sink") or molecular repeller (i.e., "source") so that a hypothetical molecule (e.g., a partially-hydrated sodium ion) would be either attracted toward or repelled from it, respectively (see SI Section S8). The importance of the topological interplay between hydropathic "sinks" and "sources" in ion permeation dynamics and, specifically, ion selectivity was recently highlighted in [49]. Similarly to the previous section, we introduced the expression

$$\Omega^{(1)} = \bigcup_{\alpha} \{\Gamma^{(1)}(\alpha)\} \quad (\text{r2})$$

that incorporates topological information extracted from every molecular scale $l_\alpha(\mathbf{p})$ where $\Gamma^{(1)}(\alpha)$ represents the set of all detected dipole centers for a given scaling index α (see Methods). Consequently, the contour map of Figure 2(a) was partitioned into four domains; the pair $T_1^{(1)}, T_4^{(1)}$ and the pair $T_2^{(1)}, T_3^{(1)}$ corresponding to configurations of $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ with orientation toward the ES, i.e., $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$, and with orientation toward the IS, i.e., $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$, respectively. Notably, $T_3^{(1)}$ and $T_4^{(1)}$ cover the largest contour area so that the map can be roughly split into two parts, namely, an ES and an IS part, revealing a topological dichotomy of the pore (Figure 2(a)). The structural advantage behind this topological configuration is the stabilization of CC via an accumulation of dipoles around it in a direction roughly parallel to P (Figure 2(a)). Following this

line of thought, $T_1^{(1)}$ and $T_2^{(1)}$ might guarantee stabilization of hydrophilic atomic structures at the ES and IS contributing to the formation of $T_1^{(0)}$ and $T_2^{(0)}$, respectively (Figure 1(a), 2(a)). For $l_\alpha(\mathbf{p}) > \nu(\mathbf{p})$ a qualitative change in the spatial profile of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ takes place where the number of dipoles decreases, as indicated by flattening of $\psi^{(1)}$ (Figure 2(a), (b)), and the topological dichotomy of the pore is disrupted. Specifically, the structural combination of the PDs with the VSDs causes a displacement of dipole centers toward the IS, as indicated in Figure 2(a) by the "(*)" arrow, so that the size of $T_3^{(1)}$ over $T_4^{(1)}$ increases. Consequently, a unidirectional dipole field configuration is macroscopically established with

$$m_z^{(1)}(\mathbf{p}, L(\mathbf{p})) = -m_{total}^{(0)} \cdot p_z - 0.1368 \quad (\text{r3})$$

that is in accordance with the molecular field theory presented in [50] (see SI Section S9). In summary, these topological characteristics result in a skewed distribution $\psi^{(1)}$ indicating that the vast majority of dipole centers (approximately 90%) are characterized by $l_\alpha(\mathbf{s}^{(1)}) \leq \nu(\mathbf{p})$. This finding highlights the prominent role that PDs play in shaping properties of the hydropathic imbalance field acting along P .

Similarly to the previous section, we employed the smoothed profile of the distribution $\phi^{(1)}$, $sm(\phi^{(1)})$, in order to identify how geometrical and hydropathic topological characteristics correlate with each other. We found that the global maximum of $sm(\phi^{(1)})$ at $p_z = 2.8$ not only co-localizes with the *L176-I210* side chain transition but also with the local maximization of $R(\mathbf{p})$, both identifying the CC (Figure 2(a), (c) and SI Table S3). This observation illustrates how microscopic, pore-lining dipoles around the NaVAb's mass center co-aggregate into a larger structure that holds the CC open and is consistent with experimental findings presented in [24] showing that opening of a hydrophobic pore can be achieved by adding dipoles parallel to its lining.

Figure 2: Spatial profile of the hydropathic imbalance pore function. (a), Contour map of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in Q$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. Blue and red contour domains represent configurations of $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ with orientation toward the ES and toward the IS, respectively. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict structural pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ indicates the structural boundary between the PDs and the VSDs. Black arrows indicate the clustering behavior of dipole centers found in $\Omega^{(1)}$ resulting in the formation of contour domains $T_1^{(1)}$, $T_2^{(1)}$, $T_3^{(1)}$ and $T_4^{(1)}$. Black arrow labeled as "*" indicates the clustering behavior of dipole centers for $l_\alpha(\mathbf{p}) > \nu(\mathbf{p})$. (b), Histogram of $\Omega^{(1)}$ along $l_\alpha(\mathbf{p})$ -direction, $\psi^{(1)}$. (c), Histogram of $\Omega^{(1)}$ along $\mathbf{p}$ -direction, $\phi^{(1)}$ where $sm(\phi^{(1)})$ represents a smoothed version of $\phi^{(1)}$.

From the perspective of a hypothetical molecule, the global maximum of $sm(\phi^{(1)})$ indicates the formation of a molecular repeller approximately in the center of the pore as indicated by the relative orientation of $T_3^{(1)}$ with respect to $T_4^{(1)}$. This topology does not favor entrance of molecules to the CC from the SF or from the AG while, on the other hand, facilitates a pathway for waters or small, hydrophobic compounds penetrating the

lipid membrane to enter the pore through the NaVAb's side fenestrations (see Figure 4 in [21]). Accordingly, the repeller dynamics is practically reversed for the case of lipid-membrane penetrating molecules; they can easily access the center of the pore through the fenestrations as they are attracted and, consequently, "trapped" into CC under the influence of $T_3^{(1)}$ or $T_4^{(1)}$, respectively (Figure 2(a),(c)). This "trapping" mechanism would then allow small, hydrophobic drug agents to reach and interact with the T206-M209-V213 drug-binding pore site [21,51–53] (see Figure 3(b) for the location of this pore site). On the other hand, the global minimum of $sm(\phi^{(1)})$ at $p_z = -17.9$ accounts for the formation of $T_1^{(1)}$ and $T_3^{(1)}$ and co-localizes with the M181-S178 side chain transition indicating a decrease in dipole field topological variability at the ES mouth of the pore (Figure 2(a),(c) and SI Table S3). This topological configuration facilitates a passage for an incoming hydrated ion through the ES mouth toward the SF under the influence of a strong, bi-directional field effect originating from the combination of $T_1^{(1)}$ and $T_3^{(1)}$. The local maximum of $sm(\phi^{(1)})$ at $p_z = -11.7$ co-localizes with the narrowing of the pore thus capturing the local minimization of $R(\mathbf{p})$, as well as, with the global minimum of $sm(\phi^{(0)})$ (Figure 1(a), 2(a),(c)). Given that dipole accumulation around $p_z = -11.7$ occurs roughly parallel to P , the relative orientation of $T_1^{(1)}$ with respect to $T_3^{(1)}$ generates a local attractor favoring binding and, consequently, dehydration of ions arriving at the extracellular end of the SF (Figure 1(a), 2(a),(c)). Formation of local ion-binding sites (such as of the site_{HFS} [21]) is attributed to the spatial organization of the SF T175-L176-E177-S178-W179 residue complex [21] and, especially, of E177 side chains playing a key-role in ion selectivity in sodium and calcium channels [54]. The local maximum of $sm(\phi^{(1)})$ at $p_z = 22.7$ identifies with high precision the pore point where pore occlusion takes place revealing a microscopic dipole accumulation mechanism occurring roughly parallel to P , as we can visually deduce from Figure 2(a) for $p_z = 22.7$ in the vicinity of $D(\mathbf{p})$ line, which would favor molecular localization. Thus pore occlusion of the pre-open NaVAb might not only be the outcome of a purely structural effect but also of a hydropathic gating process. The local minima at $p_z = -7.1$

and $p_z = 10.8$ reflect the formation of $T_3^{(1)}$ and $T_4^{(1)}$ thus identifying locations on P around which dipole field topological variability is minimized, respectively (Figure 2(a),(c)). In that way, molecules resting or entering CC through its fenestrations (see Figure 4 of [21]) would be transported toward the SF and the AG under the influence of the strong, unidirectional field effect originating from $T_3^{(1)}$ and $T_4^{(1)}$, respectively.

What are the lessons to be learned from cumulative hydropathic topological analysis? A summarizing read-out of the cumulative hydropathic topological analysis is presented in Figure 3 alongside with targeted conservation analysis [55,56], as well as, 3D hydropathicity mapping of the pore walls. Specifically, a condensed output of our method was obtained in terms of a topological quasiprobability score (TQS) (see SI eq. s12) that quantifies the probability along the pore for a molecular attractor (or repeller) to occur (Figure 3(c)). TQS essentially provides with a qualitative description of the relative size and location of energetic barriers to molecular permeation, i.e., "gates" [29], along the pore (see SI Section S8). According to this interpretation scheme, a central gating mechanism keeps the hydrophobic CC "closed" by imposing large energy barriers to sodium ions entering from the ES or the IS side while facilitating an "opening" to small hydrophobes and waters entering via the side fenestrations. Hence, hydrated sodium ions escaping from the SF's local binding sites have to overcome a large energy barrier in order to arrive at the center of the pore (Figure 3(c)). This barrier stems from the molecular influx into CC through the fenestrations which is in opposite direction to the ion escape trajectory thus preventing (or, at least, not favoring) ion entrance into the CC. Escape trajectory is thus likely to involve multiple, partial re- and dehydration cycles, i.e., an interplay of attractors-vs-repellers sites, so that ions are gradually transported from the narrow, strongly-hydrophilic and highly-conserved SF environment toward the wide, hydrophobic center of the pore (see Figure 3(a),(b), and (c) where TQS attains its global negative and positive, respectively, maximum and in-between fluctuations of the weakly-smoothed TQS trace indicating the attractors-vs-repellers interplay).

The advantage behind a long escape trajectory is that NaVAb's speci-

ficity is optimized as entrance into CC is permitted only to ions that can re- and dehydrate according to the local gradients of the energy landscape.

Figure 3: **Summary of cumulative hydropathic topological analysis; topological description of single-molecule permeation energy landscape.** (a), Residue-level 3D hydropathicity mapping of a single PD structural unit (residue

sequence: M130:M221) based on the Kapcha-Rosky hydropathic scale performed in pymol [41] computational environment. **(b)**, Conservation analysis of a PD structural unit. Analysis was performed on the ConSurf Server (consurf.tau.ac.il) according to the algorithmic implementations described in [55, 56]. **(c)**, A summary read-out of cumulative hydropathic topological analysis is provided in terms of the strongly- and weakly-smoothed TQS plotted for $\mathbf{p} \in Q$. Geometrical pore characteristics are represented in terms of the pore radius trace, $R(\mathbf{p})$, plotted for $\mathbf{p} \in Q$. Pore points belonging to the subsets $Q_{2 \cdot i + 1, i=0,1,\dots,7}$ minimize their distance from the pore-lining residue side chains M181, S178, E177, L176, I210, V213, C217 and M221, respectively. Vertical green zones in (c) highlight the locations of persistent attractors and repellers along P corresponding to negative-value and positive-value maxima of TQS, respectively, and indicating the location and relative size of energy barriers imposed to permeating molecules. ES stands for extracellular side, EF for extracellular funnel, SF for selectivity filter, CC for central cavity and AG for activation gate. Note that the SF residue side chains complex T175-L176-E177-S178-W179 averages the highest conservation score with 2.4 (out of 4.0) with pore-lining residue conservation scores confined within $[0, 2.0]$.

Ions arriving in the center of the pore are fully-hydrated thus they can rapidly diffuse toward the hydropathically-diverse AG where a local attractor site is formed at $p_z \approx 19.0$ (see Figure 3(a), and (c) where TQS attains its local negative maximum, and previous Section for properties of the atomic environment around $p_z \approx 19.0$). Finally, a weak, repeller mechanism at $p_z \approx -20.3$ attributed to the hydrophobic-to-hydrophilic interplay among the M121 and S178 side chains is likely to regulate ion access to the SF from the ES (see Figure 3(a), and (c) where TQS attains its local positive maximum, and previous Section for properties of the atomic environment around $p_z \approx -20.3$).

Discussion

In this study we followed the trace of earlier works on cumulative hydropathic analysis of protein systems (see [36–39, 50]) and presented a compu-

tational scheme that allows for visual detection and analysis of hydrophobic pattern formation around a NaVCh's pore. Specifically, considering that structural symmetries play a fundamental role in biological pore (for a review see [57]), we developed a scaling analysis methodology that utilizes the tools of hydrophobic density and hydrophobic imbalance pore functions (see eq. m8 and m9, respectively) and returns two-dimensional spatial profiles (i.e., contour maps) of these experimental quantities where the relevant dimensions are unfolding perpendicular and parallel to the membrane surface. In the absence of structural pore symmetries, e.g., in the case of heteromeric eukaryotic NaVCh channel such as the Nav1.7, our algorithm is still applicable but radial contributions to the hydrophobic dipole field need to be taken into account. This can be done by introducing an appropriate color scheme for illustrating changes in orientation and magnitude of the radial hydrophobic dipole field component (see eq. m7) so that a third contour map is constructed. For the sake of simplicity and due to limitations in computing time, we treated the pre-open I217C NaVAb as a rigid entity by neglecting thermal atomic fluctuations. However, introduction of fluctuations at room temperature is not expected to affect contour pattern formation due to the linear additive scheme used for estimating cumulative effects. The same argument applies also for the choice of an atomic hydrophobic value, namely, choosing a set of different atomic hydrophobic indices, as well as, increasing the noise amplitude χ_i up to a threshold value is not expected to affect our findings as long as grouping of atoms into hydrophobic and hydrophilic ones remains unchanged.

We demonstrated that the spatial profiles of hydrophobic density and hydrophobic imbalance pore functions exhibit pseudo-symmetrical characteristics with respect to an axis parallel to the membrane surface so that the pore is roughly dichotomized with respect to the CC. This becomes evident if we focus on the relative positioning of the pairs $\{T_1^{(0)}, T_2^{(0)}\}$, $\{T_1^{(1)}, T_2^{(1)}\}$ and $\{T_3^{(1)}, T_4^{(1)}\}$ with respect to a symmetry axis that is placed perpendicular to the principal pore axis at $p_z \approx 2.8$ (Figure 1(a), and 2(a)). Accordingly, two qualitatively-different stability mechanisms coexist and co-determine I217C NaVAb's functional placement within the membrane;

at the microscopic regime, hydropathically-diverse pore-lining components are stabilized via an increase in their topological symmetry (see Figure 2(a) formation of $\{T_1^{(1)}, T_2^{(1)}\}$ and $\{T_3^{(1)}, T_4^{(1)}\}$), while, toward the macroscopic regime, stabilization of the weakly-hydrophilic ensemble of PD and VSD atoms requires a gradual decrease in their topological similarity (see Figure 2(a), shrinkage of $T_4^{(1)}$). The biophysical principle underlying this mode of organization is the optimization strategy dictating that burying of hydrophobic atoms around a protein’s core contributes to protein stability [58] (the pre-open I217C NaVAb’s hydrophobic core was identified as the CC (see Figure 1(a))).

Persistent extrema of the distributions $\phi^{(0)}$ and $\phi^{(1)}$ were utilized in order to investigate hydropathic gating properties of the pore. In particular, we demonstrated that the maxima of $sm(\phi^{(1)})$ identify pore locations where hydropathic topological and geometrical characteristics of the atomic environment favor the formation of a central hydrophobic gate (at $p_z \approx 2.8$) regulating ion transport between the hydrophilic SF attractor (at $p_z \approx -11.7$) and the hydropathically-diverse activation gate (AG) [21] (at $p_z \approx 22.7$). Conservation analysis of the PDs highlighted the highly-conserved nature of the SF as already pointed out in [21] (see Figure 3). Accordingly, the SF attractor topological configuration reflects an evolutionary-driven functional optimization of the PDs that might be relevant for other NaVCh species as well. Given that these locations account for a smoothed, multi-scale hydropathic effect, thermal fluctuations are expected to induce only small dislocations.

In summary, the strength of the presented methodologies is that they allow not only for a detailed mapping of multiscale hydropathic characteristics lining a NaVCh’s pore but also for elucidating how their underlying topology correlates with the pore’s geometry at atomic resolution. This is a crucial step toward understanding where along a NaVCh’s pore hydropathic gates are likely to form and how they contribute to its functioning. A weaknesses, however, is that multiscale hydropathic mapping at atomic resolution is computationally expensive in comparison to traditional structural biology methods that operate at residue-level and that interpretation

of observables can be far from trivial. To circumvent these complexities statistical summary measures, such as the distributions ψ and ϕ , can be employed alongside with targeted conservation analysis.

Supplementary Information

S1. Atomic properties

Each atom is assigned with a van der Waals (*vdW*) radius and a hydrophobic index according to Tables A1 and A2, respectively, shown in Dissertation Appendices.

S2. Structural organization of pore domains (PDs) and voltage-sensing domains (VSDs)

The radial distribution function (RDF) of a collection of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around a pore point $\mathbf{p}$ was approximated by

$$G(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta N(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho(\mathbf{p})} \text{ with } l'_\alpha(\mathbf{p}) = l_\alpha(\mathbf{p}) - \Delta l_\alpha(\mathbf{p}) \quad (\text{s1})$$

where $\Delta l_\alpha(\mathbf{p}) = \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha}$ is the thickness of a spherical shell around $\mathbf{p}$, $\Delta N(\mathbf{p}, l_\alpha(\mathbf{p})) = N(\mathbf{p}, l_\alpha(\mathbf{p})) - N(\mathbf{p}, l'_\alpha(\mathbf{p}))$ is the number of atoms found within the spherical shell of thickness $\Delta l_\alpha(\mathbf{p})$ centered at $\mathbf{p}$ and $\rho(\mathbf{p}) = \frac{N_c}{V(\mathbf{p})}$ is the average atomic density around $\mathbf{p}$ with $V(\mathbf{p}) = \frac{4}{3} \cdot \pi \cdot (L(\mathbf{p}) - R(\mathbf{p}))^3$ being the volume of a sphere of radius $l_\alpha(\mathbf{p})$ centered at $\mathbf{p}$.

In order to investigate the structural positioning of the PDs relatively to the VSDs, we divided all N_c atoms into two groups, one containing only PD atoms, i.e., atoms forming the PDs, and one containing only VSD atoms, i.e., atoms forming the VSDs. Note that the total number of PD atoms is $N_{PD} = 6068$ and they belong to the residue sequence M130:M221 that forms the structural units of the S5 domain, the pore turret loop, the P-helix, the selectivity filter, the P2-helix and the S6 domain (see SI Figure S1 of [21]). On the other hand, the total number of VSD atoms is $N_{VSD} = 8708$ and

they belong to the residue sequence M001:G129 that forms the structural units of the S1N helix, the S1 domain, the S1-S2 loop, the S2 domain, the S2-S3 loop, the S3 domain, the S3-S4 loop, the S4 domain and the S4-S5 linker (see SI Figure S1 of [21]).

By using equation s1 we approximated the RDFs of the PD and VSD atoms around $\mathbf{p}$ with

$$G_{PD}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{PD}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{PD}(\mathbf{p})} \quad (\text{s2})$$

and

$$G_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{VSD}(\mathbf{p})} \quad (\text{s3})$$

where $\Delta n_{PD}(\mathbf{p}, l_\alpha(\mathbf{p}))$ and $\Delta n_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))$ are the numbers of PD and VSD atoms found within a spherical shell of thickness $\Delta l_\alpha(\mathbf{p})$ around $\mathbf{p}$, and $\rho_{VSD}(\mathbf{p}) = \frac{N_{VSD}}{V(\mathbf{p})}$ and $\rho_{PD}(\mathbf{p}) = \frac{N_{PD}}{V(\mathbf{p})}$ are the atomic densities of PD and VSD atoms around $\mathbf{p}$, respectively. Next, we introduced the equilibrium RDF

$$e(\mathbf{p}, l_\alpha(\mathbf{p})) = sm(G_{PD}(\mathbf{p}, l_\alpha(\mathbf{p})) - G_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))) \quad (\text{s4})$$

allowing us to investigate how the RDF of PD atoms behaves relatively to the RDF of VSD atoms where $sm(\cdot)$ introduces a weak smoothing along $l_\alpha(\mathbf{p})$ -direction (see caption of Figure S1). What is of interest in equation s4 is its zero-crossing behavior of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing $l_\alpha(\mathbf{p})$ as it indicates whether the probability of finding a PD atoms is larger, i.e., $e(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$, or smaller, i.e., $e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$, than that of finding a VSD atom within the spherical shell of thickness $\Delta l_\alpha(\mathbf{p})$. The zero-crossing behavior of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing $l_\alpha(\mathbf{p})$ was investigated by detecting for every $\mathbf{p} \in Q$ the pair $\{l'_\alpha(\mathbf{p}), l_\alpha(\mathbf{p})\}$ for which the zero-crossing condition $e(\mathbf{p}, l'_\alpha(\mathbf{p})) \cdot e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ is satisfied and approximating the molecular scale for which $e(\mathbf{p}, l_\alpha(\mathbf{p})) \approx 0$ with

$$\nu(\mathbf{p}) = l'_\alpha(\mathbf{p}) - \frac{e(\mathbf{p}, l'_\alpha(\mathbf{p}))}{e(\mathbf{p}, l_\alpha(\mathbf{p})) - e(\mathbf{p}, l'_\alpha(\mathbf{p}))} \cdot \Delta l_\alpha(\mathbf{p}) \quad (\text{s5})$$

As we demonstrate in Figure S1, $\nu(\mathbf{p})$ exhibits a smooth profile so that the contour map can be split into two domains colored as green and ma-

genta contours accounting for $e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ and $e(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$, respectively. Accordingly, $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ serves as a rough indication of the structural boundary between the PDs and the VSDs given that for $l_\alpha(\mathbf{p}) < \nu(\mathbf{p})$ the number of sampled PD atoms is always larger than that of VSD atoms. The global maximum of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ occurs for $p_z = 19.0$ and $l_\alpha(\mathbf{p}) = 4.3$ Å where the geometrical closing of the channel is initiated. This finding suggests that an increased density of PD atoms is required in order to stabilize the closed AG. On the other hand, the local maximum of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ occurs for $p_z = -13.1$ and $l_\alpha(\mathbf{p}) = 38.1$ Å indicating that the density of VSD atoms tends to increase deviating from the CC toward the ES.

Figure S1: **Spatial profile of the equilibrium RDF.** Contour map of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in Q$ and $\alpha = 1, 2, \dots, K_\alpha$. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict structural pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ highlights the molecular scales for which $e(\mathbf{p})$ crosses zero. Grey shaded circles, as well as, dashed vertical and horizontal black lines, highlight contour map locations where extrema of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ are located. Smoothing of expression s4 was performed by using the built-in R [59] function of the Nadaraya-Watson kernel regression method with a bandwidth parameter value of $bw = 2$.

S3. Zero-crossing behavior of hydropathic density pore function along the principal pore axis

We here show that when $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ crosses zero along $\mathbf{p}$ - direction for a given α , it is expected, in a statistical sense, to become vanishingly small. Zero-crossing properties of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ were calculated by statistically analyzing the set of values

$$V^{(0)}(\alpha) = \{m^{(0)}(\bigcup_{\alpha} \{\Gamma^{(0)}(\alpha)\})\} \quad (\text{s6})$$

For $\alpha > 10$ the median and mode of $V^{(0)}(\alpha)$ remain close to zero as indicated by their traces (Figure S2). Furthermore, the 95%-interval values are shown to gradually decrease for increasing α indicating that the width of the distribution gradually decreases.

Figure S2: **Vanishing of hydropathic density pore function.** Statistical properties of $V^{(0)}(\alpha)$ for $\alpha = 10, 20, \dots, K_\alpha = 800$. Error bars represent 95%-interval values of $V^{(0)}(\alpha)$. $median(\cdot)$ and $mode(\cdot)$ return the median and the mode of $V^{(0)}(\alpha)$, respectively.

In particular, 95% of the values found in $V^{(0)}(K_\alpha = 800)$ are contained in $[-0.006297, 0.006977]$ with $median(V^{(0)}(K_\alpha = 800)) = -6.2e - 06$ and $mode(V^{(0)}(K_\alpha = 800)) = -6.2e - 05$ where $median(\cdot)$ and $mode(\cdot)$ return the median and mode of a set of values, respectively. Consequently,

$m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ is expected to obtain a vanishingly small value when it crosses zero. This finding indicates that $\|\Delta\mathbf{p}\|$ (see Methods) is sufficiently small, i.e., N_p is sufficiently large, to guarantee that no "sudden jumps" of $|m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))|$ occur between two consecutive pore points for a given α . Deviations from zero of the mode and median, as well as, large 95% interval values observed for small α , e.g., for $\alpha < 100$, are due to low atomic sampling (Figure S2).

S4. Construction and smoothing of zero-crossing points distributions

Distributions ψ were obtained by placing bins of width 0.1 at bin centers 0, 0.1, 0.2, .., $\text{round}(\max(L(\mathbf{p})), 1)$ and, then, counting the number of $l_\alpha(\mathbf{s})$ (see main text, eq. m10) that fall within each bin.

Distributions ϕ were obtained by placing bins of width 0.1 at every p_z , and, then, counting the number of s_z that fall within, s_z represents the z-coordinate of a zero-crossing point $\mathbf{s}$ (see main text, eq. m10).

Smoothing of distributions was performed by utilizing the built-in R [59] function that implements the Nadaraya-Watson kernel regression smoothing method for a bandwidth parameter value $bw=7$.

S5. Geometrical correlation between pore-lining residues and cumulative hydropathic topological characteristics

A residue mass center was calculated by

$$\mathbf{b} = \frac{1}{M_b} \sum_i^{n_b} m_i^b \cdot \mathbf{c}_i^b \quad (\text{s7})$$

where $\mathbf{c}_i^b$ is the center and m_i^b is the mass of the i -th atom belonging to the residue under scrutiny, $M_b = \sum_i^{n_b} m_i^b$ is the total residue mass and n_b is the total number of atoms forming the residue. The mass centers of pore-lining residues were mapped on P by determining the four nearest neighbor residue mass centers around every $\mathbf{p} \in Q$. For that, we repeated four times

the calculation

$$Y(\mathbf{p}) = j \mid \min_{j=1,2,\dots,N_b} \{||\mathbf{b}_j - \mathbf{p}||\} \quad (\text{s8})$$

where each time the residue mass center corresponding to the solution index $Y(\mathbf{p})$ was removed from the collection of the $N_b = 884$ residues. In that way, we specified the subsets of pore points $Q_{2 \cdot i+1, i=0,1,\dots,7}$ that minimize their distance from four pore-lining residues of the same type, i.e., from a pore-lining residue side chain, and the subsets of pore points $Q_{2 \cdot i, i=1,2,\dots,7}$ that minimize their distance from four pore-lining residues of different type thus accounting for a structural transition from one pore-lining residue side chain to the next one (Table S3). Based on the total residual hydropathicity (see caption of Table S2), a structural transition from one pore-lining residue side chain to the next one along $\hat{\mathbf{z}}$ -direction is categorized as hydrophobic-to-hydrophobic, hydrophilic-to-hydrophilic, hydrophobic-to-hydrophilic or hydrophilic-to- hydrophobic (Table S3). Extrema of smoothed topological distributions $sm(\phi^{(0)})$ and $sm(\phi^{(1)})$ (see Figure 1 and 2, respectively) were assigned to a $Q_{i=1,2,\dots,15}$ subset based on the calculation

$$W(p_z^{ext.}) = i \mid \min_{i=1,2,\dots,15} \{|p_z^{ext.} - median(Q_{i,z})|\} \quad (\text{s9})$$

where $p_z^{ext.}$ represents the location on P of an extremum of $sm(\phi^{(0)})$ (see Figure 1(a)) or $sm(\phi^{(1)})$ (see Figure 2(a)) and $Q_{i,z}$ is the set that contains only the z -coordinate of a $\mathbf{p} \in Q_i$. We consider an extremum of $sm(\phi^{(0)})$ or $sm(\phi^{(1)})$ found at $p_z^{ext.}$ to be geometrically correlated, i.e., co-localize, with the $W(p_z^{ext.})$ -th subset (Table S3).

Pore points	Near. neigh.	Ext. of	Ext. of
	residues	$sm(\phi^{(0)})$	$sm(\phi^{(1)})$
$Q_1 = \{\mathbf{p} \mid p_z \in [-26.8, -18.7]\}$	M181 M402 M623 M844 •		

$Q_2 = \{\mathbf{p} p_z \in (-18.7, -18.4)\}$	M181 M623 S399 S841 $\bullet \rightarrow \bullet$	loc. max. at -20.3	gl. min. at -17.9
$Q_3 = \{\mathbf{p} p_z \in [-18.4, -11.9]\}$	S178 S399 S623 S841 $\bullet$		
$Q_4 = \{\mathbf{p} p_z \in (-11.9, -11.5)\}$	S399 S841 E398 E840 $\bullet \rightarrow \bullet$	gl. min. at -12.0	loc. max. at -11.7
$Q_5 = \{\mathbf{p} p_z \in [-11.5, -7.4]\}$	E177 E398 E619 E840 $\bullet$		
$Q_6 = \{\mathbf{p} p_z \in (-7.4, -7.2)\}$	E398 E840 L397 L839 $\bullet \rightarrow \bullet$	loc. max. at -4.4	loc. min. at -7.1
$Q_7 = \{\mathbf{p} p_z \in [-7.2, 1.9]\}$	L176 L397 L618 L839 $\bullet$		
$Q_8 = \{\mathbf{p} p_z \in (1.9, 2.1)\}$	L176 L618 I431 I873 $\bullet \rightarrow \bullet$		gl. max. at 2.8
$Q_9 = \{\mathbf{p} p_z \in [2.1, 5.1]\}$	I210 I431 I652 I873 $\bullet$		
$Q_{10} = \{\mathbf{p} p_z \in (5.1, 5.5)\}$	I431 I873 V434 V876 $\bullet \rightarrow \bullet$	loc. min. at 7.3	
$Q_{11} = \{\mathbf{p} p_z \in [5.5, 11.7]\}$	V213 V434 V876 V655 $\bullet$		
$Q_{12} = \{\mathbf{p} p_z \in (11.7, 12.2)\}$	V213 V655 C217 C559 $\bullet \rightarrow \bullet$		loc. min. at 10.8
$Q_{13} = \{\mathbf{p} p_z \in [12.2, 19.0]\}$	C217 C438 C559 C880 $\bullet$		
$Q_{14} = \{\mathbf{p} p_z \in (19.0, 19.5)\}$	C217 C438 M221 M663 $\bullet \rightarrow \bullet$	gl. max. at 19.0	
$Q_{15} = \{\mathbf{p} p_z \in [19.5, 27.1]\}$	M221 M442 M663 M884 $\bullet$		loc. max. at 22.7

Table S3: **Geometrical correlations between mappings of pore-lining residues on P and extrema of the smoothed topological distributions $sm(\phi^{(0)})$ and $sm(\phi^{(1)})$.** Pore points $\mathbf{p} \in Q$ belonging to the subsets $Q_{2 \cdot i + 1, i=0,1,\dots,7}$ are nearest neighbors (near. neigh.) to the pore-lining residue side chains M181, S178, E177, L176, I210, V213, C217 and M221, respectively. Pore points belonging to the sub-

sets $Q_{2-i,i=1,2,\dots,7}$ account for side chain transitions M181-S178, S178-E177, E177-L176, L176-I210, I210-V213, V213-C217 and C217-M221, respectively, and are nearest neighbors to a combination of different pore-lining residues. (•) and (•) indicate that the near. neigh. residues are hydrophobic and hydrophilic, respectively (for classification of residues into hydrophobic and hydrophilic see caption of Table A2 at Dissertation Appendices). Accordingly, (•→•), (•→•), (•→•) and (•→•) indicate that a hydrophobic-to- hydrophobic, a hydrophilic-to-hydrophilic, a hydrophobic-to-hydrophilic and a hydrophilic-to- hydrophobic structural transition from one pore-lining residue side chain to the next with respect to $\hat{\mathbf{z}}$ takes place, respectively. An extremum (ext.) of $sm(\phi^{(0)})$ or $sm(\phi^{(1)})$ is assigned to $Q_{i=1,2,\dots,15}$ based on a minimal distance geometrical criterion (see equation s9).

S6. Impact of radial structural symmetries on the cumulative hydropathic dipole moment field

A consequence of the homologous tetrameric conformation of the NaVAb channel is the existence of radial structural symmetries with respect to P . In order to investigate their impact on the cumulative hydropathic dipole field, we introduced the ratios

$$\begin{aligned}\Lambda(\mathbf{p}, l_\alpha(\mathbf{p})) &= \frac{\|\vec{h}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))\|}{\|\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))\|} \\ K(\mathbf{p}, l_\alpha(\mathbf{p})) &= \frac{\|\vec{h}_{xy}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))\|}{\|\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))\|}\end{aligned}\tag{s10}$$

Next, we calculated statistical properties of the sets of values $\lambda(\mathbf{p}) = \{\Lambda(\mathbf{p}, l_\alpha(\mathbf{p})) \mid \alpha = 10, 20, \dots, K_\alpha = 800\}$ and $\kappa(\mathbf{p}) = \{K(\mathbf{p}, l_\alpha(\mathbf{p})) \mid \alpha = 10, 20, \dots, K_\alpha = 800\}$ for every $\mathbf{p} \in Q$. As we can see in Figure S3, the medians and modes of the distributions $\lambda(\mathbf{p})$ and $\kappa(\mathbf{p})$ remain close to one and zero for $p_z > 0$, respectively. On the other hand, for $p_z < 0$, the modes of $\lambda(\mathbf{p})$ and $\kappa(\mathbf{p})$ deviate from one and zero, respectively, indicating that symmetric imperfections of the 3D structure are amplified toward the ES (Figure S3).

Fluctuations of the 95%-interval values serve as markers of symmetric imperfections and, as we can see in Figure S3, they tend to increase in am-

plitude toward the ES. Nevertheless, given that the median and the mode of $\lambda(\mathbf{p})$ are always larger than 0.984 and 0.99, respectively, we consider the contribution of radial hydrophathic imbalance effects (see xy -component appearing in equation [m7]) to be negligible. Accordingly, throughout this study we consider the cumulative hydrophathic dipole field to be adequately described in terms of $\vec{h}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ only.

Figure S3: **Impact of radial structural pore symmetries on the hydrophathic dipole field.** Statistical properties of $\lambda(\mathbf{p})$ and $\kappa(\mathbf{p})$ are plotted for every $\mathbf{p} \in Q$. Error bars represent 95%-interval values of $\lambda(\mathbf{p})$ and $\kappa(\mathbf{p})$. $median(\cdot)$ and $mode(\cdot)$ return the median and the mode, respectively, of $\lambda(\mathbf{p})$ and $\kappa(\mathbf{p})$.

S7. Zero-crossing behavior of hydrophathic imbalance pore function along the principal pore axis

We here followed the same methodology as in Section S2 and calculated the set of values

$$V^{(1)}(\alpha) = \{m_z^{(1)}(\bigcup_{\alpha} \{\Gamma^{(1)}(\alpha)\})\} \quad (\text{s11})$$

As we can see in Figure S4, the distribution of $V^{(1)}(\alpha)$ exhibits similar characteristics to the distribution of $V^{(0)}(\alpha)$ (Figure S2, S4). Specifically, it is narrowly distributed around zero for every $\alpha = 10, 20, \dots, K_\alpha = 800$ and its width gradually decreases as indicated by the decreasing tendency of

the 95%-interval values for increasing α (Figure S4). Given that 95% of the values found in $V^{(1)}(K_\alpha = 800)$ are bounded within $[-0.0955, 0.08939]$ with $\text{median}(V^{(1)}(K_\alpha = 800)) = -3.7e - 05$ and $\text{mode}(V^{(1)}(K_\alpha = 800)) = -2.7e - 05$, $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ is expected to obtain a vanishingly small value when it crosses zero.

Figure S4: **Vanishing of hydropathic imbalance pore function.** Statistical properties of $V^{(1)}(\alpha)$ for $\alpha = 10, 20, \dots, K_\alpha = 800$. Error bars represent 95%-interval values of $V^{(1)}(\alpha)$. $\text{median}(\cdot)$ and $\text{mode}(\cdot)$ appearing return the median and the mode of $V^{(1)}(\alpha)$, respectively.

S8. Sinks-vs-sources field topology

A zero-crossing point of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$, $\mathbf{s}$ (see main text, eq. m10), is classified either as "sink" or "source" depending on the sign of the slope at $\mathbf{s}$. Specifically, if $\text{sign}(\frac{f(\mathbf{p}, l_\alpha(\mathbf{p})) - f(\mathbf{p}', l_\alpha(\mathbf{p}'))}{\Delta \mathbf{p}})$ returns negative (positive), then $\mathbf{s}$ is classified as "sink" ("source").

The orientation of the field at a "sink" changes from pointing-toward-the-IS to pointing-toward-the-ES along $\hat{\mathbf{z}}$ -direction so that a hypothetical molecule transported along P under the field's influence would be attracted toward $\mathbf{s}$ (Figure S5). This dynamics is reversed for the case of a "source" so that a hypothetical molecule transported along P under the field's influence would be repelled from $\mathbf{s}$ (Figure S5).

Based on the sink-vs-sources classification we can decompose the topological distribution $\phi^{(1)}$ into its additive "sink" and "source" terms, i.e., $\phi^{(1)} = \phi^{(1,-)} + \phi^{(1,+)}$ where $\phi^{(1,-)}$ and $\phi^{(1,+)}$ are the distributions of "sinks" and "sources" along P , respectively. In order to obtain a relative measure of the sink-vs-sources distribution, we considered the difference $\epsilon^{(1)} = \phi^{(1,+)} - \phi^{(1,-)}$. In particular, a positive (negative) sign of $\epsilon^{(1)}$ indicates that "sources" ("sinks") locally outnumber "sinks" ("sources"). Accordingly, positive (negative) maximization of $\epsilon^{(1)}$ identifies locations on P where cumulative field topology is shaped by "sources" ("sinks").

Figure S5: **An example of sinks-vs-sources field topology.** Trace of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ is plotted for $\alpha = 190$. The locations where $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ crosses zero and, consequently, the orientation of the field is changing, are highlighted with magenta and cyan circles. A negative and positive zero-crossing slope along $\hat{\mathbf{z}}$ -direction correspond to a "sink" and "source", respectively. Arrows around the zero-crossing locations indicate how the orientation of the field is locally changing.

Based on that, we introduced a discrete topological quasiprobability score (TQS)

$$\tau^{(1)} = \frac{\epsilon^{(1)}}{\sum_{\mathbf{p} \in Q} |\epsilon^{(1)}|} \text{ with } 0 \leq |\tau^{(1)}| \leq 1 \text{ and } \sum_{\mathbf{p} \in Q} |\tau^{(1)}| = 1 \quad (\text{s12})$$

that relaxes the notion of probability as it allows for negative values in order to distinguish between persistent, i.e., cumulative, "sinks" and "sources"

topological configurations along P . Given that "sinks" ("sources") correspond to minima (maxima) of the hydropathic imbalance potential landscape for a given α , the interplay of extrema of TQS along P provides with a qualitative description of the single-molecule permeation energy landscape in terms of the location and relative size of TQS extrema. In particular, a large and narrow positive(negative)-value maximum of TQS indicates the existence of a persistent repeller (attractor) topological configuration imposing a local energy barrier to single-molecule permeation [29] and is assigned with a positive(negative)-value probability score. If TQS attains values close to zero, then the local topology is shaped by a combination of "sinks" and "sources" indicating mixed attractors-vs-repellers dynamics. Note that we borrow this interpretation scheme from the field of computational topology and thermodynamics where similar topological approaches have been used in describing energy landscapes of complex atomic systems [60]. Finally, in order to obtain a smooth approximation of the energy landscape profile we performed a strong and weak smoothing of TQS by utilizing the built-in R [59] function that implements the Nadaraya-Watson kernel regression smoothing method for a bandwidth parameter value $bw=7$ and $bw=0.5$, respectively.

S9. Linear behavior of the hydropathic imbalance pore function

The linear trace of $\mathbf{m}_z^{(1)}(\mathbf{p}, L(\mathbf{p}))$ was best fitted with

$$\tilde{m}_z^{(1)}(\mathbf{p}, L(\mathbf{p})) = -0.00487 \cdot p_z - 0.13638 \quad (\text{s13})$$

by using the built-in R [59] function for linear regression with a residual standard fitting error of $9.3e - 29$ (see Figure S6). This is in agreement with equation (2.15) presented in [50] indicating that the displacement of the macroscopic dipole center, let it be $\mathbf{s}^{(1)} = (0, 0, s_z^{(1)})$, along P , i.e., the

vector from any $\mathbf{p}$ to $\mathbf{s}^{(1)}$ for $\alpha = K_\alpha$, is given by

$$\begin{aligned} \vec{d}(\mathbf{p}, L(\mathbf{p})) &= \frac{\vec{m}_z^{(1)}(\mathbf{p}, L(\mathbf{p}))}{m^{(0)}(\mathbf{p}, L(\mathbf{p}))} \implies \|\vec{d}(\mathbf{p}, L(\mathbf{p}))\| = \frac{\|\vec{m}_z^{(1)}(\mathbf{p}, L(\mathbf{p}))\|}{|m^{(0)}(\mathbf{p}, L(\mathbf{p}))|} \\ \implies |d_z(\mathbf{p}, L(\mathbf{p}))| &= \frac{|m_z^{(1)}(\mathbf{p}, L(\mathbf{p}))|}{|m^{(0)}(\mathbf{p}, L(\mathbf{p}))|} \implies |m_z^{(1)}(\mathbf{p}, L(\mathbf{p}))| = \quad (s14) \\ |m^{(0)}(\mathbf{p}, L(\mathbf{p}))| \cdot |d_z(\mathbf{p}, L(\mathbf{p}))| &= \begin{cases} |m_{total}^{(0)}| \cdot p_z - \gamma & \text{for } p_z \geq s_z^{(1)} \\ \gamma - |m_{total}^{(0)}| \cdot p_z & \text{for } p_z < s_z^{(1)} \end{cases} \end{aligned}$$

with $\gamma = |m_{total}^{(0)}| \cdot s_z^{(1)} = -0.1368$ as the macroscopic dipole center is found at $(\mathbf{s}^{(1)} = (0, 0, s_z^{(1)} = -28.0041))$ and $m^{(0)}(\mathbf{p}, L(\mathbf{p})) = m_{total}^{(0)} = 0.00487$ (see main text, eq. r1).

Figure S6: **Linear behavior of $m_z^{(1)}(\mathbf{p}, L(\mathbf{p}))$.** Trace of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ is plotted for $\mathbf{p} \in Q$ and for $\alpha = K_\alpha$.

Author contributions

M.N.X. conceived the idea, designed the study and performed the analysis; R.W. and P.L. helped with the refinement of algorithmic steps; R.W., P.L., D.K., Y.Y., J.H., B.H.S. and S.G.W provided with critical feedback and

helped with the interpretation of the results; Y.Y. and S.G.W encouraged M.N.X. to focus on specific aspects of the findings; B.H.S. helped with the supervision of the project; G.L. and C.G.F. were in charge of overall direction; M.N.X. wrote the manuscript in consultation with all the co-authors; PROPANE study group provided with technical and scientific support.

Acknowledgments

The study was partly funded by the European Union 7th Framework Programme (grant n602273).

Competing Interests

The authors declare that there is no conflict of interest.

References

- [1] Hille, B. *Ion channels of excitable membranes*, 3rd ed. (Sinauer Associates Inc., Sunderland MA (2001).
- [2] Catterall, W.A. Voltage-gated sodium channels at 60: structure, function and pathophysiology. *J Physiol* **590**, 2577-2589 (2012).
- [3] Cummins, T.R., Dib-Hajj, S.D., Waxman, S.G. Electrophysiological properties of mutant Nav1.7 sodium channels in a painful inherited neuropathy. *J Neurosci* **24**, 8232-8236 (2014).
- [4] Han, C., Rush, A.M., Dib-Hajj, S.D. et al. Sporadic onset of erythromelgia: a gain-of-function mutation in Nav1.7. *Ann Neurol* **59**, 553-558 (2006).
- [5] Harty, T.P., Dib-Hajj, S.D., Tyrrell, L., Blackman, R., Hisama, F.M., Rose, J.B., Waxman, S.G. Nav1.7 mutant A863P in erythromelgia: effects of altered activation and steady-state inactivation on excitability of nociceptive dorsal root ganglion neurons. *J Neurosci* **26**, 12566-12575 (2006).
- [6] Lampert, A., Dib-Hajj, S.D., Eastman, E.M., Tyrrell, L., Lin, Z., Yang, Y., Waxman, S.G. Erythromelgia mutation L823R shifts activation and inactivation of threshold sodium channel Nav1.7 to hyperpolarized potentials. *Biochem Biophys Res Commun* **390**, 319-324 (2006).
- [7] Stadler, T., O'Reilly, A.O., Lampert, A. Erythromelgia mutation Q875E stabilizes the activated state of sodium channel Nav1.7. *J Biol Chem* **290**, 6316-6325 (2015).
- [8] Yang, Y., Wang, Y., Li, S. et al. Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythromelgia. *J Med Genet* **41**, 171-174 (2004).

- [9] Dib-Hajj, S.D., Rush, A.M., Cummins, T.R. et al. Gain-of-function mutation in Nav1.7 in familial erythromelalgia induces bursting of sensory neurons. *Brain* **128**, 1847-1854 (2005).
- [10] Drenth, J.P., te Morsche, R.H.M., Guillet, G., Taieb, A., Lee Kirby, R., Jansen, J.B.M. SCN9A mutations define primary erythromelalgia as a neuro-pathic disorder of voltage gated sodium channels. *J Invest Dermatol* **124**, 1333-1338 (2005).
- [11] Lee, M.J., Yu, H.S., Hsieh, S.T. et al. Characterization of a familial case with primary erythromelalgia from Taiwan. *J Neurol* **254**, 210-214 (2007).
- [12] Drenth, J.P., Waxman, S.G. Mutations in sodium-channel gene SCN9a cause a spectrum of human genetic pain disorders. *J Clin Invest* **117**, 3603-3609 (2007).
- [13] Fertleman, C.R., Baker, M.D., Parker, K.A. et al. SCN9A mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52**, 767-774 (2006).
- [14] Jarecki, B.W., Sheets, P.L., Jackson, J.O. 2nd, Cummins, T.R. Paroxysmal extreme pain disorder mutations within the D3/S4-S5 linker of Nav1.7 cause moderate destabilization of fast inactivation. *J Physiol* **586**, 4137-4153 (2008).
- [15] Dib-Hajj, S.D., Estacion, M., Jarecki, B.W. et al. Paroxysmal extreme pain disorder M1627K mutation in human Nav1.7 renders DRG neurons hyperexcitable. *Mol Pain* **4**, 37 (2008).
- [16] Theile, J.W., Jarecki, B.W., Piekarz, A.D., Cummins, T.R. Nav1.7 mutations associated with paroxysmal extreme pain disorder, but not erythromelalgia, enhance Nav β 4 peptide-mediated resurgent sodium currents. *J Physiol* **589**, 597-608 (2011).
- [17] Faber, C.G., Hoeijmakers, J.G., Ahn, H.S. et al. Gain of function Nav1.7 mutations in idiopathic small fiber neuropathy. *Ann. Neurol.* **71**, 26-39 (2012).
- [18] Hoeijmakers, J.G., Han, C., Merkies, I.S. et al. Small nerve fibres, small hands and small feet: a new syndrome of pain, dysautonomia and acromesomelia in a kindred with a novel Nav1.7 mutation. *Brain* **135**, 345-358 (2012).
- [19] Lampert, A., O'Reilly, A.O., Dib-Hajj, S.D., Tyrrell, L., Wallace, B.A., Waxman, S.G. A Pore-blocking Hydrophobic Motif at the Cytoplasmic Aperture of the Closed-state Nav1.7 Channel Is Disrupted by the Erythromelalgia-associated F1449V Mutation. *J Biol Chem* **283**, 24118-24127 (2008).

- [20] Yang, Y., Estacion, M., Dib-Hajj, S.D., Waxman, S.G. Molecular architecture of a sodium channel S6 Helix: radial tuning of the voltage-gated sodium channel 1.7 activation gate. *J Biol Chem* **288**, 13741-13747 (2013).
- [21] Payandeh, J., Scheuer, T., Zheng, N., Catterall, W.A. The crystal structure of a voltage-gated sodium channel. *Nat* **475**, 353-358 (2011).
- [22] Yonkunas, M., Kurnikova, M. The hydrophobic effect contributes to the closed state of a simplified ion channel through a conserved Hydrophobic patch at the pore-helix crossing. *Front Pharmacol* **6**, 284 (2015).
- [23] Kitaguchi, T., Sukhareva, M., Swartz, K.J. Stabilizing the closed S6 gate in the Shaker K_v channel through modification of a hydrophobic seal. *J Gen Physiol* **124**, 319-332 (2004).
- [24] Beckstein, O., Biggin, P.C., Sansom, M.S.P. A Hydrophobic Gating Mechanism for Nanopores. *J. Phys. Chem. B* **105**, 12902-12905 (2001).
- [25] Beckstein, O., Sansom, M.S.P. Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc Natl Acad Sci USA* **100**, 7063-7068 (2003).
- [26] Allen, R., Hansen, J.-P. Molecular dynamics investigation of water permeation through nanopores. *J Chem Phys* **119**, 3905-3919 (2003).
- [27] Beckstein, O., Sansom, M.S.P. The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Phys Biol* **1**, 42-52 (2004).
- [28] Jensen, M.Ø. et al. Principles of conduction and hydrophobic gating in K^+ channels. *Proc Natl Acad Sci USA* **107**, 5833-5838 (2010).
- [29] Aryal, P., Sansom M.S.P, Tucker, S.J. Hydrophobic Gating in Ion Channels. *J Mol Biol* **427**, 121-130 (2015).
- [30] Israelachvili, J., Pashley, R.M. The hydrophobic interaction is long range, decaying exponentially with distance. *Nat* **300**, 341-342 (1982).
- [31] Hammer, M.U., Anderson, T.H., Chaimovich, A., Shell, M.S., Israelachvili, J. The search for the hydrophobic force law. *Faraday Discuss* **146**, 299-308 (2010).
- [32] Tabor, R.F., Grieser, F., Dagastine, R.R., Chan, D.Y.C. The hydrophobic force: measurements and methods. *Phys Chem Chem Phys* **16**, 18065-18075 (2014).

- [33] Eisenberg, D., Weiss, R.M., Terwilliger, T.C., Wilcox, W. Hydrophobic moments and protein structure. *Faraday Symp Chem Soc* **17**, 109-120 (1982).
- [34] Reißer, S., Strandberg, E., Steinbrecher, T., Ulrich, A.S. 3D hydrophobic moment vectors as a tool to characterize the surface polarity of amphiphilic peptides. *Biophys J* **106**, 2385-2394 (2014).
- [35] Meirovitch, H., Scheraga, H.A. Empirical studies of hydrophobicity. 3. Radial distribution of clusters of hydrophobic and hydrophilic amino acids. *Macromolecules* **14**, 340-345 (1981).
- [36] Silverman, B.D. Hydrophobic moments of protein structures: Spatially profiling the distribution. *Proc Natl Acad Sci USA* **98**, 4996-5001 (2001).
- [37] Zhou, R., Silverman, B.D., Royyuru, A.K., Athma, P. Spatial profiling of protein hydrophobicity: Native vs. decoy structures. *Proteins* **52**, 561-572 (2003).
- [38] Silverman, B.D. Hydrophobicity of transmembrane proteins: Spatially profiling the distribution. *Protein Sci* **12**, 586-599 (2003).
- [39] Silverman, B.D. Hydrophobic Moments of Tertiary Protein Structures. *Proteins* **53**, 880-888 (2003).
- [40] Doyle, D.A., Cabral, J.M., Pfuetzner, R.A. The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science* **280**, 69-77 (1998).
- [41] DeLano, W.L. PyMOL. DeLano Scientific, San Carlos, CA, 700 (2002).
- [42] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **13**, 1605-12 (2004).
- [43] Vriend, G. WHAT IF: A molecular modeling and drug design program. *J Mol Graph* **8**, 52-56 (1980).
- [44] Hooft, R.W., Sander, C., Vriend, G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins* **26**, 363-376 (1996).
- [45] Humphrey, W., Dalke, A., Schulten, K. VMD-Visual Molecular Dynamics. *J Molec Graph* **14**, 33-38 (1996).
- [46] Smart, O.S., Neduvetil, J.G., Wang, X., Wallace, B.A., Sansom, M.S.P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* **14**, 354-360 (1996).

- [47] Kapcha, L.H., Rossky, P.J. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J Mol Biol* **426**, 484-498 (2014).
- [48] Freitas, J.A., Tobias, D.J. Voltage sensing in membranes: From macroscopic currents to molecular motions. *J Membr Biol* **3**, 419-430 (2015).
- [49] Westra, R.L. Resonance-driven ion transport and selectivity in prokaryotic ion channels. *Phys Rev E* **100**, 062410 (2019).
- [50] Silverman, B.D. Three-dimensional moments of molecular property fields. *J Chem Inf Comput Sci* **40**, 1470-1476 (2000).
- [51] Ragsdale, D.S., McPhee, J.C., Scheuer, T., and Catterall, W.A. (1994) Molecular determinants of state-dependent block of Na⁺ channels by local anesthetics. *Science* **265**, 1724-1728 (1994).
- [52] Catterall, W.A. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron* **26**, 13-25 (2000).
- [53] Yarov-Yarovoy, V., McPhee, J.C., Idsvoog, D., Pate, C., Scheuer, T., Catterall, W.A. Role of amino acid residues in transmembrane segments IS6 and IIS6 of the Na⁺ channel α subunit in voltage-dependent gating and drug block. *J Biol Chem*, **277**, 35393-35401 (2002).
- [54] Heinemann, S.H., Terlau, H., Stühmer, W., Imoto, K., Numa, S. Calcium channel characteristics conferred on the sodium channel by single mutations. *Nat* **356**, 441-443 (1992).
- [55] Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., Ben-Tal, N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucl Acids Res* **33**, W299-W302 (2005).
- [56] Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., Ben-Tal, N. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics* **19**, 163-164 (2003).
- [57] Forrest, L.R. Structural symmetry in membrane proteins. *Annu. Rev. Biophys.* **44**, 311-337 (2015).
- [58] Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'Brien, R., Sturtevant, J.M., Regan, L. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Science* **5**, 1584-1593 (1996).

- [59] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2014)
<http://www.R-project.org/>.
 - [60] Carr, J.M., Mazauric, D., Cazals, F., Wales, D.J. Energy landscapes and persistent minima. *J Chem Phys* **144**, 054109 (2016).
-

Chapter 3

Non-extensivity and criticality of atomic hydrophobicity around a voltage-gated sodium channel's pore; a modeling study

Markos N. Xenakis, Dimos Kapetis, Yang Yang, Jordi Heijman, Stephen G. Waxman, Giuseppe Lauria, Catharina G. Faber, Hubert J. Smeets, Patrick J. Lindsey, Ronald L. Westra

J Biol Phys, Volume 4, Issue 1, Pages 61-77, March 2021

Introduction

The main function of voltage-gated sodium channels (NaVChs) is to control initiation and propagation of electrical signals along neuronal tissues [1]. Enhancing our mechanistic understanding of NaVChs physics is important not only for biology and protein science but also for medicine. In humans, mutations in NaVCh-encoding genes can perturb structural and functional properties of the expressed protein thus leading to abnormal electrical cellular signaling. These abnormalities may sum up and manifest at an organism level as ion channel diseases known as "channelopathies" [2]. Numerous studies have shown that mutations in the *textitSCN9A* gene encoding the Nav1.7 channel are causally related to inherited forms of chronic pain [3–18], cardiac arrhythmias [19, 20] and epilepsy [21].

The structural motif of a NaVCh is that of a tetramer where each of the four subunits comprises a pore domain (PD) interlinked with a voltage-sensing domain (VSD). This conserved architecture facilitates a central pore through which sodium ions are conducted with high specificity. Barriers imposed to ion transport through the pore stem from the energy required for dehydrating them, i.e., removing water molecules from their hydration shell [22–24]. Dehydration of ions is achieved via interaction with the pore walls; ions have to fit closely in the pore's confined geometry so that removed waters are immediately replaced by waters provided by the walls [22]. It is hence the interplay of geometrical and hydrophobic (i.e., hydrophobic or hydrophilic) characteristics along a NaVCh's pore essentially determines its "gating" [25–30] behavior, i.e., its permeability to specific ions. For example, in the case of the pre-open NaVAb bacterial channel [31], a hydrophobic and wide cavity favors rapid diffusion of hydrated ions while a hydrophilic and narrow pore region acts as a selectivity filter (SF) by binding with and partially-dehydrating ions. The crucial role of hydrophobic effects is also manifested at scales larger than these of the pore's microenvironment as they contribute to a NaVCh's stability (e.g., via

formation of interaction networks between pore-lining structures and their surroundings [32, 33]). However, despite the plethora of NaVCh structures available, modeling the synergy of hydrophobic effects across different NaVCh scales remains an open problem. This is due to the spatial complexity underlying hydrophobic interactions [34, 35] posing significant challenges in development of suitable theoretical vehicles and computational tools for multiscale hydrophobic analysis of protein structures.

Non-extensive systems [36] are characterized by thermodynamic functions that do not scale in a linear fashion for increasing system size. A hallmark of non-extensivity is the emergence of power-law scalings reflecting the long-range nature of interactions among system components. In this study we test the hypothesis of non-extensivity for the hydrophobic property of a NaVCh protein system. Prior knowledge supporting this hypothesis is; (a) hydrophobic effects are long-range with underlying interactions laws varying across different scales [37], (b) system-size-dependent, i.e., cumulative, hydrophobic characteristics of globular proteins exhibit universality [38], (c) spatial inhomogeneities of the atomic mixture around a NaVCh's pore can generate non-random patterns and topologies [39]. Accordingly, we present here a computational framework for modeling cumulative hydrophobic characteristics of the atomic environment around a NaVCh's pore. For that, we utilized the finite-size scaling analysis methodologies presented in [39] and embedded them into a thermodynamic framework.

In continuation to [39], we focus on the pre-open NaVAb channel [31]. We first show that accumulation of atoms around its pore can be accurately described in terms of the Richards model [40] that is known to admit an algebraic q -generalization [41]. A direct implication of this finding is the existence of evolutionary-conserved inflection points matching closely the geometrical crossover from the pore domains to the voltage-sensing domains. Thermodynamic properties of the pore were then investigated in terms of the hydrophobic imbalance

effect [42]. For that, we analyzed the pre- and post-inflection scaling behavior of a hydrophobic imbalance pore function [39] inferring about system-size-dependent changes in interatomic hydrophobic interactions. This yielded a bi-phasic and scale-invariant hydrophobic interactions model applying with high accuracy to the extracellular side of the NaVAb’s SF. Although hydrophobic scaling mechanisms in proteins have been previously investigated (e.g., see [43–45]), our work provides some first insight on how these phenomena emerge in a prototype NaVCh system and might be contributing to energetic optimization of its functional architecture.

Methods

3D structure preparation

We employed the pre-open I217C NaVAb structural model (PDB code: 3RVY) providing with a crystallography of the NaVAb at the highest available resolution (2.7 Å resolution). Protonation of the structure was performed with the WHAT IF software [46, 47] and the principal axes of the protonated structure were estimated with the VMD software [48]. Next, the protonated structure was embedded into a coordinate system $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ with origin O so that its principal pore axis, i.e., the axis approximating the direction of the channel’s pore, was aligned with the z -axis with orientation from the extracellular side (ES) to the intracellular side (IS) with respect to $\hat{\mathbf{z}}$. The molecular mass center $\mathbf{e} = \frac{1}{M} \sum_{i=1}^{N_c} m_i \cdot \mathbf{c}_i$ of the structure was set to coincide with O , where $\mathbf{c}_i = (c_{x,i}, c_{y,i}, c_{z,i})$ is the atomic center of the i -th atom, m_i is the mass of the i -th atom, $N_c = 14776$ is the total number of atoms and $M = \sum_{i=1}^{N_c} m_i$ is the total molecular mass (see Dissertation Appendices, Table A1 for atomic masses values). An illustration of the pre-open I217C NaVAb structural model is presented in Supplementary Information (SI), Section S1.

Geometrical characteristics of the pore

Let us consider P to represent the principal pore axis with $\mathbf{p} = (p_x, p_y, p_z) \in Q \subset P$ being a pore point (see SI Section S2 for construction of Q). Then, the pore radius at $\mathbf{p}$ is given by [49]

$$R(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| - vdW_i \} \quad (\text{m1})$$

where $\|\cdot\|$ is the euclidean norm and vdW_i is the Van der Waals radius of the i -th atom (see Appendices, Table A1) and the distance between $\mathbf{p}$ and its nearest neighbor atom corresponds to

$$D(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| \} \quad (\text{m2})$$

Moreover, the outer surface radius at $\mathbf{p}$ is given by [39]

$$L(\mathbf{p}) = \max_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| + vdW_i \} \quad (\text{m3})$$

where the unit of measurement for $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ is expressed in $[\text{\AA}]$.

Atomic sampling around the pore

We consider nested sampling spheres centered at $\mathbf{p}$ with radii given by [39]

$$l_\alpha(\mathbf{p}) = D(\mathbf{p}) + \alpha \cdot \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha} \text{ for } \alpha = 1, 2, \dots, K_\alpha \rightarrow \infty \quad (\text{m4})$$

where K_α is the total number of sampling spheres and α denotes the index of the sampling sphere (thus, from now on, we will refer to it as "scaling index"). Accordingly, $l_\alpha(\mathbf{p})$ indicates the size (or, equivalently, the scale) of the spherical cluster of atoms around $\mathbf{p}$ in $[\text{\AA}]$. Then, the number of atoms around $\mathbf{p}$ is given by [39]

$$N(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \quad (\text{m5})$$

where $\theta(\cdot)$ is the heaviside function.

Candidate models of atomic accumulation

Atomic accumulation around $\mathbf{p}$ was modeled by employing the GROFIT routine [50] and applying it on equation m5. Accordingly, a collection of candidate models including re-parametrized algebraic forms [51] of the Logistic model [52]

$$n_{LOG}(\mathbf{p}, l_{\alpha}(\mathbf{p})) = A(\mathbf{p}) \cdot \{1 + \exp\left(\frac{4 \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_{\alpha}(\mathbf{p})) + 2\right)\}^{-1} \quad (\text{m6})$$

, of the Gompertz model [53]

$$n_{GOM}(\mathbf{p}, l_{\alpha}(\mathbf{p})) = A(\mathbf{p}) \cdot \exp\left(-\exp\left(\frac{e \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_{\alpha}(\mathbf{p})) + 1\right)\right) \quad (\text{m7})$$

with $e = \exp(1)$, of the modified Gompertz model [50]

$$\begin{aligned} n_{MGOM}(\mathbf{p}, l_{\alpha}(\mathbf{p})) = & A(\mathbf{p}) \cdot \exp\left(-\exp\left(\frac{e \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_{\alpha}(\mathbf{p})) + 1\right)\right) \\ & + A(\mathbf{p}) \cdot \exp(w(\mathbf{p}) \cdot (l_{\alpha}(\mathbf{p}) - l_{shift}(\mathbf{p}))) \end{aligned} \quad (\text{m8})$$

and of the Richards model [40]

$$\begin{aligned} n_{RIC}(\mathbf{p}, l_{\alpha}(\mathbf{p})) = & A(\mathbf{p}) \cdot \{1 + \tilde{q}(\mathbf{p}) \cdot b(\mathbf{p}) \cdot \exp(-d(\mathbf{p}) \cdot l_{\alpha}(\mathbf{p}))\}^{-1/\tilde{q}(\mathbf{p})} \\ \text{with } b(\mathbf{p}) = & \exp(1 + \tilde{q}(\mathbf{p}) + d(\mathbf{p}) \cdot s(\mathbf{p})) \\ \text{and } d(\mathbf{p}) = & \frac{t(\mathbf{p})}{A(\mathbf{p})} \cdot (1 + \tilde{q}(\mathbf{p}))^{1+1/\tilde{q}(\mathbf{p})} \end{aligned} \quad (\text{m9})$$

were fitted on $N(\mathbf{p}, l_{\alpha}(\mathbf{p}))$ along $l_{\alpha}(\mathbf{p})$ -direction where $\{A(\mathbf{p}), t(\mathbf{p}), s(\mathbf{p}), \tilde{q}(\mathbf{p}), w(\mathbf{p}), l_{shift}(\mathbf{p})\}$ are model parameters.

In order to proceed with model parameters interpretation we introduced the inflection point for an atomic accumulation model with

$$\xi(\mathbf{p}) = \{l_\alpha(\mathbf{p}) \mid \frac{\partial^2 n(\mathbf{p}, l_\alpha(\mathbf{p}))}{\partial l_\alpha(\mathbf{p})^2} = 0\} \quad (\text{m10})$$

that determines the location along $l_\alpha(\mathbf{p})$ -direction where the atomic radial distribution function (RDF), $\frac{\partial n(\mathbf{p}, l_\alpha(\mathbf{p}))}{\partial l_\alpha(\mathbf{p})}$, is maximized. Then, parameter $t(\mathbf{p})$ is interpreted as the slope of $n(\mathbf{p}, l_\alpha(\mathbf{p}))$ at $\xi(\mathbf{p})$ thus accounting for the maximum atomic accumulation rate (or, equivalently, for the maximum atomic density) around $\mathbf{p}$. Note that $t(\mathbf{p})$ can be written as $t(\mathbf{p}) = \frac{A(\mathbf{p})}{os(\mathbf{p}) := o(\mathbf{p}) - s(\mathbf{p})}$ with $o(\mathbf{p})$ determining the location along $l_\alpha(\mathbf{p})$ -direction where the asymptote domain begins as $A(\mathbf{p})$ is the asymptote value of the fitted model, i.e., $n(\mathbf{p}, l_\alpha(\mathbf{p}) \rightarrow \infty) = A(\mathbf{p})$, and with $s(\mathbf{p})$ determining the location along $l_\alpha(\mathbf{p})$ -direction where the lag domain ends, i.e., the size of the lag atomic accumulation domain (in analogy to the "lag phase" parameter appearing in [51]). Parameter $\tilde{q}(\mathbf{p})$ determines the shape of the Richards model curve, as well as, the location of the inflection point along $l_\alpha(\mathbf{p})$ -direction thus serving as a "summary parameter". Finally, parameters $w(\mathbf{p})$ and $l_{shift}(\mathbf{p})$ of the modified Gompertz model determine the location and the slope, respectively, of a second increase in the modified Gompertz model curve (see [50]). Note that the Logistic and the Gompertz model are special cases of the Richards model for $\tilde{q}(\mathbf{p}) = 1$ and $\tilde{q}(\mathbf{p}) \rightarrow 0$, respectively, as shown in [54]. For a graphical interpretation of Richards model parameters see SI Section S3.

Thermodynamic modeling of the pore

Two types of interactions are relevant for NaVChs; (a), an interaction of the atomic structure with itself, i.e., a stabilizing interaction, and, (b), an interaction of the atomic structure with a pore-permeating species of hydration radius $\approx R(\mathbf{p})$ which can "fit" in the pore. A

molecular index that can efficiently summarize these interactions is the hydrophathic dipole moment [55]. Accordingly, we utilize the cumulative hydrophathic dipole field (CHDF) at $\mathbf{p}$, i.e., the hydrophathic dipole field effect at $\mathbf{p}$ originating from the spatial arrangement of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around it, given by [39]

$$\begin{aligned} \vec{\mathbf{h}}(\mathbf{p}, l_\alpha(\mathbf{p})) = & \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \cdot HI_i^\chi \cdot \vec{\mathbf{r}}_{\mathbf{p},i} = \\ & \underbrace{h_x(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{x}} + h_y(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{y}}}_{\vec{\mathbf{h}}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))} + \underbrace{h_z(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{z}}}_{\vec{\mathbf{h}}_z(\mathbf{p}, l_\alpha(\mathbf{p}))} \end{aligned} \quad (\text{m11})$$

where $HI_i^\chi = HI_i + \chi_i$ is the i -th atomic hydrophathic index corresponding to the Kapcha-Rossky atomic hydrophathic indices [56] (see Appendices, Table A2) with additive Gaussian noise $\chi_i \in \mathcal{N}(\mu=0, \sigma=0.001)$ and $\vec{\mathbf{r}}_{i,p}$ is a vector from $\mathbf{p}$ to $\mathbf{c}_i$. Introduction of the weak noise source χ_i practically guarantees that $\|\vec{\mathbf{h}}(\mathbf{p}, l_\alpha(\mathbf{p}))\|$ is non-zero for every combination of $\mathbf{p}$ and $l_\alpha(\mathbf{p})$ while not affecting its spatial behavior. Moreover, the measurement unit of $\vec{\mathbf{h}}(\mathbf{p}, l_\alpha(\mathbf{p}))$ is given in Debye units [kcal·Å/mol] [55] with 1 mol representing a cluster of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms.

Due to preservation of structural NaVAb symmetries with respect to P , the magnitude of the radial field component, $\|\vec{\mathbf{h}}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))\|$, is vanishingly small and, consequently, $\|\vec{\mathbf{h}}(\mathbf{p}, l_\alpha(\mathbf{p}))\|$ is adequately represented in terms of the magnitude of the axial field component, $\|\vec{\mathbf{h}}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\|$, i.e.,

$$\|\vec{\mathbf{h}}(\mathbf{p}, l_\alpha(\mathbf{p}))\| = \sqrt{\|\vec{\mathbf{h}}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))\|^2 + \|\vec{\mathbf{h}}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\|^2} \approx \|\vec{\mathbf{h}}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\| \quad (\text{m12})$$

where a computational proof for the validity of m12 is given in SI of [39]. Consequently, the orientation of the field is described by only two states; an "in" state which is characterized by $\vec{\mathbf{h}}_z(\mathbf{p}, l_\alpha(\mathbf{p}))$

pointing towards the IS, i.e., by $h_z(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$, and an "out" state which is characterized by $\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))$ pointing towards the ES, i.e., by $h_z(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ (for a detailed analysis of the CHDF topology emerging from the "in"-vs-"out" interplay across different scales is presented in [39]).

Testing the non-extensivity hypothesis for atomic hydrophobicity around NaVAb's pore

The hydrophobic non-extensivity hypothesis was tested by investigating how the CHDF behaves relative to the atomic number, $N(\mathbf{p}, l_\alpha(\mathbf{p}))$, for increasing molecular scale $l_\alpha(\mathbf{p})$. For that, we scrutinized the scaling behavior of the hydrophobic imbalance [42] (or, equivalently, of the interatomic hydrophobic interaction strength) given by [39]

$$I(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{||\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))||}{N(\mathbf{p}, l_\alpha(\mathbf{p}))} \text{ [kcal}\cdot\text{\AA/mol]} \quad (\text{m13})$$

by searching for sets of pore points for which $I(\mathbf{p}, l_\alpha(\mathbf{p}))$ can be accurately approximated by a power-law function along $l_\alpha(\mathbf{p})$ -direction. Note that 1 *mol* in equation m13 represents 1 *atom*. Interpretation of the findings was based on the rationale that a power-law scaling behavior of $I(\mathbf{p}, l_\alpha(\mathbf{p}))$ implies that the interatomic hydrophobic interaction laws around $\mathbf{p}$ remain invariant under scale transformations or, equivalently, its range spans up to $\approx l_\alpha(\mathbf{p})$ Å. Accordingly, a positive (negative) power-law exponent indicates that the range of hydrophobic interatomic interactions increases (decreases), i.e., quantifies the rate at which non-local (local) bonds are formed within the atomic structure. The interaction energy associated with bond formation stabilizing the ensemble of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around $\mathbf{p}$, i.e., the atom-packing energy (AE), is then be approximated by $||\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))||/l_\alpha(\mathbf{p})$ measured in [kcal/mol].

Results

Prior knowledge required in order to interpret the findings of this study is retrieved from [39] and summarized in Table R1.

Pore region	Location	Geometry	Hydropath
selectivity filter (SF)	$p_z = -11.7$	narrow	hydrophilic
central cavity (CC)	$p_z = 2.8$	wide	hydrophobic
activation gate (AG)	$p_z = 22.7$	occluded	hydropathically-diverse

Table R1: **Geometrical and hydropathic characteristics of the pore’s microenvironment.** We present a summary of geometrical and hydropathic characteristics of the selectivity filter, central cavity and activation gate pore regions in accordance to [31, 39].

Bi-phasic structural organization of the atomic environment around NaVAb’s pore

In this section, we present a mathematical model for describing atomic accumulation around NaVAb’s pore. This modeling approach is based on the ansatz that the scaling behavior of the atomic number (see eq. m5 and Figure 1(a) for an illustration), corresponds to a sigmoid function with inflection points correlating with the evolutionary-conserved NaVCh architectural motif that dictates the structural and functional separation of the pore domains (PDs) from the voltage-sensing domains (VSDs). The validity of this hypothesis was tested by executing the GROFIT [50] algorithm on the normalized atomic number traces (see caption of Figure 1). The mathematical model that best fitted normalized atomic number traces along $l_\alpha(\mathbf{p})$ -direction was selected from a pool of candidate accumulation models (see Methods, Section 4) based on minimization of an Akaike information criterion (see [50] for algorithmic details).

Our modeling procedures demonstrate that atomic accumulation around NaVAb’s pore can be accurately described by the Richards

model (see eq. m9) and special cases of it, such as the Logistic model (see eq. m6) and the Gompertz model (see eq. m7). This becomes obvious if we focus on the trace of the summary model parameter, $\tilde{q}(\mathbf{p})$, approaching unity and zero for $\{p_z \rightarrow -5.5^-, p_z \rightarrow 1.2^+\}$ and for $\{p_z \rightarrow 10.0^-, p_z \rightarrow 10.9^+\}$ thus indicating a smooth transition from the Richards model pore regions towards the Gompertz and Logistic model pore regions, respectively. Note that discontinuities arising at the boundaries $[5.5, 1.2]$ and $[10.0, 10.9]$ are due to numerical artifacts introduced by the GROFIT solver (Figure 1(b)). The Gompertz and Logistic model pore regions are located on the left and right of the hydrophobic central cavity (CC) thus capturing the transition from the pore center towards the SF and the AG (see Table 1 and Figure 1(b)). The accuracy of the Richards model modeling approximation is demonstrated not only by the smallness of absolute fitting error but also by the small uncertainties in the estimation of model parameter (see Figure 1(c)).

A direct consequence of the Richards model applicability is that the atomic environment around the pore can be partitioned into three consecutive geometrical domains, namely, a lag domain (for $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$), an inflection domain (for $s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq o(\mathbf{p})$) and an asymptote domain (for $l_\alpha(\mathbf{p}) > o(\mathbf{p})$) (Figure 1(a) and SI Section S3). Equivalently, we may consider only two phases, i.e., a pre-inflection and a post-inflection phase realized for $l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ and $l_\alpha(\mathbf{p}) > \xi(\mathbf{p})$, respectively (see eq. m10 and Figure 1(a)). Size variations of the inflection domain described by $os(\mathbf{p})$ (see Methods) indicate how atomic density is changing along P . In particular, within the Gompertz model pore region the size of the inflection domain maximizes while the atomic accumulation rate, $t(\mathbf{p})$ (see Methods), is down-regulated so that a weakly-dense atomic environment is formed at the extracellular side (ES) of the CC (Figure 1(a),(b)). In contrast, deviating from the Gompertz model pore region towards the intracellular side (IS), $os(\mathbf{p})$ decreases while $t(\mathbf{p})$ increases so that a strongly-dense atomic environment is formed around the AG implementing pore occlusion (Figure

1(a),(b)). This qualitative change in atom-packing conditions is captured by the Logistic model pore region as the summary parameter, $\tilde{q}(\mathbf{p})$ (see Methods), becomes larger (smaller) than one for $p_z > 10.9$ ($p_z < 10.0$) (Figure 1(b)). Deviating from the Gompertz model pore region towards the extracellular side (ES), both $os(\mathbf{p})$ and $t(\mathbf{p})$ exhibit a weakly-decreasing tendency indicative of a structural tendency for uniform atom packing conditions.

Figure 1: **Mathematical modeling of atomic accumulation around NaVAb's pore.** (a), Contour map of the normalized traces $\tilde{N}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in Q$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict pore's geometrical characteristics. Magenta dashed line $\nu(\mathbf{p})$ serves as a geometrical representation of the geometrical crossover from the PDs to the VSDs. Line $\xi(\mathbf{p})$ depicts the trace

of inflection points, and lines $s(\mathbf{p})$ and $o(\mathbf{p})$ depict the ending and beginning of the lag and asymptote domain, respectively. **(b)**, Traces of normalized model parameters $\bar{A}(\mathbf{p})$, $\bar{t}(\mathbf{p})$, $\bar{o}s(\mathbf{p})$ and of summary model parameter $\tilde{q}(\mathbf{p})$ for $\mathbf{p} \in Q$. **(c)**, MAFE of the mathematical accumulation model approximation on $\bar{N}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in Q$. Normalizations were performed with respect to the maximum values of the corresponding traces.

Biophysical importance of the inflection mechanism becomes evident if we consider how our observations correlate with the spatial organization of the PDs relatively to the VSDs. For that, we approximated the scales at which the spatial transition, i.e., the geometrical crossover, from the PDs to the VSDs takes place (see SI Section S4). We found that the traces of inflection points and of the PDs-VSDs geometrical crossover follow closely each other with their mutual distance on the contour map of Figure 1(a) given by $1.82 \pm 0.94 \text{ \AA}$ and with Spearman correlation scores of 0.92 and 0.99 for $p_z > 0$ and $p_z \leq 0$, respectively. Hence, maximization of atomic density around the pore implements the structural separation of the PDs from the VSDs or, equivalently, splits the atomic environment into a pre-inflection and post-inflection phase accounting for the coupled sub-structures of the PDs and of the VSDs, respectively. Densening of atoms for $l_\alpha(\mathbf{p}) \approx \xi(\mathbf{p})$ can result in strengthening of interatomic interactions which, in turn, favors coupling and relative stability of the PDs and VSDs atomic sub-systems. Experimental evidence for this kind of mechanism in voltage-gated potassium channels is presented in [57–60].

Non-extensivity of atomic hydrophobicity around NaVAb's pore

What do observations of the previous section imply for the nature of underlying interatomic stabilizing interactions? To answer this question let us remind here that stability of atom-packing around protein

cores is largely determined by hydrophobic effects [61]. Following this rationale, we examined the scaling behavior of the hydrophobic inter-atomic interactions strength (HIIS) (see eq. m13) with respect to the pre- and post-inflection geometrical intervals which essentially tests the validity of the non-extensivity hypothesis.

Coarse-grained computational evidence for the validity of the non-extensivity hypothesis is presented in Figure 2. We show that maximization of HIIS is expected to take place within the narrow interval $\nu(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ thus identifying the PDs-VSDs geometrical crossover as a strong-interactions intra-channel site. Moreover, the spatial transition from the lag domain to the inflection domain is associated with a local minimization of HIIS followed by its up-regulation and, eventually, maximization up to $l_\alpha(\mathbf{p}) \approx \xi(\mathbf{p})$ (Figure 2). Below this regime, HIIS is down-regulated toward its global minimization at the end of the asymptote domain (see caption of Figure 2). These observations indicate a "tuning" of HIIS scaling behavior with respect to the structural classification scheme presented in Section 1. From a stability viewpoint, the main advantage of this configuration is that the HIIS between the PDs and the VSDs is maximized within a narrow window around $\nu(\mathbf{p})$ thus optimizing the PDs-VSDs coupling. Within the lag domain our modeling procedures appear to be much less efficient in describing cumulative atomic properties as indicated by the increase in deviations between fitted models and empirical NaVAb data (Figure 2(a),(b)). This modeling impotence indicates that within lag-domain scales, HIIS scaling analysis suffers from small-size sampling which, in turn, suggests the existence of a cut-off scale $l_\alpha(\mathbf{p}) \approx s(\mathbf{p})$ above which cumulative hydrophobic effects become an important driving force of atomic accumulation. Equivalently, within the lag-domain HIIS exhibits non-trivial short-range dependencies which might originate from water structuring effects at the wall surfaces prevailing within scales smaller than 15 Å [37]. However, given that HIIS exhibits a pseudo-symmetric spatial profile with respect to the center of the pore coinciding with the CC center, hydrophobic effects orig-

inating from scales smaller than $s(\mathbf{p})$ have a major contribution to pore gating behavior and, hence, to ion de-/re-hydration dynamics as demonstrated in [39].

Figure 2: Non-extensivity of atomic hydrophobicity around the NaVAb's pore. (a), Trace of a statistical representation of the normalized atomic number, $\langle \tilde{N}(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, and of its best-fitted model, $\langle n(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, are plotted in log-scale. $\langle n(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$ corresponds to the Richards model. (b), Trace of a statistical representation of the hydrophobic imbalance magnitude, $\langle I(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, is plotted with its linear fittings for $\langle \alpha_s \rangle < \alpha \leq \langle \alpha_\xi \rangle$ and $\alpha > \langle \alpha_\xi \rangle$, respectively, in log-scale. Up- and down-regulation tendencies were quantified in terms of the linear fittings $\gamma \cdot \log[\alpha] + \beta$ with $\gamma_{pre} = 1.41$ and $\beta_{pre} = -8.4$, $\gamma_{post} = -2.76$ and $\beta_{post} = 16.43$. The corresponding Pearson coefficients are $PC_{pre} = 0.99$ and $PC_{post} = -0.97$, respectively, indicating the "goodness" of the regulation. Error bars in (a) and (b) represent 95%-interval values. For clarity, the size of error bars in (b) is reduced

by a factor of 0.5. Calculation of $\langle \bar{N}(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, $\langle n(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, $\langle I(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, $\langle \alpha_s \rangle$, $\langle \alpha_\xi \rangle$, $\langle \alpha_\nu \rangle$ and $\langle \alpha_o \rangle$ was performed according to the statistical scheme presented in SI Section S5. All normalizations were performed with respect to the maximum values of the corresponding traces.

Fine-grained analysis of the scaling behavior of HIIS revealed a clear enhancement of its scale-invariant character at the ES of the highly-conserved SF (see SI Section S6 and Figure 3(a),(b)). Specifically, out of all possible atomic accumulation paths initiated along P , scale-invariance occurs only for $-18.0 \leq p_z \leq -16.5$ corresponding to a pore region which is directly-neighbored by the hydrophobic M181 residue side chain belonging to the P2-helix, as well as, by the hydrophilic S178 residue side chains belonging to the SF (see Figure 3(a) and [39] for direct-neighboring residues to $\mathbf{p}$). From the perspective of a pore-permeating ion species, this atom-packing configuration implies a large-amplitude, short-range interaction exerted upon incoming ions from NaVAb's walls as indicated by large values of HIIS for $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$ (Figure 3(b)). Indeed, the SF residue sequence T175:W179 is part of this atomic configuration as the sampling radius $s(\mathbf{p})$ encloses its structure as a whole (Figure 3(a)). Below the lag scales, the it is straightforward to approximate the amplitude of the CHDF (see m12) for $-18.0 \leq p_z \leq -16.5$ with the following power-law scheme

$$\begin{aligned}
 ||\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))|| = & \\
 \begin{cases} -h_z(\mathbf{p}, l_\alpha(\mathbf{p})) \sim n_{RIC}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot l_\alpha(\mathbf{p})^{\gamma_{pre}(\mathbf{p})} & \text{for } s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p}) \\ -h_z(\mathbf{p}, l_\alpha(\mathbf{p})) \sim n_{RIC}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot l_\alpha(\mathbf{p})^{\gamma_{post}(\mathbf{p})} & \text{for } l_\alpha(\mathbf{p}) > \xi(\mathbf{p}) \end{cases} & \quad (r1)
 \end{aligned}$$

with $\gamma_{pre}(\mathbf{p}) = 2.27 \pm 0.1$ and $\gamma_{post}(\mathbf{p}) = -4.36 \pm 0.09$ being the pre- and post-inflection scaling exponents (see Figure 3(b), and SI Figure S6), and $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ replaced with its best-fitted model, namely, the Richards model for parameters $\{A(\mathbf{p}) \approx 1.05, t(\mathbf{p}) \approx 0.2, s(\mathbf{p}) \approx$

18.02, $\tilde{q}(\mathbf{p}) \approx 0.54\}$. Topological invariance of $\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))$ (as indicated by the negative sign of $h_z(\mathbf{p}, l_\alpha(\mathbf{p}))$ at the rhs of r1) allows for an interpretation of the observed scaling laws. The atom-packing energy (AE) associated with this CHDF configuration is retrieved by dividing r1 with $l_\alpha(\mathbf{p})$ (see Figure 3(c) and its caption for the accuracy of the approximation). This estimation scheme predicts a PDs-VSDs coupling energy of ~ 282.1 kcal as AE maximizes within the narrow window $[\nu(\mathbf{p}), \xi(\mathbf{p})]$ (Figure 3(c)). Microscopic building-blocks of the CHDF are dipole-dipole interactions (e.g., hydrogen bonds) formed between neighboring atoms. Each of these short-range interactions can be assigned with a dipole moment. The physical meaning attributed then to equation r1 is that the mean-field amplitude and orientation of these dipole moments remain invariant for increasing system-size. Consequently, in the pre-inflection phase, a long-range stabilizing effect emerges reflecting a non-random alignment of dipoles with respect to $-18.0 \leq p_z \leq -16.5$ across different NaVAb scales. Equivalently, we may think of this stabilizing mechanism as an interactions-network [31] expanding throughout the PDs with a rate that is quantified in terms of $\gamma_{pre}(\mathbf{p})$. The E177 side chains were previously identified as the interactions-network center as they form bonds with their direct atomic environment which extend throughout the P- and P2-helices and, most likely, throughout the PDs as a whole (see SI Figure 10 of [31]).

Our findings closely match this anticipation although they predict a network-initiator role for a pore-lining cluster of M181 and S178 atomic components found at the ES of the E177. In the post-inflection phase, spatial dynamics is reversed as the network contracts with a rate of $|\gamma_{post}(\mathbf{p})|$ indicative of short-rangeness rebound and energy dissipation due to bond breaking. Notably, $|\gamma_{post}(\mathbf{p})|$ is almost double the size as $\gamma_{pre}(\mathbf{p})$ indicating that breaking of long-range bonds in favor of short-range ones is a "faster" process than creating them.

Figure 3: Non-extensive modeling of the interatomic hydropathic interactions around NaVAb’s pore. (a), Cartoon illustration of a single PD structural unit illustrating packing of atoms around the pore region $p_z \in [-18.0, -16.5]$. The radii $s(p)$ and $\xi(p)$ centered at the middle of $p_z \in [-18.0, -16.5]$ roughly account for the scales at which the lag domain and the pre-inflection domains end, i.e., for the sizes of the lag and pre-inflection domain, respectively. Residues T175, L176, E177, S178 and W179 forming the selectivity filter are colored according to their hydropathic score based on the Kapcha-Rossky scale. $p_z \in [-18.0, -16.5]$ is directly neighbored by atomic components forming the M181 and S178 residue side chains. (b), Trace of HIIS for a randomly-chosen pore point from the pore region $p_z \in [-18.0, -16.5]$. The best-fitting pre- and post-inflection power-law approximations of HIIS are also plotted with their mean absolute relative fitting errors

being 0.089 ± 0.004 and 0.21 ± 0.008 , respectively. (c), Trace of the corresponding AE. Pre- and post-inflection modeling approximations of AE are also plotted with their mean absolute relative modeling errors being 0.096 ± 0.005 and 0.21 ± 0.009 , respectively. Model extrapolation toward $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$ results in a mean absolute relative fitting error of 1.77 ± 0.47 . $\nu(\mathbf{p})$ and $o(\mathbf{p})$ account for the scales at which the PDS-VSDs geometrical crossover occurs and the asymptote domain beings, respectively.

This fast rebound of bonding-locality provides with energy surplus for stabilization of the NaVAb within the membrane environment as the atom-packing energy of the largest atomic cluster (≈ 12.3 kcal) is larger than that of the initial atomic cluster (≈ 2.5 kcal) (Figure 3(c)).

In summary, the main advantage behind the CHDF configuration described by r1 is that the ratio $\frac{\|\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\|}{N(\mathbf{p}, l_\alpha(\mathbf{p}))}$ does not depend anymore to microscopic details of the atomic environment which implies that HIIS stabilizing the SF is insensitive to small-amplitude perturbations. For example, let us consider the case of an unexpected structural modification introducing a perturbation $N(\mathbf{p}, l_\alpha(\mathbf{p})) \rightarrow N(\mathbf{p}, l_\alpha(\mathbf{p})) + \epsilon(\mathbf{p}, l_\alpha(\mathbf{p}))$ where $\epsilon(\mathbf{p}, l_\alpha(\mathbf{p}))$ is a small-amplitude perturbation additive source. This will result in a perturbation of the CHDF magnitude $\|\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\| \rightarrow \|\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\| + \zeta(\mathbf{p}, l_\alpha(\mathbf{p}))$ that, due to r1, is regulated by $\|\vec{h}_z(\mathbf{p}, l_\alpha(\mathbf{p}))\| + \zeta(\mathbf{p}, l_\alpha(\mathbf{p})) \sim l_\alpha(\mathbf{p})^{\gamma(\mathbf{p})} \cdot \{N(\mathbf{p}, l_\alpha(\mathbf{p})) + \epsilon(\mathbf{p}, l_\alpha(\mathbf{p}))\}$ so that the HIIS law remains intact.

Discussion

An algebraic relationship between $\tilde{q}(\mathbf{p})$ and the non-extensive entropic index, $q(\mathbf{p})$ [36], is established with $\tilde{q}(\mathbf{p}) = 1 - q(\mathbf{p})$ [41]. Hence, equation r1 illustrates how the CHDF emerges from the spatial arrangement of atoms around the pore and establishes a direct link with the non-extensive entropic $q(\mathbf{p})$ -index. This is the main finding of this

study and highlights the fact that hydropathic forces are of entropic nature; observations presented in Figure 3 can be understood as an entropic re-configuration of the atomic environment, i.e., phase transition of the crystal structure (thought of as a glass system [43]), due to re-organization in the interactions range. The critical scale $\xi(\mathbf{p})$ determines then the maximum bond-length of the underlying interactions. Symmetry breaking associated with the critical transition is shown in Figure 2 of [39] where pseudo-symmetric topological pairs formed within the pre-inflection phase gradually dissolve below $\xi(\mathbf{p})$.

The long-rangeness of the HIIS around the SF might be crucial for the stable functioning of the biological machinery interacting with selected ions. In particular, ions entering the NaVAb's SF pore region from the ES are expected to strongly interact with pore-lining components at subsequent interactions sites [31] corresponding to negative minima of the energy landscape [39]. These interactions will result in their partial dehydration and, consequently, heat exchange with the surrounding atomic environment. To counteract for this local increase in pore pressure due to water-removal, a non-local interactions network provides peripheral structural support to the biological machinery implementing dehydration. Indeed, long-range interactions are thought to play a role in selectivity by indirectly influencing the stability of SF's dehydration binding sites [62, 63].

Mutations affecting the SF of NaVChs can have dramatic effects on ion selectivity (e.g., see [64]). The long-rangeness of the HIIS provides with a "shield" against mutations as HIIS would remain intact in occurrence of perturbations. Energetic perturbations originating from mutations affecting the PDs are expected to be damped-out toward the lag domain according to $\zeta(\mathbf{p}, l_\alpha(\mathbf{p})) \sim (\xi(\mathbf{p}) - l_\alpha(\mathbf{p}))^{-\gamma_{pre}(\mathbf{p})} \cdot \epsilon(\mathbf{p}, l_\alpha(\mathbf{p}))$ thus not affecting pore-lining SF components. Similarly, energetic perturbations originating from mutations affecting the VSDs are expected to be damped-out toward the asymptote domain thus not affecting channel's energetic contact with the membrane. What is at risk, however, here? Destabilizing energetic perturbations originating either

from the PDs or the VSDs can be amplified toward the inflection points which might jeopardize the relative stability of the PDs and VSDs coupled subsystems. Thus some structural flexibility of the S4-S5 units linking the PDs with the VSDs (see [31]) is necessary so that potential energy surpluses would be absorbed.

Conclusions

Critical points correspond to extrema in some property (or properties) with respect to which the protein system under scrutiny has been optimized via long-range and highly-cooperative interactions such as these described under the umbrella-term "hydropathic effects" [43,44]. In this study, we provided with computational evidence for the validity of the atomic hydropathicity non-extensivity hypothesis for the pre-open I217C NaVAb, which directly implies the existence of critical points associated with the stability of the pre-open I217C NaVAb state. Specifically, we showed that the atomic environment around the pore of the pre-open I217C NaVAb structural model is organized with respect to highly-conserved molecular locations associated with the modular PDs-VSDs architectural motif and with the SF atomic structure. A direct consequence of this is that hydropathic interactions strength exhibits a bi-phasic, scale-invariant behavior than can be modeled in terms of the Richards model parametric space $\{A(\mathbf{p}), t(\mathbf{p}), s(\mathbf{p}), \tilde{q}(\mathbf{p}) = 1 - q(\mathbf{p})\}$ and critical exponents $\{\gamma_{pre}(\mathbf{p}), \gamma_{post}(\mathbf{p})\}$. Hence, our modeling procedures can provide with a reduction of the spatial complexity of the NaVCh atomic system under scrutiny under meaningful biological and evolutionary constraints. This is important for NaVCh design, mutation-response analysis (e.g., see [65]) and pharmacological control where spatial complexity of the atomic environment poses major challenges.

Supplementary Information

S1. The pre-open I217C NaVAb structural model

Figure S1: **Cartoon illustration of the pre-open I217C NaVAb structural model.** (a), top-view from the extracellular side (ES) towards the intracellular side (IS). (b), side-view. Helices forming the pore domains are illustrated in blue color (residue sequence: M130:M221). Helices forming the voltage-sensor domains are illustrated in red color (residue sequence: M001:G129). Structural details of the pre-open I217C NaVAb structural model can be found in [31].

S2. Discretization of the principal pore axis

Following [39] we introduced the line grid $Q = \{\mathbf{p}_1, \mathbf{p}_1 + \Delta\mathbf{p}, \dots, \mathbf{p}_{N_p} - \Delta\mathbf{p}, \mathbf{p}_{N_p}\} \subset P$, where N_p is the total number of grid pore points, $\|\Delta\mathbf{p}\|$ is the sampling distance between two consecutive pore points and $\mathbf{p}_1 = (0, 0, p_{z,1})$, $\mathbf{p}_{N_p} = (0, 0, p_{z,N_p})$ are boundary pore points. Q is constructed by setting $p_{z,1} = \text{round}(\min_{i=1,2,\dots,N_c}(c_{z,i}), 1) = -27.1$ and $p_{z,N_p} = \text{round}(\max_{i=1,2,\dots,N_c}(c_{z,i}), 1) = 26.8$ with $\text{round}(x \in \mathbb{R}, n \in \mathbb{Z}^+)$ returning the value of x rounded up to the n -th decimal digit and by setting $N_p = 540$ so that $\|\Delta\mathbf{p}\| = 0.1 \text{ \AA}$.

S3. Graphical interpretation of atomic accumulation parameters

Figure S2: **Richards model exemplary traces.** The Richards model curve is illustrated for $A = 1$, $t = 1$, $s = 1.5$ and for $\tilde{q} \approx 0.0$, $\tilde{q} = 1.0$ and $\tilde{q} = 2.0$, where A determines the model's asymptotic behavior, t the slope at the inflection point ξ , s the size of the lag domain and $\tilde{q}$ the shape of the curve, as well as, the location of ξ . The model curve is split into three domain parts; a lag domain part ($l \leq s$), an inflection domain part ($s < l \leq o$), and an asymptote domain part ($l > o$). The inflection point ξ is shown for the case $\{A = 1, t = 1, s = 1.5, \tilde{q} \approx 1.0\}$.

S4. Modeling of the PDs-VSDs geometrical crossover

We approximated the radial distribution function (RDF) of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around $\mathbf{p}$ by

$$G(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta N(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho(\mathbf{p})} \quad \text{with } l'_\alpha(\mathbf{p}) = l_\alpha(\mathbf{p}) - \Delta l_\alpha(\mathbf{p}) \quad (\text{A1})$$

where $\Delta l_\alpha(\mathbf{p}) = \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha}$ is the thickness of a spherical shell around $\mathbf{p}$, $\Delta N(\mathbf{p}, l_\alpha(\mathbf{p})) = N(\mathbf{p}, l_\alpha(\mathbf{p})) - N(\mathbf{p}, l'_\alpha(\mathbf{p}))$ is the number of atoms found within the spherical shell of thickness $\Delta l_\alpha(\mathbf{p})$ centered at $\mathbf{p}$ and $\rho(\mathbf{p}) = \frac{N_c}{V(\mathbf{p})} = \frac{N_c}{\frac{4}{3} \cdot \pi \cdot (L(\mathbf{p}) - R(\mathbf{p}))^3}$ is the average atomic density around $\mathbf{p}$.

Next, we divided all N_c atoms into two groups, one containing only PD atoms, i.e., atoms forming the PDs, and one containing only VSD

atoms, i.e., atoms forming the VSDs. The total number of PD atoms is $N_{PD}=6068$ and they belong to the residue sequence M130:M221 that forms the structural units of the S5 domain, the pore turret loop, the P-helix, the selectivity filter, the P2-helix and the S6 domain (see SI Figure 1 of [31]). On the other hand, the total number of VSD atoms is $N_{VSD}=8708$ and they belong to the residue sequence M001:G129 that forms the structural units of the S1N helix, the S1 domain, the S1-S2 loop, the S2 domain, the S2-S3 loop, the S3 domain, the S3-S4 loop, the S4 domain and the S4-S5 linker (see SI Figure 1 of [31]). Note that $N_c = N_{PD} + N_{VSD}$.

Using equation A1 we approximated the RDFs of the PD and VSD atoms around $\mathbf{p}$ with

$$G_{PD}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{PD}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{PD}(\mathbf{p})} \quad (\text{A2})$$

and

$$G_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{VSD}(\mathbf{p})} \quad (\text{A3})$$

, respectively, where $N(\mathbf{p}, l_\alpha(\mathbf{p})) = N_{PD}(\mathbf{p}, l_\alpha(\mathbf{p})) + N_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))$, $\rho_{PD}(\mathbf{p}) = \frac{N_{PD}}{V(\mathbf{p})}$ and $\rho_{VSD}(\mathbf{p}) = \frac{N_{VSD}}{V(\mathbf{p})}$.

In order to obtain a relative measure of how the RDF of PD atoms varies with respect to the RDF of VSD atoms, and vice versa, we introduced the equilibrium PD-VSD RDF function [39]

$$e(\mathbf{p}, l_\alpha(\mathbf{p})) = sm(G_{PD}(\mathbf{p}, l_\alpha(\mathbf{p})) - G_{VSD}(\mathbf{p}, l_\alpha(\mathbf{p}))) \quad (\text{A4})$$

where $sm(\cdot)$ implements the Nadaraya-Watson kernel regression function with a bandwidth parameter value of $bw=2$. $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ is interpreted as the smoothed probability of finding a PD atom instead of a VSD atom, and vice-versa, at distance $l_\alpha(\mathbf{p})$ from $\mathbf{p}$. Specifically, if $e(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$ ($e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$), then the smoothed probability of finding a PD atom within the spherical shell of width $\Delta l_\alpha(\mathbf{p})$ is larger than that of finding a VSD (PD) atom. Accordingly, what is of interest

here is the sign-change behavior of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing $l_\alpha(\mathbf{p})$. We investigated it by detecting for every $\mathbf{p} \in Q$ the pair $\{l'_\alpha(\mathbf{p}), l_\alpha(\mathbf{p})\}$ for which the sign-change condition $e(\mathbf{p}, l'_\alpha(\mathbf{p})) \cdot e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ is satisfied and calculating the sampling radius

$$\nu(\mathbf{p}) = l'_\alpha(\mathbf{p}) - \frac{e(\mathbf{p}, l'_\alpha(\mathbf{p}))}{e(\mathbf{p}, l_\alpha(\mathbf{p})) - e(\mathbf{p}, l'_\alpha(\mathbf{p}))} \cdot \Delta l_\alpha(\mathbf{p}) \quad (\text{A5})$$

for which $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ changes sign along $l_\alpha(\mathbf{p})$ -direction. Equation A5 roughly quantifies the size of the PDs thus indicating the location around the pore of the geometrical crossover from the PDs to the VSDs takes place.

S5. Statistical representation of scalar functions

Let $\langle \cdot \rangle$ be a statistical operator returning the median of the data set upon which it operates. Note that the choice of the median as a statistical measure reflects the fact that no assumption has been made for the distribution of data set values.

Statistical representation of scalar function $f(\mathbf{p}, l_\alpha(\mathbf{p}))$. A statistical representation of the scalar function $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ for a given α was obtained in terms of $\langle f(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$ where the subscript α indicates that the statistical operator acts for a given α , i.e., on the set of values $\Gamma(\alpha) = \{f(\mathbf{p}, l_\alpha(\mathbf{p})) \mid \mathbf{p} \in Q\}$.

Statistical representations of scalar functions $s(\mathbf{p})$, $\nu(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$. Statistical representations of the functions $s(\mathbf{p})$, $\nu(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$ were obtained by the following algorithmic scheme:

Step 1. Let $f(\mathbf{p})$ represent one of the aforementioned scalar functions, then, for every $\mathbf{p}$ find the α indexes for which $|f(\mathbf{p}) - l_\alpha(\mathbf{p})|$ is minimized, i.e., $\alpha_f = \{\alpha \mid \min_{\alpha \in A} (|f(\mathbf{p}) - l_\alpha(\mathbf{p})|)\}$.

Step 2. Calculate the median $\langle \alpha_f \rangle$ of the data set α_f . Note that the minimum and the maximum value contained in α_f is given by $\min(\alpha_f)$ and $\max(\alpha_f)$.

S6. Scaling behavior of the hydrophobic imbalance magnitude along the pore

We scanned all different paths leading from the NaVAb's pore axis toward its exterior. The scaling behavior of the HIIS for each path corresponding to a $\mathbf{p} \in Q$ (see SI Section S2 for construction of Q) was checked for its power-law property. That was done by linear approximation on $\log[I(\mathbf{p}, l_\alpha(\mathbf{p}))]$ for the lag domain interval, i.e., $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$, and for the pre- and post-inflection phase intervals, i.e., $s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ and $l_\alpha(\mathbf{p}) > \xi(\mathbf{p})$, respectively. Accordingly, for each interval we obtained a linear fitting of the form $\gamma(\mathbf{p}) \cdot \log[l_\alpha(\mathbf{p})] + \beta(\mathbf{p})$ where the "goodness" of the fitting was evaluated in terms of the Pearson coefficient, as well as, in terms of the uncertainty in estimation of $\gamma(\mathbf{p})$ retrieved from the mean absolute fitting error (MAFE).

Figure S3: **Scaling analysis of the interatomic hydrophobic interaction**

strength. (a), Traces of the power-law exponents $\gamma_{lag}(\mathbf{p})$, $\gamma_{pre}(\mathbf{p})$ and $\gamma_{post}(\mathbf{p})$ describing the scaling behavior of the interatomic hydrophathic interaction strength within lag interval $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$, a pre-inflection interval $s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ and the post-inflection interval $l_\alpha(\mathbf{p}) > \xi(\mathbf{p})$, respectively, are plotted for every pore point $\mathbf{p} \in Q$. Shaded area around $\gamma_{lag}(\mathbf{p})$, $\gamma_{pre}(\mathbf{p})$ and $\gamma_{post}(\mathbf{p})$ indicate uncertainty in their calculation in terms of the mean absolute fitting error. (b), Traces of the corresponding Pearson coefficients $PC_{lag}(\mathbf{p})$, $PC_{pre}(\mathbf{p})$ and $PC_{post}(\mathbf{p})$ are plotted. Green-shaded and red-shaded areas indicate pore regions where $|PC_{pre}(\mathbf{p})|$ and $|PC_{post}(\mathbf{p})|$ attain values larger than 0.95, respectively. The intersection of the green-shaded and red-shaded areas the pore region $-18.0 \leq p_z \leq -16.5$ located at the ES of the SF.

As we demonstrate in Figure S3, HIIS profiles along $\mathbf{p}$ -direction are organized with respect to $p_z \approx 2.8$ corresponding to the center of the CC, as well as, to the hydrophathic topological center of the pore controlling its gating behavior [39]. Pre-inflection hydrophathic interactions exhibit an odd pseudo(i.e., non-exact)-symmetric behavior with respect to $p_z \approx 2.8$ as indicated by the behavior of the $\gamma_{pre}(\mathbf{p})$ trace (see Figure 3(a)). On the other hand, $\gamma_{lag}(\mathbf{p})$ and $\gamma_{post}(\mathbf{p})$ exhibit an even pseudo(i.e., non-exact)-symmetric behavior with respect to an interactions-symmetry-axis placed at $p_z \sim 2.8$ perpendicular to P (see Figure S3(a)). This symmetry scheme suggests that what is actually differentiating among the IS and the ES is the range of hydrophathic interactions within the PDs; both $\gamma_{lag}(\mathbf{p})$ and $\gamma_{post}(\mathbf{p})$ attain negative values toward the IS and the ES but, in striking contrast, $\gamma_{pre}(\mathbf{p})$ exhibits a strong, positive increase at the ES and a weak decrease towards negative values at the IS.

The transition to negative values of the triplet $\{\gamma_{lag}(\mathbf{p}), \gamma_{pre}(\mathbf{p}), \gamma_{post}(\mathbf{p})\}$ at the IS appears to initiate at the right of the Logistic model pore region (i.e., for $p_z > 10.9$). On the other hand, the transition of $\gamma_{pre}(\mathbf{p})$ to a large, positive value initiates on the left of the Gompertz model pore region (i.e., for $p_z < -5.5$). Taken together, these observations suggest that the range of interactions stabilizing the closed AG is complexly

depending on scale and, most likely, stays short while long-range interactions emerge within the PDs toward the ES.

Figure S4: **Conservation analysis of a PD domain.** Conservation analysis of a single PD domain. The SF, and the P- and P2-helices are annotated. The highest conservation score is attributed to the SF residue complex T175-L176-E177-S178-W179 scoring an average of 2.4 (out of 4.0) with pore-lining residue conservation scores confined within $[0, 2.0]$. Analysis was performed on the ConSurf Server (<http://consurf.tau.ac.il/>) according to the algorithmic implementations described in [66, 67].

Indeed, a clear enhancement of the power-law scaling of $I(\mathbf{p}, l_\alpha(\mathbf{p}))$ occurs on the left of the Gompertz model pore region, i.e., for $p_z < -5.5$. In particular, $PC_{post}(\mathbf{p})$ attains values smaller than -0.95 for $-18.0 \leq p_z \leq -7.4$ (see red-shaded area in Figure S3(a),(b)) so that it minimizes at the SF, i.e., for $p_z \approx -11.7$, while $PC_{pre}(\mathbf{p})$ attains values higher than 0.95 for $-18.0 \leq p_z \leq -16.5$ so that it maximizes at the ES of the SF (see green-shaded area in Figure S3(a),(b)). Accordingly, within the narrow interval $-18.0 \leq p_z \leq -16.5$ both $|PC_{pre}(\mathbf{p})|$ and $|PC_{post}(\mathbf{p})|$ attain their maximum values while the uncertainty in their calculation remains small (see shaded area around $\gamma_{pre}(\mathbf{p})$ and $\gamma_{post}(\mathbf{p})$ in Figure S3(a)) allowing for an accurate approximation of HHS in

terms of

$$I(\mathbf{p}, l_\alpha(\mathbf{p})) = \begin{cases} l_\alpha(\mathbf{p})^{\gamma_{pre}(\mathbf{p})} & \text{for } s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p}) \\ l_\alpha(\mathbf{p})^{\gamma_{post}(\mathbf{p})} & \text{for } l_\alpha(\mathbf{p}) > \xi(\mathbf{p}) \end{cases} \quad (\text{A6})$$

The $-18.0 \leq p_z \leq -16.5$ pore region is directly surrounded by the M181 and the S178 residue side chains as shown in [39]. The highly-conserved nature of the pore-forming atomic compartments around $-18.0 \leq p_z \leq -16.5$ is shown in Figure S4 (see also [31]).

Author contributions

M.N.X. designed the study, performed computational experiments and curated the findings; R.W. and P.L. contributed to the refinement of algorithmic steps; R.W., P.L., D.K., Y.Y., J.H., B.H.S. and S.G.W. provided with critical feedback and helped with the interpretation of the findings; Y.Y. and S.G.W. encouraged M.N.X. to focus on specific aspects of the findings; B.H.S., R.W. and P.L. supervised the project; G.L. and C.G.F. were in charge of overall research direction and supervision; M.N.X. wrote the manuscript in consultation with all the co-authors; PROPANE study group provided with technical and scientific support.

Acknowledgments

The study was partly funded by the European Union 7th Framework Programme (grant n602273).

Competing Interests

The authors declare that there is no conflict of interest.

References

- [1] Hille, B. *Ionic channels of excitable membranes*, 3rd ed. (Sinauer Associates Inc., Sunderland MA (2001).
- [2] Hübner, C.A. & Jentsch, T.J. Ion channel diseases. *Hum Mol Genet* **11**, 2435-2445 (2002).
- [3] Cummins, T.R., Dib-Hajj, S.D., Waxman, S.G. Electrophysiological properties of mutant Nav1.7 sodium channels in a painful inherited neuropathy. *J Neurosci* **24**, 8232-8236 (2014).
- [4] Han, C., Rush, A.M., Dib-Hajj, S.D. et al. Sporadic onset of erythromelgia: a gain-of-function mutation in Nav1.7. *Ann Neurol* **59**, 553-558 (2006).
- [5] Harty, T.P., Dib-Hajj, S.D., Tyrrell, L., Blackman, R., Hisama, F.M., Rose, J.B., Waxman, S.G. Nav1.7 mutant A863P in erythromelgia: effects of altered activation and steady-state inactivation on excitability of nociceptive dorsal root ganglion neurons. *J Neurosci* **26**, 12566-12575 (2006).
- [6] Lampert, A., Dib-Hajj, S.D., Eastman, E.M., Tyrrell, L., Lin, Z., Yang, Y., Waxman, S.G. Erythromelgia mutation L823R shifts activation and inactivation of threshold sodium channel Nav1.7 to hyperpolarized potentials. *Biochem Biophys Res Commun* **390**, 319-324 (2006).
- [7] Stadler, T., O'Reilly, A.O., Lampert, A. Erythromelgia mutation Q875E stabilizes the activated state of sodium channel Nav1.7. *J Biol Chem* **290**, 6316-6325 (2015).
- [8] Yang, Y., Wang, Y., Li, S. et al. Mutations in SCN9A, encoding a sodium channel alpha subunit, in patients with primary erythromelgia. *J Med Genet* **41**, 171-174 (2004).

- [9] Dib-Hajj, S.D., Rush, A.M., Cummins, T.R. et al. Gain-of-function mutation in Nav1.7 in familial erythromelalgia induces bursting of sensory neurons. *Brain* **128**, 1847-1854 (2005).
- [10] Drenth, J.P., te Morsche, R.H.M., Guillet, G., Taieb, A., Lee Kirby, R., Jansen, J.B.M. SCN9A mutations define primary erythromelalgia as a neuropathic disorder of voltage gated sodium channels. *J Invest Dermatol* **124**, 1333-1338 (2005).
- [11] Lee, M.J., Yu, H.S., Hsieh, S.T. et al. Characterization of a familial case with primary erythromelalgia from Taiwan. *J Neurol* **254**, 210-214 (2007).
- [12] Drenth, J.P., Waxman, S.G. Mutations in sodium-channel gene SCN9a cause a spectrum of human genetic pain disorders. *J Clin Invest* **117**, 3603-3609 (2007).
- [13] Fertleman, C.R., Baker, M.D., Parker, K.A. et al. SCN9A mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52**, 767-774 (2006).
- [14] Jarecki, B.W., Sheets, P.L., Jackson, J.O. 2nd, Cummins, T.R. Paroxysmal extreme pain disorder mutations within the D3/S4-S5 linker of Nav1.7 cause moderate destabilization of fast inactivation. *J Physiol* **586**, 4137-4153 (2008).
- [15] Dib-Hajj, S.D., Estacion, M., Jarecki, B.W. et al. Paroxysmal extreme pain disorder M1627K mutation in human Nav1.7 renders DRG neurons hyperexcitable. *Mol Pain* **4**, 37 (2008).
- [16] Theile, J.W., Jarecki, B.W., Piekarz, A.D., Cummins, T.R. Nav1.7 mutations associated with paroxysmal extreme pain disorder, but not erythromelalgia, enhance Nav β 4 peptide-mediated resurgent sodium currents. *J Physiol* **589**, 597-608 (2011).
- [17] Faber, C.G., Hoeijmakers, J.G., Ahn, H.S. et al. Gain of function Nav1.7 mutations in idiopathic small fiber neuropathy. *Ann. Neurol.* **71**, 26-39 (2012).
- [18] Hoeijmakers, J.G., Han, C., Merkies, I.S. et al. Small nerve fibres, small hands and small feet: a new syndrome of pain, dysautonomia and acromesomelia in a kindred with a novel Nav1.7 mutation. *Brain* **135**, 345-358 (2012).
- [19] Qu, Y., Rogers, J., Tanada, T., Scheuer, T. Catterall, W.A. Molecular determinants of drug access to the receptor site for antiarrhythmic drugs in the cardiac Na⁺ channel. *Proc Natl Acad Sci USA* **270**, 25696-25701 (1995).

- [20] Keating, M.T. Sanguinetti M.C. Molecular and cellular mechanisms of cardiac arrhythmias. *Cell* **104**, 569-580 (2001).
- [21] Catterall, W.A., Kalume, F., Oakley, J.C. NaV1.1 channels and epilepsy. *J Physiol* **588**, 1849-1859 (2010).
- [22] Mullins, L.J. The permeation of some cations into muscle. *J Gen Physiol* **42**, 817-829 (1959).
- [23] Mullins, L.J. An analysis of conductance changes in squid axon. *J Gen Physiol* **42**, 1013-1035 (1959).
- [24] Mullins, L.J. The macromolecular properties of excitable membranes. *Ann NY Acad Sci* **94**, 390-404 (1961).
- [25] Beckstein, O., Biggin, P.C., Sansom, M.S.P. A Hydrophobic Gating Mechanism for Nanopores. *J Phys Chem B* **105**, 12902-12905 (2001).
- [26] Beckstein, O., Sansom, M.S.P. Liquid-vapor oscillations of water in hydrophobic nanopores. *Proc Natl Acad Sci USA* **100**, 7063-7068 (2003).
- [27] Allen, R., Hansen, J.-P. Molecular dynamics investigation of water permeation through nanopores. *J Chem Phys* **119**, 3905-3919 (2003).
- [28] Beckstein, O., Sansom, M.S.P. The influence of geometry, surface character, and flexibility on the permeation of ions and water through biological pores. *Phys Biol* **1**, 42-52 (2004).
- [29] Jensen, M.Ø, Borhani, D.W., Lindorff-Larsen, K. Principles of conduction and hydrophobic gating in K^+ channels. *Proc Natl Acad Sci USA* **107**, 5833-5838 (2010).
- [30] Aryal, P., Sansom M.S.P, Tucker, S.J. Hydrophobic Gating in Ion Channels. *J Mol Biol* **427**, 121-130 (2015).
- [31] Payandeh, J., Scheuer, T., Zheng, N., Catterall, W.A. The crystal structure of a voltage-gated sodium channel. *Nat* **475**, 353-358 (2011).
- [32] Yonkunas, M., Kurnikova, M. The hydrophobic effect contributes to the closed state of a simplified ion channel through a conserved hydrophobic patch at the pore-helix crossing. *Front Pharmacol* **6**, 284 (2015).
- [33] Kitaguchi, T., Sukhareva, M., Swartz, K.J. Stabilizing the closed S6 gate in the Shaker K_v channel through modification of a hydrophobic seal. *J Gen Physiol* **124**, 319-332 (2004).

- [34] Hummer, G., Garde, S., García, A.E., Paulaitis, M.E., Pratt, L.R. Hydrophobic Effects on a Molecular Scale. *J Phys Chem B* **102**, 10469-10482 (1998).
- [35] Lum, K., Chandler, D., Weeks, J.D. Hydrophobicity at Small and Large Length Scales. *J Phys Chem B* **103**, 4570-4577 (1999).
- [36] Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J Stat Phys* **52**, 479 (1988).
- [37] Hammer, M.U., Anderson, T.H., Chaimovich, A., Shell, M.S., Israelachvili, J. The search for the hydrophobic force law. *Faraday Discuss* **146**, 299-308 (2010).
- [38] Zhou, R., Silverman, B.D., Royyuru, A.K., Athma, P. Spatial profiling of protein hydrophobicity: Native vs. decoy structures. *Proteins* **52**, 561-572 (2003).
- [39] Xenakis, M.N., Kapetis, D., Yang, Y. et al. Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution. *Proteins* **88**, 1319-1328 (2020).
- [40] Richards, F.J. A flexible growth function for empirical use. *J Exp Bot* **10**, 290-300 (1959).
- [41] Martinez, A.S., González, R.S., Terçariol, C.A.S. Continuous growth models in terms of generalized logarithm and exponential functions. *Physica A* **387**, 5679-5687 (2008).
- [42] Silverman, B.D. Hydrophobic Moments of Tertiary Protein Structures *Proteins* **53**, 880-888 (2003).
- [43] Phillips, J.C. Scaling and self-organized criticality in proteins I. *Proc Natl Acad Sci USA* **106**, 3107-3112 (2009).
- [44] Phillips, J.C. Scaling and self-organized criticality in proteins II. *Proc Natl Acad Sci USA* **106**, 3113-3118 (2009).
- [45] Morett, M. Self-organized critical model for protein folding. *Physica A* **390**, 3055-3059 (2011).
- [46] Vriend, G. WHAT IF: A molecular modeling and drug design program. *J Mol Graph* **8**, 52-56 (1980).
- [47] Hooft, R.W., Sander, C., Vriend, G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins* **26**, 363-376 (1996).

- [48] Humphrey, W., Dalke, A., Schulten, K. VMD-Visual Molecular Dynamics. *J Mol Graph* **14**, 33-38 (1996).
- [49] Smart, O.S., Neduvilil, J.G., Wang, X., Wallace, B.A., Sansom, M.S.P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* **14**, 354-360 (1996).
- [50] Kahm, M. and Hasenbrink, G., Lichtenberg-Fraté, H., Ludwig, J., Kschischo, M. Fitting Biological Growth Curves with R. *J Stat Softw* **33**, 1-21 (2010).
- [51] Zwietering, M.H., Jongenburger, I., Rombouts, F.M., van 't Riet, K. Modeling of the bacterial growth curve. *Appl Environ Microbiol* **56**, 1875-1881 (1990).
- [52] Blumberg, A.A. Logistic growth rate functions. *J Theoret Biol* **21**, 42 (1968).
- [53] Gompertz, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc London* **115**, 513-585 (1825).
- [54] Tjørve, E., Tjørve, K.M.C. A unified approach to the Richards-model family for use in growth analyses: Why we need only two model forms. *J Theor Biol* **267**, 417-425 (2010).
- [55] Eisenberg, D., Weiss, R.M., Terwilliger, T.C., Wilcox, W. Hydrophobic moments and protein structure. *Faraday Symp Chem Soc* **17**, 109-120 (1982).
- [56] Kapcha, L.H., Rossky, P.J. A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure. *J Mol Biol* **426**, 484-498 (2014).
- [57] Carvalho-de-Souza, J.L., Bezanilla, F. Noncanonical mechanism of voltage sensor coupling to pore revealed by tandem dimers of Shaker. *Nat Commun* **10**, 3584 (2019).
- [58] Carvalho-de-Souza, J.L., Bezanilla, F. Non-canonical interactions between voltage sensors and pore domain in Shaker K⁺-channel. *Biophys J* **112**, 162a (2017).
- [59] Conti, L., Renhorn, J., Gabrielsson, A., Tureson, A., Liin, S.I., Lindahl, E., Elinder, F. Reciprocal voltage sensor-to-pore coupling leads to potassium channel C-type inactivation. *Sci Rep* **6**, 27562 (2016).

- [60] Fernández-Mariño, A. I., Harpole, T. J., Oelstrom, K., Delemotte, L., Chanda, B. Gating interaction maps reveal a noncanonical electromechanical coupling mode in the Shaker K⁺ channel. *Nat Struct Mol Biol* **25**, 320-326 (2018).
 - [61] Gaines, J.C., Clark, A.H., Regan, L., O'Hern C.S. Packing in Protein Cores *J Phys: Condens Matter* **29**, 293001 (2017).
 - [62] Yu, H., Noskov, S.Y., Roux, B. Two mechanisms of ion selectivity in protein binding sites. *Proc Natl Acad Sci USA* **107**, 20329-20334 (2010).
 - [63] Roux, B., Bernèche, S., Egwolf, B., Lev, B., Noskov, S.Y., Rowley, C.N., Yu, H. Ion selectivity in channels and transporters. *J Gen Physiol* **137**, 415-426 (2011).
 - [64] Heinemann, S.H., Terlau, H., Stuhmer, W., Imoto, K., Numa, S. Calcium channel characteristics conferred on the sodium channel by single mutations. *Nat* **356**, 441-443 (1992).
 - [65] Kapetis, D., Yang, Y., Sassone, J. et al. Network topology of NaV1.7 mutations in sodium channel-related painful disorders. *BMC Systems Biology* **11**, 28 (2017).
 - [66] Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., Ben-Tal, N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucl Acids Res* **33**, W299-W302 (2005).
 - [67] Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., Ben-Tal, N. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. *Bioinformatics* **19**, 163-164 (2003).
-

Chapter 4

Hydropathicity-based prediction of pain-causing NaV1.7 variants

Markos N. Xenakis, Dimos Kapetis, Yang Yang, Monique M. Gerrits,
Jordi Heijman, Stephen G. Waxman, Giuseppe Lauria, Catharina G.
Faber, Ronald L. Westra, Patrick J. Lindsey, Hubert J. Smeets

Accepted at *BMC Bioinformatics*

Abstract

Mutation-induced variations of the NaV1.7 channel protein's functional architecture are causally related to a broad spectrum of human pain disorders. Predicting *in silico* the phenotype of NaV1.7 variant is of major clinical importance; it can aid in reducing costs of *in vitro* pathophysiological characterization of NaV1.7 variants, as well as, in the design of drug agents for counteracting adhere pain symptoms. In this work, we utilize spatial complexity of hydrophobic effects toward predicting which NaV1.7 variants cause pain (and which are neutral) based on the location of its mutation site within the NaV1.7 structure. For that, we analyze topological and scaling hydrophobic characteristics of the atomic environment around NaV1.7's pore and probe their spatial correlation with mutation sites. We show that pain-related mutation sites occupy structural locations in proximity to a hydrophobic patch lining the pore while clustering at a critical hydrophobic-interactions distance from the selectivity filter (SF). Taken together, these observations can differentiate pain-related NaV1.7 variants from neutral ones, i.e., NaV1.7 variants not causing pain disease, with 80.5% sensitivity and 93.7% specificity [area under the receiver operating characteristics curve = 0.872]. Our findings suggest that maintaining hydrophobic NaV1.7 interior intact, as well as, a finely-tuned (dictated by hydrophobic interactions) distance from the SF might be necessary molecular conditions for physiological NaV1.7 functioning. The main advantage for using the presented predictive scheme is its negligible computational cost, as well as, hydrophobicity-based biophysical rationalization.

Introduction

Voltage-gated sodium channels (NaVChs) are pore-forming proteins spanning the cell membrane. They are members of the ion channels

superfamily and their main physiological role is to control transport of sodium ions across cell membranes. The human NaV1.7 channel is encoded by the *SCN9A* gene and is preferentially expressed in peripheral neurons (e.g., dorsal root ganglion (DRG) nociceptors) responsible for networking pain signals. The structure of the NaV1.7 α -subunits is that of a pore-forming tetramer via assembly of four heterogeneous domains (DI-DIV). Three intracellular loops (L1-L3) form structural interconnections among subsequent domains. Each domain comprises six transmembrane helices (S1-S6) organized into a pore module (PM) forming an ion-conduction pathway coupled with a voltage-sensor (VS). Mechanistic description of NaV1.7's function is that VSs react to extracellular changes in ionic concentrations by moving outwards thus exerting a pulling force upon the PM which opens the channel pore. Closed-to-open gating transition leads to channel activation, i.e., renders it conductive to sodium ions. Missense mutations in the *SCN9A* gene can destabilize the NaV1.7's functional architecture thus disrupting physiological gating and, consequently, deregulate flow of sodium ions through the pore. At a cellular level, these genetically-caused destabilizations can affect neuronal excitability by inducing a gain-of-function (GOF) effect, i.e., by increasing the net ionic flow, thus triggering a wide spectrum of pain diseases such as inherited erythromelalgia (IEM) [?, 1–30], paroxysmal extreme pain disorder (PEPD) [31–37] and small fiber neuropathy (SFN) [38–42]. A proof of concept for the GOF-pain correlation hypothesis came from identification of missense *SCN9A*-gene mutations inducing a loss-of-function (LOF) effect, i.e., decreasing sodium-ions membrane currents, that is causally related to clinical symptoms of loss of pain sensation [43–45].

Hydropathic interactions (HIs) represent a summary of fundamental molecular interactions [46] driving molecular phenomena such as protein folding, protein hydrophobic-core stability, self-assembly of amphiphilic molecules and "dewetting" transitions occurring at the nanoscale (for a review in HIs-driven phenomena see [47]). Within the field of ion channels research, experimental and computational

studies have shown that HIs are crucial regulators of gating processes occurring within the pore’s microenvironment (termed as hydrophobic gating phenomena [48]), as well as, of channel stability via formation of hydrogen bonds networks [49,50] initiated by pore-lining hydrophobic residue patches [51]. Computational modeling of HIs combined with biophysical observations extracted from *in vitro* NavCh pathophysiological characterization can propel our understanding of mechanistic linkages between mutation-induced perturbations and human pain pathophysiology. Key-studies toward this direction were [52] and [53] demonstrating how the F1449V mutation and the in-frame-deletion L955Del, respectively, can disrupt a hydrophobic ring stabilizing the putative activation gate (AG) of the Nav1.7 thus acting as disease-causing molecular triggers. Moreover, computational modeling successfully deduced an energetic coupling between two different IEM-related mutations foreseen by their geometrical proximity in Nav1.7 structure [7]. A question that naturally arose from these studies was whether a detailed examination of HIs network characteristics within a Nav1.7 structure can reveal statistically-significant but also biophysically-relevant differentiations among the WT structure and its variants. This question was probed in [54]; a network-theoretical computational framework was introduced in order to capture changes in interatomic HIs within a Nav1.7 WT structure induced by pain-related mutations. The study reported on a betweenness centrality network measure achieving a statistically-important differentiation of pain-related variants from a collection of neutrals, i.e., variants not causing pain disease. Notably, this approach highlighted the prominent role that HIs play in Nav1.7’s stability and reported on plausible mutation-mechanism scenarios disrupting hydrophobic contacts among neighboring and distant residues. Another remark on the multi-scale nature of HIs was made by the authors of [55] suggesting that a pathogenic mutation in the *KCNA1* gene encoding the human voltage-gated potassium channel KV1.1 can de-tune HIs equilibrium (and, consequently, destabilize KV1.1’s pre-open conforma-

tion) implying that mutation-induced perturbation effects can destroy finely-tuned network-like HIs expanding throughout the structure as a whole. Interestingly, the fine-tuning hypothesis was proposed also for the NaV1.7; a recent study employing a machine learning (MLE) computational pipeline for predicting NaV1.7's variant pathogenicity suggested that the fine-tuning of the NaV1.7 channel is so delicate that limits classification accuracy of practically any computational approach [56]. Taken together, these observations highlight the highly-cooperative nature of HIs [46] and suggest that even small changes in the hydropathic spatial distribution profile of a channel structure can have a detrimental impact on the functional architecture which, in turn, might induce clinically-observed alterations of electrophysiology.

Following [54, 56], this study aims at probing the finely-tuned hypothesis for the NaV1.7's atomic hydropathic environment in order to predict whether a NaV1.7 variant causes pain or not. We utilize a closed-state structural model of the NaV1.7 retrieved from homology modeling based on the pre-open NaVA b [57] template (first presented in [53] and later used also in [7]) and investigate cumulative, i.e., scale-dependent, hydropathic properties of its porous atomic environment in relation to structural locations of missense *SCN9A*-gene mutations. In order to tackle spatial complexities emerging from the highly-cooperative nature of HIs we adopt a modeling approach rooted in the hypothesis that proteins can be represented as self-organized criticality (SOC) [58] archetypes; protein structures are thought to have been evolutionary optimized with respect to extrema in some thermodynamic property (or properties) capturing a qualitative reorganization of the atomic environment [59, 60]. The intra-channel locations where these macroscopic thermodynamic changes take place correspond to so-called critical points of the atomic structure [59, 60]. The highly-cooperative nature of HIs has placed structure-retrieved hydropathic properties in the epicenter of SOC hypothesis [59–62]). It is important to note that computational evidence for a universal hydrophobic-to-hydrophilic (or inside-outside with "inside" referring

to the hydrophobic core and "outside" referring to the hydrophilic exterior) spatial transition in protein systems was first provided before the formulation of the SOC hypothesis (see [63]). Departing from this phenomenological basis, we here utilize the finite-size scaling analysis methodologies presented in [64, 65] for screening hydropathic morphology around NaV1.7's pore [64] toward identification of critical points associated with NaV1.7's functional architecture. Biophysical relevance of retrieved observations is justified not only in terms of the scale-invariance of a carefully-chosen cumulative hydropathicity-property function but also with respect to conserved structural NaV1.7 features such as the PM-VSs spatial transition and the location of the selectivity filter (SF). In particular, we demonstrate that the atomic cumulative distribution function around NaV1.7's pore exhibits a sigmoid profile with inflection points matching closely the conserved PM-VSs spatial transition. This provides a rigorous description of atom-packing geometry and, consequently, a macroscopic partitioning of the atomic environment around the pore allowing for mapping the spatial profile of the atomic cumulative hydropathicity-property function and mutation sites on two dimensions. The SOC hypothesis is then accepted (or rejected) depending on whether the cumulative hydropathicity-property function is globally maximized and exhibits power-law-like scaling behavior in the vicinity of the inflection point (or not).

Our mapping procedures reveal the formation of a hydrophobic patch (HP) incorporating NaV1.7's central cavity (CC) and activation gate (AG), as well as, a hydrophobic pore wall at the extracellular side (ES) of the SF. We report on two "hot" map areas attracting pain-related mutation sites which are distributed either inside or along HP's periphery. Probing the SOC hypothesis for the NaV1.7 structure reveals that "hot" structural locations tend to cluster at a distance of 33.4 Å from the SF. Stability implications of these observations can be concertized by considering that in the vicinity of the critical point the range and intensity of HIs increase in a power-law fashion thus favoring

amplification and propagation of mutation-induced perturbations only peripherally to the HP and at critical HIs-distance from the SF thus not directly affecting neither of them. The clinical translational value of our findings is tested by predicting pathogenicity of 84 NaV1.7 variants; a weighted average of HP- and SF-related distance metrics can classify up to 29 (out of 36) pain-related variants and 45 (out of 48) neutral variants correctly.

Methods

All computations were performed in R [66] environment unless stated differently.

3D structure preparation

We fetch the NaV1.7 atomic structure model constructed via homology modeling procedures based on the closed-state NaVAb [57] template (for model construction details see [7]). Its principal axes were estimated by using the VMD software [67]. A global coordinate system $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ was introduced with its center at O and the NaV1.7's principal pore axis, i.e., the axis approximating the direction of the channel's pore, was aligned with the z -axis with orientation from the ES toward the intracellular side (IS) with respect to $\hat{\mathbf{z}}$. The atomic center $\mathbf{e} = \frac{1}{M} \sum_{i=1}^{N_c} m_i \cdot \mathbf{c}_i$ of the 3D structure was set to overlap with O , where $\mathbf{c}_i = (c_{x,i}, c_{y,i}, c_{z,i})$ is the atomic center of the i -th atom, m_i is the mass of the i -th atom, $N_c = 18567$ is the total number of atoms and $M = \sum_{i=1}^{N_c} m_i$ is the total molecular mass (see Dissertation Appendices, Table A1 for atomic masses values).

Geometrical characteristics of the pore

We navigated through the skewed NaV1.7's pore by introducing pore points $\mathbf{p}$ (see Supplementary Information (SI), S1). The pore radius

at $\mathbf{p}$ is given by [71]

$$R(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| - vdW_i \} \quad (\text{m2})$$

where $\|\cdot\|$ is the euclidean norm and vdW_i is the van der Waals radius of the i -th atom (values of van der Waals radii are presented in Table A1 of Dissertation Appendices). The distance between $\mathbf{p}$ and its nearest neighbor atom corresponds then to

$$D(\mathbf{p}) = \min_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| \} \quad (\text{m3})$$

and the outer surface radius at $\mathbf{p}$ is given by [64, 65]

$$L(\mathbf{p}) = \max_{i=1,2,\dots,N_c} \{ \|\mathbf{c}_i - \mathbf{p}\| + vdW_i \} \quad (\text{m4})$$

where the unit of measurement for $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ is expressed in $[\text{\AA}]$.

Finite-size sampling around the pore

The atomic environment around $\mathbf{p}$ is sampled with concentric spheres placed at $\mathbf{p}$ of increasing radius [64, 65]

$$l_\alpha(\mathbf{p}) = D(\mathbf{p}) + \alpha \cdot \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha} \text{ for } \alpha = 1, 2, \dots, K_\alpha \rightarrow \infty \quad (\text{m5})$$

where K_α is the total number of sampling spheres and α denotes the index of the sampling sphere. $l_\alpha(\mathbf{p})$ indicates the size, i.e., molecular scale, of the spherical cluster of atoms around $\mathbf{p}$ in $[\text{\AA}]$ and $L(\mathbf{p})$ the finite channel size measured with respect to $\mathbf{p}$. Accordingly, the atomic cumulative distribution function (CDF) at $\mathbf{p}$ is given by [64, 65]

$$N(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \quad (\text{m6})$$

where $\theta(\cdot)$ is the heaviside function. Note that $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ essentially describes how atoms are packed around $\mathbf{p}$. In computational practice K_α is set to be "large enough" approximating the continuous case via dense sampling.

Mathematical modeling of atomic accumulation

Modeling of the CDF was performed by employing the GROFIT routine [72]. A collection of candidate models including re-parametrized algebraic forms [73] of the Logistic model [74]

$$n_{LOG}(\mathbf{p}, l_\alpha(\mathbf{p})) = A(\mathbf{p}) \cdot \{1 + \exp(\frac{4 \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_\alpha(\mathbf{p})) + 2)\}^{-1} \quad (\text{m7})$$

, of the Gompertz model [75]

$$n_{GOM}(\mathbf{p}, l_\alpha(\mathbf{p})) = A(\mathbf{p}) \cdot \exp(-\exp(\frac{e \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_\alpha(\mathbf{p})) + 1)) \quad (\text{m8})$$

with $e = \exp(1)$, of the modified Gompertz model [72]

$$\begin{aligned} n_{MGOM}(\mathbf{p}, l_\alpha(\mathbf{p})) = & A(\mathbf{p}) \cdot \exp(-\exp(\frac{e \cdot t(\mathbf{p})}{A(\mathbf{p})} \cdot (s(\mathbf{p}) - l_\alpha(\mathbf{p})) + 1)) \\ & + A(\mathbf{p}) \cdot \exp(w(\mathbf{p}) \cdot (l_\alpha(\mathbf{p}) - l_{shift}(\mathbf{p}))) \end{aligned} \quad (\text{m9})$$

and of the Richards model [76]

$$\begin{aligned} n_{RIC}(\mathbf{p}, l_\alpha(\mathbf{p})) = & A(\mathbf{p}) \cdot \{1 + \tilde{q}(\mathbf{p}) \cdot b(\mathbf{p}) \cdot \exp(-k(\mathbf{p}) \cdot l_\alpha(\mathbf{p}))\}^{-1/\tilde{q}(\mathbf{p})} \\ \text{with } b(\mathbf{p}) = & \exp(1 + \tilde{q}(\mathbf{p}) + k(\mathbf{p}) \cdot s(\mathbf{p})) \text{ and} \\ k(\mathbf{p}) = & \frac{t(\mathbf{p})}{A(\mathbf{p})} \cdot (1 + \tilde{q}(\mathbf{p}))^{1+1/\tilde{q}(\mathbf{p})} \end{aligned} \quad (\text{m10})$$

were fitted on $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $l_\alpha(\mathbf{p})$ -direction where $\{A(\mathbf{p}), t(\mathbf{p}), s(\mathbf{p}), \tilde{q}(\mathbf{p}), w(\mathbf{p}), l_{shift}(\mathbf{p})\}$ are model parameters. The mathematical model that best fitted $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ traces along $l_\alpha(\mathbf{p})$ -direction

was selected based on minimization of an Akaike information criterion (see [72] for algorithmic details).

Following [65], model parameters interpretation was performed with respect to the inflection point $\xi(\mathbf{p}) = \{l_\alpha(\mathbf{p}) \mid \frac{\partial^2 n(\mathbf{p}, l_\alpha(\mathbf{p}))}{\partial l_\alpha(\mathbf{p})^2} = 0\}$ that determines the location along $l_\alpha(\mathbf{p})$ -direction where the radial distribution function (RDF), $\frac{\partial n(\mathbf{p}, l_\alpha(\mathbf{p}))}{\partial l_\alpha(\mathbf{p})}$, maximizes. The RDF maximum value is given by the parameter $t(\mathbf{p})$ accounting for the maximum atomic accumulation rate (or, equivalently, for the maximum atomic density) around $\mathbf{p}$. Parameter $A(\mathbf{p})$ is the asymptote value of the fitted model, i.e., $n(\mathbf{p}, l_\alpha(\mathbf{p}) \rightarrow \infty) = A(\mathbf{p})$, describing what happens when $L(\mathbf{p})$ becomes arbitrary large. Parameter $s(\mathbf{p})$ determines the location along $l_\alpha(\mathbf{p})$ -direction where the lag domain ends, i.e., the size of the lag atomic accumulation domain. Interpretation of parameter $t(\mathbf{p})$ is retrieved in terms of the ratio $t(\mathbf{p}) = \frac{A(\mathbf{p})}{os(\mathbf{p})=o(\mathbf{p})-s(\mathbf{p})}$ with $o(\mathbf{p})$ determining the location along $l_\alpha(\mathbf{p})$ -direction where the asymptote atomic accumulation domain begins. Parameter $\tilde{q}(\mathbf{p})$ affects the shape of the Richards model curve, as well as, the location of the inflection point along $l_\alpha(\mathbf{p})$ -direction thus plays the role of the summary atom-packing parameter. Parameters $w(\mathbf{p})$ and $l_{shift}(\mathbf{p})$ of the modified Gompertz model indicate the location and the slope, respectively, of a second increase in the modified Gompertz model curve (see [72]). The Logistic and the Gompertz model are retrieved from the Richards model for $\tilde{q}(\mathbf{p}) = 1$ and $\tilde{q}(\mathbf{p}) \rightarrow 0$, respectively, as shown in [77], thus they are considered as special cases of the Richards model.

Cumulative hydropathicity-property functions

The hydropathic density of the atomic environment around $\mathbf{p}$ was approximated in terms of [64]

$$m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))}{N(\mathbf{p}, l_\alpha(\mathbf{p}))} \sim \text{kcal}/(\text{mol} \equiv \text{atom}) \quad (\text{m12})$$

with $h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \cdot HI_i^w$

where $\mathbf{h}^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ corresponds to the cumulative zero-order hydropathic pore moment function [78] with $HI_i^w = HI_i + w_i$ representing the i -th atomic hydrophobic index in accordance with the corrected Kapcha-Rossky atomic hydrophobic scale presented in [64] (note that the Kapcha-Rossky atomic hydrophobic scale was initially published in [79] where a HIS atom was missing and was later added in [64]) with additive Gaussian noise $w_i \in \mathcal{N}(\mu=0, \sigma=0.001)$. The superscript "(0)" indicates the moment order.

The hydropathic interatomic interaction strength (HIIS) at $\mathbf{p}$, i.e., the average interaction strength between an atomic component found within the cluster of size $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ and its surroundings, was approximated in terms of the hydropathic imbalance (or interaction strength) pore function [64, 65]

$$\vec{m}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))}{N(\mathbf{p}, l_\alpha(\mathbf{p}))} \sim \text{kcal} \cdot \text{\AA}/(\text{mol} \equiv \text{atom})$$

with $\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \sum_{i=1}^{N_c} \theta(l_\alpha(\mathbf{p}) - \|\mathbf{c}_i - \mathbf{p}\|) \cdot HI_i^w \cdot \vec{r}_{\mathbf{p},i} \quad (\text{m13})$

$$= \underbrace{h_x^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{x}} + h_y^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{y}}}_{\vec{h}_{xy}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))} + \underbrace{h_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{z}}}_{\vec{h}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))}$$

where $\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ corresponds to cumulative first-order hydropathic

pore moment function [78] quantifying the hydrophobic inter-cluster interaction strength (HIcIS) with $\vec{r}_{\mathbf{p},i}$ being the vector from $\mathbf{p}$ to $\mathbf{c}_i$. The superscript "(1)" indicates the moment order.

Introduction of the weak noise source w_i practically guarantees that $|h^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))|$ and $||\vec{h}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))||$ are non-zero for every combination of $\mathbf{p}$ and $l_\alpha(\mathbf{p})$ while their scaling behavior remains practically unaffected. Throughout this study we consider pore's physicochemical field characteristics to be adequately described in terms of the axial component, $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) = \vec{h}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))/N(\mathbf{p}, l_\alpha(\mathbf{p})) = m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot \hat{\mathbf{z}}$ given that the magnitude of the radial component, $||\vec{m}_{xy}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))|| = ||\vec{h}_{xy}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))||/N(\mathbf{p}, l_\alpha(\mathbf{p}))$, is vanishingly small relative to $||\vec{m}^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))||$ after a cut-off, lag-domain scale (see SI, S2). Hence, we focus throughout this study only on the scaling behavior and topology of the HIIS axial field component which can occupy only two states; an "in" state which is characterized by $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ pointing towards the intracellular side (IS), i.e., by $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$, and an "out" state which is characterized by $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ pointing towards the extracellular side (ES), i.e., by $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$. Topological changes in HIIS axial field component are detected according to the algorithmic scheme presented in [64] (see SI, S3).

Finite-size scaling of HIIS

In accordance to [65], a scale-invariant interval of the HIIS axial component corresponds to combinations of $\mathbf{p}$ with α for which the power-law approximation

$$||\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))|| \sim l_\alpha(\mathbf{p})^{\gamma(\mathbf{p})} \quad (\text{m14})$$

is accurately satisfied indicating that HIIS stabilizing the cluster of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around $\mathbf{p}$ span a range up to $\sim l_\alpha(\mathbf{p})$ Å. Sign of $\gamma(\mathbf{p})$ quantifies the rate at which intensity and range of HIIS increase or decrease for increasing atomic cluster size. From a HIIS-network stand-

point, $\gamma(\mathbf{p})$ indicates whether HIs network interconnectivity is up- or down-regulated, i.e., whether HIs contacts, e.g., hydrogen bonds, are created or destroyed within the structure. The energy levels associated with HIs contacts creation (or destruction) guaranteeing stability of the atomic cluster can then be approximated by

$$U(\mathbf{p}, l_\alpha(\mathbf{p})) := ||\vec{h}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))||/l_\alpha(\mathbf{p}) \sim N(\mathbf{p}, l_\alpha(\mathbf{p})) \cdot l_\alpha(\mathbf{p})^{\gamma(\mathbf{p})-1} \quad (\text{m15})$$

measured in kcal/(mol $\equiv$ atomic cluster).

Results

In Figure 1 we demonstrate that atomic accumulation around NaV1.7's pore follows a sigmoid profile which can be accurately described by the Richards model (SI, S4). Inflection behavior of atomic accumulation can thus be described in terms of the inflection points, $\xi(\mathbf{p})$, so that the atomic environment around the pore is partitioned into four consecutive domains spanning NaV1.7 from the inside to the outside (Figure 1(b),(c)). Structural locations of the inflection points correspond to intra-channel regions where the atomic density maximizes and were found to closely follow the PM-VSs spatial transition (see SI, S5 for calculation of PM-VSs spatial transition characteristics). Accordingly, $\xi(\mathbf{p})$ serves as a macroscopic boundary line splitting the atomic environment around NaV1.7's pore into two phases, namely, a pre-inflection phase for $l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ and a post-inflection phase for $l_\alpha(\mathbf{p}) > \xi(\mathbf{p})$ accounting for atomic sub-environments containing mainly structural components belonging to the PM and VSs, respectively (Figure 1(b)).

Figure 1: **Atomic accumulation around NaV1.7's pore.** (a),(b), Cartoon illustration of the NaV1.7 human channel (side and top views). The atomic environment around the pore is partitioned into three consecutive domains; a lag domain, an inflection domain, and an asymptote domain. The upper boundaries of the lag domain, of the pre-inflection domain and of the post-inflection domain are also shown in terms of their statistical representations $\langle \alpha_s \rangle$, $\langle \alpha_\xi \rangle$, $\langle \alpha_\xi \rangle \approx \langle \alpha_\nu \rangle$ and $\langle \alpha_o \rangle$, respectively. Helical structures in (a) and (b) forming the PM and the VSs are colored with red and blue color, respectively. (c), Traces of statistical representations of the normalized (with respect to N_c) atomic CDF, $\langle \bar{N}(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$, and of its best-fitted Richards model $\langle n(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$ are plotted in log-scale with shaded areas around $\langle N(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$ indicative of 95% confidence intervals. Statistical representations of scalars are calculated according to SI, S6.

Based on the geometrical partition scheme summarized in Figure 1

we proceeded with mapping of missense *SCN9A*-gene mutation intra-channel structural locations. For that, we utilized a set of well-studied GOF NaV1.7 mutations phenotypically related with IEM, SFN and PEPD pain disease (total number of pain-related mutation sites: 36) and a set of neutral NaV1.7 mutations, i.e., NaV1.7 variants not causing disease, (total number of neutral mutations: 48) and retrieved their corresponding two-dimensional map site (SI, S7).

Figure 2: Geometrical mapping of missense *SCN9A*-gene mutation sites. (a), Two sets of missense *SCN9A*-gene mutation are employed; a pain-related set containing IEM, PPD and SFN mutation sites and a neutral set containing mutation sites not related to pain disease (SI, S7). Lines α_s , α_ξ and α_o highlight the boundaries among consecutive domains. Map areas "a1", "a2" and "a3" indicate maximizations in density of mutation sites. (b), Histogram of mutation sites along α -direction. (c), Histogram of mutation sites along $\mathbf{p}$ -direction. Red- and blue-colored histograms account for distribution of pain-related and neutral mutation sites, respectively.

The majority (i.e., 54%) of pain-related mutations sites are sparsely

distributed within the inflection domain and, more precisely, attracted toward a centrally-located map area along the inflection-points line α_ξ (see Figure 2(a), area "a2"). On the other hand, the majority (i.e., 75%) of neutral mutation sites are distributed within the second part of the inflection domain so that their map density tends to maximize approximately in the middle of the second part of the inflection domain (see Figure 2(a), area "a3"). The rest 46% of pain-related mutation sites are found within the lag domain toward the intracellular side of the NaV1.7. In particular, their map density maximizes in the vicinity of the boundary line α_s toward the IS (see Figure 2(a), area "a1"). Taken together, these observations indicate that for decreasing molecular scale the probability of a missense *SCN9A*-gene mutation to translate into a pain-related phenotype increases which is of little surprise considering that mutations affecting NaV1.7's interior are more likely to perturb packing of S5-S6 pore-forming helices and, consequently, affect pore's gating behavior.

Mapping of hydropathic density profile reveals the formation of a large HP incorporating the central cavity (CC) and the AG (Figure 3). Specifically, we demonstrate that the center of the pore is lined by predominantly hydrophobic atomic components expanding toward the IS where occlusion of the pore takes place by the ring of Y405 (DI), F960 (DII) F1449 (DIII) and F1752 (DIV) residues which are known to form the NaV1.7's activation gate (AG) (Figure 3). Macroscopically, the wide CC translates into a structural contraction event as the outer surface radius is locally minimized so that the channel is split into two funnel-like structural compartments (see trace of $L(\mathbf{p})$ on Figure 3). Hydropathic density variations can be summarized in terms of three visually-distinguishable contour domains; the $T_2^{(0)}$ domain accounting for HP's formation, the $T_1^{(0)}$ domain accounting for a hydrophobic pore wall placed between the hydrophilic selectivity filter (SF) and the hydrophilic extracellular side (ES), and the hydrophilic $T_3^{(0)}$ domain covering the largest contour map area and incorporating

both the SF and the ES mouth. The SF's microenvironment is formed by the residues D361 (DI), E930 (DII), K1406 (DIII) and A1698 (DIV) where a bare sodium ion of radius $\approx 1.8 \text{ \AA}$ can exactly fit in (Figure 3). Strikingly, approximately 53% of pain-related mutation sites are found within $T_2^{(0)}$ while map areas "a1" and "a2" are distributed along HP's boundary, i.e., occupy contour map area that belongs to both domains $T_2^{(0)}$ and $T_3^{(0)}$ (Figure 3). On the other hand, only 10% of neutral mutations sites are located within the $T_2^{(0)}$ while area III is distributed solely within $T_3^{(0)}$ (Figure 3).

Figure 3: Spatial profile of the hydropathic density around NaV1.7's pore. Contour map of the hydropathic density pore function, $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$, for $\mathbf{p} \in P$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. Blue and red color contour domains represent hydrophobic and hydrophilic domains around the pore, respectively. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict geometrical pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ represents the PM-VSs spatial transition. Dashed black lines $s(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$ account for the boundaries between different domains (see Methods). Zero-crossing points of $m^{(0)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ collected in $\Omega^{(0)}$ describe the boundaries among hydropathic domains $T_1^{(0)}$, $T_2^{(0)}$ and $T_3^{(0)}$. Pain-related and neutral mutation sites are represented by different symbols (see Legend). Mutations sites highlighted

with red color correspond to misclassified events. Grey-shaded areas "a1", "a2" and "a3" highlight map areas where map density of mutation sites maximizes.

Given that mutations affecting a protein's hydrophobic interior pose a high risk for detrimental destabilizations [80,81], we hypothesize that mutations occurring at structural locations in proximity to HP are more likely to be related with pain phenotype as their perturbing effect can directly affect tight hydrophobic packing around channel's mass center. We test this hypothesis by calculating the distance between each mutation structural location and HP's boundary (SI, S8a) and fed retrieved distances into a binary classifier. We achieved to classify correctly 29 (out of 36) and 38 (out of 48) of pain-related and neutral, respectively, mutations correctly with a cut-off distance of 18.13 Å (Figure 4). This translates to an area under receiver operating characteristics (ROC) curve of 0.787 and pathogenicity prediction with specificity of 0.791 and sensitivity of 0.805 (Figure 4(a)).

Misclassified pain-related mutations S211P, L823R, W1538R, I720K, I739V and T1596I are found outside of $T_2^{(0)}$ thus not in proximity to HP (Figure 3 and 4(b)). Only a single pain-related misclassification is found within $T_2^{(0)}$, namely, R185H (Figure 3 and 4(b)). This striking misclassification is due to the elementary statistical approach adopted for calculating distance scores which fails to capture the complex geometry of $T_2^{(0)}$ (SI, S8a). Misclassified neutral mutations are V1428I, T920N, V194I, V1613I, T1398N, I1399D, S1419N, D1662A, D1674A and K1700A are found either inside $T_2^{(0)}$ or in close proximity to HP's boundary with a tendency to cluster around the SF (Figure 3 and 4(b)).

Figure 4: **Binary classification of NaV1.7 variants based on their distance from the HP's boundary.** (a), ROC curve constructed based on the HP's-boundary distance data (for construction of data set see SI, S8a). Optimal threshold value corresponds to specificity and sensitivity values of 0.791 and 0.805, respectively. Area under ROC curve is 0.787. (b), Visualization of ROC curve data. Optimal threshold value 18.13 Å is marked with black dashed line. Shaded area around distance values indicates the 95% confidence intervals. ROC curve is constructed in R [66] by using the pROC package [82].

HI's underlying atom packing around NaV1.7's pore can be adequately described in terms of the HIIS axial field component illustrated in Figure 5(a). HIIS axial field topology is organized into five domains, namely, $T_1^{(1)}$, $T_2^{(1)}$, $T_3^{(1)}$, $T_4^{(1)}$ and $T_5^{(1)}$ (Figure 5(a)). The centrally-located $T_5^{(1)}$ domain covers the largest map area and roughly dichotomizes the contour map into two pseudo-symmetric

parts, namely, an ES part incorporating $T_1^{(1)}$ and $T_3^{(1)}$, and an IS part incorporating $T_2^{(1)}$ and $T_4^{(1)}$. Pain-related mutation sites are solely found within the $T_4^{(1)}$ (58%) and $T_5^{(1)}$ (42%) domains. On the other hand, neutral sites are found within the $T_3^{(1)}$ (14%), $T_4^{(1)}$ (19%) and $T_5^{(1)}$ (67%) domains.

In order to decode mutation sites clustering behavior on the contour map of Figure 5(a) we adopted a phenomenological perspective based on the existence of a critical point, $\xi(\mathbf{p}_{crit.})$, associated with the SF (SI, S9). A crucial result that motivated us to adopt such an approach is that pain-related mutation sites are attracted toward the critical point in sheer contrast to neutral mutation sites which are repelled from it (Figure 5(b)). We term this phenomenon critical clustering. Geometrically, the formation of the critical mutation sites cluster reflects the tendency of structural locations of pain-related mutations to minimize their distance from the surface of the critical sphere of radius $\xi(\mathbf{p}_{crit.}) \approx 33.4 \text{ \AA}$; intuitive graphical representation of this phenomenon is provided in Figure 5(a) where we show that "hot" areas "a1" and "a2" intersect with the dashed green line $\xi(\mathbf{p}_{crit.})$ representing critical sphere's surface (Figure 5(a)).

Armed with the observation that pain-related mutation sites (in contrast to neutrals) tend to occupy locations at critical hydropathic-interactions distance from the SF, we focused on the scaling behavior of HIIS around the critical point that is adequately described in terms of the power-law scheme

$$m_z^{(1)}(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \sim \begin{cases} l_\alpha(\mathbf{p}_{crit.})^{\gamma_{partI}(\mathbf{p}_{crit.})} & \text{for } s(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \xi(\mathbf{p}_{crit.}) \\ l_\alpha(\mathbf{p}_{crit.})^{\gamma_{partII}(\mathbf{p}_{crit.})} & \text{for } \xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq o(\mathbf{p}_{crit.}) \end{cases} \quad (r1)$$

accounting for a HIs-network expansion and contraction within the pre- and post-inflection phase intervals $s(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \xi(\mathbf{p}_{crit.})$ and $\xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq o(\mathbf{p}_{crit.})$ with rates of $\gamma_{partI}(\mathbf{p}_{crit.}) = 2.27 \pm$

0.18 and $\gamma_{partII}(\mathbf{p}_{crit.}) = -5.18 \pm 1.02$, respectively (Figure 5(c)). On the left of the interval $\xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \nu(\mathbf{p}_{crit.})$, both, the range and intensity of HIIS maximize as the HIs-network configuration exceeds its critical size marking the transition from the pre-inflection phase toward the post-inflection phase. The energy levels associated with this phase transition are given by

$$U(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \sim \begin{cases} N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \cdot l_\alpha(\mathbf{p}_{crit.})^{\gamma_{partI}(\mathbf{p}_{crit.})-1} & \text{for } s(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \xi(\mathbf{p}_{crit.}) \\ N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \cdot l_\alpha(\mathbf{p}_{crit.})^{\gamma_{partII}(\mathbf{p}_{crit.})-1} & \text{for } \xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \nu(\mathbf{p}_{crit.}) \end{cases} \quad (r2)$$

where $N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$ can be replaced with its best-fitted Richards model, $n_{ric}(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$ (see caption of Figure 5 for Richards model parameters), providing with an estimation of the atom-packing energy (AE) (Figure 5(c)). Similarly to the NaVAb case [65], AE maximization occurs in the vicinity of the interval $\xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \nu(\mathbf{p}_{crit.})$ so that energetic coupling of the PM with the VSs is dictated by the phase transition (Figure 5(c),(d)).

Equation r1 indicates that interatomic HIs law is robust to microscopic modifications of the atomic structure, e.g., addition, removal or deletion of a small number of atoms corresponding to small-amplitude perturbations of $N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$ [83]. This happens however at the cost of re-tuning HICIS and, hence, also of AE that, in the case of small-amplitude perturbations of $N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$, are expected to be up- and down-regulated toward and away from the critical point, respectively, in a power-law fashion described by r2 (Figure 5(c)). Mutation-induced perturbations propagating throughout the structure are thus expected to be amplified in the vicinity of the critical point while, on the other hand, to be damped out toward the interior (i.e., toward the HP and the SF) and toward channel exterior bounded by outer pore surface radius. Given that mutations in the structural proximity of the SF are highly likely to have a deleterious LOF effect [84], observed

damping-out mechanism might act as a shield protecting SF's biological machinery from mutations occurring within the pre-inflection phase. On the other hand, mutations occurring in the post-inflection phase are unlikely to affect the SF as they have to overcome a large energy barrier in order to reach channel interior. We thus hypothesize that critical clustering of pain-related mutations might actually reflect a trade-off between the two extremes; a destructive destabilization and an insignificant one.

Figure 5: **Spatial profile of HIIS along NaV1.7's pore.** (a), Contour map of HIIS axial part, $m_z^{(1)}(p, l_{\alpha}(p))$, for $p \in P$ and $\alpha = 1, 2, \dots, K_{\alpha} = 800$. Blue and red color contour domains represent configurations of $\bar{m}_z^{(1)}(p, l_{\alpha}(p))$ with orientation "out" and "in", respectively. Black lines $R(p)$, $D(p)$ and $L(p)$ depict geometri-

cal pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ represents the PM-VSs spatial transition. Dashed black lines $s(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$ account for the boundaries between different domains. $\xi(\mathbf{p}_{crit.})$ represents the critical radius. Black arrows [a], [b], [c], [d] and [e] indicate the clustering behavior of the zero-crossing points of $m_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ collected in $\Omega^{(1)}$ describing boundaries among contour domains $T_1^{(1)}$, $T_2^{(1)}$, $T_3^{(1)}$, $T_4^{(1)}$ and $T_5^{(1)}$. Mutations sites highlighted with red color correspond to misclassified events. Grey-shaded areas "a1", "a2" and "a3" highlight map areas where map density of mutation sites maximizes. **(b)**, Traces of normalized (with respect to corresponding maximum values) median distances between pain-related and neutral mutation sites from the critical point, $\xi(\mathbf{p})$, represented as $\bar{D}_\xi^{pain} = \bar{D}_\xi(\mathbf{p})(V_{pain})$ and $\bar{D}_\xi^{neut.} = \bar{D}_\xi(\mathbf{p})(V_{neut.})$, respectively, are plotted for $\mathbf{p} \in P$ (see SM, S8b for calculation of $D_{\xi(\mathbf{p})}(V_{pain})$ and $D_{\xi(\mathbf{p})}(V_{neut.})$). **(c)**, Power-law scaling of HHS axial part for $\mathbf{p} = \mathbf{p}_{crit.}$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. The best-fitting first-part-inflection-domain and second-part-inflection-domain power-law approximations are also plotted with mean absolute relative fitting errors being 0.09 ± 0.01 and 0.15 ± 0.03 , respectively. **(d)**, Power-law-like behavior of AE, $U(\mathbf{p}, l_\alpha(\mathbf{p}))$, for $\mathbf{p} = \mathbf{p}_{crit.}$ and $\alpha = 1, 2, \dots, K_\alpha = 800$. First- and second-part-inflection-domain modeling approximations of AE are also plotted with their mean absolute relative modeling errors being 0.11 ± 0.02 and 0.14 ± 0.03 , respectively. AE model extrapolation toward the lag-domain $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$ and the asymptote-domain $l_\alpha(\mathbf{p}) > o(\mathbf{p})$ result in a mean absolute relative fitting error of 6.06 ± 16.0 and 1.55 ± 6.39 , respectively. Richards model parameters used for modeling AE are $\{A(\mathbf{p}_{crit.}) = 1.03, t(\mathbf{p}_{crit.}) = 0.03, s(\mathbf{p}_{crit.}) = 18.16, \tilde{q}(\mathbf{p}_{crit.}) = 0.47\}$.

We tested the critical-clustering hypothesis by calculating the distance each mutation structural location and SF's critical point (SI, S8b), and feeding retrieved distances into a binary classifier. We achieved to classify correctly 28 (out of 36) and 39 (out of 48) of pain-related and neutral mutations correctly with a cut-off distance of $5.8 \text{ \AA}$. This translates to an area under receiver operating characteristics (ROC) curve of 0.824 and pathogenicity prediction with specificity of 0.812 and sensitivity of 0.777 (Figure 6(a)). Intuitive geometrical depiction of this result requires to think of a "hot" spherical shell

squeezed between the spheres of radii $\xi(\mathbf{p}_{crit.})+5.8 \text{ \AA}$ and $\xi(\mathbf{p}_{crit.})-5.8 \text{ \AA}$ centered at $\mathbf{p}_{crit.}$ incorporating areas I and II thus containing the majority of pain-related structural locations. This tendency can be deduced from Figure 5 where we can see that correctly-classified pain-related and neutral sites tend to minimize and maximize, respectively, their distance from the critical radius $\xi(\mathbf{p}_{crit.})$.

The opposite holds for misclassified mutations. Note however that due to the pore points offset, distances of sites from $\xi(\mathbf{p}_{crit.})$ line on Figure 5 are not equal with the distances of their structural locations from the surface of the sphere of $\xi(\mathbf{p})$ (discrepancies are of order $\approx 3.13 \pm 4.63 \text{ \AA}$). Misclassified pain-related mutations are I136V, W1538R, I1461T, R185H, I228M, I720K, I739V and T1596I indicating that sensitivity output is qualitatively similar to the HP-based classification attempt. On the other hand, quality of specificity differs significantly among classification attempts as critical-point distance criterion misclassified neutrals M145L, M146S, R1207K, T1210N, V1613I, D890N, K1412I, K1415I and S1419N are clustering within the "hot" spherical shell in proximity to HP's boundary.

Finally, in order to harvest the classification power of both predictors, we linearly combined distance metrics by calculating a weighted distance average (SI, S10). The weighted distance average achieved to classify correctly 29 (out of 36) pain-related mutations and 45 (out of 48) neutrals, i.e., sensitivity = 0.805, specificity = 0.937, area under ROC curve = 0.872 (Figure 7(a)). The threshold weighted distance value is $9.6 \text{ \AA}$ and it indicates which mutations are found in proximity to SF's critical point and HP's boundary.

Figure 6: **Binary classification of NaV1.7 variants based on their distance from SF's critical point.** (a), ROC curve constructed based on SF's-critical-point distance data (for construction of data set see SI, S8b). Optimal threshold value 5.8 Å corresponds to specificity and sensitivity values of 0.812 and 0.777, respectively. Area under ROC curve is 0.824. (b), Visualization of ROC curve data. Optimal threshold value 5.8 Å is marked with black dashed line. ROC curve is constructed in R [66] by using the pROC package [82].

The relatively-low sensitivity of the weighted distance average is not surprising if we consider that both classifications attempts failed in correctly classifying pain-related mutations found far away from the HP and from the SF; misclassified pain-related mutations are I136V, W1538R, I1461T, R185H, I720K, I739V and T1596I and all of them are found within the post-inflection phase with the exception of I1461T which is located within the lag domain but still far away from the HP

and from the SF (Figure 3, 5(a) and 7). On the other hand, misclassified neutrals R1207K, V1613I and S1419N occupy "hot" spots located in proximity to SF's critical point and HP's boundary (Figure 3, 5(a) and 7).

Figure 7: **Binary classification of NaV1.7 variants based on the weighted distance average.** (a), ROC curve constructed based on weighted distance data (for construction of data set see SI, S10). Optimal threshold value 9.6 Å corresponds to specificity and sensitivity values of 0.805 and 0.937, respectively. Area under ROC curve is 0.872. (b), Visualization of ROC curve data. Optimal threshold value 9.6 Å is marked with black dashed line. ROC curve is constructed in R [66] by using the pROC package [82].

Discussion and concluding remarks

Criticality hypothesis in biology aims at explaining how emergence of power-laws increases biological system's robustness and efficiency hand-in-hand with evolution. Empirical evidence for complex biological systems operating near critical points include cases of gene expression [85], DNA sequences [86], protein structures [59–62], cell growth [87] and neuronal dynamics underlying brain activity [88]. In practice, criticality implies that system dynamics are delicately balanced between an ordered state where perturbations are damped-out and a disordered state where perturbations are amplified. Consequences of critical dynamics are associated with optimal information processing [89], enhanced network stability [90] and maximal sensitivity to external stimuli [91].

In this work, instead of trying to predict the effect of missense *SCN9A*-gene mutations via comparing mutant NaV1.7 structures *in silico*, we extracted hydropathic features of the wild-type atomic environment encoding NaV1.7's response to mutation-induced variations. Stated differently, we hypothesized that some regions of the atomic environment around NaV1.7's pore exhibit higher sensitivity to mutation-induced perturbations due to the long-range nature of HIs guaranteeing their stability; a hallmark of SOC is that avalanche-like perturbing effects are amplified and fast-spreading throughout critical network locations [92]. To test this hypothesis we mapped mutation structural locations on their corresponding mutation sites and probed topological and scaling hydropathic characteristics of the atomic bulk around the pore. Importantly, this is possible due to the relatively-large number of pain-related mutations providing with the opportunity of structure-based mutation statistics and, consequently, identification of densely-populated (by mutation sites) structural domains.

The starting point of the presented procedures was the approximation of the atomic cumulative distribution function around NaV1.7's pore demonstrating that packing of atoms follows a sigmoid accumu-

lation pattern. The generality of the Richards model was found to be adequate for this modeling purpose verifying the sigmoid accumulation hypothesis and, consequently, revealing a biphasic spatial organization of the atomic environment around the pore dictated by the spatial transition from the PM from the VSs. We showed that the pore is lined by a HP dominating within channel interior and that HIs stabilizing atom-packing around the SF are critically tuned with respect to the local inflection points. This NaV1.7 feature is shared with its evolutionary-ancestor, namely, with the pre-open NaVA channel, suggesting that HIs scale-invariance might be conserved from NaVChs of bacterial homomers to NaVChs of mammalian heteromers [65].

Pain-related mutations tend to occupy structural locations in proximity to the HP while maintaining a critical HIs-distance from the SF. Geometrically, this result indicates that the majority of pain-related mutations are found within a spherical shell around the SF incorporating parts of the HP. What might be the evolutionary principle underlying this non-random mutation distribution around NaV1.7's pore? Given that the DEKA SF sequence is conserved among human and non-human NaVCh templates [93], we propose that expression of mutations at critical hydropathic-interactions distance from the SF might reflect an evolutionary trade-off between potentially-deleterious destabilizations occurring too close to the SF and insignificant polymorphisms occurring far away from it. According to this rationale, mutations occupying critical hydropathic-interactions network locations lead to a GOF effect by increasing channel's configuration space and, consequently, expanding physiological range of ionic currents, while not risking structure deletion or severe destabilizations that can induce a LOF effect [84].

Misclassification of seven pain-related events found within the post-inflection phase (namely, of I136V, W1538R, I1461T, R185H, I720K, I739V and T1596I) suggests that the destabilizing mutation effect within the post-inflection phase and, specifically, within the VSs needs to be locally investigated. In particular, misclassified pain-related

events are likely to affect local properties of the VSs which are however crucial for physiological gating and ionic currents. It might therefore be useful for future studies to consider a decoupling of the PM from the VSs in order to focus solely on the cumulative hydropathic topology and HIs-networking within the VSs. Moreover, a goal for follow-up studies is to include biophysical characteristics of substituted amino acids (e.g., size, charge, hydropathicity-property, degree of conservation) into our model as this might not only improve classification accuracy but also provide with a more detailed picture of the mutation effect.

Admittedly, a limitation of this study is the small (from a statistics point of view) number of available mutation events. To resolve this issue and provide with stronger statistical validation, we may consider in future studies to increase number of neutral and pain-related mutations, for example, by introducing NaV1.7 variants found in the Genome Aggregation Database (gnomAD) [94]. A methodological weakness is that we neglected radial hydropathic effects. In particular, even if the amplitude of the radial HIIS component is small in comparison to the amplitude of the axial HIIS component, its role might be non-negligible for interactions between ions and pore walls. Moreover, it is important to remind at this point that we chose to work with a closed-state structural model of the NaV1.7 which most likely corresponds to a pre-open state of NaV1.7's gating cycle. The rationale underlying this choice was that we sought to establish a connection (if any) between the closed-state NaV1.7 and the closed-state NaVAb before we focus on the human NaV1.7 structure captured at an inactivated state [93].

Our findings suggest that pathogenicity of NaV1.7 variants can be predicted if corresponding mutation structural locations are in structural and, hence, also hydropathic proximity to either the HP or the SF. Prediction of pathogenicity can then be performed with negligible computational effort and similar or even higher accuracy to [54] (reported accuracy: 0.81) but also to the more recent study of [56] where

a MLE computational pipeline was employed (reported accuracy on the human NaV1.7 template: 63.5%). In an era where MLE pipelines are becoming more and more popular, presented phenomenological framework could provide biophysical rationalization to MLE-retrieved observations suggesting that hydropathicity property is a key-feature for predicting functional effects of missense variants in NaVCh variants [84].

Supplementary Information

S1. Navigating through NaV1.7's pore

The HOLE routine [71] was called $n = 100$ times where the control parameters are fixed, but the searching path is randomly chosen according to the rule $\text{RASEED} = i \cdot 1000$ with $i = 1, 2, \dots, n$ (Table S1). After the i -th call we obtain a set of points, let it be Q , where each point, $\mathbf{q}_i = (q_{x,i}, q_{y,i}, q_{z,i}) \in Q$, is located on a membrane-parallel plane sampled with distance $0.1 \text{ \AA}$ (Table S1). After the completion of the n calls, we collected the membrane-parallel planes containing n points and, under the Gaussian assumption, the geometrical location of a pore point was approximated by

$$\mathbf{p} = \left(\frac{1}{n} \sum_{i=1}^n q_{x,i}, \frac{1}{n} \sum_{i=1}^n q_{y,i}, \frac{1}{n} \sum_{i=1}^n q_{z,i} \right) \in P \quad (\text{S1})$$

where P contains $N_p = 920$ pore points (note that N_p depends on the choice of HOLE parameters (see Table S1)). Due to the skewness of the pore, x and y coordinates of pore points are non-zero. Hence, pore points are radially displaced from the z -axis with an average offset of $3.13 \pm 4.63 \text{ \AA}$.

Parameter	Value
CVECT	0.0 0.0 1.0
CPOINT	e
SAMPLE	0.1
MCDISP	0.03
MCKT	0.05
MCSTEP	1000
ENDRAD	25.
RASEED	$i \cdot 1000$

Table S1: **HOLE routine parameters.** Parameter CVECT specifies a vector which lies in the direction of the channel’s pore. Parameter CPOINT corresponds to the NaV1.7’s atomic center $\mathbf{e} = (e_x, e_y, e_z)$ of the 3D structure (see Methods). SAMPLE parameter determines the distance between two sampling plains in Å. Parameter MCKT specifies the initial "Boltzmann factor" to be used by Monte Carlo algorithm. Parameter MCSTEP specifies the number of steps performed by the Monte Carlo algorithm. Parameter ENDRAD specifies the pore radius in Å value above which the HOLE routine considers that the end of pore has been reached. Parameter RASEED resets the seed for the monte carlo algorithm of HOLE routine after the i -th call.

S2. Radial component of the hydropathic imbalance pore function

The field topology of the HIIS axial component, $\vec{m}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))$ (see eq. m13), is summarized in Figure S1. We neglected the impact of the radial component as it becomes significantly small already within the lag domain. This however does not necessarily imply that radial hydropathic effects are irrelevant for ion conduction and selectivity but rather that their contribution to NaV1.7’s stability after some cut-off scale is marginal.

As we may deduce from visual investigation of Figure S1, the magnitude of the radial component of the HIIS, $||\vec{m}_{xy}(\mathbf{p}, l_\alpha(\mathbf{p}))||$ (see eq. m13), attains its maximum values for small values of the α scaling

parameter at the ES and the IS. This observation reflects partially a finite-size effect introduced at the ES and IS channel boundaries. For scales larger than $l_{\alpha=55}(\mathbf{p})$ radial HIIS component becomes significantly smaller than the axial component as the statistical index $\langle ||\vec{m}_{xy}(\mathbf{p}, l_{\alpha}(\mathbf{p}))|| / ||\vec{m}(\mathbf{p}, l_{\alpha}(\mathbf{p}))|| \rangle_{\alpha > 55}$ remains close to ≈ 0.18 .

Figure S1: **Spatial profile of the magnitude of the radial components of the HIIS along NaV1.7's pore.** Contour map of the radial HIIS component, $||\vec{m}_{xy}(\mathbf{p}, l_{\alpha}(\mathbf{p}))||$, for $\mathbf{p} \in P$ and $\alpha = 1, 2, \dots, K_{\alpha} = 800$. Black lines $R(\mathbf{p})$, $D(\mathbf{p})$ and $L(\mathbf{p})$ depict geometrical pore characteristics. Magenta dashed line $\nu(\mathbf{p})$ represents the PM-VSs spatial transition. Black dashed lines $s(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$, respectively, represent boundaries between consecutive geometrical domains. Grey-shaded areas "a1", "a2" and "a3" highlight map areas where map density of mutation sites maximizes.

S3. Extracting topological information from cumulative hydropathicity-property function

Topological information from hydropathicity-property functions was extracted according to the following algorithmic scheme [64]:

- Let $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ be a hydropathicity-property function, then if for a given scaling index α there is a pair $\{\mathbf{p}' = \mathbf{p} - \Delta\mathbf{p}, \mathbf{p}\} \in P$ for which the sign-change condition $f(\mathbf{p}', l_\alpha(\mathbf{p}')) \cdot f(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ is satisfied, calculate the four-dimensional point

$$(\mathbf{s} = \mathbf{p}' + \frac{|f(\mathbf{p}', l_\alpha(\mathbf{p}'))|}{|f(\mathbf{p}', l_\alpha(\mathbf{p}'))| + |f(\mathbf{p}, l_\alpha(\mathbf{p}))|} \cdot \Delta\mathbf{p}, l_\alpha(\mathbf{s})) \quad (\text{S2})$$

that represents a zero-crossing point of $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction with $|\cdot|$ returning the absolute value of f and $\mathbf{s}$ obtained by linear interpolation. The set of all detected zero-crossing points of $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ along $\mathbf{p}$ -direction for a given scaling index α is represented as $\Gamma(\alpha)$ so that hydropathic topological information is summarized as

$$\Omega = \bigcup_{\alpha} \{\Gamma(\alpha)\} \quad (\text{S3})$$

S4. Geometric parameters determining atomic accumulation around NaV1.7's pore

Richards model parameters traces determining atomic accumulation around NaV1.7's pore are shown in Figure S2. Atomic accumulation is clearly differentiated among the ES and the IS so that a densening of the atomic environment occurs at the AG. This preferential increase in atomic density is initiated after the CC pore region resulting maximization of atomic accumulation rate takes place in combination with a shrinkage of the inflection domain, i.e., minimization of the distance between the lag domain and the asymptote domain, at the AG pore region (Figure S2(a)). In between the CC and the AG we mark the location of the Logistic model pore region capturing the sign-change of $\tilde{q}(\mathbf{p})$ where a local maximization of the mean absolute fitting error (MAFE) is observed (Figure S2(a),(b)). Notably, MAFE remains always smaller than 2% indicative of the goodness of Richards model fittings (Figure S2(b)).

Figure S2: **Geometric parameters of atomic accumulation around NaV1.7's pore.** (a), Traces of normalized Richards model parameters $\bar{A}(\mathbf{p})$, $\bar{t}(\mathbf{p})$, $\bar{o}s(\mathbf{p})$, and $\bar{q}(\mathbf{p})$ for $\mathbf{p} \in P$ (see Methods for model parameters definition). (b), Mean absolute fitting error (MAFE) of the Richards model approximation on the normalized CDF trace $\bar{N}(\mathbf{p}, l_\alpha(\mathbf{p}))$ for $\mathbf{p} \in P$. All normalizations were performed with respect to the maximum values of the corresponding traces.

S5. Modeling the PM-VSs spatial transition

We consider the NaV1.7 atomic structure to consist of two coupled atomic subsystems, namely, the PM and the VSs subsystems. Accordingly, we divided all N_c atoms into two groups, one containing only PM atoms including the PM-VS linker, i.e., atoms forming the PM and the linker helices, and one containing only VS atoms, i.e., atoms forming the VS helices. The total number of PM atoms is $N_{PM}=10109$ and are sampled from the residue sequences P229:K417 (DI), P839:T972 (DII), M1296:G1461 (DIII) and K1617:T1763 (DIV) (note that we consider the S4-S5 linker, as well as, extensions of the S6 helices to be part of the PM group). On the other hand, the total number of VS atoms is $N_{VS}=8458$ and are sampled from the residue

sequences L109:I228 (DI), P717:W838 (DII), K1165:G1295 (DIII) and P1485:A1616 (DIV).

The discretized RDF of $N(\mathbf{p}, l_\alpha(\mathbf{p}))$ atoms around $\mathbf{p}$ reads

$$G(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta N(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho(\mathbf{p})} \quad \text{with } l'_\alpha(\mathbf{p}) = l_\alpha(\mathbf{p}) - \Delta l_\alpha(\mathbf{p}) \quad (\text{S4})$$

where $\Delta l_\alpha(\mathbf{p}) = \frac{L(\mathbf{p}) - D(\mathbf{p})}{K_\alpha}$ is the thickness of a spherical shell around $\mathbf{p}$, $\Delta N(\mathbf{p}, l_\alpha(\mathbf{p})) = N(\mathbf{p}, l_\alpha(\mathbf{p})) - N(\mathbf{p}, l'_\alpha(\mathbf{p}))$ is the number of atoms found within the spherical shell of thickness $\Delta l_\alpha(\mathbf{p})$ centered at $\mathbf{p}$ and $\rho(\mathbf{p}) = \frac{N_c}{V(\mathbf{p})} = \frac{N_c}{\frac{4}{3} \cdot \pi \cdot (L(\mathbf{p}) - R(\mathbf{p}))^3}$ is the average atomic density around $\mathbf{p}$.

Using equation S4 we approximated the RDFs of the PM and VS atoms around $\mathbf{p}$ with

$$G_{PM}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{PM}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{PM}(\mathbf{p})} \quad (\text{S5})$$

and

$$G_{VS}(\mathbf{p}, l_\alpha(\mathbf{p})) = \frac{\Delta n_{VS}(\mathbf{p}, l_\alpha(\mathbf{p}))}{\frac{4}{3} \cdot \pi \cdot (l_\alpha(\mathbf{p})^3 - (l'_\alpha(\mathbf{p}))^3) \cdot \rho_{VS}(\mathbf{p})} \quad (\text{S6})$$

, respectively, where $N(\mathbf{p}, l_\alpha(\mathbf{p})) = N_{PM}(\mathbf{p}, l_\alpha(\mathbf{p})) + N_{VS}(\mathbf{p}, l_\alpha(\mathbf{p}))$, $\rho_{PM}(\mathbf{p}) = \frac{N_{PM}}{V(\mathbf{p})}$ and $\rho_{VS}(\mathbf{p}) = \frac{N_{VS}}{V(\mathbf{p})}$.

In order to obtain a relative measure of how the RDF of PM atoms varies with respect to the RDF of VS atoms, and vice versa, we used the PM-VSs equilibrium RDF [64, 65]

$$e(\mathbf{p}, l_\alpha(\mathbf{p})) = sm(G_{PM}(\mathbf{p}, l_\alpha(\mathbf{p})) - G_{VS}(\mathbf{p}, l_\alpha(\mathbf{p}))) \quad (\text{S7})$$

where $sm(\cdot)$ implements the Nadaraya-Watson kernel regression R-function [66] with a bandwidth parameter value of $bw = 15$. $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ is interpreted as the smoothed probability of finding a PM atom instead of a VS atom, and vice-versa, at distance $l_\alpha(\mathbf{p})$ from $\mathbf{p}$. Specifically, if $e(\mathbf{p}, l_\alpha(\mathbf{p})) > 0$ ($e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$) then the smoothed probability of finding a PM atom within the spherical shell of width $\Delta l_\alpha(\mathbf{p})$

is larger than that of finding a VS (PM) atom. Accordingly, what is of interest here is the sign-change behavior of $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ for increasing $l_\alpha(\mathbf{p})$. We investigated it by detecting for every $\mathbf{p}$ the pair $\{l'_\alpha(\mathbf{p}), l_\alpha(\mathbf{p})\}$ for which the sign-change condition $e(\mathbf{p}, l'_\alpha(\mathbf{p}))e(\mathbf{p}, l_\alpha(\mathbf{p})) < 0$ is satisfied and approximating with linear interpolation the root location

$$\nu(\mathbf{p}) = l'_\alpha(\mathbf{p}) - \frac{e(\mathbf{p}, l'_\alpha(\mathbf{p}))}{e(\mathbf{p}, l_\alpha(\mathbf{p})) - e(\mathbf{p}, l'_\alpha(\mathbf{p}))} \cdot \Delta l_\alpha(\mathbf{p}) \quad (\text{S8})$$

where $e(\mathbf{p}, l_\alpha(\mathbf{p}))$ changes sign along $l_\alpha(\mathbf{p})$ -direction.

S6. Statistical representation of scalar functions

Let $\langle \cdot \rangle$ be a statistical operator returning the median of the data set upon which it operates. Note that the choice of the median as a statistical measure reflects the fact that no assumption has been made for the distribution of data set values.

Statistical representation of scalar function $f(\mathbf{p}, l_\alpha(\mathbf{p}))$. A statistical representation of the scalar function $f(\mathbf{p}, l_\alpha(\mathbf{p}))$ for a given α was obtained in terms of $\langle f(\mathbf{p}, l_\alpha(\mathbf{p})) \rangle_\alpha$ where the subscript α indicates that the statistical operator acts for a given α , i.e., on the set of values $\Gamma(\alpha) = \{f(\mathbf{p}, l_\alpha(\mathbf{p})) \mid \mathbf{p} \in P\}$.

Statistical representations of scalar functions $s(\mathbf{p})$, $\nu(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$. Statistical representations of the functions $s(\mathbf{p})$, $\nu(\mathbf{p})$, $\xi(\mathbf{p})$ and $o(\mathbf{p})$ were obtained by the following algorithmic scheme;

Step 1. Let $f(\mathbf{p})$ represent one of the aforementioned scalar functions, then, for every $\mathbf{p}$ find the α indexes for which $|f(\mathbf{p}) - l_\alpha(\mathbf{p})|$ is minimized, i.e., $\alpha_f = \{\alpha \mid \min_{\alpha \in A} (|f(\mathbf{p}) - l_\alpha(\mathbf{p})|)\}$.

Step 2. Calculate the median $\langle \alpha_f \rangle$ of the data set α_f . Note that the minimum and the maximum value contained in α_f is given by $\min(\alpha_f)$ and $\max(\alpha_f)$.

S7. Missense *SCN9A*-gene mutations

A collection of well-studied gain-of-function *SCN9A* mutations are introduced which have been shown to be causally related with IEM, SFN and PEPD phenotypes (Table S2). A collection of neutrals *SCN9A* mutations is incorporated from [54] containing variants not causing biophysical abnormalities (nABN) and homologous single-nucleotide polymorphism (hSNPs) (Table S3). In addition, we introduce also four variants which have been previously classified as neutrals according to classification procedures described in [95] (see caption of Table S3). Selection criterion for hSNPs is that they share at least 90% nucleotide sequence identity with the *SCN9A* homologous genes and was implemented based on the NCBI HomoloGene Database [96]. All nABN cases are associated with *in vitro* observations.

What is of interest for this study is the site where the mutation occurs within the structure, i.e., the mutation structural location, given by

$$\mathbf{v} = \frac{1}{M_b} \sum_i^{n_b} m_i^b \cdot \mathbf{c}_i^b \quad (\text{S9})$$

where $\mathbf{c}_i^b$ is the center and m_i^b is the mass of the i -th atom belonging to the mutated residue, $M_b = \sum_i^{n_b} m_i^b$ is the total residue mass and n_b is the total number of atoms forming the residue.

The mutation structural location is mapped on two dimensions by rounding its z -coordinate, v_z , to SAMPLE accuracy (see Table S1) so that it can be assigned to a pore point. Then, the distance between the rounded mutation structural location and its assigned pore point determines the location of the mutation site on contour maps of Figures 2, 3 and 5(a).

Mutation	Reference	Mutation	Reference
I136V (IEM)	[1, 2]	A1632T (IEM)	[29]
S211P (IEM)	[3]	A1632G (IEM)	[30]

F216S (IEM)	[4]	A1746G (IEM)	[28]
I234T (IEM)	[5]		
S241T (IEM)	[6, 7]	V1298F (PEPD)	[25, 31]
L245V (IEM)	[8]	V1299F (PEPD)	[31, 32]
N395K (IEM)	[9]	I1461T (PEPD)	[33]
V400M (IEM)	[10]	G1607R (PEPD)	[34]
L823R (IEM)	[11]	L1612P (PEPD)	[35]
F826Y (IEM)	[12]	M1627K (PEPD)	[32, 33, 36]
I848T (IEM)	[2, 13–15]	A1632E (PEPD)	[37]
G856R (IEM)	[16]		
G856D (IEM)	[17]	R185H (SFN)	[38]
L858H (IEM)	[15, 18, 19]	I228M (SFN)	[39]
L858F (IEM)	[20]	I720K (SFN)	[40]
A863P (IEM)	[21]	I739V (SFN)	[38]
V872G (IEM)	[22]	M932L (SFN)	[40]
Q875E (IEM)	[23]	R1279P (SFN)	[41]
L955Del (IEM)	[24]	T1596I (SFN)	[42]
P1308L (IEM)	[25]		
V1316A (IEM)	[3, 26]		
F1449V (IEM)	[27]		
W1538R (IEM)	[28]		

Table S2: **Pain-related NaV1.7 mutations.** NaV1.7 mutations causing IEM, PEPD and SFN symptoms. References hint to the published study describing biophysical attributes of the corresponding mutation.

Mutation	Type	Mutation	Type
S126A	hSNP	L127A	hSNP
M145L	hSNP	N146S	hSNP
V194I	hSNP	L201V	hSNP

N206D	hSNP	E759D	hSNP
A766T	hSNP	A766V	hSNP
I767V	hSNP	V795I	hSNP
A815S	hSNP	K1176R	hSNP
R1207K	hSNP	T1210N	hSNP
I1235V	hSNP	A1505V	hSNP
S1509T	hSNP	S1509A	hSNP
Q1530P	hSNP	Q1530K	hSNP
Q1530D	hSNP	H1531Y	hSNP
M1532V	hSNP	E1534D	hSNP
Y1537N	hSNP	T1548S	hSNP
H1560Y	hSNP	H1560C	hSNP
V1565I	hSNP	I1577L	hSNP
D1586E	hSNP	T1590K	hSNP
T1590R	hSNP	V1613I	hSNP
D890E	hSNP	D890V	hSNP
T1398M	hSNP	I1399D	hSNP
D1411S	hSNP	D1411N	hSNP
K1412I	hSNP	K1415I	hSNP
D1662A	hSNP		
K1700A	hSNP	N1245S [97]	nABN
D1674A	hSNP	L1267V [97]	nABN
S1419N	hSNP	V1428I	nABN
K1412E	hSNP	T920N	nABN

Table S3: **List of incorporated neutral NaV1.7 mutations.** Neutral mutations consist of two subgroups; a hSNPs group and a nABNs group incorporated from [54]. V1428I and T920N nABNs represent unpublished observations of the PROPANE consortium. In addition, we consider as neutrals four variants that were previously classified as neutrals in accordance to [95]; M130I, M787V, S802G and V810M.

S8. Distance metrics for mutation structural locations

S8a. Estimating distance between mutation structural locations and HP's boundary

The distance between a mutation structural location, $\mathbf{v}$, and the HP's boundary, $\Omega^{(0)}$, is estimated by

$$D_{HP}(\mathbf{v}) = \langle \{ | \|\mathbf{v} - \mathbf{s}^{(0)}\| - l_\alpha(\mathbf{s}^{(0)}) | \mid (\mathbf{s}^{(0)}, l_\alpha(\mathbf{s}^{(0)})) \in \Omega^{(0)} \} \rangle \quad (\text{S10})$$

representing the statistical-median value of the shortest euclidean distances between $\mathbf{v}$ and the surfaces of the spheres of radius $l_\alpha(\mathbf{s}^{(0)})$ centered at $\mathbf{s}^{(0)}$ (see eq. S2).

S8b. Estimating distance between mutation structural locations and inflection points

The distance between a mutation structural location, $\mathbf{v}$, and an inflection point, $\xi(\mathbf{p})$, is estimated by

$$D_{\xi(\mathbf{p})}(\mathbf{v}) = | \|\mathbf{v} - \mathbf{p}\| - \xi(\mathbf{p}) | \quad (\text{S11})$$

corresponding to the shortest euclidean distance of $\mathbf{v}$ from the surface of the inflection sphere of radius $\xi(\mathbf{p})$ centered at $\mathbf{p}$.

For a collection of mutation structural locations (e.g., for the collection of GOF mutation structural locations), let it be V , equation S11 is computed for every $\mathbf{v} \in V$, and the statistical-median value $D_{\xi(\mathbf{p})}(V) = \langle \{ | \|\mathbf{v} - \mathbf{p}\| - \xi(\mathbf{p}) | \mid \mathbf{v} \in V \} \rangle$ is employed in order to statistically describe the distance of mutation structural locations from $\xi(\mathbf{p})$.

S9. Scaling analysis of HIIS axial part

In order to detect critical signatures of HIs around NaV1.7's pore, we quantified linearity of $\log[|\vec{\mathbf{m}}_z^{(1)}(\mathbf{p}, l_\alpha(\mathbf{p}))|]$ -vs- $\log[l_\alpha(\mathbf{p})]$ within the lag

domain interval, i.e., for $D(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$, and within the first- and second-part of the inflection domain, i.e., for $s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$ and $\xi(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq o(\mathbf{p})$, respectively, by fitting the line $\gamma(\mathbf{p}) \cdot \log[l_\alpha(\mathbf{p})] + \beta(\mathbf{p})$ where the goodness of the fitting for each domain interval was estimated in terms of the corresponding Pearson coefficient score, $PC(\mathbf{p})$, and corresponding fitting error.

Figure S3: **Scaling analysis of HIIS axial field component.** (a), Traces of power-law exponents $\gamma_{lag}(\mathbf{p})$, $\gamma_{partI}(\mathbf{p})$ and $\gamma_{partII}(\mathbf{p})$ describing the scaling behavior of the HIIS axial field component, (see Methods, eq. [m13]), within the lag domain (i.e., for $l_\alpha(\mathbf{p}) \leq s(\mathbf{p})$), the first part of the inflection domain (i.e., for $s(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq \xi(\mathbf{p})$), and the second part of the inflection domain (i.e., for

$\xi(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq o(\mathbf{p})$) respectively, are plotted for $\mathbf{p} \in P$. Shaded area around the curves indicates uncertainty in exponent quantification based on fitting error analysis. (b), Traces of corresponding Pearson coefficients $PC_{lag}(\mathbf{p})$, $PC_{partI}(\mathbf{p})$ and $PC_{partII}(\mathbf{p})$ are plotted for $\mathbf{p} \in P$. Note that $0 \leq |PC(\mathbf{p})| \leq 1$ always holds with a value close to 1.0 indicating a "good" power-law approximation. Green-shaded area around $SF_{crit.}$ pore point indicates the pore region $-13.4 \leq p_z \leq -5.8$ where the product $|PC_{partI}(\mathbf{p})| \cdot |PC_{partII}(\mathbf{p})|$ attains its largest values as, both, $|PC_{partI}(\mathbf{p})|$ and $|PC_{partII}(\mathbf{p})|$ attain values higher than 0.95. This is indicative of maximization of the "goodness" of the power-law approximation. ES, $SF_{crit.}$, CC, AG, and IS labels mark the locations of the extracellular side, of the critical pore point $\mathbf{p}_{crit.}$, of the central cavity, of the activation gate, and of the intracellular side, respectively.

Noteworthy, for cases where limited data are available (such as the case presented in this study) usage of fitting techniques upon well-defined intervals is advised in order to eliminate numerical biases [98]. To do that, we fetched the least square fitting algorithm implemented within the R [66] environment and proceeded with extracting power-law exponents from log-vs-log plots. Linear fittings on asymptote domain intervals (i.e., for $o(\mathbf{p}) < l_\alpha(\mathbf{p}) \leq L(\mathbf{p})$) were not considered as asymptote domain does not contain any mutations sites (see Main Text). Traces of Pearson coefficients are shown in Figure S3 where we show that, both, $|PC_{partI}(\mathbf{p})|$ and $|PC_{partII}(\mathbf{p})|$ attain values higher than 0.95 within the SF interval $-13.4 \leq p_z \leq -5.8$ where corresponding scaling exponents $\gamma(\mathbf{p})_{partI}$ and $\gamma(\mathbf{p})_{partII}$ reveal an up- and down-regulation of HIIS axial field component before and after the inflection point, respectively. Accordingly, the power-law scheme described by equation [m14] (see main text, Methods) is accurately reproduced for pore points $-13.4 \leq p_z \leq -5.8$ where maximization of the product $|PC_{partI}(\mathbf{p})| \cdot |PC_{partII}(\mathbf{p})|$ occurs at $\mathbf{p}_{crit.} = (p_x \approx -2.58, p_y \approx -0.06, p_z = -12.1)$ (see Figure S3, and Main Text, Figure 5(a)). Note that the trace of $\gamma_{lag}(\mathbf{p})$ remains close to zero while the magnitude of its Pearson coefficient attains values

higher than 0.95 for a short interval in the vicinity of the CC thus not allowing for a robust power-law approximation of HIIS axial field component within the lag domain.

S10. Weighted distance average classification

We calculated the weighted topological-distance average $w_{HP} \cdot D_{HP}(\mathbf{v}) + w_{\xi} \cdot D_{\xi(\mathbf{p}_{crit.})}(\mathbf{v})$, where w_{HP} and w_{ξ} are weights, for all the pain-related and neutral mutation structural locations, and fed retrieved distances into a binary classifier. The quality of retrieved classifications is demonstrated in Figure S4 in terms of the area under the ROC curve for $w_{\xi} = 1 - w_{HP} = i \cdot 0.001$, $i = 0, 1, \dots, 1000$. The linear combination of $D_{HP}(\mathbf{v})$ with $D_{\xi(\mathbf{p}_{crit.})}(\mathbf{v})$ which maximized the area under the ROC curve corresponds to $w_{\xi} = 0.618$ and $w_{HP} = 0.382$ (Figure S4).

Figure S4: **Area under ROC curve retrieved from linear combination of topological distances $D_{HP}(\mathbf{v})$ and $D_{\xi(\mathbf{p}_{crit.})}(\mathbf{v})$.** The area under the ROC curve is plotted for different weights $w_{\xi} = 1 - w_{HP} = i \cdot 0.001$, $i = 0, 1, \dots, 1000$ determining the weighted topological-distance average $w_{HP} \cdot D_{HP}(\mathbf{v}) + w_{\xi} \cdot D_{\xi(\mathbf{p}_{crit.})}(\mathbf{v})$ where $D_{HP}(\mathbf{v})$ and $D_{\xi(\mathbf{p}_{crit.})}(\mathbf{v})$ are given by equations S10 and S11, respectively.

Author contributions

M.N.X. designed the study, performed computations, and analyzed the data; D.K., R.W. and P.L. contributed to refinement of algorithmic procedures; Y.Y. provided with the 3D structural model of the NaV1.7 channel; M.M.G. contributed to variants selection and classification; R.W., P.L., D.K., Y.Y., J.H., H.J.S. and S.G.W. provided with critical feedback and helped with the interpretation of the results; Y.Y. and S.G.W. encouraged M.N.X. to focus on specific aspects of the findings; H.J.S. supervised the study; G.L. and C.G.F. were in charge of overall direction; M.N.X. wrote the manuscript in consultation with all the co-authors; All co-authors have critically revised the manuscript.

Acknowledgments

The study was partly funded by the European Union 7th Framework Programme (grant n602273). We acknowledge technical and scientific support provided by the PROPANE study group.

Competing Interests

The authors declare that there is no conflict of interest.

References

- [1] Cheng, X., Dib-Hajj, S.D., Tyrrell, L., Waxman, S.G. Mutation I136V alters electrophysiological properties of the Na(v)1.7 channel in a family with onset of erythromelalgia in the second decade. *Mol Pain* **4**, 1 (2008).
- [2] Estacion, M., Choi, J.S., Eastman, E.M. et al. Can robots patch-clamp as well as humans? Characterization of a novel sodium channel mutation. *J Physiol* **588**, 1915-1927 (2010).
- [3] Wu, M.T., Huang, P.Y., Yen, C.T., Chen, C.C., Lee, M.J. A novel SCN9A mutation responsible for primary erythromelalgia and is resistant to the treatment of sodium channel blockers. *PLoS One* **8**, e55212 (2013).
- [4] Choi, J.S., Dib-Hajj, S.D., Waxman, S.G. Inherited erythromelalgia: limb pain from an S4 charge-neutral Na channelopathy. *Neurology* **67**, 1563-1567 (2006).
- [5] Ahn, H.S., Dib-Hajj, S.D., Cox, J.J. et al. A new Nav1.7 sodium channel mutation I234T in a child with severe pain. *Eur J Pain* **14**, 944-950 (2010).
- [6] Lampert, A., Dib-Hajj, S.D., Tyrrell, L., Waxman, S.G. Size matters: Erythromelalgia mutation S241T in Nav1.7 alters channel gating. *J Biol Chem* **281**, 36029-36035 (2006).
- [7] Yang, Y., Dib-Hajj, S.D., Zhang, J. et al. Structural modelling and mutant cycle analysis predict pharmacoresponsiveness of a Na(V)1.7 mutant channel. *Nat Commun* **3**, 1186 (2012).
- [8] Emery, E.C., Habib, A.M., Cox, J.J. et al. Novel SCN9A mutations underlying extreme pain phenotypes: unexpected electrophysiological and clinical phenotype correlations. *J Neurosci* **35**, 7674-7681 (2015).

- [9] Sheets, P.L., Jackson, J.O. 2nd, Waxman, S.G., Dib-Hajj, S.D., Cummins, T.R. A Nav1.7 channel mutation associated with hereditary erythromelalgia contributes to neuronal hyperexcitability and displays reduced lidocaine sensitivity. *J Physiol* **581**, 1019-1031 (2007).
- [10] Fischer, T.Z., Gilmore, E.S., Estacion, M. et al. A novel Nav1.7 mutation producing carbamazepine-responsive erythromelalgia. *Ann Neurol* **65**, 733-741 (2009).
- [11] Lampert, A., Dib-Hajj, S.D., Eastman, E.M., Tyrrell, L., Lin, Z., Yang, Y., Waxman, S.G. Erythromelalgia mutation L823R shifts activation and inactivation of threshold sodium channel Nav1.7 to hyperpolarized potentials. *Biochem Biophys Res Commun* **390**, 319-324 (2006).
- [12] Wu, B., Zhang, Y., Tang, H. et al. A Novel SCN9A Mutation (F826Y) in Primary Erythromelalgia Alters the Excitability of Nav1.7. *Curr Mol Med* **17**, 450-457 (2017).
- [13] Cummins, T.R., Dib-Hajj, S.D., Waxman, S.G. Electrophysiological properties of mutant Nav1.7 sodium channels in a painful inherited neuropathy. *J Neurosci* **24**, 8232-8236 (2004).
- [14] Han, C., Dib-Hajj, S.D., Lin, Z. et al. Early- and late-onset inherited erythromelalgia: genotype-phenotype correlation. *Brain* **132**, 1711-1722 (2009).
- [15] Theile, J.W., Cummins, T.R. Inhibition of Nav β 4 peptide-mediated resurgent sodium currents in Nav1.7 channels by carbamazepine, riluzole, and anandamide. *Mol Pharmacol* **80**, 724-734 (2011).
- [16] Tanaka, B.S., Nguyen, P.T., Zhou, E.Y. et al. Gain-of-function mutation of a voltage-gated sodium channel NaV1.7 associated with peripheral pain and impaired limb development. *J Biol Chem* **292**, 9262-9272 (2017).
- [17] Hoeijmakers, J.G., Han, C., Merkies, I.S. et al. Small nerve fibres, small hands and small feet: a new syndrome of pain, dysautonomia and acromesomelia in a kindred with a novel NaV1.7 mutation. *Brain* **135**, 345-358 (2012).
- [18] Cummins, T.R., Dib-Hajj, S.D., Waxman, S.G. Electrophysiological properties of mutant Nav1.7 sodium channels in a painful inherited neuropathy. *J Neurosci* **24**, 8232-8236 (2004).
- [19] Rush, A.M., Dib-Hajj, S.D., Liu, S., Cummins, T.R., Black, J.A., Waxman, S.G. A single sodium channel mutation produces hyper- or hypoexcitability in different types of neurons. *Proc Nat Acad Sci USA* **103**, 8245-8250 (2006).

- [20] Han, C., Rush, A.M., Dib-Hajj, S.D. et al. Sporadic onset of erythromelgia: a gain-of-function mutation in Nav1.7. *Ann Neurol* **59**, 553-558 (2006).
- [21] Harty, T.P., Dib-Hajj, S.D., Tyrrell, L., Blackman, R., Hisama, F.M., Rose, J.B., Waxman, S.G. Nav1.7 mutant A863P in erythromelgia: effects of altered activation and steady-state inactivation on excitability of nociceptive dorsal root ganglion neurons. *J Neurosci* **26**, 12566-12575 (2006).
- [22] Choi, J.S., Zhang, L., Dib-Hajj, S.D. et al. Mexiletine-responsive erythromelgia due to a new Na(v)1.7 mutation showing use-dependent current fall-off. *Exp Neurol* **216**, 383-389 (2009).
- [23] Stadler, T., O'Reilly, A.O., Lampert, A. Erythromelgia mutation Q875E Stabilizes the activated state of sodium channel Nav1.7. *J Biol Chem* **290**, 6316-6325 (2015).
- [24] Cheng, X., Dib-Hajj, S.D., Tyrrell, L., Te Morsche, R.H., Drenth, J.P., Waxman, S.G. Deletion mutation of sodium channel Na(V)1.7 in inherited erythromelgia: enhanced slow inactivation modulates dorsal root ganglion neuron hyperexcitability. *Brain* **134**, 1972-1986 (2011).
- [25] Cheng, X., Dib-Hajj, S.D., Tyrrell, L., Wright, D.A., Fischer, T.Z., Waxman, S.G. Mutations at opposite ends of the DIII/S4-S5 linker of sodium channel Nav1.7 produce distinct pain disorders. *Mol Pain* **6**, 24 (2010).
- [26] Estacion, M., Yang, Y., Dib-Hajj, S.D. et al. A new Nav1.7 mutation in an erythromelgia patient. *Biochem Biophys Res Commun* **432**, 99-104 (2013).
- [27] Dib-Hajj, S.D., Rush, A.M., Cummins, T.R. et al. Gain-of-function mutation in Nav1.7 in familial erythromelgia induces bursting of sensory neurons. *Brain* **128**, 1847-1854 (2005).
- [28] Cregg, R., Laguda, B., Werdehausen, R. et al. Novel mutations mapping to the fourth sodium channel domain of Nav1.7 result in variable clinical manifestations of primary erythromelgia. *Neuromolecular Med* **15**, 265-278 (2013).
- [29] Eberhardt, M., Nakajima, J., Klinger, A.B. et al. Inherited pain: sodium channel Nav1.7 A1632T mutation causes erythromelgia due to a shift of fast inactivation. *J Biol Chem* **289**, 1971-1980 (2014).
- [30] Yang, Y., Huang, J., Mis, M.A. et al. Nav1.7-A1632G Mutation from a Family with Inherited Erythromelgia: Enhanced Firing of Dorsal Root Ganglia Neurons Evoked by Thermal Stimuli. *J Neurosci* **36**, 7511-7522 (2016).

- [31] Jarecki, B.W., Sheets, P.L., Jackson, J.O. 2nd, Cummins, T.R. Paroxysmal extreme pain disorder mutations within the D3/S4-S5 linker of Nav1.7 cause moderate destabilization of fast inactivation. *J Physiol* **586**, 4137-4153 (2008).
- [32] Theile, J.W., Jarecki, B.W., Piekarz, A.D., Cummins, T.R. Nav1.7 mutations associated with paroxysmal extreme pain disorder, but not erythromelalgia, enhance Nav β 4 peptide-mediated resurgent sodium currents. *J Physiol* **589**, 597-608 (2011).
- [33] Fertleman, C.R., Baker, M.D., Parker, K.A. et al. SCN9A mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52**, 767-774 (2006).
- [34] Choi, J.S., Boralevi, F., Brissaud, O. et al. Paroxysmal extreme pain disorder: a molecular lesion of peripheral neurons. *Nat Rev Neurol* **7**, 51-55 (2011).
- [35] Suter, M.R., Bhuiyan, Z.A., Laedermann, C.J. et al. p.L1612P, a novel voltage-gated sodium channel Nav1.7 mutation inducing a cold sensitive paroxysmal extreme pain disorder. *Anesthesiology* **122**, 414-423 (2015).
- [36] Dib-Hajj, S.D., Estacion, M., Jarecki, B.W. et al. Paroxysmal extreme pain disorder M1627K mutation in human Nav1.7 renders DRG neurons hyperexcitable. *Mol Pain* **4**, 37 (2008).
- [37] Estacion, M., Dib-Hajj, S.D., Benke, P.J. et al. NaV1.7 gain-of-function mutations as a continuum: A1632E displays physiological changes associated with erythromelalgia and paroxysmal extreme pain disorder mutations and produces symptoms of both disorders. *J Neurosci* **28**, 1079-11088 (2008).
- [38] Han, C., Hoeijmakers, J.G., Liu, S. et al. Functional profiles of SCN9A variants in dorsal root ganglion neurons and superior cervical ganglion neurons correlate with autonomic symptoms in small fibre neuropathy. *Brain* **135**, 2613-2628 (2012).
- [39] Estacion, M., Han, C., Choi, J.S. et al. Intra- and interfamily phenotypic diversity in pain syndromes associated with a gain-of-function variant of NaV1.7. *Mol Pain* **7**, 92 (2011).
- [40] Faber, C.G., Hoeijmakers, J.G., Ahn, H.S. et al. Gain of function Nav1.7 mutations in idiopathic small fiber neuropathy. *Ann. Neurol* **71**, 26-39 (2012).
- [41] Huang, J., Yang, Y., Dib-Hajj, S.D. et al. Depolarized inactivation overcomes impaired activation to produce DRG neuron hyperexcitability in a Nav1.7 mutation in a patient with distal limb pain. *J Neurosci* **34**, 12328-12340 (2014).

- [42] Blesneac, I., Themistocleous, A.C., Fratter, C. et al. Rare Nav1.7 variants associated with painful diabetic peripheral neuropathy. *Pain* **159**, 469-480 (2018).
- [43] Cox J.J., Reimann, F., Nicholas, A.K. et al. An SCN9A channelopathy causes congenital inability to experience pain. *Nature* **444**, 894-898 (2006).
- [44] Goldberg, Y.P., MacFarlane, J., MacDonald, M.L. et al. Loss-of-function mutations in the Nav1.7 gene underlie congenital indifference to pain in multiple human populations. *Clin Genet* **71**, 311-319 (2007).
- [45] Nilsen, K.B., Nicholas, A.K., Woods, C.G., Mellgren, S.I., Nebuchennykh, M., Aasly, J. Two novel SCN9A mutations causing insensitivity to pain. *Pain* **143**, 155-158 (2009).
- [46] Hammer, M.U., Anderson, T.H., Chaimovich, A., Shell, M.S., Israelachvili, J. The search for the hydrophobic force law. *Faraday Discuss* **146**, 299-308 (2010).
- [47] Hummer, G., Garde, S., García, A.E., Paulaitis, M.E., Pratt, L.R. Hydrophobic Effects on a Molecular Scale. *J Phys Chem B* **102**, 51 (1998).
- [48] Aryal, P., Sansom M.S.P, Tucker, S.J. Hydrophobic Gating in Ion Channels. *J Mol Biol* **427**, 121-130 (2015).
- [49] Li, E., Wimley, W.C., Hristova, K. Transmembrane helix dimerization: beyond the search for sequence motifs. *Biochim Biophys Acta* **1818**, 183-193 (2012).
- [50] Lin, J., Motylinski, J., Krauson, A. J., Wimley, W. C., Searson, P.C., Hristova, K. Interactions of membrane active peptides with planar supported bilayers: an impedance spectroscopy study. *Langmuir* **28**, 6088-6096 (2012).
- [51] Yonkunas, M., Kurnikova, M. The hydrophobic effect contributes to the closed state of a simplified ion channel through a conserved hydrophobic patch at the pore-helix crossing. *Front Pharmacol* **6**, 284 (2015).
- [52] Lampert, A., O'Reilly, A.O., Dib-Hajj, S.D., Tyrrell, L., Wallace, B.A., Waxman, S.G. A Pore-blocking Hydrophobic Motif at the Cytoplasmic Aperture of the Closed-state Nav1.7 Channel Is Disrupted by the Erythromelalgia-associated F1449V Mutation. *J Biol Chem* **283**, 24118-24127 (2008).
- [53] Yang, Y., Estacion, M., Dib-Hajj, S.D., Waxman, S.G. Molecular Architecture of a Sodium Channel S6 Helix: radial tuning of the voltage-gated sodium channel 1.7 activation gate. *J Biol Chem* **288**, 13741-13747 (2013).

- [54] Kapetis, D., Yang, Y., Sassone, J. et al. Network topology of NaV1.7 mutations in sodium channel-related painful disorders. *BMC Systems Biology* **11**, 28 (2017).
- [55] Hasan S., Hunter T., Hunter G., Pessia M., D’Adamo M.C. Commentary: A channelopathy mutation in the voltage-sensor discloses contributions of a conserved phenylalanine to gating properties of Kv1.1 channels and ataxia. *Front Cell Neurosci* **12**, 174 (2018).
- [56] Toffano, A., Chiarot, G., Zamuner, S. et al. Computational Pipeline to probe NaV1.7 gain-of-functions variants in neuropathic painful syndromes. *Sci Rep* **10**, 17930 (2020).
- [57] Payandeh, J., Scheuer, T., Zheng, N., Catterall, W.A. The crystal structure of a voltage-gated sodium channel. *Nat* **475**, 353-358 (2011).
- [58] Bak, P., Tang, C., Wiesenfeld, K. Self-organized criticality: an explanation of 1/f noise. *Phys Rev Lett* **59**, 381-384 (1987).
- [59] Phillips, J.C. Scaling and self-organized criticality in proteins I. *Proc Natl Acad Sci USA* **106**, 3107-3112 (2009).
- [60] Phillips, J.C. Scaling and self-organized criticality in proteins II. *Proc Natl Acad Sci USA* **106**, 3113-3118 (2009).
- [61] Morett, M. Self-organized critical model for protein folding. *Physica A* **390**, 3055-3059 (2011).
- [62] Phillips, J.C. Self-organized criticality in proteins: Hydropathic roughening profiles of G-protein-coupled receptors. *Phys Rev E* **87**, 032709 (2013).
- [63] Zhou, R., Silverman, B.D., Royyuru, A.K., Athma, P. Spatial profiling of protein hydrophobicity: Native vs. decoy structures. *Proteins* **52**, 561-572 (2003).
- [64] Xenakis, M.N., Kapetis, D., Yang, Y. et al. Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution. *Proteins* **88**, 1319-1328 (2020).
- [65] Xenakis, M.N., Kapetis, D., Yang, Y. et al. Non-extensivity and criticality of atomic hydropathicity around a voltage-gated sodium channel’s pore; a modeling study. *J Biol Phys* **47**, 61-77 (2021).
- [66] R Core Team *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/> (2014).

- [67] Humphrey, W., Dalke, A., Schulten K. VMD-Visual Molecular Dynamics. *J Molec Graph* **14**, 33-38 (1996).
- [68] Loss, R.D. Atomic weights of the elements 2001 (IUPAC technical report). *Pure Appl Chem* **73**, 667-683 (2001).
- [69] Bondi, A. Van der Waals volumes and radii. *J Phys Chem* **68**, 441-51 (1964).
- [70] Rowland, R.S., Taylor, R. Intermolecular nonbonded contact distances in organic crystal structures: comparison with distances expected from van der Waals radii. *J Phys Chem* **100**, 7384-7391 (1996).
- [71] Smart, O.S., Neduvilil, J.G., Wang, X., Wallace, B.A., Sansom, M.S.P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* **14**, 354-360 (1996).
- [72] Kahm, M., Hasenbrink, G., Lichtenberg-Fraté, H., Ludwig, J., Kschischo M. Fitting Biological Growth Curves with R. *J Stat Softw* **33**, 1-21 (2010).
- [73] Zwietering, M.H., Jongenburger, I., Rombouts, F.M., van't Riet, K. Modeling of the bacterial growth curve. *Appl Environ Microbiol* **56**, 1875-1881 (1990).
- [74] Blumberg, A.A. Logistic growth rate functions. *J Theoret Biol* **21**, 42 (1968).
- [75] Gompertz, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos Trans R Soc London* **115**, 513-585 (1825).
- [76] Richards, F.J. A flexible growth function for empirical use. *J Exp Bot* **10**, 290-300 (1959).
- [77] E., Tjørve, Tjørve, K.M.C. A unified approach to the Richards-model family for use in growth analyses: Why we need only two model forms. *J Theor Biol* **267**, 417-425 (2010).
- [78] Eisenberg, D., Weiss, R.M., Terwilliger, T.C., Wilcox, W. Hydrophobic moments and protein structure. *Faraday Symp Chem Soc* **17**, 109-120 (1982).
- [79] Kapcha, L.H., Rossky, P.J. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J Mol Biol* **426**, 484-498 (2014).
- [80] Gromiha, M.M., Oobatake, M., Kono, H., Uedaira, H., Sarai, H. Role of structural and sequence information in the prediction of protein stability changes: comparison between buried and partially buried mutations. *Protein Eng* **12**, 549-555 (1999).

- [81] Gromiha, M.M. Factors influencing the thermal stability of buried protein mutants. *Polymer* **44**, 4061-4066 (2003).
- [82] Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J., Müller, M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, **12**, 77 (2011).
- [83] Following [65], let us consider the case of a missense SCN9A mutation expressed in the perturbation form of $N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \rightarrow N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) + \epsilon(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$, with $\epsilon(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.}))$ being a small-amplitude perturbation source describing addition and/or removal of a small number of atoms due mutation translation. Then, HicIS is regulated according to $h_z^{(1)}(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) + \zeta(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) \sim l_\alpha(\mathbf{p}_{crit.})^{\gamma(\mathbf{p}_{crit.})} \cdot (N(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})) + \epsilon(\mathbf{p}_{crit.}, l_\alpha(\mathbf{p}_{crit.})))$ with $\gamma(\mathbf{p}_{crit.}) = \gamma_{partI}(\mathbf{p}_{crit.})$ for $s(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq \xi(\mathbf{p}_{crit.})$ and $\gamma(\mathbf{p}_{crit.}) = \gamma_{partII}(\mathbf{p}_{crit.})$ for $\xi(\mathbf{p}_{crit.}) < l_\alpha(\mathbf{p}_{crit.}) \leq o(\mathbf{p}_{crit.})$ so that the HicIS law described by r_l remains intact around the critical point.
- [84] Heyne, H.O., Baez-Nieto, Iqba, S. et al. A machine learning method can predict loss- versus gain-of-function effects of human genetic variants in disease-associated ion channels. *Sci Transl Med* **12**, 556 (2020).
- [85] Tsuchiya, M., Giuliani, A., Hashimoto, M., Erenpreisa, J., Yoshikawa, K. Self-Organizing Global Gene Expression Regulated through Criticality: Mechanism of the Cell-Fate Change. *PLoS ONE* **11**, 12 (2016).
- [86] Pavlos, G.P., Karakatsanis, L.P., Iliopoulos, A.C. et al. Measuring complexity, nonextensivity and chaos in the DNA sequence of the Major Histocompatibility Complex. *Physica A* **438**, 188-209 (2015).
- [87] Furusawa, C., Kaneko, K. Adaptation to optimal cell growth through self-organized criticality. *Phys. Rev. Lett.* **108**, 208103 (2012).
- [88] Janina, H., Gross, T. Self-organized criticality as a fundamental property of neural systems. *Front Syst Neurosci* **8**, 166 (2014).
- [89] Beggs, J.M. The criticality hypothesis: How local cortical networks might optimize information processing. *Phil Trans R Soc A* **366**, 329-343 (2008).
- [90] Bertschinger, N., Natschlager, T. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comput* **16**, 1413-1436 (2004).
- [91] Kinouchi, O., Copelli, M. Optimal dynamical range of excitable networks at criticality. *Nat Phys* **2**, 348-351 (2006).

- [92] Watkins, N.W., Pruessner, G., Chapman, S.C. et al. 25 Years of Self-organized Criticality: Concepts and Controversies. *Space Sci Rev* **198**, 3-44 (2016).
- [93] Shen, H., Liu, D., Wu, K., Lei, J., Yan, N. Structures of human Na_V1.7 channel in complex with auxiliary subunits and animal toxins. *Science* **363**, 1303-1308 (2019).
- [94] Karczewski, K.J., Francioli, L.C., Tiao, G. et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*, 531210 (2019).
- [95] Wallis, Y., Payne, S., McAnulty, C. et al. Practice guidelines for the evaluation of pathogenicity and reporting of sequence variants in clinical molecular genetics. *ACGS* (2013). http://www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf (accessed May 25, 2013).
- [96] NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* Jan 4;44(D1):D7-19 (2016). doi: 10.1093/nar/gkv1290.Epub2015Nov28.PMID:26615191;PMCID:PMC4702911
- [97] Brouwer, B.A., Merkies, I.S.J., Gerrits, M.M., Waxman, S.G, Hoeijmakers J.G.J., Faber, C.G. Painful neuropathies: the emerging role of sodium channelopathies. *J Peripher Nerv Syst* **19**, 53-65 (2014).
- [98] Moretti, P., Muñoz, M.A. Griffiths phases and the stretching of criticality in brain networks. *Nat Commun* **4**, 2521 (2013).

Chapter 5

General Discussion

Aspects of molecular complexity

Voltage-gated sodium channels (NaVChs) are remarkable examples of molecular complexity; they are optimally embedded within cell membranes, process incoming ion species information with high specificity, and can spontaneously re-organize in the presence of external fields. How can NaVChs achieve these high levels of functional robustness and structural adaptation under ambiguous environmental conditions? A hallmark of complexity is that the degree of collective integration of individual system components does not depend on system size which in turn increases robustness and adaptivity of the system as a whole. This phenomenon is termed scale-invariance (for a comprehensive introduction see [1]) and, in some cases, can be attributed to spatiotemporal system dynamics operating within a self-organized criticality (SOC) regime.

Since the seminal work of P. Bak and co-workers [2] SOC-based modeling approaches have been applied in various scientific disciplines such as space physics (e.g. see [3]), seismology (e.g., see [4]), sociology (e.g., see [5]), electrical grid engineering (e.g., see [6]) ecology (e.g., see [7]), neuroscience (e.g., see [8]), material science (e.g., see [9]) and,

of course, protein science [10, 11], just to name a few. From a statistical physics viewpoint, SOC is associated with the existence of a critical point that mathematically corresponds to a bifurcation of the macroscopic dynamics. Accordingly, at the critical point, the system undergoes a phase transition, i.e., a qualitative change in its macroscopic behavior that is detectable via scaling analysis of some carefully-chosen property (or properties). It is important to mention here that for finite-size systems (such as the cases presented in this Dissertation) the notion of the critical point relaxes to that of the critical region. Hence, finite-size artifacts (e.g., power-laws truncation [12], subsampling [13] and misleading goodness-of-fit power-law testing [14]) might strongly influence scaling-related observables. Sizes of sampling spaces investigated throughout this Dissertation were of order $\sim 10^4$ providing with a relatively (with respect to structural NaVCh resolution) dense grid for scaling laws observation. Moreover, scaling intervals were not arbitrarily selected but based on key-structural features of the atomic environment extracted from atomic radial distribution functions. Accordingly, we proceeded with a structure-guided scaling analysis which allowed for rationalization of goodness-of-fit scores.

As anticipated from [10, 11], we demonstrated that atomic hydrophobicity can be used as a SOC-index for NaVCh structures. A crucial step toward this direction was the selection of an appropriate hydrophobic scale for atomic-level hydrophobic characterization. There are more than 90 hydrophobic scales available in the literature [15–17] and although they attempt to describe similar physicochemical molecular properties they may have different separation capacities [18]. Recently, L.H. Kapcha and P.J. Rossky introduced a hydrophobic scale allowing for binary classification of atomic components [19]. The Kapcha-Rossky atomic hydrophobic scale was shown to accurately discriminate among polar and non-polar moieties within each amino acid including its backbone atom groups. In particular, each atom type found in the 20 naturally-occurring amino acids is classified as either hydrophobic or hydrophilic based on its polar or non-polar nature quantified

in terms of partial charges retrieved from the Optimized Potentials for Liquid Simulations [20] parameter set. A crucial property of this binary hydropathic classification approach is that it attempts to capture a natural transition between hydrophobic and hydrophilic atoms at the microscopic level, i.e., within a single amino acid. Choosing to work with the Kapcha-Rossky atomic hydrophobic scale throughout this Dissertation was an crucial step toward increasing sample size and investigating properties of the atomic bulk incorporating atomic components of several amino acids. Intrinsic molecular noise (e.g., atomic thermal fluctuations) is not expected to affect reported observations unless it exceeds a threshold value above which binary hydropathic classification of atomic components is destroyed.

Spatial inhomogeneities in the distribution profile of hydropathic moment functions around a NaVCh's pore is the connecting link between topology and scale-invariance. Landau was the first one to realize the physical importance of inhomogeneities in crystal-like materials under stress which he treated as ensembles of thermodynamically-open, small volumes $dV(\mathbf{r}, t)$ placed at crystal-structure locations $\mathbf{r}$ [21]. Following the thread of this early observation, we introduced a computational framework of concentric spherical volumes placed at pore points $\mathbf{r} \equiv \mathbf{p}$ with site-specific radii $t \equiv l_\alpha(\mathbf{p})$. This set the basis for investigating the relationship between topology and scaling; Where and why do topological events occur along the ion conduction pathway? And how do they relate to the scaling behavior of hydropathic moment functions around it? We systematically addressed these questions, first, by introducing the concept of cumulative hydropathic topology, second, by retrieving underlying scaling laws and, third, by combining both, topology and scaling, toward mutation-response prediction.

Gaining insight into structure, function and mutation response

We presented a computational and phenomenological framework for detecting and analyzing cumulative hydropathic field topology, as well as, spatial signatures of hydropathic-field phase transitions around a NaVCh's pore. This shed light upon a central hydrophobic-gate mechanism of the pre-open NaVAb pore [22] by implementing a sinks-vs-sources topological decomposition of the cumulative hydropathic field [23]. When topology was correlated with underlying scaling laws, it revealed a scale-invariant behavior of the hydropathic interactions field strength around the highly conserved selectivity filter (SF) capturing the structural transition from the pore domains (PDs) to the voltage-sensing domains (VSDs) [24]. Evolutionary constraints guiding NaVAb structure optimization could then be directly read-out suggesting that SOC in NaVAb might be a consequence of slowly-driving conservation dynamics dictating the structural separation of the pore module from the voltage sensors around the narrow and hydrophilic SF pore region. Stability importance of these observations is the estimation of the coupling energy between the PDs and the VSDs, as well as, of the relative stability of structural components found in the interior and exterior of the NaVAb.

There is a large demand for effective drugs treating pain-related sodium channelopathies. Successful pharmacological targeting of the NaV1.7 remains, however, a challenge due to several reasons ranging from methodological biases and limitations underlying clinical studies to heterogeneous pharmacoresponsiveness of individuals within the drug-targeted population [25]. Atomic-level computational modeling combined with genomic analysis and *in vitro* biophysical analysis can pave the way to more efficient pharmacological design and precision medicine [26]. A necessary intermediate step in order to achieve that, is to better understand how mutations can perturb and propagate

through NaV1.7's structure.

Modeling mutation-induced perturbing effects in protein structures is usually based on experimental data retrieved from *in-silico* mutagenesis. Then, a phenomenological model is constructed providing with an estimation of the amplitude order of the perturbation. Following this paradigm, a network-based phenomenological model for understanding mutation-induced perturbations on the NaV1.7 structure was presented by PROPANE consortium in order to evaluate stability of NaV1.7 mutants simulated via *in silico* mutagenesis [27]. This study demonstrated that the network measure of betweenness centrality can differentiate between pathogenic NaV1.7 variants and non-pathogenic NaV1.7 variants with satisfactory accuracy. Disruption of hydrophobic interactions in NaV1.7 mutants was observed not only locally, i.e., in the direct vicinity of a mutated amino acid, but also non-locally, i.e., in relation to distant amino acids, suggesting that maintaining NaV1.7 hydrophobic-interactions network intact might be the key for minimizing mutation-induced destabilization risk. This Dissertation elaborated on these observations by treating the NaV1.7 wild-type as an atomic network resting at a hydrophobically-tuned SOC state. In analogy to the NaVAb case, we reported on a scale-invariance of the hydrophobic field around NaV1.7's SF following closely the structural separation of the pore module (PM) from the voltage-sensors (VSs) [28].

A consequence of NaV1.7's SOC-like tuning is that gain-of-function (GOF) NaV1.7 mutations around the pore occupy non-random locations. We termed this phenomenon as critical clustering [28]; pain-related NaV1.7 mutation sites are distributed within and around a hydrophobic patch (HP) lining NaV1.7's pore while maintaining a critical hydrophobic-interactions distance from the SF incorporating the highly-conserved DEKA amino-acid sequence. Specifically, pain-related mutation sites appear to be "attracted" toward NaV1.7's critical point mediating the coupling of the PM with the VSs so that their perturbing stimuli is expected to be maximized (in terms of range and

intensity) only in the vicinity of the critical point thus not sacrificing tight packing of pore-lining atomic components forming the SF and the HP. "Attraction" refers here to an evolutionary-driven clustering phenomenon probing mutation effect on the NaV1.7 structure while ensuring a delicate balance between network stability [29] and maximal sensitivity to stimuli [30]. Accordingly, we propose that a GOF effect might be the outcome of a perturbation that is initiated neither too close nor too far away from the highly-conserved DEKA sequence or from the HP so that a trade-off between potentially-deleterious and insignificant perturbations occurring in NaV1.7's interior and exterior, respectively, is achieved. It is important to note here that mutations in the vicinity of the SF were shown to strongly correlate with a loss-of-function (LOF) effect at a genetic level (i.e., decrease protein expression) or at a cellular level (i.e., decrease net ionic currents) [31]. The critical clustering hypothesis supports this idea and, moreover, provides a structural basis for elaborating upon distinction between LOF and GOF effects in terms of the critical hydropathic-interactions distance of corresponding mutations sites from the SF.

Applicability of the critical clustering hypothesis in phenotype prediction differentiated between pathogenic and neutral NaV1.7 variants with high fidelity and at minimum computational cost as, in contrast to [27], no mutagenesis experiments were performed. Specifically, we demonstrate that methodologies presented in this Dissertation can distinguish pain-related NaV1.7 variants from neutral ones with an accuracy of 0.87 [28] while in [27] classification accuracy did not exceed 0.81 (accuracy corresponds to the area under the receiver operating characteristics curve). Moreover, a recent study reported on a classification accuracy of 63.5% retrieved from a machine learning (MLE) computational pipeline [32] operating on the human NaV1.7 structure (PDB code: 6J8J [33]). Similarly to [27], the authors of [32] have utilized non-negligible computational resources in order to build their predictive model. In summary, the classification approach curated in this Dissertation is by far the most inexpensive (in terms of compu-

tational power) one achieving state-of-the-art accuracy. However, our predictive scheme needs to be tested on the human NaV1.7 6J8J template, as well as, on different conformational NaV1.7 states in order to justify any claims of accuracy superiority. Hence, we choose to refrain from such claims but nevertheless point out that the main advantages of our approach is its SOC-based biophysical rationalization and its inexpensive computational implementation.

Concluding remarks and future perspectives

The functional architecture of the pre-open NaVAb channel corresponds to a hydrophathically-tuned SOC state where at least two hydrophathic gates co-exist and co-regulate ion transport through the pore. The Kapcha-Rossky hydrophathic scale can efficiently describe crucial-for-stability properties of the atomic bulk around a NaVCh's pore. Our findings suggest that the NaV1.7 shares similar SOC properties with its evolutionary bacterial ancestor, namely, the NaVAb. SOC phenomenological framework can substantially simplify NaVAb and NaV1.7 functional architecture analysis and provide with semi-analytical predictions for mutation response in terms of hydrophathic-interactions field topology and scale-invariance. The high accuracy of these predictions indicates that atomic-level hydrophathic interactions biomarkers can be of large translational value.

We are witnessing a structural revolution where numerous ion channel structures become available. In fact, the crystal structures of the NaV1.2 [34], NaV1.4 [35] and NaV1.7 [33] mammalian channels were recently resolved providing for the first time with atomic-level crystallographies of human NaVChs. Key-insights into modulation of NaV1.7 by Protoxin-II (a peptide isolated from the *Thrixopelma pruriens* tarantula) enhanced our understanding of activation gating processes and propelling design of subtype-selective NaVCh antagonists [36]. Moreover, the resting conformation of the NaVAb was

resolved supporting the "sliding helix" hypothesis for voltage-sensing but also providing a complete picture for gating, pore opening, and activation-gate closure throughout NaVAb's conformational cycle [37]. It is thus now possible to investigate topological and scaling properties of NaVAb's hydrophobic porous environment at different conformational states in a comparative way.

Mutation-response prediction can now be performed in a site-specific manner based on the 3D-structure of a protein [38], thus allowing harvesting the fruits of ongoing structural revolution. As demonstrated in [31], sequence- and structure-based data can be integrated in MLE pipelines for variants classification across the entire gene *SCNxA* family revealing shared molecular patterns associated with disease. Moreover, a recent study showed that introducing structure-retrieved information into a machine learning and simulated annealing algorithmic procedures can assist discovery of NaV1.7 inhibitors counteracting pain symptoms [39]. However, without an appropriate molecular field theory capable of interpreting data-based models, physical principles underlying these observations remain elusive or - at least - cannot be directly linked to the spatial organization of atoms around the pore. For example, in a recently published MLE-based computational study, the apparent predictor's failure was attributed to the finely-tuned nature of NaV1.7 functional architecture [32]. Taken together, these recent developments suggest that a promising future direction would be integration of hydrophobic topology and scaling metrics into MLE algorithms.

Knowing whether a missense mutations is associated with a GOF or LOF effect is crucial for the clinic as it could eliminate the risk of incorrect treatment (as an example of incorrect treatment we may think of treating epileptic individuals carrying LOF mutations with sodium channel blockers [40]). Hence, a step to be taken beyond predicting whether a variant is pathogenic or not could be to evaluate a NaV1.7 variant's functional profile in terms of its location within the structure. For that, however, it would be necessary to increase

the number of LOF missense mutations that until remains relatively low (only a few cases of LOF missense *SCN9A* mutations have been reported so far and they are related to insensitivity to pain [41–43]).

Finally, given the high similarity among *SCNx*A and *CACNA1x* genes translating alpha-1 subunit of voltage-gated calcium channels (CaVChs), methodologies and observations of this Dissertation could be relevant not only for other NaVChs but also also for CaVChs. After all, scaling laws are expected to exhibit universality [44] which practically means that protein structures of the same evolutionary background are highly likely to exhibit the similar hydropathic scaling properties as we demonstrated for the cases of the pre-open NaVAb and NaV1.7 systems. Moreover, as the authors of [31] showed, treating sequence- and structure-retrieved data of NaVChs and CaVChs jointly can improve model’s predictability power. The clinical spectrum of known channelopathies associated with *SCNx*A and *CACNA1x* genes expands and currently spans areas of neurology, cardiology, psychiatry and ophthalmology thus it is increasingly important to develop a holistic understanding of molecular mechanisms leading to neurological disease. To address these global medical needs, we will inevitably have to resolve issues of molecular complexity and better understand physical principles underlying functional architectures of ion channels.

References

- [1] Khaluf, Y., Ferrante, E., Simoens, P., Huepe ,C. Scale invariance in natural and artificial collective systems: a review. *J R Soc Interface* **14**, 20170662 (2017).
- [2] Bak, P., Tang, C. Wiesenfeld, K. Self-organized criticality: an explanation of $1/f$ noise. *Phys Rev Lett* **59**, 381-384 (1987).
- [3] Chang, T., Tam, S.W., Wu, C. et al. Complexity, Forced and/or Self-Organized Criticality, and Topological Phase Transitions in Space Plasmas. *Space Sci Rev* **107**, 425-445 (2003).
- [4] Chen, K., Bak, P., & Obukhov, S.P. Self-organized criticality in a crack-propagation model of earthquakes.s *Phys Rev A* **43**, 625 (1991).
- [5] Brunk, G. G. Why Do Societies Collapse?: A Theory Based on Self-Organized Criticality. *J Theor Politics* **14**, 195-230 (2002).
- [6] Rohden, M., Sorge, A., Timme, M., & Witthaut, D. Self-Organized Synchronization in Decentralized Power Grids. *Phys Rev Lett* **109**, 064101 (2012).
- [7] Levin, S.A. Self-organization and the Emergence of Complexity in Ecological Systems. *BioScience* **55**, 1075-1079 (2005).
- [8] Moretti, P., & Muñoz, M.A. Griffiths phases and the stretching of criticality in brain networks. *Nat Commun* **4**, 2521 (2013).
- [9] Oddershede, L., Dimon, P., & Bohr, J. Self-organized criticality in fragmenting. *Phys Rev Lett* **71**, 3107 (1993).
- [10] Phillips, J.C. Scaling and self-organized criticality in proteins I. *Proc Natl Acad Sci USA* **106**, 3107-3112 (2009).

- [11] Phillips, J.C. Scaling and self-organized criticality in proteins II. *Proc Natl Acad Sci USA* **106**, 3113-3118 (2009).
- [12] Bonachela, J., De Franciscis, S., Torres, J., & Muñoz, M. Self-organization without conservation: are neuronal avalanches generically critical? *J Stat Mech* P02015, (2010).
- [13] Priesemann, V., Munk, M., & Wibral, M. Subsampling effects in neuronal avalanche distributions recorded in vivo. *BMC Neurosci* **10**, 40 (2009).
- [14] Newman, M. E. Power laws, pareto distributions and zipf's law. *Contemp Phys* **46**, 323-351 (2005).
- [15] Juretić, D. Protein secondary structure conformations and associated hydrophobicity scales. *J Math Chem* **14**, 35-45 (1993).
- [16] Kawashima, S., Pokarowski, P., Pokarowska, M., Kolinski, A., Katayama, T., Kanehisa, M. AAindex: amino acid index database, progress report 2008. *Nucleic Acids Res* **36**, D202-5 (2008).
- [17] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. *Protein Identification and Analysis Tools on the ExPASy Server*. In: Walker J.M. (eds) *The Proteomics Protocols Handbook*. Springer Protocols Handbooks. Humana Press (2005).
- [18] Simm, S., Einloft, J., Mirus, O., et al. 50 years of amino acid hydrophobicity scales: revisiting the capacity for peptide classification. *Biol Res* **49**, 31 (2016).
- [19] Kapcha, L.H., Rossky, P.J. A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure. *J Mol Biol* **426**, 484-498 (2014).
- [20] Jorgenson, W.L., Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* **110**, 1657-66 (1988).
- [21] Landau, L.D., Lifshitz, E.M. *Statistical Physics*, trans. by E. Peierls and R.F. Peierls, (Pergamon Press, London, 1958).
- [22] Xenakis, M.N., Kapetis, D., Yang, Y. et al. Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution. *Proteins* **88**, 1319-1328 (2020).
- [23] Westra, R.L. Resonance-driven ion transport and selectivity in prokaryotic ion channels. *Phys Rev E* **100**, 062410 (2019).

- [24] Xenakis, M.N., Kapetis, D., Yang, Y. et al. Non-extensivity and criticality of atomic hydrophobicity around a voltage-gated sodium channel's pore; a modeling study. *J Biol Phys* **47**, 61-77 (2021).
- [25] Waxman, S.G. Rational Drug Design for Pain Medicine: A New Nav1.7 Inhibitor. *Anesthesiology* **133**, 497-499(2020).
- [26] Yang, Y., A Mis, M., Estacion, M., Dib-Hajj, S.D., Waxman, S.G. Nav1.7 as a Pharmacogenomic Target for Pain: Moving Toward Precision Medicine. *Trends Pharmacol Sci* **39**, 258-275 (2018).
- [27] Kapetis, D., Yang, Y., Sassone, J. et al. Network topology of Nav1.7 mutations in sodium channel-related painful disorders. *BMC Syst Biol* **11**, 28 (2017).
- [28] Xenakis, M.N., Kapetis, D., Yang, Y., et al. Hydrophobicity-based prediction of pain-causing Nav1.7 variants. *Accepted*.
- [29] Bertschinger, N., Natschlager, T. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comput* **16**, 1413-1436 (2004).
- [30] Kinouchi, O., Copelli, M. Optimal dynamical range of excitable networks at criticality. *Nat Phys* **2**, 348-351 (2006).
- [31] Heyne, H.O., Baez-Nieto, Iqba, S. et al. A machine learning method can predict loss- versus gain-of-function effects of human genetic variants in disease-associated ion channels. *Sci Transl Med* **12**, 556 (2020).
- [32] Toffano, A., Chiarot, G., Zamuner, S. et al. Computational Pipeline to probe Nav1.7 gain-of-functions variants in neuropathic painful syndromes. *Sci Rep* **10**, 17930 (2020).
- [33] Shen H., Liu D., Wu K., Lei J., Yan N. Structures of human Nav1.7 channel in complex with auxiliary subunits and animal toxins. *Science* **363**, 1303-1308 (2019).
- [34] Pan, X., Li, Z., Huang, X. et al. Molecular basis for pore blockade of human Na⁺ channel Nav1.2 by the μ -conotoxin KIIIA. *Science* **363**, 1309-1313 (2019).
- [35] Pan, X., Li, Z., Zhou, Q. et al. Structure of the human voltage-gated sodium channel Nav1.4 in complex with β 1. *Science* **362**, eaau2486 (2018).
- [36] Xu, H., Li, T., Rohou, A. et al. Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin. *Cell* **176**, 702-715 (2018).

- [37] Wisedchaisri, G., Tonggu, L., McCord, E. Resting-State Structure and Gating Mechanism of a Voltage-Gated Sodium Channel. *Cell* **178**, 993-1003 (2019).
 - [38] Iqbal, S., Pérez-Palma, E., Jespersen, J.B., et al. Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. *Proc Nat Acad Sci USA* **117**, 28201-28211 (2020).
 - [39] Kong, W., Tu, X., Huang, W., Yang, Y., Xie, Z., Huang, Z. Prediction and Optimization of NaV1.7 Sodium Channel Inhibitors Based on Machine Learning and Simulated Annealing. *J Chem Inf Model* **60**, 2739-2753 (2020).
 - [40] Wolff, M., Johannesen, K.M., Hedrich, U.B.S. et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain* **140**, 1316-1336 (2017).
 - [41] Cox, J.J., Reimann, F., Nicholas, A.K., et al. An SCN9A channelopathy causes congenital inability to experience pain. *Nat* **444**, 894-898 (2006).
 - [42] Goldberg, Y.P., MacFarlane, J., MacDonald, M.L. et al. Loss-of-function mutations in the NaV1.7 gene underlie congenital indifference to pain in multiple human populations. *Clin Genet* **71**, 311-319 (2007).
 - [43] Nilsen, K.B., Nicholas, A.K., Woods, C.G., Mellgren, S.I., Nebuchennykh, M., Aasly, J. Two novel SCN9A mutations causing insensitivity to pain. *Pain* **143**, 155-158 (2009).
 - [44] Kadanoff, L.P. Scaling and universality in statistical physics. *Physica A* **163**, 1-14 (1990).
-

Appendices

Each atom is represented as a hydrophobic sphere of mass and radius corresponding to IUPAC atomic mass indices [1] and to van der Waals radii [2], respectively (Table A1). Hydrophobic property of each atom is determined according to the weighted Kapcha-Rosky hydrophobic scale [3] (Table A2).

Atom type	AM [<i>amu</i>]	<i>vdW</i> radii [Å]
C	12.0107	1.7
N	14.0067	1.55
O	15.9994	1.52
S	32.0650	1.8
H	1.00794	1.09

Table A1: **Atomic masses (AM) and van der Waals (*vdW*) radii.** Atomic masses correspond to the IUPAC indices [1]. *vdW* radii for C, N, O, S atoms were assigned according to [2]. The *vdW* radius of H atoms was taken from [4].

Appendices

Res.	Atom	HI	Res.	Atom	HI	Res.	Atom	HI	Res.	Atom	HI
ALA	N	+1	ASN	N	+1	CYS	HA	-0.5	GLU	HB3	-0.5
ALA	CA	-0.5	ASN	CA	-0.5	CYS	HB2	-0.5	GLU	HG2	-0.5
ALA	C	+1	ASN	C	+1	CYS	HB3	-0.5	GLU	HG3	-0.5
ALA	O	+1	ASN	O	+1	CYS	HG	+1			
ALA	CB	-0.5	ASN	CB	-0.5				GLY	N	+1
ALA	H	+1	ASN	CG	+1	GLN	N	+1	GLY	CA	-0.5
ALA	HA	-0.5	ASN	OD1	+1	GLN	CO	-0.5	GLY	C	+1
ALA	HB1	-0.5	ASN	ND2	+1	GLN	C	+1	GLY	O	+1
ALA	HB2	-0.5	ASN	H	+1	GLN	O	+1	GLY	H	+1
ALA	HB3	-0.5	ASN	HA	-0.5	GLN	CB	-0.5	GLY	HA2	-0.5
			ASN	HB2	-0.5	GLN	CG	-0.5	GLY	HA3	-0.5
			ASN	HB3	-0.5	GLN	CD	+1			
ARG	N	+1	ASN	HD21	+1	GLN	OE1	+1	HIS	N	+1
ARG	CA	-0.5	ASN	HD22	+1	GLN	NE2	+1	HIS	CA	-0.5
ARG	C	+1				GLN	H	+1	HIS	C	+1
ARG	O	+1	ASP	N	+1	GLN	HA	-0.5	HIS	O	+1
ARG	CB	-0.5	ASP	CA	-0.5	GLN	HB2	-0.5	HIS	CB	-0.5
ARG	CG	-0.5	ASP	C	+1	GLN	HB3	-0.5	HIS	CG	-0.5
ARG	CD	-0.5	ASP	O	+1	GLN	HG2	-0.5	HIS	ND1	+1
ARG	NE	+1	ASP	CB	-0.5	GLN	HG3	-0.5	HIS	CD2	-0.5
ARG	CZ	+2	ASP	CG	+2	GLN	HE21	+1	HIS	CE1	+1
ARG	NH1	+2	ASP	OD1	+2	GLN	HE22	+1	HIS	NE2	+1
ARG	NH2	+2	ASP	OD2	+2				HIS	H	+1
ARG	H	+1	ASP	H	+1	GLU	N	+1	HIS	HA	-0.5
ARG	HA	-0.5	ASP	HA	-0.5	GLU	CA	-0.5	HIS	HB2	-0.5
ARG	HB2	-0.5	ASP	HB2	-0.5	GLU	C	+1	HIS	HB3	-0.5
ARG	HB3	-0.5	ASP	HB3	-0.5	GLU	O	+1	HIS	HD1	+1
ARG	HG2	-0.5				GLU	CB	-0.5	HIS	HD2	-0.5
ARG	HG3	-0.5				GLU	CG	-0.5	HIS	HE1	+1
ARG	HD2	-0.5	CYS	N	+1	GLU	CD	+2			
ARG	HD3	-0.5	CYS	CA	-0.5	GLU	OE1	+2	ILE	N	+1
ARG	HE	+1	CYS	C	+1	GLU	OE2	+2	ILE	CA	-0.5
ARG	HH11	+2	CYS	O	+1	GLU	H	+1	ILE	C	+1
ARG	HH12	+2	CYS	CB	-0.5	GLU	HA	-0.5	ILE	O	+1
ARG	HH21	+2	CYS	SG	+1	GLU	HB2	-0.5	ILE	CB	-0.5
ARG	HH22	+2	CYS	H	+1	GLU					

ILE	CG1	-0.5	LYS	N	+1	MET	HB3	-0.5	PRO	CD	-0.5
ILE	CG2	-0.5	LYS	CA	-0.5	MET	HG2	-0.5	PRO	HA	-0.5
ILE	CD1	-0.5	LYS	C	+1	MET	HG3	-0.5	PRO	HB2	-0.5
ILE	H	+1	LYS	O	+1	MET	HE1	-0.5	PRO	HB3	-0.5
ILE	HA	-0.5	LYS	CB	-0.5	MET	HE2	-0.5	PRO	HG2	-0.5
ILE	HB	-0.5	LYS	CG	-0.5	MET	HE3	-0.5	PRO	HG3	-0.5
ILE	HG12	-0.5	LYS	CD	-0.5				PRO	HD2	-0.5
ILE	HG13	-0.5	LYS	CE	-0.5	PHE	N	+1	PRO	HD3	-0.5
ILE	HG21	-0.5	LYS	NZ	+2	PHE	CA	-0.5			
ILE	HG22	-0.5	LYS	H	+1	PHE	C	+1	SER	N	+1
ILE	HG23	-0.5	LYS	HA	-0.5	PHE	O	+1	SER	CA	-0.5
ILE	HD11	-0.5	LYS	HB2	-0.5	PHE	CB	-0.5	SER	C	+1
ILE	HD12	-0.5	LYS	HB3	-0.5	PHE	CG	-0.5	SER	O	+1
ILE	HD13	-0.5	LYS	HG2	-0.5	PHE	CD1	-0.5	SER	CB	-0.5
			LYS	HG3	-0.5	PHE	CD2	-0.5	SER	OG	+1
LEU	N	+1	LYS	HD2	-0.5	PHE	CE1	-0.5	SER	H	+1
LEU	CA	-0.5	LYS	HD3	-0.5	PHE	CE2	-0.5	SER	HA	-0.5
LEU	C	+1	LYS	HE2	-0.5	PHE	CZ	-0.5	SER	HB2	-0.5
LEU	O	+1	LYS	HE3	-0.5	PHE	H	+1	SER	HB3	-0.5
LEU	CB	-0.5	LYS	HZ1	+2	PHE	HA	-0.5	SER	HG	-0.5
LEU	CG	-0.5	LYS	HZ2	+2	PHE	HB2	-0.5			
LEU	CD1	-0.5	LYS	HZ3	+2	PHE	HB3	-0.5	THR	N	+1
LEU	CD2	-0.5				PHE	HD2	-0.5	THR	CA	-0.5
LEU	H	+1	MET	N	+1	PHE	HD3	-0.5	THR	C	+1
LEU	HA	-0.5	MET	CA	-0.5	PHE	HE2	-0.5	THR	O	+1
LEU	HB2	-0.5	MET	C	+1	PHE	HE3	-0.5	THR	CB	-0.5
LEU	HB3	-0.5	MET	O	+1	PHE	HZ	-0.5	THR	OG1	+1
LEU	HG	-0.5	MET	CB	-0.5				THR	OG2	-0.5
LEU	HD11	-0.5	MET	CG	-0.5	PRO	N	-0.5	THR	H	+1
LEU	HD12	-0.5	MET	SD	+1	PRO	CA	-0.5	THR	HA	-0.5
LEU	HD13	-0.5	MET	CE	-0.5	PRO	C	+1	THR	HB	-0.5
LEU	HD21	-0.5	MET	H	+1	PRO	O	+1	THR	HG1	+1
LEU	HD22	-0.5	MET	HA	-0.5	PRO	CB	-0.5	THR	HG21	-0.5
LEU	HD23	-0.5	MET	HB2	-0.5	PRO	CG	-0.5	THR	HG22	-0.5

THR	HG23	-0.5	TRP	HA	-0.5	TYR	CD2	-0.5	VAL	C	+1
			TRP	HB2	-0.5	TYR	CE1	-0.5	VAL	O	+1
TRP	N	+1	TRP	HB3	-0.5	TYR	CE2	-0.5	VAL	CB	-0.5
TRP	CA	-0.5	TRP	HD1	-0.5	TYR	CZ	-0.5	VAL	CG1	-0.5
TRP	C	+1	TRP	HE1	+1	TYR	OH	+1	VAL	CG2	-0.5
TRP	O	+1	TRP	HE3	-0.5	TYR	H	+1	VAL	H	+1
TRP	CB	-0.5	TRP	HZ2	-0.5	TYR	HA	-0.5	VAL	HA	-0.5
TRP	CG	-0.5	TRP	HZ3	-0.5	TYR	HB2	-0.5	VAL	HB	-0.5
TRP	CD1	-0.5	TRP	HH2	-0.5	TYR	HB3	-0.5	VAL	HG11	-0.5
TRP	CD2	-0.5				TYR	HD1	-0.5	VAL	HG12	-0.5
TRP	NE1	+1	TYR	N	+1	TYR	HD2	-0.5	VAL	HG13	-0.5
TRP	CE2	-0.5	TYR	CA	-0.5	TYR	HE1	-0.5	VAL	HG21	-0.5
TRP	CE3	-0.5	TYR	C	+1	TYR	HE2	-0.5	VAL	HG22	-0.5
TRP	CZ2	-0.5	TYR	O	+1	TYR	HH	-0.5	VAL	HG23	-0.5
TRP	CZ3	-0.5	TYR	CB	-0.5						
TRP	CH2	-0.5	TYR	CG	-0.5	VAL	N	+1			
TRP	H	+1	TYR	CD1	-0.5	VAL	CA	-0.5			

Table A2: **Atomic hydropathic indices (HIs) for naturally-occurring amino acids.** Each atom of a naturally-occurring amino acid is assigned with a HI according to the weighted Kapcha-Rosky atomic hydropathic scale [3]. Hydrophobic atoms contribute with -0.5 , hydrophilic atoms with $+1.0$ and charged atoms with $+2.0$ to the total hydropathic score of the amino acid that they belong to. The total score of individual amino acids is calculated by adding up all the corresponding atomic HIs. Accordingly, the amino acids PHE (F), PRO (P), ILE (I), LEU (L), TRP (W), VAL (V), TYR (Y) and MET (M) are classified as hydrophobic with a total hydrophobicity of -4.0 , -4.0 , -3.5 , -3.5 , -3.0 , -2.0 , -1.5 and -1.0 , respectively. The amino acids ALA (A), THR (T), GLY (G), CYS (C), SER (S), GLN (Q), HIS (H), LYS (K), GLU (E), ASN (N), ASP (D) and ARG (R) are classified as hydrophilic with a total hydrophilicity of 1.0 , 2.0 , 2.5 , 3.5 , 3.5 , 5.0 , 5.0 , 5.0 , 6.0 , 6.5 , 7.5 and 14.5 , respectively. The unit of measurement of a HI roughly corresponds to [kcal/mol]. The HH2 atom belonging to the TRP amino acid appears with bold letters to indicate that it was added after communication with the authors of [3] as it did not appear in the original publication (see Table 1 in [3]).

References

- [1] Loss, R.D. Atomic weights of the elements 2001 (IUPAC technical report). *Pure Appl Chem* **73**, 667-683 (2001).
 - [2] Bondi, A. Van der Waals volumes and radii. *J Phys Chem* **68**, 441-51 (1964).
 - [3] Kapcha, L.H., Rossky, P.J. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J Mol Biol* **426**, 484-498 (2014).
 - [4] Rowland, R.S., Taylor, R. Intermolecular nonbonded contact distances in organic crystal structures: comparison with distances expected from van der Waals radii. *J Phys Chem* **100**, 7384-7391 (1996).
-

Dissertation Summary

How does the heart maintain its regular beating rhythm? How does contraction of muscle fibers occur? How does the human body detect temperature variations and process pain stimuli? Although physiological pathways of aforementioned processes might differ, the cellular principle underlying them is common; specialized, excitable cells initiate these processes by altering their electrical properties. Biological molecules allowing cells to alter their electrical properties are named ion channels. Ion channels are pore-forming proteins placed in the membrane of the cell thus creating an ionic passage across the membrane. Ionic flow through this passage establishes an electric current between the cell's interior and the cell's surrounding environment. A remarkable characteristic of an ion channel protein is that it can be highly selective; it permeates entrance to the cell only to specific ion species.

Voltage-gated sodium channels (NaVChs) are members of the ion channels superfamily and their main responsibility is to control the flow of sodium ions across the membrane. The human NaV1.7 channel is preferentially expressed in peripheral neurons such as the dorsal root ganglion (DRG) neurons and plays a key-role in human nociception. Studying the genetic background, and the structure and function of the NaV1.7 channel can thus not only shed light upon molecular mechanisms of pain stimuli transmission throughout the nervous system but also enhance our understanding of pathologies that arise from

their disruption. In particular, missense mutations in the *SCN9A* gene encoding the NaV1.7 channel protein can deregulate electrical properties of DRG neurons and, consequently, alter their electrophysiological activity thus triggering pain-related diseases such as inherited erythromelalgia (IEM), paroxysmal extreme pain disorder (PEPD) and small fiber neuropathy (SFN).

Computational pipelines aiming at translating NaVCh-related genetic, structural and functional data into clinically-relevant information are rare and not easy to construct as they demand considerable effort and expertise. Although the NaV1.7 channel is one of the most well-studied biophysical cases in pain medicine, basic questions regarding its function and structure are still open. How do missense *SCN9A*-gene mutations affect the functional architecture of the NaV1.7? Which missense *SCN9A*-gene mutations lead to a disease phenotype? Answering these questions could not only provide with insight into disease-related molecular mechanisms and therapeutic targets but also reveal fundamental principles underlying conformational stability of NaVChs and ion channels in general.

The overall aim of this Dissertation was to demonstrate how theoretical and computational approaches rooted in complex systems science might contribute to our understanding of NaVCh functional architectures toward predicting phenotypic response in the occurrence of a mutation. The term "molecular complexity" was used throughout this Dissertation as an umbrella term to describe observations of function-relevant hydropathic self-similarities and topologies within a NaVCh structure. Stated differently, we hypothesized that hydropathicity property plays a key-role in dictating how atomic NaVCh components are organized, i.e., packed, around the pore, as well as, in guaranteeing stable channel functioning within the membrane. This led us to the construction of a phenomenological framework for modeling the spatial organization of hydropathic structural components around a NaVCh's pore at atomic resolution. The clinical relevance and potential of presented phenomenology was demonstrated on a

structural model of the NaV1.7 where clinical phenotypes of NaV1.7 variants were predicted.

Although we utilized tools and concepts from different disciplines, this Dissertation does not represent a multidisciplinary piece of work but, rather, an interdisciplinary approach as it was conducted at the intersection of disciplines such as structural biology, statistical physics, neurology and medicine. Hence, besides its direct epistemological impact, this Dissertation investigated also how aforementioned disciplines may or may not intersect with each other and what kind of conceptual frameworks are required in order to fluidly cross their boundaries.

Chapter 1 aimed at familiarizing the reader with the field of NaVChs research from the viewpoint of complex systems science and to briefly cover background information of biophysical and computational modeling approaches that are essential for this Dissertation. Content of general introduction was drawn from scientific articles published not later than the starting time-point of this Dissertation, i.e., October of 2014, in accordance to Maastricht University's rules. We first presented basic aspects of NaVCh structure and function, as well as, the role of NaVChs in health and disease. We focused on the human NaV1.7 channel, its role in pathophysiology of pain and shortly reviewed how missense *SCN9A*-gene mutations can lead to gain-of-function (GOF) effects associated with IEM, PEPD and SFN disease. In relation to pain disease, we also covered what is the current status of pharmacological research targeting the NaV1.7. Next, we presented key-concepts and developments in computational modeling that serve as a methodological basis for this Dissertation. We highlighted the importance of hydropathicity property in studying stability of protein structures, as well as, modeling challenges arising from spatial inhomogeneities of the hydropathic-interactions laws driving molecular processes. We presented evidence for the ability of hydropathicity-based measures to capture crucial alternations of the NaV1.7 functional architecture associated with neuropathic pain disorders. Emphasis was

given on the notion of cumulative hydrophobicity and its potential as a hydrophobic-interactions summary index. The last section of **Chapter 1** outlines the main aim and scope of this Dissertation and briefly introduces the complex-systems phenomenological framework from which methods of this Dissertation stem from.

Chapter 2 presents a novel concept for analyzing NaVCh functional architectures, namely, the concept of cumulative hydrophobic topology (CHT). CHT analysis aims at mapping hydrophobic density variations and topological configurations of the hydrophobic dipole field around a NaVCh's pore on two dimensions and, consequently, translate them into stability- and function-relevant information. In order to demonstrate usefulness of CHT analysis, we utilized a prototype NaVCh structure, namely, the bacterial NaVAb channel at pre-open conformation available at 2.7 Å resolution (PDB code: 3RVY). Interpretation of CHT observables provided insight into spatial organization and, hence, relative stability of components forming the NaVAb's pore module (PM) and voltage-sensors (VSs), as well as, into hydrophobic gating mechanisms regulating ion conduction. A probabilistic summary of CHT pore characteristics was introduced in order to quantify the relative energetic size of hydrophobic gates along the pore. We found that a central gating mechanism renders NaVAb's hydrophobic central cavity (CC) impermeable by imposing large energy barriers to ions entering from the extracellular side (ES) or from the intracellular side (IS), while allowing entrance to small hydrophobes and waters into the pore through NaVAb's side fenestrations. Accordingly, hydrated sodium ions escaping from NaVAb's selectivity filter (SF) have to overcome a large energy barrier in order to arrive at the center of the pore and rest in the cavity. Escape trajectory of selected sodium ions is thus likely to be long and gradual, i.e., occur in multiple steps involving subsequent partial de- and rehydration events. We proposed that the advantage behind a long escape trajectory from the SF toward the CC could be that NaVAb's specificity is enhanced as entrance into the CC is permitted only to ions that can re- and de-

hydrate according to the local gradients of the energy ion-permeation landscape.

Chapter 3 built upon findings of **Chapter 2** and provided with computational evidence demonstrating that the pre-open NaVAb structure can be described as a hydrophatically-tuned self-organized critical (SOC) molecular system. Although the SOC hypothesis has been previously proposed as a possibility for summarizing and modeling molecular complexity of a protein, **Chapter 2** is the first study to systematically probe the SOC hypothesis for an ion channel protein. In order to do that, we constructed a computational toolbox for rigorous modeling of cumulative atomic characteristics around a NaVCh's pore. We found that the atomic environment around NaVAb's pore exhibits a bi-phasic spatial profile closely following the structural separation of the PM from the VSs. The Richards growth model provided with accurate fittings on the atomic cumulative distribution function traces around the pore revealing the existence of structural inflection points separating the two phases. Next, we investigated the scaling behavior of the amplitude of the hydrophatic dipole field with respect to detected inflection points. This led to the discovery of a critical-points region around the SF and, hence, verified the SOC hypothesis for the pre-open NaVAb state. A direct consequence of this finding is that interatomic hydrophatic interactions regulating SF's stability are of long-range nature with their intensity regulated by a pair of pre- and post-inflection critical exponents. We proposed that the long-rangeness of the interatomic hydrophatic interactions might be crucial for NaVAb's stable functioning. In particular, expansion of hydrophatic interactions network around the SF could provide peripheral structural support to the SF's binding sites implementing ion dehydration. Moreover, we estimated that the coupling energy between the PM and the VSs is ~ 282.1 kcal, while the energy cost for maintaining the pre-open NaVAb conformation stable in the membrane is ~ 12.3 kcal. From a mutation-response perspective, this energy difference implements a "shield" against potentially destabilizing

events as mutation-induced energetic perturbations are expected to be damped-out toward channel's interior or exterior thus not affecting the SF neither sacrificing channel's stability within the membrane.

Chapter 4 utilizes the methodologies presented in **Chapter 2** and **Chapter 3** in order to predict whether a NaV1.7 variant causes pain disease or not. To achieve that, we mapped locations of missense *SCN9A*-gene mutations upon cumulative NaV1.7 hydropathic characteristics and probed their relationship to underlying hydropathic topological and scaling laws. We demonstrated that pain-related mutation sites tend to occupy structural locations in proximity to a hydrophobic patch (HP) lining the pore and extending from the CC toward the AG. In addition, we found that pain-related mutation sites cluster within a spherical shell around the SF so that the vast majority of them are found at a distance of ≈ 33.4 Å from a pore point located at the ES of the SF. The biophysical importance of this observation became clear in the light of a local hydropathic scale-invariance mechanism which, similarly to the NaVA β case, identified the local inflection point as a critical point. Armed with this finding, and in accordance to findings of **Chapter 2**, we proposed that clustering of mutations at critical hydropathic-interactions distance from the SF might reflect a delicate trade-off between potentially-deleterious destabilizations occurring too close to the SF and insignificant polymorphisms occurring far away from it. According to this rationale, mutations leading to GOF effects associated with pain phenotypes can increase channel's configurational space and, consequently, expand the physiological range of ionic currents, while not risking structure deletion or severe destabilization of the SF that would impede selectivity process. We combined observations associated with the distance of mutations sites from the HP and from the SF by introducing a weighted distance average. The weighted distance average classified correctly the phenotype of 29 (out of 36) pain-related variants and 45 (out of 48) neutral variants [area under the receiver operating characteristics curve = 0.872]. This result suggested that predicting the pathogenicity of a missense *SCN9A*-gene

mutation can be done with relatively-high accuracy purely based on the location of corresponding mutation sites within the NaV1.7 structure. Accordingly, we argued that the main advantage of the presented variants classification scheme in comparison to state-of-the-art approaches is that despite its negligible computational cost does not sacrifice classification accuracy.

Chapter 5 discussed and reviewed key-points and -findings of this Dissertation. We did so through the prism of complex systems science and highlighted notions of complexity that served as building blocks for this Dissertation. We pointed out that a crucial element of this Dissertation was utilization of an atomic-level binary hydropathic classification approach allowing for treating each atoms in the structure as a single hydropathic component. In fact, shifting our attention from amino acids to atoms, allowed for increasing the size of sampled components around the pore so that spatial inhomogeneities of the distribution profiles of their hydropathic characteristics became evident. A promising future research direction rooted in the phenomenological framework that this Dissertation presented could be attempting to predict whether missense mutations in genes translating human NaVChs are associated with GOF or loss-of-function effects. This development would be of great clinical importance as it could increase efficiency of personalized clinical interventions. Given that the crystallographies of the human NaV1.2 and NaV1.4 channels are now available, methodologies and observations of this Dissertation can be immediately applied on channelopathies associated with *SCN2A* and *SCN4A* genes, respectively. Moreover, due to the high structural similarity between calcium-gated sodium channels and NaVChs channels findings of this Dissertation could be relevant for *CACNA1x*-genes-associated channelopathies as well. Finally, we acknowledge the increasing importance and influence of machine learning (MLE) in the field of NaVChs translational research, and discuss why and how molecular complexity metrics can be valuable assets for MLE-based computational pipelines.

Impact Paragraph

Ion channels are information transmitters of living cells. Ion channel defects appear in a broad range of diseases including epilepsy, migraine, diabetes, hypertension, cardiac arrhythmia, and neuropathic pain. The socioeconomic burden coming from these diseases is significant and affects all aspects of a patient's everyday life, as well as, of their family members. One of the scopes of biomedical research targeting ion channels is to improve diagnostics and develop therapeutic agents and strategies for counteraction, or at least, management of aforementioned diseases. In this Dissertation, we focus on a specific type of ion channels, namely, on voltage-gated sodium channels (VGSCs), which are known to play a key-role in human pain pathophysiology. Results and findings of this Dissertation provide with novel insight into how VGSCs remain functionally-stable within the cell membrane, as well as, into how genetically-caused defects affecting VGSC structure relate to neuropathic pain disease. The main impact of this Dissertation is an improvement in the diagnostic pipeline of inherited neuropathic pain diseases achieved by increasing speed and accuracy of diagnostics while lowering their cost. This is crucial for patients and their family members as it can lead to faster and more accurate genetic counseling and disease diagnosis. Finally, a secondary impact of this Dissertation is that it enriches our knowledge of ion channels as complex systems which, in turn, can prepare the ground for theoretical and computational approaches investigating fine-tuning phenomena in pore-forming molecules to flourish.

Acknowledgments

I would like to start by thanking my initial supervision team, Dr. Patrick Lindsey, Dr. Ronald Westra, Dr. Enno de Lange, and, Dr. Radek Szklarczyk, as well as, Dr. Hubert Smeets, and, Dr. Catharina Faber, for making it possible for me to engage with this project by selecting be for the PhD candidate position.

Special thanks go to Dr. Patrick Lindsey, Dr. Ronald Westra, and, Dr. Hubert Smeets for supervising by constantly supporting, guiding, and co-shaping the material of this Dissertation, especially, during turbulent times. Without your mentoring, this Dissertation would not have been materialized.

I would also like to thank Dr. Pieter Collins, Dr. Pietro Bonizzi, and, Dr. Ralph Peeters for stimulating discussions upon selected topics and ideas, especially, methods and applications of dipole-moments expansion.

I am grateful to all scientific, as well as, non-scientific stuff in the Team of Clinical Genomics and Cell Biology, led by Dr. Hubert Smeets, for contributing to conceive, but, also to materialize this Dissertation. Special thanks go to Dr. Monique Gerrits, and to MSc Milena Molasy for contributing to collecting and interpreting genetic data.

I would also like to express my gratitude to the PROPANE ensemble for providing with scientific support, as well as, with an excellent network to develop as a scientist within the field of ion channels. Special thanks go to Dr. Dimos Kapetis, Dr. Yang Yang, Dr. Jordi Heijman, Dr. Stephen Waxmann, and Dr. Giuseppe Lauria

for supporting, encouraging, and always being available for critically key-points/ideas presented in this Dissertation.

Gratitude also to Dr. Angelika Lampert for fruitful discussions upon ion channels, and for sharing electrophysiology data at the beginning of this PhD journey.

Least but not last, I acknowledge the crucial role that my family, especially, my mother Korinna, my friends, and my partner, Meri Linna, played in shaping the content of this Dissertation via engaging in continuous exchange of ideas, intellectual discourses, and, providing with unconditional love and continuous support throughout this journey.

Curriculum Vitae

Markos Xenakis was born 5th of May, 1986, in Viersen NRW, DE. He holds the greek nationality, and speaks four languages; greek, german, english, and finnish. Markos's research is focused on computational biology and medicine. He is currently interested in acquiring a deeper understanding of functional architectures of voltage-gated (especially sodium) channels toward pharmacological control and mutation-response prediction in pain disorders. His broader research interest is rooted in complex systems, both, phenomenological theory and applications.

Contact details:

Address: Villa Snäcksund 1, Raasepori 10600, Finland

Phone: +35 8401994016

E-mail: markos.xenakis@maastrichtuniversity.nl

Scholar Profile: <https://scholar.google.com/citations?user=5i1dhz8AAAAJ&hl=en>

Current position

PhD Researcher, Department of Toxicogenomics, Section of Clinical Genomics, Maastricht University

Areas of specialisation

Computational Biology and Neurology; Ion Channels; Complex Systems

Appointments held

- Mar. 2011 - Oct. 2011 *Research Associate*
Complexity Research Team, Democritus University of Thrace, GR
- May 2012 - Oct. 2013 *Research Assistant*
Institute of Medicine and Neuroscience, Neuromodulation (INM-7), Forschungszentrum Juelich, DE
- Oct. 2014 - present *PhD Researcher*
Department of Toxicogenomics, Section Clinical Genomics, Maastricht University, NL
- May 2016 - July 2016 *Scientific Visitor*
Department of Neurology, and Center for Neuroscience and Regeneration Research, School of Medicine, Yale University, US

Education

- Sep. 1997 - Jun. 2004 *Abitur (German Baccalaureate)*, German School of Athens, GR
- Sep. 2004 - Mar. 2011 *Dipl-Ing Electrical and Computer Engineering*, Democritus University of Thrace, GR
- Oct. 2011 - Nov. 2013 *MSc in Simulation Sciences*, RWTH Aachen University and Forschungszentrum Juelich, German Research School for Simulation Sciences, DE

Selected Journal Publications

- **Xenakis, M.N.**, Kapetis, D., Yang, Y., et al. Hydropathicity-based prediction of pain-causing NaV1.7 variants. *In press*.
- **Xenakis, M.N.**, Kapetis, D., Yang, Y., et al. Non-extensivity and criticality of atomic hydropathicity around a voltage-gated sodium channel's pore; a modeling study. *J Biol Phys* **47**, 61-77 (2021).
- **Xenakis, M.N.**, Kapetis, D., Yang, Y., et al. Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution. *Proteins* **88**, 1319-1328 (2020).
- Kapetis, D., Sassone, J., Yang, Y., Galbardi, B., **Xenakis, M.N.**, [...] Network topology of NaV1.7 mutations in sodium channel-related painful disorders. *BMC Systems Biology* **11**, 28 (2017).
- Patsourakos, S., Georgoulis, M.K., Vourlidas, A., [...], **Xenakis, M.N.**, Sarris, E., Tsinganos, K., Vlahos, L. The major geoeffective solar eruptions of 2012 March 7: comprehensive sun-to-earth analysis. *Astrophys J* **817**, 14 (2016).
- Pavlos, G.P., Karakatsanis, L.P., Iliopoulos, A.C., Pavlos, E.G., **Xenakis, M.N.**, Clark, P., Duke J., Monos, D.S. Measuring complexity, nonextensivity and chaos in the DNA sequence of the Major Histocompatibility Complex. *Physica A* **438**, 188-209 (2015).
- Popovych, O.V., **Xenakis, M.N.**, Tass, P.A. The Spacing Principle for Unlearning Abnormal Neuronal Synchrony *PLoS ONE* **10**, e0117205 (2015).
- Pavlos, G.P., Karakatsanis, L.P., **Xenakis, M.N.** Tsallis non-extensive statistics, intermittent turbulence, SOC and chaos in the solar plasma. Part two: Solar flares dynamics. *Physica A* **392**, 3920-3944 (2013).
- Pavlos, G.P., Karakatsanis, L.P., **Xenakis, M.N.** Tsallis non-extensive statistics, intermittent turbulence, SOC and chaos in the solar plasma. Part one: Sunspot dynamics *Physica A* **391**, 6287-6319 (2012).

Selected Conference Attendances

- **Xenakis, M.N.**, Westra, R., Lindsey, P. *Modeling spatial complexity of a voltage-gated sodium channel* 25th Summer School of Dynamical Systems and Complexity, National Center of Scientific Research "Demokritos", Athens, GR (July 2018).
- **Xenakis, M.N.**, Westra, R., Lindsey, P. *Investigating the cumulative hydrophobic topology of a bacterial voltage-gated sodium channel* Revolutions in Structural Biology: Celebrating the 100th Anniversary of Sir John Kendrew, EMBL, Heidelberg, DE (Nov. 2017).
- **Xenakis, M.N.**, Westra, R., Lindsey, P. *Hydropathic topology of an ion channel's pore* Annual MHeNS Conference, Maastricht University, Maastricht, NL (Dec. 2016).
- Monos, D.S., Pavlos, G.P., Karakatsanis, L.P., Iliopoulos, A.C., **Xenakis, M.N.**, Pavlos, E.G. *A Multidisciplinary Approach for the Understanding of DNA Sequences and Organization: Where Genetics Meets Physics and the High Throughput Sequencing Technologies Meet Complexity Theory* 3rd Summer School in Medical & Biosciences Research & Management, Mani Lakonias, GR (May 2014).
- Popovych, O.V., **Xenakis, M.N.**, Tass, P.A. *Desynchronization by Spaced Coordinated Reset Neuromodulation* NeuroVisionen 10, Forschungszentrum Juelich, DE (Sep. 2014).
- **Xenakis, M.N.**, Tass, P.A., Popovych, O.V. *Cumulative desynchronizing effects of weak stop-and-go coordinated reset stimulation*, INM Retreat, Forschungszentrum Juelich, Juelich, DE (Jul. 2013).

Inventions

- O.V. Popovych (1/3), P.A. Tass (1/3), **M.N. Xenakis** (1/3) *Device and method for effective non-invasive desynchronizing neurostimulation* (WO2016071236A1), application filed by Forschungszentrum Jülich GmbH (2016).

-
- O.V. Popovych (1/3), P.A. Tass (1/3), **M.N. Xenakis** (1/3) *Apparatus and method for effective invasive desynchronizing neurostimulation* (DE102014115994A1), application filed by Forschungszentrum Jülich GmbH (2014).

Teaching and Supervision

- 26th of Feb. 2016 *Journal club presentation for master course in "Systems Biology*, Maastricht University, NL
Presentation title: "Computational and analytical tools for molecular moments analysis."
- April 2017 - June 2017 *Tutoring bachelor students in "Linear Algebra" course*, Maastricht University, NL
Weekly-based tutoring of 13 bachelor students by using a problem-based-learning oriented teaching method.
- Sep. 2018 - May 2018 *Co-supervision of Bachelor Thesis research project conducted by Nicholas Verzè*, Maastricht University, NL
Thesis title: "Building a model using hydrophobic properties to identify the likelihood of genetic variants of Nav1.7 being pathogenic."

Professional Activity

- Jan. 2019 - present *Consultancy services in Applied Sciences and Arts*, Providing scientific support and consultancy services to academic and non-academic actors.
- Current Project: Simulation & Complex Systems Science Consultant within the DementiaXR initiative @ Fabric, Helsinki, FIN.