Learning molecular traits of human pain disease via voltage-gated sodium channel structure renormalization

Markos N. Xenakis[1,2*] and Angelika Lampert[1,2]

[1]Institute of Neurophysiology, Uniklinik RWTH Aachen, Pauwelsstraße 30, Aachen, 52074, NRW, Germany.
[2]Scientific Center for Neuropathic Pain Research Aachen, SCN^{AACHEN}, Uniklinik RWTH Aachen, Roermonderstraße 110a, Aachen, 52072, NRW, Germany.

*Corresponding author(s). E-mail(s): mrksxenakis@gmail.com;
Contributing authors: alampert@ukaachen.de;

Abstract

Mammalian neurophysiology vitally depends on the stable functioning of transmembrane, pore-forming voltage-sensing proteins known as voltage-gated sodium channels (NaVChs). Deciphering the principles of NaVCh spatial organization can illuminate fundamental structure-function aspects of pore-forming proteins and offer new opportunities for pharmacological treatment of associated diseases such as chronic pain. Here, we introduce a renormalization group flow paradigm permitting a formal investigation of NaVCh thermostability properties. Our procedures are solidified by deriving a q-deformed statistical mechanical entropy and validated over 121 experimentally resolved NaVCh structures of prokaryotic and eukaryotic origin. We uncover the universality of a critical inflection point regulating the thermostability of the pore domain relative to the voltage sensors, summarized in terms of a generalized Widom scaling law. A machine learning algorithm, rationalized in terms of the 'loosening' of inertia and conductivity channel constraints, identifies pain-disease-associated mutation hotspots in the human NaV1.7 channel. Our work illustrates how first-principles-based machine learning approaches can deliver accurate insights for human pain medicine and clinicians at a reduced computational cost, while clarifying the self-organized critical nature of NaVChs.

Keywords: Voltage-gated Sodium Channel, Renormalization, Criticality, Human Pain Disease, Machine Learning

1 Introduction

The primary neurophysiological role of voltage-gated sodium channels (NaVChs) is to initiate and propagate action potentials along neuronal tissue [1]. Accordingly, NaVChs act as 'traffic controllers' for sodium ions crossing the cell membrane, thereby shaping the upstroke phase of action potentials [2, 3].

The key property that makes sodium ion transport entropically favorable across the hydrophobic membrane environment is hydrophilicity [4]. Hydrophilic groups coordinate the dehydration of sodium ions along a NaVCh pore, facilitating the ion's passage through the membrane [5–7]. The physical principle underlying NaVCh selectivity is thought to arise from a delicate balancing of strong and potentially long-range interactions between the selectivity filter (SF) and surrounding residue clusters [8], as inferred from early electrophysiological studies [9].

Members of the NaVCh superfamily follow a common structural organization principle: they form a porous-like environment composed of four radially arranged homologous domains (labeled DI-IV) [10]. The first NaVCh structure to be crystallized was the prokaryotic sodium channel from Arcobacter butzleri (NaVAb), captured in a pre-open state [11]. As anticipated from earlier crystallographic studies on potassium channels [12, 13], the NaVAb crystal verified that each domain comprises a pore module (PM) interlinked with a voltage sensor domain (VSD) [11]. The PM is composed of two antiparallel alpha-helical segments (labeled S5-6) connected by an extracellular loop and the SF [11]. The spatial arrangement of the four PMs inside a NaVCh gives rise to the pore domain (PD) sub-architecture, where a narrow, sieve-like extracellular SF-lined environment, followed by a hydrophobic central cavity, leads to an intracellular pore narrowing where the putative activation gate (AG) is located [11]. On the other hand, the VSD consists of four transmembrane alpha-helices (labeled S1-4) connected to the PM by a linker bridging S4 with S5 [11]. Positively charged residues in the S4 helix can sense the membrane depolarization stimuli [14]. Their outward movement exerts a pulling force that opens the pore [14–17]. In contrast to their prokaryotic ancestors, eukaryotic NaVChs display a more diverse structural and functional profile, mirroring the breaking of radial structural symmetries [10]. Consequently, the dynamic repertoire of eukaryotic NaVChs is richer, yet more specialized and allosterically efficient compared to their prokaryotic counterparts [18–20].

Complex biomolecules, such as NaVChs, have evolved to balance sensitivity and robustness to perturbations of environmental and genetic origin [21, 22]. Concretely, a search in the gnomAD [23] database reveals thousands of benign variants associated with genes encoding NaVChs in humans (e.g., 729 benign variants are found in the *SCN9A* gene [24]), encoding the human NaV1.7 channel molecule which is crucial for human pain perception [25]. This implies that NaVChs, albeit complex, exhibit substantial structural and functional tolerance to random variations in the encoding genetic material. What physical principle guarantees this?

It turns out that protein systems leverage self-similarity (or scale-invariance [26]) to optimally distribute internal energies and external stresses throughout their structure [27, 28]. This phenomenon is known as self-organized criticality (SOC) [29]. SOC is believed to be evolution's answer to the need for maintaining an internal molecular state of high functional sensitivity without sacrificing mutational robustness

[21, 22, 27, 28, 30, 31]. In logical connection with early studies demonstrating the universal characteristics of hydropathy plots in globular proteins [32–35], SOC signatures manifest in protein molecules as a pattern of extrema (peaks and valleys) of intrinsic thermostability cost functions associated with water-mediated interactions [30, 31].

Scaling analysis of the hydropathic dipole field acting along the pore shed light on SOC phenomenology in the NaVAb molecule captured at a pre-open state [36] and in a computationally optimized NaV1.7 molecule found in humans [37]. It was demonstrated that the distribution of atoms around the pore of the NaVAb molecule is predominantly unimodal, meaning the corresponding cumulative distribution function is characterized by a prominent inflection point [36, 37]. Also, the amplitude of the hydropathic dipole field exhibited a clear self-similar increment about the pore axis, leading to a distinct peak near the inflection point [36, 37]. Moreover, the radial location of this prominent inflection point marked the characteristic size of the PD, i.e., the distance from the current pore point where the structural transition from the PD to the VSDs takes place [36, 37]. Based on this triple coincidence, it is compelling to conceptually parallel the spatial organization of atoms around the pore with the archetypal SOC sandpile model [29] as a means to develop a toy model for NaVCh molecular complexity [38].

Accordingly, residue structural positions are abstracted as lattice sites, with mutations analogous to externally added sand grains perturbing the lattice structure. Slowly adding sand grains to the lattice leads to the formation of a theoretical sandpile. Here, 'theoretical' indicates that the sandpile itself is not a direct observable. Rather, its defining characteristic, defined by its slope, can be inferred from the scaling (power-law) exponents [30, 31, 36, 37], characterizing the self-similar incrementing of the underlying hydropathic dipole field about a pore point. Each pore point thus serves as an initial condition for traversing the sandpile at a varying slope determined by the self-similar property of the locally acting field [36, 37]. If a sand grain lands on a site where the slope of the sandpile is too steep, it is expected to adversely affect the stability of the molecule. Stated differently, a steep sandpile slope prescribes a high likelihood of mutation-induced destabilization. This insight provided a rationale for distinguishing between pain-disease-associated and benign structural locations in NaV1.7 [37].

A self-consistent scaling theory for proteins should be grounded in the renormalization group (RG) framework [39]. Here, we examine this concept within the NaVCh superfamily and demonstrate how it can yield biomedically relevant insights concerning inherited forms of human pain disease. Specifically, starting from the simplest possible radial differential equation that can rationally support a pore-forming molecular architecture with a PD/VSD interface, we derive the RG flow equation facilitating a substantial yet biophysically meaningful reduction of molecular degrees of freedom (DoF) about a NaVCh pore. This yields universal insights into dimensionality reduction, entropy constraints, and symmetry breaking of hydropathic dipole fields across the NaVCh superfamily, as revealed by atomic-level analysis of 121 sufficient-resolution NaVCh structures of prokaryotic and eukaryotic origin. At a single-NaV1.7-molecule level, we rigorously describe mutation clustering relative to the PD/VSD interface,

establishing a transparent machine learning framework for detecting NaV1.7 structural locations where the steepening of the sandpile slope correlates with a human pain disease phenotype.

2 Methods

2.1 Analytical procedures

We assume that the NaVCh molecule has settled into a reasonably long-lived or frequently visited metastable state.

2.1.1 Spatial organization principles

We consider the cloud of atomic coordinates, $\mathbf{a} \in \mathcal{A}$, around a pore point, $\mathbf{p} \in \mathcal{P} \subseteq \mathbb{R}^3$ (SI Eq. (S1) and Fig. 1(a)).

We introduce an ensemble of nested spherical subsystems about $\mathbf{p}$:

$$\mathcal{B}(\mathbf{p}, l) = \{\mathbf{a} \in \mathcal{A} \quad | \quad \mathrm{r_a} < l\}, \quad \mathrm{r_a} = \|\overbrace{\mathbf{p} - \mathbf{a}}^{=\vec{\mathbf{r}}_\mathrm{a}}\| \quad [\text{Å}] \tag{1}$$

by continuously varying the radius parameter $l > 0$ [Å]. Here, $\vec{\mathbf{r}}_\mathrm{a}$ is the vector from $\mathbf{p}$ to $\mathbf{a}$. $\|\cdot\|$ is the Euclidean norm. $\partial\mathcal{B} = \{(x, y, z) \in \mathbb{R}^3 \quad | \quad (x - \mathrm{p}_x)^2 + (y - \mathrm{p}_y)^2 + (z - \mathrm{p}_\perp)^2 = l^2\}$ denotes the surface (or interface) of $\mathcal{B}$, where $\mathrm{p}_\perp$ is the membrane-perpendicular pore point coordinate, i.e., the pore point coordinate indicating the direction of the principal pore axis.

Functions and parameters appearing henceforth depend on $(\mathbf{p}, l)$ and $\mathbf{p}$, respectively. For clarity, this dependence is omitted.

2.1.1.1 Atom packing in the continuum

The number of atoms residing on $\partial\mathcal{B}$ is continuously approximated with the following generalized logistic differential equation [36, 40]:

$$\begin{aligned} &\partial n = \partial l\Big(\overbrace{n/\zeta}^{\text{repulsion term}} - \underbrace{n^{\nu+1}/(\zeta K^\nu)}_{\text{stabilization term}}\Big) \quad [\text{atom}] \quad \text{for} \quad n_0 > 0, \\ &\text{with} \quad K > 0 \quad [\text{atom}] \quad \text{and} \quad \nu = \frac{\zeta}{\xi}, \quad \zeta > 0 \quad [\text{Å}], \quad \xi > 0 \quad [\text{Å}]. \end{aligned} \tag{2}$$

n is the number of atoms residing in $\mathcal{B}$. ζ represents the effective range over which atoms push each other apart (i.e., repel each other). On the other hand, ξ represents the effective range over which atoms pull each other together (i.e., attract each other). $n_0 = n(l_0)$ is the atom packing initial condition realized for some initial l-value, l_0. Meaning n_0 is the number of atoms residing in the initial spherical subsystem, $\mathcal{B}_0 = \mathcal{B}|_{l=l_0}$. K [atom] is the carrying capacity of the NaVCh structure. It delimits the number of atoms residing in $\mathcal{B}$ for $l \to \infty$ (see also (3)), thereby setting an upper bound on the number of atoms the structure can accommodate.

We propose an effective phenomenology supporting (2) based on the interplay between hydrophobic attraction and hydrophilic repulsion [41]. The nonlinear term in the right hand side (RHS) of (2) accounts for attractive, hydrophobicity-driven atomic interactions stabilizing an n-cluster. Namely, it reflects the tendency of hydrophobic constituents to 'hide' inside $\mathcal{B}$. The term $n^{\nu+1}$ indicates that any group of $\nu+1$ atoms can form an interaction subnetwork. Since no spatial constraints apply to the formation of such a group, $n^{\nu+1}$ accounts also for long-range interatomic interactions. Hence, $\nu = \zeta/\xi$ explains the emergence of long-range interactions through bond length adjustment driven by atomic repulsivity.

Accordingly, the linear term appearing in the RHS of Eq. (2) accounts for repulsive, hydrophilicity-driven interactions pushing atoms toward $\partial\mathcal{B}$. This is attributed to space-filling water structuring effects that increase the energy required for efficient atom packing, according to the thermodynamic self-assembly principles outlined in [42]. Additionally, ζ incorporates the Pauli exclusion principle, as discussed in Ref. [43].

2.1.1.2 The characteristic size of a pore domain

Solving (2) for n gives:

$$n = K\Big(1+\nu\frac{\varepsilon}{\varepsilon_i}\Big)^{-1/\nu} \quad [\text{atom}], \quad n_{l\to\infty} = K \quad [\text{atom}], \tag{3}$$

where

$$\begin{aligned} &\varepsilon = \varepsilon_0 \exp\left(-l/\xi\right) \quad [\text{kcal/atom}] \\ &\text{with} \quad \varepsilon_i = \varepsilon|_{l=l_i} = \varepsilon_0 \exp\left(-l_i/\xi\right) \quad [\text{kcal/atom}], \quad l_i = \xi \ln\left(\frac{(n_0/K)^{-\nu}-1}{\nu}\right) + l_0 \quad [\text{Å}]. \end{aligned} \tag{4}$$

ε imposes an upper bound on the absolute hydropathic energy of an individual atom. On the basis of simplicity, ε_0 [kcal/atom] is chosen to be a constant. Accordingly, $n_0\varepsilon_0$ [kcal] delimits the absolute hydropathic energy stored in the initial spherical subsystem, $\mathcal{B}_0$. $1/\xi$ [Å^{-1}] can then be interpreted as the rate at which hydrophobic energy surpluses, stored in $\mathcal{B}_0$, are consumed in useful (stabilizing) interactions as the n-cluster size increases. We note that this interpretation is simply a reformulation of our initial understanding about $1/\xi$ deduced from (2).

We do not seek a biophysical interpretation for the logarithmic argument, $\frac{(n_0/K)^{-\nu}-1}{\nu}$. Instead, we focus on l_i.

We emphasize that l_i marks the l-value for which the second-order derivative of n with respect to l, i.e., the curvature of n along the radial direction,

$$\frac{\partial^2 n}{\partial l^2} = \frac{n}{\xi^2}\frac{\left(\left(\frac{\varepsilon}{\varepsilon_i}\right)^{-1}-1\right)}{\left(\left(\frac{\varepsilon}{\varepsilon_i}\right)^{-1}+\nu\right)^2} \quad [\text{atom/Å}^2], \tag{5}$$

becomes zero. Stated differently, l_i is an inflection point of n. It follows that for $l = l_i$, (2) is globally maximized, i.e.,:

$$\max_{l} \left\{ \frac{\partial n}{\partial l} \right\} = \frac{\partial n}{\partial l}\Big|_{l_i}. \tag{6}$$

The biophysical significance of (6) becomes apparent when considering empirical evidence suggesting that $\partial\mathcal{B}_i = \partial\mathcal{B}|_{l=l_i}$ serves as an approximation of the interface mediating the structural transition from the PD to the VSDs [36, 37]. Accordingly, we treat l_i as the characteristic length scale of the PD sub-architecture, i.e., the maximum size at which the PD retains its essential functional and structural characteristics, independent of VSD influence (for an illustration see Fig. 1(b), left). Namely, beyond l_i (i.e., for $l > l_i$), the coupling interactions between a pore module and a radially succeeding voltage sensor cannot be neglected, driving the structural transition from the former towards the latter.

The irregular cylindrical surface emerging from the dense arrangement of $\mathcal{B}_i$ spheres along $\mathcal{P}$ is parameterized by:

$$\partial \bigcup_{\mathbf{p}} \mathcal{B}_i. \tag{7}$$

(7) serves as the characteristic cover for the PD, in the sense that it incorporates essential structural elements of the PD sub-architecture.

2.1.1.3 Mean-field conditioning of the atomic environment

The length scale

$$l_{i,\mathrm{b}} \coloneqq l_i - \xi \ln\left(\frac{2^{\nu} - 1}{\nu}\right) \quad [\text{Å}], \tag{8}$$

denotes the l-value for which the atom packing condition $n/K = 0.5$ is satisfied, meaning exactly half out of the K atoms are residing in $\mathcal{B}_{i,\mathrm{b}} = \mathcal{B}|_{l=l_{i,\mathrm{b}}}$. 'b' stands here for 'bound', since $l_{i,\mathrm{b}}$ imposes an upper and lower bound on l_i, depending on whether the attractive or repulsive interaction range prevails, i.e., whether $\nu < 1$ or $\nu > 1$, respectively. Accordingly, the negative and positive sign of $\ln\left(\frac{2^{\nu}-1}{\nu}\right)$ decides whether the length scale bound applies from below or from above, i.e., whether $l_{i,\mathrm{b}} > l_i$ or $l_{i,\mathrm{b}} < l_i$, respectively.

If the attractive interaction range equals the repelling interaction range (i.e., $\nu = 1$), then $\ln\left(\frac{2^{\nu}-1}{\nu}\right) = 0$ and $l_{i,\mathrm{b}} = l_i$, and (3) becomes isomorphic with the Fermi-Dirac (i.e., logistic) distribution. Hydrophobicity-driven attraction and hydrophilicity-driven repulsion are then delicately balanced, imposing a mean-field conditioning on the atomic environment around a pore point. Across an evolutionary time scale, $\nu = 1$ promotes isotropic space exploration in both the PD and the VSDs, as an equal number of atoms is distributed in the PD and the VSDs (i.e., half of the K atoms are found in $\mathcal{B}_i = \mathcal{B}_{i,\mathrm{b}}$ and the other half outside of it). Generally, this favors the emergence of globular-like molecular shapes, as demonstrated in Ref. [43], which, in the context of this study, is driven by a roughly equal and uniform allocation of masses inside and outside $\mathcal{B}_i$.

2.1.1.4 Unit mass fractal dimension

Expressing (2) in terms of (3), reveals that packing of atoms around a pore point satisfies the self-similarity relationship:

$$\frac{1}{K}\frac{\partial n}{\partial l} = \frac{\varepsilon}{\varepsilon_i \xi}\left(\frac{n}{K}\right)^{\nu+1} \quad [\text{Å}^{-1}], \tag{9}$$

which implies that:

$$\begin{cases} \int_{l_0}^{L} \mathrm{d}l\, \frac{\partial n}{\partial l} \to K \quad [\text{atom}] \quad (\text{a}) \\ \int_{l_0}^{L} \mathrm{d}l\, \left(\frac{n}{K}\right)^{\nu+1} \varepsilon \to \varepsilon_i \xi \quad [\text{kcal} \times \text{Å/atom}] \quad (\text{b}) \end{cases} \quad \text{for} \quad l_0 \to 0 \quad \text{and} \quad L \to \infty, \tag{10}$$

where $\mathcal{B}_L = \mathcal{B}|_{l=L}$ is the smallest sphere that accommodates all NaVCh structure atoms (SI Eq. (S4) and Fig. 1(b) right part).

(10) introduces two distinct yet interrelated constraints. Specifically, (10)(a) and (10)(b) ensure that even if the molecular radial size, s [Å] (SI Eq. (S5)), becomes infinitely large, the total number of atoms cannot exceed K, and the hydropathic interaction energy per atom converges to $\varepsilon_i \xi$ [kcal×Å/atom], respectively.

The intrinsic dimension (i.e., the unit mass fractal dimension [44]) of an n-cluster is continuously measured with:

$$d_f = \frac{\partial \ln(n)}{\partial \ln(l)} = \frac{l}{\xi\left(\left(\frac{\varepsilon}{\varepsilon_i}\right)^{-1} + \nu\right)}. \tag{11}$$

(11) describes how 'intensively' (or 'compactly', as discussed in Ref. [45]) n atoms fill space in $\mathcal{B}$. Generally, the fraction of empty space in $\mathcal{B}$ increases with decreasing d_f.

2.1.2 Statistical mechanical entropy

Let us now consider the following discretization of l:

$$l_\alpha = \bar{R} + \frac{\alpha}{N_\alpha} s \quad [\text{Å}], \quad \alpha = 1, 2, \ldots, N_\alpha \in \mathbb{Z}_{>0}, \tag{12}$$

where α is a scale index [36, 40]. N_α is the total number of scale iterations, determining the resolution of the discretization. $\bar{R}$ [Å] is the minimum Euclidean distance separating $\mathcal{A}$ from $\mathbf{p}$, while s [Å] remains the molecular radial size (first considered below (10) and given by SI Eq. (S5)). Note that $\mathcal{B}|_{l=\bar{R}}$ is empty of atoms.

The probability that an atom resides in $\mathcal{B}_\alpha = \mathcal{B}|_{l=l_\alpha}$ is given by:

$$p_\alpha = \frac{\left(1 + \nu \frac{\varepsilon_\alpha}{\varepsilon_i}\right)^{-1/\nu}}{Z_p} = \frac{n_\alpha / K}{Z_p}, \quad Z_p = \sum_\alpha n_\alpha / K, \tag{13}$$

where $(1 + \nu x)^{-1/\nu}$, $x \geq 0$, is identified with the survival function of a Pareto Type-II distribution [46]. Z_p is the partition sum guaranteeing that $\sum_\alpha p_\alpha = 1$.

The probability that $(\nu+1)$-atoms reside in $\mathcal{B}_\alpha$ is given by the Escort probability [47]:

$$\pi_\alpha = \frac{p_\alpha^{\nu+1}}{Z_\pi} = \frac{(n_\alpha/K)^{\nu+1}}{Z_\pi Z_p^{\nu+1}}, \quad Z_\pi = \sum_\alpha p_\alpha^{\nu+1}, \tag{14}$$

where Z_π is the corresponding partition sum.

π's biophysical importance stands out when we consider the following finite-size corrected version of (10(b)):

$$\begin{aligned} C\sum_\alpha \Delta l_\alpha \pi_\alpha \varepsilon_\alpha &= \varepsilon_i \overbrace{(\xi + c_s/s + c_\alpha/N_\alpha)}^{=:\Xi \quad [\text{Å}]} \quad [\text{kcal}\times\text{Å/atom}] \\ \Delta l_\alpha &= l_{\alpha+1} - l_\alpha \quad [\text{Å}], \quad C = Z_p^{\nu+1} Z_\pi, \end{aligned} \tag{15}$$

where Ξ is a finite-size corrected version of ξ. c_s [Å^2] and c_α [Å] account for interfacial and radial geometric distortions due to s's and N_α's finiteness, respectively. Note that Δl_α is constant. C is a normalization factor whose specific value is irrelevant.

Crucially, for $N_\alpha \to \infty$, the constraint (10)(b) can be recovered:

$$C\sum_\alpha \Delta l_\alpha \pi_\alpha \varepsilon_\alpha \overset{N_\alpha\to\infty}{\to} \int_{l_0}^{L} \mathrm{d}l \left(\frac{n}{K}\right)^{\nu+1} \varepsilon \overset{l_0\to 0, L\to\infty}{\to} \varepsilon_i \xi \quad [\text{kcal}\times\text{Å/atom}]. \tag{16}$$

An entropy functional whose maximization explicitly incorporates (13)-(15) is given by:

$$\mathcal{S} = k\frac{1-\sum_\alpha p_\alpha^q}{q-1} \quad [\text{kcal}/(\Theta\times\text{atom})], \quad q = 1+\nu, \tag{17}$$

where q is known as the nonextensivity index [48, 49] and k [kcal/($\Theta\times$ atom)] is the equivalent of Boltzman's constant. Note that [Θ] is the physical unit at which the temperature of the current pore point is measured (SI 1.4).

From an information theoretical viewpoint, $\mathcal{S}$ can be trivially interpreted as the average amount of surprise (or 'unexpectedness') associated with determining how many atoms reside in $\mathcal{B}_\alpha$.

From an evolutionary perspective, $\mathcal{S}$ illustrates how the competition between hydrophobicity-driven attraction and hydrophilicity-driven repulsion determines the information content of the atomic environment around a pore point.

Local scarcity of hydrophilicity can reduce associated repulsion effects (i.e., $\zeta\to 0$), promoting the formation of a hydrophobic core. $\mathcal{S}$ approaches then the following upper bound:

$$\mathcal{S}_{q\to 1} = -k\sum_\alpha p_{\alpha,\nu\to 0}\ln(p_{\alpha,\nu\to 0}) \quad [\text{kcal}/(\Theta\times\text{atom})], \tag{18}$$

where $p_{\alpha,\zeta\to 0} \propto \frac{n_{\zeta\to 0}}{K}$ is a common exponential (SI Eqs. (S6),(S10)).

Maximizing $\mathcal{S}$ toward $\mathcal{S}_{q\to 1}$ expands the NaVCh configuration space volume, since the amount of water required to solvate a hydrophobic n-cluster packed in $\mathcal{B}$ is typically less than that required for solvating n individually dispersed hydrophobic atoms. The remaining, non-solvating water molecules can then uncoordinately engage in hydrogen bond interactions with NaVCh constituents on $\partial\mathcal{B}$, effectively increasing the NaVCh

configuration DoF as water structuring effects diminish. Simultaneously, we notice that the intrinsic dimension of the PD,

$$d_f|_{l_i} \overset{(11)}{=} \frac{l_i}{\xi q}, \tag{19}$$

is decreased, since

$$\frac{l_i}{\xi q} \overset{(4)}{=} \frac{1}{q}\left(\ln\left(\frac{(K/n_0)^{q-1}-1}{q-1}\right)+\frac{l_0}{\xi}\right) \overset{\zeta\to 0}{\to} \ln(\ln(K/n_0))+l_0/\xi \tag{20}$$

is monotonically and positively correlated with ζ for fixed parameter values $\{l_0, n_0, K, \xi\}$ (SI Fig. S1 inset).

(20) illustrates how the PD sub-architecture can mitigate the potentially disordering effect of excess entropy by lowering its intrinsic dimension.

2.1.3 Thermostability

2.1.3.1 Empirical measure

The thermostability of an n-cluster is probed with the hydropathic moments toolbox [32–37, 40]:

$$\int_{\mathcal{B}} \mathrm{d}\vec{\mathbf{r}} \sum_{\mathbf{a}\in\mathcal{A}} \delta(\vec{\mathbf{r}}-\vec{\mathbf{r}}_{\mathrm{a}})(\vec{\mathbf{r}}_{\mathrm{a}})^j w_{\mathrm{a},\iota} = \begin{cases} h_j, \ j=2k, \ k\in\mathbb{Z}_{\geq 0} \\ \vec{\boldsymbol{h}}_j = \overbrace{h_{j,\parallel}\hat{\mathbf{e}}_\parallel}^{=\vec{\boldsymbol{h}}_{j,\parallel}} + \overbrace{h_{j,\perp}\hat{\mathbf{e}}_\perp}^{=\vec{\boldsymbol{h}}_{j,\perp}}, \ j=2k+1 \end{cases} \quad [\mathrm{kcal}\times\mathrm{\AA}^j], \tag{21}$$

where $\delta(\cdot)$ is a generally defined Dirac delta function, and $w_{\mathrm{a},\iota} = w_{\mathrm{a}} + \iota$ [kcal], $\iota \in \mathcal{N}(0, \sigma_\iota = 1\mathrm{e}-03)$, are noise-perturbed hydropathic weights originating from the Kapcha-&-Rossky atomic hydropathic scale [50]. Noise accounts here for randomly occurring water density fluctuations.

Note that $\vec{\boldsymbol{h}}_{j,\parallel}$ and $\vec{\boldsymbol{h}}_{j,\perp}$ act parallel and perpendicularly, respectively, to the membrane, surface with $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ being the corresponding unit vectors (SI S1.2.3). Conventionally, $\hat{\mathbf{e}}_\perp$ points towards the extracellular side (ES).

The subscript 'j' indicates the moment order determining the dimension of the space within which water-mediated interactions are probed. j-parity determines whether we probe inertia ($j=2k$) or conductivity ($j=2k+1$) constraints, concerning rotational and translational DoF expressed around and along the pore point path, $\mathcal{P}$, as detailed in SI 1.5.1-2, respectively.

2.1.3.2 Theoretical model

Normalizing h_j over n yields the n-cluster hydropathic density [32–35] whose oscillatory behavior is described by:

$$\begin{aligned} &\frac{h_j}{n} = A_j \cos(f_j) \quad [\text{kcal} \times \text{Å}^j/\text{atom}] \\ &\text{with} \quad A_j = (1+\kappa_j \xi/s)\varepsilon l^j \overset{(4)}{=} (1+\kappa_j \xi/s)\varepsilon_0 \exp(-l/\xi) l^j, \end{aligned} \tag{22}$$

where f_j determines the sign changing behavior of h_j and A_j envelopes the underlying wave packet. The factor $(1+\kappa_j \xi/s) > 0$ introduces a finite-size correction with $\kappa_j \xi$ [Å] being the adsorption bond length. κ_j is dimensionless and purely phenomenological, as it can only be deduced by fitting $A_j \cos(f_j)$ to experimental h_j/n-traces. It generally describes the degree of adhesion of atoms on $\partial\mathcal{B}$ in response to interfacial tensions.

Notably, A_j preserves the general form that the hydropathic energy has been assigned in Refs. [51–53]. Moreover, consistent with Eq. (4), for $j=0$, we verify that $A_0 \overset{s\to\infty}{\to} \varepsilon \geq |h_0|/n$, illustrating that ε serves as an upper bound for the absolute hydropathic energy of an individual atom.

2.1.3.3 Scaling of hydropathic dipole field amplitude

Hydropathic energy wave packet self-modulation effects persisting over sufficiently large l-intervals covering in the PD and the VSDs are described by [36, 37]:

$$|h_j| \leq nA_j \propto \begin{cases} l^{\eta_{j,<}}, & l \in [l_t, l_i) \\ l^{\eta_{j,>}}, & l \in (l_i, l_q] \end{cases}, \tag{23}$$

where l_t [Å] and l_q [Å] denote the l-values for which the n-curvature (Eq. (5)) is positively maximized and negatively minimized, respectively.

$(\eta_{j,<}, \eta_{j,>}) \in \mathbb{R}^2$ are the corresponding scaling exponents. $|\eta_{j,<}|$ and $|\eta_{j,>}|$ set an upper bound on how 'intensively' j-th-order dipole-dipole hydropathic interactions can fill space in the PD and the VSDs, respectively.

The sign of η_j indicates the direction of hydropathic interaction network intensification, either inwards (i.e., $\eta_j < 0$) or outwards (i.e., $\eta_j > 0$).

2.1.4 Mutational robustness

2.1.4.1 Decomposability

Sign-changes of h_j mark n-clusters of diminished thermostability. Mechanofunctional properties of these n-clusters are expected to be sensitive to perturbations.

To understand how the multiscale competition between hydrophobic attraction and hydrophilic repulsion can lead to varying thermostability and emergent

mechanofunctional properties, we consider the following decomposition ansatz:

$$\begin{aligned}
&h_j = h_{j,+} + h_{j,-} \quad \text{for} \quad l > l_{\text{cutoff}} \\
&\text{with} \quad h_{j,+} > 0, \quad h_{j,-} < 0, \quad n = n_+ + n_-, \\
&\text{and} \\
&h_{j,+} = h_{j,\text{philic}}, \quad h_{j,-} = h_{j,\text{phobic}}, \quad n_+ = n_{\text{philic}}, \quad n_- = n_{\text{phobic}} \quad \text{(case A)}, \\
&\text{or} \\
&h_{j,+} = h_{j,\text{phobic}}, \quad h_{j,-} = h_{j,\text{philic}}, \quad n_+ = n_{\text{phobic}}, \quad n_- = n_{\text{philic}} \quad \text{(case B)},
\end{aligned} \tag{24}$$

where l_{cutoff} is the ansatz cutoff scale.

We verify that for vanishingly small noise, i.e., $\sigma_\iota \to 0$ (see (22)), and $j = 2k$, case A holds. Namely, $h_{j,+} = h_{j,\text{philic}}$ and $h_{j,-} = h_{j,\text{phobic}}$ are the contributions of only hydrophilic and hydrophobic atoms, respectively, with $n_+ = n_{\text{philic}}$, $n_- = n_{\text{phobic}}$ and $l_{\text{cutoff}} \to l_0$.

Whether case A or B holds for the pair $\{h_{2k+1,\perp,\text{philic}}, h_{2k+1,\perp,\text{phobic}}\}$ requires computational investigation, as there is no general argument supporting this operation.

Since $h_{j,+}$ and $h_{j,-}$ are cumulative, and thus slowly changing, sign-preserving functions over the l-range, we can straightforwardly compute their continuously varying scaling exponents:

$$\psi_{j,+} = \frac{\partial \ln(h_{j,+})}{\partial \ln(l)} = \frac{\partial h_{j,+}}{\partial l} \frac{l}{h_{j,+}} \tag{25}$$

and

$$\psi_{j,-} = \frac{\partial \ln(-h_{j,-})}{\partial \ln(l)} = \frac{\partial h_{j,-}}{\partial l} \frac{l}{h_{j,-}}, \tag{26}$$

respectively.

The physical meaning of $\psi_{j,+}$ and $\psi_{j,-}$ complements that of η. Namely, $|\psi_{j,+}|$ and $|\psi_{j,-}|$ account for the intensification of the j-th-order dipole-dipole interaction network stabilizing the subgroups of n_+ and n_- atoms, respectively, along the radial direction.

The signs of $\psi_{j,+}$ and $\psi_{j,-}$ indicate the direction of interaction network intensification, being either inwards (if $\psi_j < 0$) or outwards (if $\psi_j > 0$).

2.1.4.2 Logarithmic composite susceptibility

The difference of $\psi_{j,-}$ from $\psi_{j,+}$,

$$\Delta\psi_j = \psi_{j,+} - \psi_{j,-}, \tag{27}$$

illustrates how the interaction networks, stabilizing the subgroups of n_+ and n_- atoms, 'compete' to fill space in $\mathcal{B}$.

If $\Delta\psi_j = 0$, the two networks are in a state of balance, where perturbations affecting the n_+ and n_- subgroups tend to induce oppositely directed $\partial\mathcal{B}$-fluctuations, effectively canceling each other out. On the other hand, $\Delta\psi_j \neq 0$ implies an interaction network imbalance, determining also the prevailing direction of a $\partial\mathcal{B}$-fluctuation. Accordingly, the sign of $\Delta\psi_j$ informs about the direction of perturbation-induced $\partial\mathcal{B}$-responses.

Dividing $\Delta\psi_j$ by l defines the following size-normalized measure of interaction network imbalance:

$$\mathcal{I}_j \coloneqq \frac{\Delta\psi_j}{l} \quad [\text{Å}^{-1}], \tag{28}$$

assuming that the interaction network imbalance is uniformly distributed over the radial extent of an n-cluster.

Integrating $\mathcal{I}_j$ over an l-range, yields the logarithm of the j-th-order composite susceptibility, i.e.,:

$$\begin{aligned} \phi_j &\coloneqq \ln(\chi_j) = \int_{l_0}^{L} \mathrm{d}l\, \mathcal{I}_j \\ \text{with} \quad \chi_j &\coloneqq -\frac{h_{j,+}}{h_{j,-}}, \end{aligned} \tag{29}$$

with χ_j being identified with the j-th-order composite susceptibility. Generally, χ_j can be interpreted as a descriptor of how internal hydrophobicity-generated fields couple with internal hydrophilicity-generated fields and vice versa.

Treating $-\frac{h_{j,+}}{h_{j,-}}$ as an effective temperature ratio reveals that $|\phi_j|$ is the equivalent of a thermostability geodesic (i.e., shortest distance measure) (SI S1.6.1). Simply put, $|\phi_j|$ quantifies how 'far' an n-cluster's thermostability state is from a state of minimal thermostability (or, maximal sensitivity to perturbations).

We emphasize that although both $|\phi_j|$ and $|h_j|$ qualify as thermostability 'distance' measures (since $h_j = 0 \overset{(24)}{\Longrightarrow} \phi_j = 0$), the geodesic property applies exclusively to $|\phi_j|$. In fact, while η_j (Eq. (23)) provides information at an envelope level, $\Delta\psi_j$ (27) reflects the actual sculpting (i.e., the radial interplay of successive peaks and valleys) of the hydropathic energy wave packet.

On the ground of consistency, $\mathcal{I}_j$ is considered to provide equivalent information with ϕ_j but at an interface level. Namely, $|\mathcal{I}_j|$ measures how 'far' an ∂n-shell's thermostability state is from a state of minimal thermostability. We term $\mathcal{I}_j$ the 'interfacial coupling strength'.

2.1.5 Renormalizability

The RG flow equation has the general form [54]:

$$\frac{\partial \mathcal{O}}{\partial \ln(\tau)} = \Lambda(\mathcal{O}), \quad \tau = \begin{cases} \frac{\varepsilon}{\varepsilon_i}, & \partial\mathcal{B}_i \mapsto \partial\mathcal{B} \text{ (perturbed from } \partial\mathcal{B}_i \text{ onto } \partial\mathcal{B}) \\ \frac{\varepsilon_i}{\varepsilon}, & \partial\mathcal{B} \mapsto \partial\mathcal{B}_i \text{ (perturbed from } \partial\mathcal{B} \text{ onto } \partial\mathcal{B}_i) \end{cases}, \tag{30}$$

where $\mathcal{O} \coloneqq \mathcal{O}(\mathbf{p}, l)$ is a NaVCh structure log-observable (referred to as the *effective coupling* [54]). τ is the normalized hydropathic energy *scale* [54], representing the relative temperature change in response to PD/VSD interface fluctuations (SI Eq. (S13)). Λ describes the dependence of $\mathcal{O}$ on τ and is the equivalent of the *beta function* [54]. It represents a continuously varying scaling exponent attaining its critical value precisely at $\tau = 1$ (SI 1.7).

The conceptual basis upon which (30) is built, is illustrated in Fig. 1(b) along with its caption.

From (30), we straightforwardly obtain the renormalizability relationship:

$$\frac{\partial \mathcal{O}}{\partial \ln(l)} = \pm \frac{l}{\xi} \Lambda(\mathcal{O}) \tag{31}$$

Eq. (31) establishes that changes in $\mathcal{O}$ over a logarithmic radial range (on the left hand side of (31)) can be equivalently obtained by considering changes in $\mathcal{O}$ driven by PD-subarchitecture deformations reflected upon PD/VSD interface fluctuations (on the right hand side of (31)).

2.2 Computational Procedures

2.2.1 Structure preparation

NaVCh structures are 'cleaned', protonated, and their orientation is fixed, following the procedures outlined in SI S1.2.

2.2.2 Curve fitting and model selection

Candidate n-models were fitted to the dimensionless empirical cumulative atom number:

$$\overbrace{\int_{\mathcal{B}} \mathrm{d}\vec{\boldsymbol{r}} \sum_{\mathbf{a} \in \mathcal{A}} \delta(\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}_{\mathrm{a}})}^{=N} / N_{\mathrm{total}}, \tag{32}$$

where $N_{\mathrm{total}} = \mathrm{card}(\mathcal{A})$ is the total number of atoms found in a 'clean', protonated NaVCh structure. Algorithmic implementation details are enclosed in SI 1.11.

3 Results

3.1 NaVCh spatial organization

We compiled a dataset comprising 71 prokaryotic NaVCh structures (subtypes: NaChBac, NaVAb, NaVMs, NaVAe, NaVRh) and 50 eukaryotic structures (subtypes: NaVPas, NaVEe1, NaVEh, NaV1-8) of sufficient resolution (SI S1.1).

3.1.1 Geometry

3.1.1.1 Inflection point universality

We report an excellent agreement between the empirical cumulative atom number, N (Eq. (32)), and its theoretical counterpart, n (Eq. (3)), across all 121 structures (Figs. 2(a),(g), SI Figs. S5-S10(a), S12-S22(a)), supported by small mean absolute fitting errors (SI S2.1). P-values remained inconclusive, as expected, due to the coarse-grained nature of our model and their sensitivity to noise (SI S2.1).

To appreciate the statistical significance of this finding, we emphasize that our fitting algorithm had to converge approximately $121 \times 690 = 83{,}490$ times, where 121 is the total number of NaVCh structures considered and 690 gives the average number

of investigated pore points per NaVCh structure, separated by a sampling distance of 0.1 Å (SI Tab. S2).

These results indicate that NaVCh atom packing physics can be compressed into the set of parameters $\{\zeta, \xi, K\}$ (Eqs. (2),(3)), in a pore-point-specific manner. Key to our understanding is that ξ and ζ are indirect measures of molecular attraction and repulsion, influencing molecular compression and expansion, respectively. Accordingly, the NaVCh molecule is conceptualized as a compressible, fluid-like material, whose mechanofunctional properties vary along its principal pore axis, approximated with the pore point path $\mathcal{P}$. In turn, this implies that the atomic environment around each pore point requires a distinct statistical mechanical description or, equivalently, admits a local temperature linked to the PD hydropathic energy levels (SI S1.4).

The formation of interfacial geometries in compressible materials requires the spontaneous cancellation of interfacial tensions, manifested as minima in the underlying interfacial free energy [55], a principle that also applies to the spatial organization of NaVChs. Structurally, this prompts the emergence of the PD/VSD interface, denoted as $\partial\mathcal{B}_i$ (Eq. (6)), mediating two qualitatively different – in terms of their thermostability character – atomic environments. This insight enables the structural annotation of atoms, determining whether they belong to the PD or the VSDs, as n/K serves as an order parameter distinguishing between two coexisting material phases (Figs. 2(b),(h), SI Figs. S5-S10(b), S12-S22(b)).

Under mean-field conditions (when $\xi = \zeta$, see below Eq. (8)), n/K admits the standard critical exponent, $\beta = 0.5$ (SI Eqs. (S15),(S16)). Moreover, it can be readily shown that β is related to other standard critical exponents through a generalized version of Widom's scaling law – whose standard form can be recovered when the characteristic renormalization scale matches the correlation length (represented by l_i and ξ, respectively) (SI S1.6.4).

3.1.1.2 Pore domain intrinsic dimensionality

Dimensionality analysis reveals that prokaryotic and eukaryotic PDs exhibit a preference for flatness, as their intrinsic dimensions (Eq. (19)) are, on average, values only slightly greater than 2 (Figs. 3(a),(b)).

By tuning the intrinsic dimension of the PD near 2, PD sub-architecture is effectively mapped onto consecutive membrane-parallel planes, with roughness reciprocally related to $|d_f|_{l_i} - 2|$, favoring a smooth cylindrical shape. As inferred from Eq. (20), this reflects a self-regulation mechanism that ensures the PD maintains its functional and structural integrity, despite the excess entropy generated by the tightening of the atomic environment, driven by hydrophilic deficiencies (or equivalently, a surplus of hydrophobicities) causing $\zeta \to 0$. Assuming ξ is fixed at a value well above zero, decreasing ζ effectively reduces interatomic distances, which, in turn, counteracts the expansion of the configuration space introduced by excess entropy. Hence, by lowering its intrinsic dimension, the PD increases its spatial organization efficiency, as the likelihood for PD-forming atoms to adopt a nearly water-free, planar configuration increases.

Eukaryotic PDs exhibit greater efficiency in spatial organization, as evidenced by their smaller mean and standard deviation values for intrinsic dimensions compared

to prokaryotic PDs (see Fig. 3 caption). This finding is consistent with the general trends reported in Ref. [56], where lower intrinsic dimensions were found to correlate with higher organismal complexity.

3.1.2 Energy and entropy

3.1.2.1 An exponentially decaying hydrophobic 'force' stabilizes the structure

As illustrated in Figs. 2(c),(i), a decrease in n-cluster size implies an increase in its average hydrophobicity, $h_{0,-}/N$ (Eqs. (22),(24)). Specifically, $h_{0,-}/N$ converges towards the value -0.5 kcal for $l \to l_0$ (decreasing scale index), denoting the hydropathic score that an individual hydrophobic atom is assigned with based on the Kapcha&Rossky hydropathic scale [50]. In contrast, the n-cluster average hydrophilicity, $h_{0,+}/N$ (Eqs. (22),(24)), tends to vanish for $l \to l_0$ (Figs. 2(c),(i)).

This trend showcases the entropically favorable arrangement where hydrophobicities 'hide' inside $\mathcal{B}$ rather than expose on its surface, $\partial\mathcal{B}$. As hydrophobic constituents fill the available space in $\mathcal{B}$, hydrophilic ones are displaced outwards, preferentially settling near $\partial\mathcal{B}$. This redistribution creates an initial surplus of hydrophobic energy available to pore-lining constituents, as evidenced by the peak in the experimentally observed absolute hydropathic energy, $|h_0|/N$, when $l \to l_0$ (Figs. 2(c),(i), SI Figs. S5-S10(c), S12-S22(c)).

However, as the n-cluster size grows (increasing scale index), the initial hydrophobic energy surplus is exponentially depleted at a rate governed by ξ, converted into stabilizing interactions that bind the nested $\mathcal{B}$'s together. Intriguingly, this would imply that interfacial adsorption effects do not substantially distort atom packing conditions (Eq. (22)). In turn, this would imply that NaVChs have evolved near a thermodynamic state where they mimic the spatial organization traits of larger, bulkier systems. This maximizes NaVCh mechanofunctional efficiency by allowing perturbations to be processed nearly adiabatically (SI S1.6.1). Our analysis supports this intuition, as shown by the good agreement between $|h_0|/N$ and ε [kcal/atoms] (Figs. 2(c),(i), SI Figs. 5-10(c), 12-22(c)).

Evidently, if the repulsive interaction range, ζ, grows too large, structural integrity could be compromised due to excessive interatomic spacing. To avoid that, the histogram of ζ is located closer to zero compared to that of ξ (Figs. 3(c),(d)). Also, in contrast to ξ, ζ accepts a smaller mean value for eukaryotic NaVChs compared to prokaryotic NaVChs (Fig. 3 caption), supporting that eukaryotic NaVChs, despite being more diverse – both structurally and functionally –, are spatially organized in a more efficient manner.

3.1.2.2 Entropy constraints along the pore

Moving from the IS to the ES along $\mathcal{P}$, we observe that the statistical mechanical entropy, $\mathcal{S}$ (Eq. (17)), is smoothly up-regulated in both prokaryotes and eukaryotes (Figs. 2(d),(j), SI Figs. S5-S10(d), S12-S22(d)), as a result of the general trend that ζ surpasses ξ at the IS. Atomic environments at the IS and ES are thus subject to qualitatively entropic constraints, likely also experiencing asymmetric perturbation responses. To sustain such a delicate metastable equilibrium, even instantaneously,

requires the contribution of external forces, since otherwise the structural integrity of the IS could be compromised. The observed nonextensivity, where the entropy is largely indifferent to changes in the radial molecular size while being slaved to the interaction range ratio, $\nu = \frac{\zeta}{\xi}$ (Eq. (2), Figs. 2(f),(l) and SI Figs. S5-S10(f), S12-S22(f)), is likely the key property ensuring that the structure can quickly react to environmental changes and visit unconventional metastable states.

$\mathcal{S}$ maximizes in response to reduction of repulsive effects ($\zeta \to 0$), indicating a tightly organized, prevalently hydrophobic atomic environment (Eqs. (17),(20)). This is typically the case in the pore center, where the equivalent of a hydrophobic core exists, namely, a prevalently hydrophobic central cavity (e.g., see Fig. 1 from [40]). To resist structural disorder caused by excessive, hydrophobicity-induced configurational freedom, the PD reduces its characteristic size in order to decrease its intrinsic dimension (SI S2.2), which, in turn, guarantees greater efficiency in spatial organization.

3.1.3 The PD/VSD structural transition as an order/disorder phase transition

Implications for the NaVCh functional architecture arising from the cancellation of interfacial tensions in the vicinity of the PD/VSD interface are best summarized by the membrane-perpendicular first-order hydropathic dipole field (HDF) amplitude, $h_{1,\perp}$ [kcal×Å] (as explained in SI 1.5.2).

The sign of $h_{1,\perp}|_{l_i}$ indicates whether the $h_{1,\perp}\hat{\mathbf{e}}_{\perp}$-field induced by the PD at the current pore point is oriented extracellularly (↑) or intracellularly (↓) (Fig. 2 caption and SI 2.4.2, 2.5.2).

As shown in Figs. 2(e),(k) and SI Figs. S5-S10(e), S12-S22(e), distinguishing between $h_{1,\perp,\uparrow}$ and $h_{1,\perp,\downarrow}$ highlights the critical [30, 31] nature of the inflection point.

Specifically, $h_{1,\perp,\downarrow}$ and $h_{1,\perp,\uparrow}$ exhibit clear negative and positive peaks, respectively, for $l \approx l_i$, indicating that $|h_{1,\perp}|$ is globally maximized. Implying that the rate of change of $h_{1,\perp}$ has stagnated, causing the interfacial free energy pair, $\{\partial h_{1,\perp,\uparrow}/\partial l, \partial h_{1,\perp,\downarrow}/\partial l\}$ [kcal], to nearly vanish. This resembles a second-order phase transition, though smeared out by NaVCh finiteness.

Regardless of its orientation, $|h_{1,\perp}|$ increments in a power-law fashion in both prokaryotic and eukaryotic PDs. The corresponding scaling exponents, $\eta_{1,\perp,<}$ (Eq. (23)), are narrowly distributed with 3.2 ± 0.6 and 3.5 ± 1.3 (see Figs. 3(e) and (f), respectively). High Pearson coefficient values verify that these observations reflect a genuine self-similar increment behavior (e.g., see Fig. 3(c) from [36]). Microscopically, this necessitates that dipole-dipole hydropathic interaction networking radial atom neighbors are nonrandomly aligned [36], as random alignment would make self-similar incrementation of their cumulative index an extraordinary evolutionary accident.

The narrowness of the $\eta_{1,\perp,<}$-distribution (Figs. 3(e),(f)), further necessitates a highly specialized nature for the water-mediated interactions established between the PD and the ion/water pore mixture. The PD operates under a narrow-banded dehydration protocol, through which ion selectivity can naturally emerge, as the radial alignment of dipoles around a pore point guarantees that water reorganization energies optimally dissipate into the atomic structure. Perturbation amplification is strictly unidirectional: perturbation amplitude attenuates inwards (towards the pore-lining

interface, $\partial\mathcal{B}_0$), while amplifies outwards (towards the PD/VSD interface), irrespective of the current pore point. This implies the existence of a directed (or ordered) allosteric network inside the PD, where any pair of perturbed pore-lining constituents can potentially influence the same distant neighbor, unlike the other way around.

Once the characteristic PD size is surpassed (i.e., for $l > l_i$), $h_{1,\perp,\downarrow}$ bends upwards, while $h_{1,\perp,\uparrow}$ bends downwards, thereafter exhibiting transient behaviors that cannot be described by a single scaling rule, as can be inferred by the broad, zero-centered $\eta_{1,\perp,>}$-exponent distribution (Figs. 3(e),(f)). $|h_{1,\perp}|$ may thus increment, decrement, or stagnate beyond the PD (note how percentile clouds increase for $l > l_i$ in Figs. 2(e),(k) and SI Figs. S5-S10(e), S12-S22(e)). Given the high Pearson coefficient values associated with distributional $\eta_{1,\perp,>}$-patterns on both sides of zero (Figs. 3(e),(f)), the observed diversification of the the scaling behavior of $|h_{1,\perp}|$ beyond the PD, supports that allosteric pathways coupling the PD to the VSDs are bidirectional: perturbations can be either attenuated or amplified from a VSD towards the PD/VSD interface in a pore-point-specific manner (SI 2.4.3, 2.5.3). This implies the multidirected (or disordered) nature of the allosteric network coupling the PD with the VSDs.

3.2 NaV1.7 mutational robustness in the context of human pain disease

Let us assume that the NaV1.7 molecule exhibits a uniform mutational robustness profile, implying that its ability to tolerate mutations is approximately the same across all of its constituent parts. Given that the PD/VSD interface is the most occupied molecular surface, it then follows that it is also a mutation hotspot, i.e., a geometric site where the likelihood of residue mutation is highest.

To scrutinize the validity of the above assumption, we examine mutation clustering in the NaV1.7 structure with the highest available resolution (PDB code: 7w9k). We consider all *SCN9A*-gene-related mutations found in gnomAD [23] and ClinVar [57] databases, yielding $N_{\text{mut}} = 1705$ events spanning three main categories: pathogenic (6.5%), non-pathogenic (11.3%), and variants of unclassified (or unknown) significance (VUS) (82.2%). The set of pathogenic mutations contains the subset of pain-disease-associated mutations (2.7%) and a substantial portion of pain-disease-unrelated pathogenic mutations (3.8%). Pain-disease-associated mutations are categorized as either gain-of-function (GoF) (2.1%) or loss-of-function (LoF) (0.5%) (SI S1.9). Non-pathogenic mutations comprise two subsets: benign and neutral, with the latter consisting of carefully selected negative controls for human pain disease, sourced from Refs. [40, 58].

To probe mutation clustering relatively to the PD/VSD interface, we consider the dataset, $\{l_{\text{mut}} - l_i\}$, where $l_{\text{mut}=1,2,\ldots,N_{\text{mut}}}$ is the Euclidean distance between a pore point and the geometric center of the residue undergoing mutation (referred to as 'structural location'). The triple vanishing of mean, mode, and median of $\{l_{\text{mut}} - l_i\}$ indicates that the likelihood for observing a mutation maximizes for $l = l_i$ and, moreover, that the underlying mutation distribution is roughly bell-shaped (Fig. (4) insets), implying the uniformity of the mutational robustness profile.

The uniform mutational robustness profile hypothesis for the NaV1.7 molecule is rejected across $\mathcal{P}$, except at the AG pore region (Fig. 4(a)). While the mean and

median of $\{l_{\text{mut}} - l_i\}$ closely agree, and the mode oscillates around them within reasonable bounds (Fig. 4(a)), their deviation from zero reflects random genetic fluctuations governed by a flipping-coin maximum entropy principle.

Accordingly, the probabilities of a mutation occurring inside or outside $\mathcal{B}$ are given by $P_{\text{in}} = N/N_{\text{total}} \approx n/n_L$ and $P_{\text{out}} = 1 - P_{\text{in}}$, respectively. Maximizing the flipping-coin entropy, $H \coloneqq -P_{\text{in}} \log_2(P_{\text{in}}) - P_{\text{out}} \log_2(P_{\text{out}})$, requires that $P_{\text{in}} = P_{\text{out}} = n_*/n_L = 0.5$ with l_* denoting the l-value for which $n_*/n_L = 0.5$ is satisfied.

As shown in Fig. 4(a), the line $l_* - l_i$ coincides with the $\{l_{\text{mut}} - l_i\}$-median, verifying that mutation clustering in the NaV1.7 is predominantly shaped by randomness. Across an evolutionary time-scale, this results in a bell-shaped mutation distribution approximately centered at $l = l_* \overset{n_L \approx K}{\approx} l_{i,\text{b}}$ (Eq. (8)), ensuring that molecular diversity is explored inside and outside $\mathcal{B}_*$ in an unbiased manner. Because the range of attractive interactions prevails over the range of repulsive interactions (SI Fig. S23), the ∂B_*-interface encapsulates the PD/VSD interface (Fig. 5(a)), acting as a buffer, dissipating mutation-induced perturbations and mitigating their impact on the PD.

This is not the case at the AG pore region, where all the statistical indices considered above tend to vanish (Fig. 4(a)), promoting a uniform mutational robustness profile as ∂B_* roughly coincides with the PD/VSD interface (Fig. 5(a)). Although the PD loses its protective buffer, it gains evolutionary momentum, allowing for greater exploration of molecular diversity. Also, mutations around the AG pore region are equally weighing on the PD and VSDs, preventing either from being disproportionately stressed.

Collapsing $\{l_{\text{mut}} - l_i\}$ into a histogram yields a roughly symmetric bell shape, well-described by the normal distribution, 4.46 ± 18.04 (inset in Fig. 4(a)). Excluding VUS, shifts the distribution to the left, with -0.47 ± 15.18 (Fig. 4(b)). Pathogenic and non-pathogenic mutation histograms admit -2.89 ± 14.9 and 0.9 ± 15.1 indicating their slight preference for the PD and the VSDs, respectively (Figs. 4(c),(d)). The pain-disease-associated mutation histogram is further left-shifted with -6.2 ± 13.0 (Fig. 4(e)), indicating a strong preference for the PD.

Distinguishing between GoF and LoF mutations reveals opposite statistical trends (Figs. 4(f),(g)). GoF mutations minimize their distance from the AG (Fig. 4(f)). On the other hand, LoF mutations converge towards the ES of the SF, where the extracellular funnel is formed (Fig. 4(g)). The molecular basis for GoF and LoF pain phenotypes can thus be underscored by qualitatively different perturbation modes. Positive $\eta_{1,\perp,>}$-exponents at the AG pore region favor GoF-mutation-triggered perturbations to amplify and propagate further into the VSDs (SI Fig. S23). On the other hand, negative $\eta_{1,\perp,>}$-exponents at the ES side of pore disfavor further amplification of LoF-mutation-triggered perturbations into the VSDs (SI Fig. S23), potentially causing them to be locally absorbed with detrimental impact on the PD.

3.2.1 Violation of inertia and conductivity constraints underpins human pain disease at the molecular level

We hypothesize that pain-disease-associated mutations exploit interaction network imbalances at an interface level to perturb the n-cluster beneath.

To explore this hypothesis for the NaV1.7, we utilize the Decomposition ansatz (24) (SI S2.3), to derive the logarithmic composite susceptibility, ϕ_j (Eq. (29)), and combine it with the interfacial coupling strength, $\mathcal{I}_j$ (Eq. (28)), in the form of the ratio $\frac{\mathcal{I}_j}{\phi_j \neq 0}$. $\frac{\mathcal{I}_j}{\phi_j \neq 0}$ establishes an upper bound for the number of available perturbation modes encoded into a structural location (SI Eq. (S24)). Simply put, $\frac{\mathcal{I}_j}{\phi_j \neq 0}$ measures how many ways there are to destabilize an n-cluster, thus serving as a perturbation potential (i.e., as a measure of the sandpile slope) associated with the structural location under scrutiny.

A mechanical analogy rationalizing $\frac{\mathcal{I}}{\phi_j \neq 0}$ is that of a maximum force or torque principle. When tightening a bolt, applying force at a stable far-end grip, where the inertia is minimum, maximizes torque. Analogously, residues distributed over strongly coupled interfaces covering clusters of vanishingly small inertia (probed with $|\phi_{2k}|$, $k = 0, 1, \ldots, 5$) are ideal mutagenesis sites; they can effortlessly perturb n-cluster rotation dynamics via allosteric shortcuts. The same idea applies to the perturbation of n-cluster conductivity (probed with $|\phi_{2k+1,\perp}|$, $k = 0, 1, \ldots, 5$) dynamics.

Contrary to non-pathogenic mutations, pain-disease associated mutations prefer interfaces covering n-clusters whose inertia decreases, as we approach the CC pore region from left and right, and, additionally, at the SF and the AG pore regions (Fig. 5(b)). Specifically, the inertia of the n-clusters whose surface acts as a hotspot for GoF and LoF mutations vanishes twice and four times, respectively, at roughly symmetric locations along $\mathcal{P}$. The arrangement and number of inertia zero-crossings may detrimentally perturb the NaV1.7 functional architecture via excess rotational DoF. Whether this leads to an increased or decreased pore-open probability and, consequently, results in a GoF- or LoF-like electrophysiological signature, is likely thresholded by the number of inertia zero-crossings, determining the tolerable excess rotational DoF before the gating cycle collapses. Targeting binding sites in the CC with anchoring molecules, e.g., local anesthetics, can thus offer a reasonable option for restoring inertia and mitigating the effects of the perturbation.

Additionally, GoF and LoF mutations prefer interfaces covering n-clusters whose conductivity is diminished past the AG pore region toward the IS (Fig. 5(c)). Also, GoF mutations cause a zero-crossing in conductivity at the ES of the SF pore region, while LoF mutations reduce conductivity at the IS of the SF pore region, without inducing a definitive sign change (Fig. 5(c)). Pain-disease-associated mutations may thus maximally perturb HDFs in the vicinity of the SF and AG, thereby potentially altering ion/water fluxes exactly at bottleneck pore regions on the left and right of the CC, where the pore radius is minimized.

At an interface level, inertia and conductivity amplitudes characterizing pain-disease-related mutations fluctuate stronger than those characterizing non-pathogenic mutations (Figs. 5(d),(e))). This suggests that GoF and LoF surface hotspots engage in substantially stronger coupling interactions with their environment compared to surfaces attracting non-pathogenic mutations. GoF and LoF surface hotspots are characterized by oppositely signed interfacial conductivities at the ES of the CC, suggesting that, once perturbed, these interfaces could generate oppositely directed ion/water flux responses (Figs. 5(e))).

3.2.2 Verification of Mechanistic Rationalization via Machine Learning Experimentation

We introduce a machine learning strategy to differentiate between pathogenic and non-pathogenic NaV1.7 locations, both locally (i.e., at $\mathbf{p}$) and globally (i.e., across $\mathcal{P}$), based on $\frac{\mathcal{I}_j}{\phi_j \neq 0}$ (SI S1.10). To avoid divergent $\frac{\mathcal{I}_j}{\phi_j \neq 0}$-values, we treat ϕ_j and $\mathcal{I}_j$ as two separate feature inputs and we also investigate the significance of inertia and conductivity constraints separately, ending up with four feature inputs, $\{\phi_{2k}, \mathcal{I}_{2k}, \phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$.

3.2.2.1 Machine learning experiment I: pain-disease-associated vs. neutrals (PDB: 7w9k)

Following previous works [37, 58], we apply our algorithm to the well-balanced dataset of pain-disease-associated mutations and neutrals.

The local performance of the classifier, evaluated in terms of the area under the curve (AUC) and f1 scores, remains generally stable along $\mathcal{P}$, indicating that the classifier can adapt well to local atom packing conditions (SI S2.6.1).

A yes/no variant classification scheme based on a linear threshold, is illustrated in Fig. 6(a). Its performance, evaluated in terms of the median of the AUC scores, reaches 0.85 (Fig. 6(e), Tab. 1).

ϕ_{2k}-feature importance shows that maintaining higher-order interactions becomes increasingly important around the SF pore region. The importance of the second and fourth-order inertia features exceed that of the zero-th-order, suggesting that rotational and vibrational modes up to at least fourth order may influence ion selectivity (Fig. 6(c)). Corroboratively, the importance of higher-order interfacial inertia features rises sharply at the SF pore region, while, strikingly, the zero-th-order term, $\mathcal{I}_0$, losses significance (SI Fig. S26(a)).

Additionally, at the SF pore region, higher-order conductivity features, particularly those of third and fifth order, become increasingly dominant, even surpassing the importance of first-order features (Fig. 6(d)). This indicates that volumetric, potentially asymmetric interactions may be crucial for the dehydration of sodium ions. In the $\mathcal{I}_{2k+1,\perp}$-feature domain, we observe similar interdependences at the ES entry point of the SF (SI Fig. S26(b)).

To clarify whether these results depend on the feature sign encoding information concerning the direction of the perturbation response (Eqs. 22,27), we repeat the machine learning experiment with the unsigned feature input, $\{|\phi_{2k}|, |\mathcal{I}_{2k}|, |\phi_{2k+1,\perp}|, |\mathcal{I}_{2k+1,\perp}|\}$. Classification results become only marginally worse (Tab. 1 vs. SI Tab. S5), suggesting that NaV1.7 functional architecture is affected by the violation of inertia and conductivity constraints shown in Fig. 5 in a largely direction-independent manner.

3.2.2.2 Machine learning experiment II: pain-disease-associated vs. non-pathogenic (PDB: 7w9k)

To demonstrate that our findings are not biased by the specific choice of neutrals, we repeat our machine learning experiment, this time focusing on the non-pathogenic mutation subset containing both neutrals and benign mutations.

Obtained AUC scores admit a median of 0.77 (Fig. 6(f), Tab. 1). Despite the considerable class imbalance, as non-pathogenic mutations greatly outnumber pain-disease-associated ones with a class ratio as low as 0.19, the algorithm maintains both accuracy and stability, showing no detrimental effects (SI Figs. S24(c), S25(c)).

These findings reinforce that pain-disease-associated mutations appear in molecular neighborhoods of significantly different perturbation potential profile compared to non-pathogenic mutations.

3.2.2.3 Machine learning experiment III: pathogenic vs. non-pathogenic (PDB: 7w9k)

Running our algorithm on the entire *SCN9A*-gene mutation dataset results in a drop of approximately 25% in performance, with the median AUC score dropping to 0.63 (Tab. 1). The reason for the observed drop is the inclusion of pain-disease-unrelated pathogenic mutations. The perturbation potential profile of structural locations whose mutability is associated with a disease phenotype distinct from painful or painless neuropathies, resembles that of non-pathogenic mutations – at least up to a first-order ϕ_j derivative. The level of confidence in this finding is high, as less than 6% of all pain-disease-unrelated pathogenic mutations are characterized as 'Pathogenic/Likely pathogenic' or 'Likely pathogenic'.

Enhancing classifier performance would involve higher-order derivatives of ϕ_j, alongside with refined scaling exponents calculation techniques as managing noise becomes progressively more challenging with higher derivative orders and greater distance from the pore. In fact, pain-disease-associated mutations affecting the VSDs are more likely to be misclassified (Fig. 6(b)). Four out of five misclassified GoF mutations are associated with an SFN phenotype and occur at VSD locations Arg185 (DI), Ile720 (DII), Ile739 (DII), and Arg1279 (DIII). The fifth misclassified GoF mutation, associated with an IEM phenotype, is found at the VSD location Trp1538 (DIV). Among the three misclassified LoF mutations, Arg99 (DI) and Leu172 (DI) are in the VSD, while Cys1719 (DIV) is in the extracellular loop connecting S5 and S6.

3.2.2.4 Repeating machine learning experiments I-III using a lower-resolution structure (PDB: 6j8j)

To assess how lower resolution impacts algorithm performance, we repeat the analysis using a widely studied NaV1.7 structure with lower resolution (PDB code: 6j8j. Despite a slight drop in performance, the results remain largely consistent (Tab. 1 vs. SI Tab. S5). Given that both the high- and low-resolution structures likely represent an inactivated state of the channel, the observed decrease in performance is due to low-resolution limitations.

3.2.2.5 Inertia-vs-conductivity: what matters more for NaV1.7 physiological functioning? (PDB: 7w9k, 6j8j)

Exchanging the feature inputs, $\{\phi_{2k}, \mathcal{I}_{2k}, \phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$, with either $\{\phi_{2k}, \mathcal{I}_{2k}\}$ or $\{\phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$, for the 7w9k or 6j8j structures, induces an algorithm's performance drop that is always less then 10% (Tab. 1, SI Tab. S5). Additionally, the AUC scores achieved with $\{\phi_{2k}, \mathcal{I}_{2k}\}$ are close to those obtained with $\{\phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$. Together,

these results suggest that the feature inputs $\{\phi_{2k}, \mathcal{I}_{2k}\}$ and $\{\phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$ contain redundant information, implying that rotational and translational NaV1.7 DoF are coupled. The corresponding perturbations modes are thus likely intertwined through feedback, with molecule rotations around the pore influencing the flow of ions through the pore, and vice versa.

4 Discussion

Although our study was motivated by an unmet need for systematically deriving machine learning features for identifying pain-disease-associated mutation hotspots in the NaV1.7, we were surprised to uncover broader implications.

By deriving – rather than postulating – the q-deformed statistical mechanical entropy directly from the underlying equation of state (Eq. (2)), we performed a formal statistical mechanical investigation of metastable NaVCh molecules. Most notably, we observed the self-organized resistance of the PD to excess entropy and found that NaVChs have evolved under the silent guidance of a generalized Widom scaling law, ultimately aiming at maximizing mechanofunctional output, as if NaVChs were highly efficient thermodynamic machines. The nonextensivity q-index works here as a descriptor of NaVCh molecular complexity, linking intrinsic dimensionality to entropy via Eq. (20).

When logically aligned, these concepts led to the derivation of thermostability geodesic measures whereas renormalizability emerged as a fundamental NaVCh structure property providing a rationale for how NaV1.7 residue mutability can lead to human pain disease via violation of inertia and/or conductivity channel constraints. Machine learning served as a verification tool rather than a predictive one, augmenting our understanding of which NaV1.7 DoF are mostly involved in the molecular makeup of human pain disease and in what manner. A unresolved challenge is that VSD hotspots are prone to misclassification, primarily due to low-resolution artifacts and the intrinsically disordered VSD nature.

Our work provides a first principles starting point for understanding design principles of NaVChs in the context of human pain disease. This is a crucial step towards cost-effectively bridging the gap between basic research and translational pain medicine. Because our procedures are entirely general, they may provide a foundational framework for analyzing diverse gene-protein relationships underlying disease in relation to interfacial molecular geometries.

Supplementary information. This article includes supplementary information (SI), which can be accessed in the accompanying PDF file: SI_RGNaVChPain.pdf.

Acknowledgements. The authors would like to sincerely acknowledge the helpful discussions and support of M.Sc. Vishal Sudha Bhagavath Eswaran and Dr. Annette Lischka. Their contributions in identifying pain-disease-associated variants and providing valuable insights during discussions have been greatly appreciated and have contributed to the development of this study.

Declarations

- Funding: This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation 363055819/GRK2415 Mechanobiology of 3D epithelial tissues (ME3T) to A.L.; 368482240/GRK2416, MultiSenses-MultiScales to A.L., LA2740/6-1 to A.L.).

- Conflict of interest/Competing interests: The authors declare no competing interests.
- Ethics approval and consent to participate: do not apply.
- Consent for publication: do not apply.
- Data availability:
  - The data generated from analyzing 71 prokaryotic NaVCh structures are available at: https://doi.org/10.5281/zenodo.14617204
  - The data generated from analyzing 50 eukaryotic NaVCh structures and machine learning experiments in the context of human pain disease are available at: https://doi.org/10.5281/zenodo.14628099
- Materials availability: do not apply
- Code availability: The code used throughout is available at: https://github.com/mnxenakis/NaVCh_Scaling
- Author contributions: M.N.X. designed the study, conceived the mathematical structure, wrote the code, analyzed the data, and prepared the manuscript. A.L. co-designed the study, contributed to the interpretation of the findings, assisted with data analysis, reviewed the manuscript, and secured the funding.

References

[1] Hille, B.: Ionic Channels of Excitable Membranes, 3rd Ed., Sunderland, MA (2001)

[2] Hodgkin, A.L., Huxley, A.F.: Currents carried by sodium and potassium ions through the membrane of the giant axon of loligo. The Journal of Physiology **116**(4), 449–472 (1952) https://doi.org/10.1113/jphysiol.1952.sp004717

[3] Catterall, W.A.: Voltage-gated sodium channels at 60: structure, function and pathophysiology. The Journal of Physiology **590**(11), 2577–2589 (2012) https://doi.org/10.1113/jphysiol.2011.224204

[4] Ahern, C.A., Payandeh, J., Bosmans, F., Chanda, B.: The hitchhiker's guide to the voltage-gated sodium channel galaxy. Journal of General Physiology **147**(1), 1–24 (2015) https://doi.org/10.1085/jgp.201511492

[5] Hille, B.: The permeability of the sodium channel to organic cations in myelinated nerve. Journal of General Physiology **58**(6), 599–619 (1971) https://doi.org/10.1085/jgp.58.6.599

[6] Hille, B.: The permeability of the sodium channel to metal cations in myelinated

nerve. Journal of General Physiology **59**(6), 637–658 (1972) https://doi.org/10.1085/jgp.59.6.637

[7] Hille, B.: Ionic selectivity, saturation, and block in sodium channels. A four-barrier model. Journal of General Physiology **66**(5), 535–560 (1975) https://doi.org/10.1085/jgp.66.5.535

[8] Roux, B., Bernèche, S., Egwolf, B., Lev, B., Noskov, S.Y., Rowley, C.N., Yu, H.: Ion selectivity in channels and transporters. Journal of General Physiology **137**(5), 415–426 (2011) https://doi.org/10.1085/jgp.201010577

[9] Heinemann, S.H., Terlau, H., Stühmer, W., Imoto, K., Numa, S.: Calcium channel characteristics conferred on the sodium channel by single mutations. Nature **356**(6368), 441–443 (1992) https://doi.org/10.1038/356441a0

[10] Yu, F.H., Catterall, W.A.: The vgl-chanome: A protein superfamily specialized for electrical signaling and ionic homeostasis. Science's STKE **2004**(253), 15 (2004) https://doi.org/10.1126/stke.2532004re15

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

[12] Doyle, A.D., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., MacKinnon, R.: The structure of the potassium channel: Molecular basis of k^+ conduction and selectivity. Science **280**(5360), 69–77 (1998) https://doi.org/10.1126/science.280.5360.69

[13] Long, S.B., Tao, X., Campbell, E.B., MacKinnon, R.: Atomic structure of a voltage-dependent k+ channel in a lipid membrane-like environment. Nature **450**, 376–382 (2007) https://doi.org/10.1038/nature06265

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

[15] Bezanilla, F.: The voltage sensor in voltage-dependent ion channels. Physiological Reviews **80**(2), 555–592 (2000) https://doi.org/10.1152/physrev.2000.80.2.555 https://doi.org/10.1152/physrev.2000.80.2.555. PMID: 10747201

[16] Horn, R.: Molecular basis for function in sodium channels. Novartis Foundation symposium **241**, 21–232 (2002)

[17] Vilin, Y.Y., Ruben, P.C.: Slow inactivation in voltage-gated sodium channels. Cell Biochemistry and Biophysics **35**(2), 171–190 (2001) https://doi.org/10.1385/CBB:35:2:171

[18] Goldschen-Ohm, M.P., Capes, D.L., Oelstrom, K.M., Chanda, B.: Multiple pore conformations driven by asynchronous movements of voltage sensors in a eukaryotic sodium channel. Nature Communications **4**, 1350 (2013)

[19] Eaholtz, G., Scheuer, T., Catterall, W.A.: Restoration of inactivation and block of open sodium channels by an inactivation gate peptide. Neuron **12**, 1041–1048 (1994) https://doi.org/10.1016/0896-6273(94)90312-3

[20] Ulbricht, W.: Sodium channel inactivation: Molecular determinants and modulation. Physiological Reviews **85**(4), 1271–1301 (2005) https://doi.org/10.1152/physrev.00024.2004 . PMID: 16183913

[21] Tang, Q.-Y., Hatakeyama, T.S., Kaneko, K.: Functional sensitivity and mutational robustness of proteins. Phys. Rev. Res. **2**, 033452 (2020) https://doi.org/10.1103/PhysRevResearch.2.033452

[22] Tang, Q.-Y., Kaneko, K.: Long-range correlation in protein dynamics: Confirmation by structural data and normal mode analysis. PLOS Computational Biology **16**(2), 1–17 (2020) https://doi.org/10.1371/journal.pcbi.1007670

[23] Gudmundsson, S., et al.: Variant interpretation using population databases: lessons from gnomad. arXiv:2107.11458 (2021)

[24] gnomAD database accessed at 20/09/2024, 17:50 pm. https://gnomad.broadinstitute.org/

[25] Körner, J., Lampert, A.: Sodium channels. 5.09-sodium channels. In: Fritzsch, B. (ed.) The Senses: A Comprehensive Reference, 2nd edn., pp. 120–141. Elsevier, Oxford (2020). https://doi.org/10.1016/B978-0-12-809324-5.24208-9

[26] Khaluf, Y., Ferrante, E., Simoens, P., Huepe, C.: Scale invariance in natural and artificial collective systems: a review. Journal of The Royal Society Interface **14**(136), 20170662 (2017) https://doi.org/10.1098/rsif.2017.0662 https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2017.0662

[27] Tang, Q.-Y., Zhang, Y.-Y., Wang, J., Wang, W., Chialvo, D.R.: Critical fluctuations in the native state of proteins. Phys. Rev. Lett. **118**, 088102 (2017) https://doi.org/10.1103/PhysRevLett.118.088102

[28] Tang, Q.-Y., Kaneko, K.: Dynamics-evolution correspondence in protein structures. Phys. Rev. Lett. **127**, 098103 (2021) https://doi.org/10.1103/PhysRevLett.127.098103

[29] Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: An explanation of the 1/f noise. Phys. Rev. Lett. **59**, 381–384 (1987) https://doi.org/10.1103/PhysRevLett.59.381

[30] Phillips, J.C.: Scaling and self-organized criticality in proteins i. Proceedings of the National Academy of Sciences **106**(9), 3107–3112 (2009) https://doi.org/10.1073/pnas.0811262106 https://www.pnas.org/doi/pdf/10.1073/pnas.0811262106

[31] Phillips, J.C.: Scaling and self-organized criticality in proteins ii. Proceedings of the National Academy of Sciences **106**(9), 3113–3118 (2009) https://doi.org/10.1073/pnas.0811308105 https://www.pnas.org/doi/pdf/10.1073/pnas.0811308105

[32] Silverman, B.D.: Hydrophobic moments of protein structures: Spatially profiling the distribution. Proceedings of the National Academy of Sciences **98**(9), 4996–5001 (2001) https://doi.org/10.1073/pnas.081086198 https://www.pnas.org/doi/pdf/10.1073/pnas.081086198

[33] Silverman, B.D.: Hydrophobicity of transmembrane proteins: Spatially profiling the distribution. Protein Science **12**(3), 586–599 (2003) https://doi.org/10.1110/ps.0214903 https://onlinelibrary.wiley.com/doi/pdf/10.1110/ps.0214903

[34] Silverman, B.D.: Hydrophobic moments of tertiary protein structures. Proteins: Structure, Function, and Bioinformatics **53**(4), 880–888 (2003) https://doi.org/10.1002/prot.10494 https://onlinelibrary.wiley.com/doi/pdf/10.1002/prot.10494

[35] Zhou, R., Silverman, B.D., Royyuru, A.K., Athma, P.: Spatial profiling of protein hydrophobicity: Native vs. decoy structures. Proteins: Structure, Function, and Bioinformatics **52**(4), 561–572 (2003) https://doi.org/10.1002/prot.10419 https://onlinelibrary.wiley.com/doi/pdf/10.1002/prot.10419

[36] Xenakis, M.N., Kapetis, D., Yang, Y., Heijman, J., Waxman, S.G., Lauria, G., Faber, C.G., Smeets, H.J., Lindsey, P.J., Westra, R.L.: Non-extensitivity and criticality of atomic hydropathicity around a voltage-gated sodium channel's pore: a modeling study. Journal of Biological Physics **47**(1), 61–77 (2021) https://doi.org/10.1007/s10867-021-09565-w

[37] Xenakis, M.N., Kapetis, D., Yang, Y., Gerrits, M.M., Heijman, J., Waxman, S.G., Lauria, G., Faber, C.G., Westra, R.L., Lindsey, P.J., Smeets, H.J.: Hydropathicity-based prediction of pain-causing nav1.7 variants. BMC Bioinformatics **22**(1), 212 (2021) https://doi.org/10.1186/s12859-021-04119-2

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

[39] Wilson, K.G.: Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. Phys. Rev. B **4**, 3174–3183 (1971) https://doi.org/10.1103/PhysRevB.4.3174

[40] Xenakis, M.N., Kapetis, D., Yang, Y., Heijman, J., Waxman, S.G., Lauria, G., Faber, C.G., Smeets, H.J., Westra, R.L., Lindsey, P.J.: Cumulative hydropathic topology of a voltage-gated sodium channel at atomic resolution. Proteins: Structure, Function, and Bioinformatics **88**(10), 1319–1328 (2020) https://doi.org/10.1002/prot.25951 . Epub 2020 Jun 13

[41] Phillips, J.: How life works: Darwinian evolution of proteins. arXiv (2021)

[42] Chandler, D.: Interfaces and the driving force of hydrophobic assembly. Nature **437**(7059), 640–647 (2005) https://doi.org/10.1038/nature04162

[43] Gomes, A.L.C., Rezende, J.R., Araújo, A.F.P., Shakhnovich, E.I.: Description of atomic burials in compact globular proteins by fermi-dirac probability distributions. Proteins: Structure, Function, and Bioinformatics **66**(2), 304–320 (2007) https://doi.org/10.1002/prot.21137 https://onlinelibrary.wiley.com/doi/pdf/10.1002/prot.21137

[44] ŠTEL, T., Fülöp, Vicsek, T.: Determination of fractal dimensions for geometrical multifractals. Physica A: Statistical Mechanics and its Applications **159**(2), 155–166 (1989) https://doi.org/10.1016/0378-4371(89)90563-3

[45] Enright, M.B., Leitner, D.M.: Mass fractal dimension and the compactness of proteins. Phys. Rev. E **71**, 011912 (2005) https://doi.org/10.1103/PhysRevE.71.011912

[46] Arnold, B.C.: Pareto Distributions. International Cooperative Publishing House, Fairland, Maryland (1983)

[47] Beck, C., Schögl, F.: Thermodynamics of Chaotic Systems: An Introduction. Cambridge Nonlinear Science Series. Cambridge University Press, Cambridge (1993)

[48] Tsallis, C.: Possible generalization of boltzmann-gibbs statistics. Journal of Statistical Physics **52**(1), 479–487 (1988) https://doi.org/10.1007/BF01016429

[49] Tsallis, C.: Introduction to Nonextensive Statistical Mechanics. Springer, Berlin (2009)

[50] Kapcha, L.H., Rossky, P.J.: A simple atomic-level hydrophobicity scale reveals protein interfacial structure. Journal of Molecular Biology **426**(2), 484–498 (2014) https://doi.org/10.1016/j.jmb.2013.09.039 . Epub 2013 Oct 10

[51] Donaldson, S.H. Jr, Lee, C.T. Jr, Chmelka, B.F., Israelachvili, J.N.: General hydrophobic interaction potential for surfactant/lipid bilayers from direct force measurements between light-modulated bilayers. Proc Natl Acad Sci U S A **108**(38), 15699–15704 (2011)

[52] Donaldson Jr., S.H., Das, S., Gebbie, M.A., Rapp, M., Jones, L.C., Roiter, Y., Koenig, P.H., Gizaw, Y., Israelachvili, J.N.: Asymmetric electrostatic and hydrophobic–hydrophilic interaction forces between mica surfaces and silicone polymer thin films. ACS Nano **7**(11), 10094–10104 (2013) https://doi.org/10.1021/nn4050112

[53] Donaldson Jr., S.H., Røyne, A., Kristiansen, K., Rapp, M.V., Das, S., Gebbie, M.A., Lee, D.W., Stock, P., Valtiner, M., Israelachvili, J.: Developing a general interaction potential for hydrophobic and hydrophilic interactions. Langmuir **31**(7), 2051–2064 (2015) https://doi.org/10.1021/la502115g

[54] Politzer, H.D.: Reliable perturbative results for strong interactions? Phys. Rev. Lett. **30**, 1346–1349 (1973) https://doi.org/10.1103/PhysRevLett.30.1346

[55] Ljunggren, S., Eriksson, J.C., Kralchevsky, P.A.: Minimization of the free energy of arbitrarily curved interfaces. Journal of Colloid and Interface Science **191**(2), 424–441 (1997) https://doi.org/10.1006/jcis.1997.4961

[56] Tang, Q.-Y., Ren, W., Wang, J., Kaneko, K.: The Statistical Trends of Protein Evolution: A Lesson from AlphaFold Database. Molecular Biology and Evolution **39**(10), 197 (2022) https://doi.org/10.1093/molbev/msac197 https://academic.oup.com/mbe/article-pdf/39/10/msac197/46424008/msac197.pdf

[57] Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chrast, J., Clark, K., Elliott, K., Faich, T., Fisher, S., al.: Clinvar: Public archive of relationships among sequence variation and human phenotype. Nucleic acids research **46**(D1), 1062–1068 (2018) https://doi.org/10.1093/nar/gkx1153

[58] Kapetis, D., Sassone, J., Yang, Y., Galbardi, B., Xenakis, M.N., Westra, R.L., Szklarczyk, R., Lindsey, P., Faber, C.G., Gerrits, M., Merkies, I.S.J., Dib-Hajj, S.D., Mantegazza, M., Waxman, S.G., Lauria, G., Taiana, M., Marchi, M., Lombardi, R., Cazzato, D., Boneschi, F.M., Zauli, A., Clarelli, F., Santoro, S., Lopez, I., Quattrini, A., Hoeijmakers, J., Sopacua, M., Greef, B., Smeets, H.J.M., Momani, R.A., Vanoevelen, J.M., Eijkenboom, I., Cestèle, S., Chever, O., Malik, R., Tavakoli, M., Ziegler, D., PROPANE Study Group: Network topology of nav1.7 mutations in sodium channel-related painful disorders. BMC Systems Biology **11**(1), 28 (2017) https://doi.org/10.1186/s12918-016-0382-0

[59] Thomas, D.J.: Concepts in protein folding. FEBS Letters **307**(1), 10–13 (1992) https://doi.org/10.1016/0014-5793(92)80892-K

[60] Krieger, E., Koraimann, G.J., Vriend, K.: YASARA: Yet Another Scientific Artificial Reality Application. YASARA Biosciences GmbH, Vienna, Austria (2002). YASARA Biosciences GmbH. http://www.yasara.org

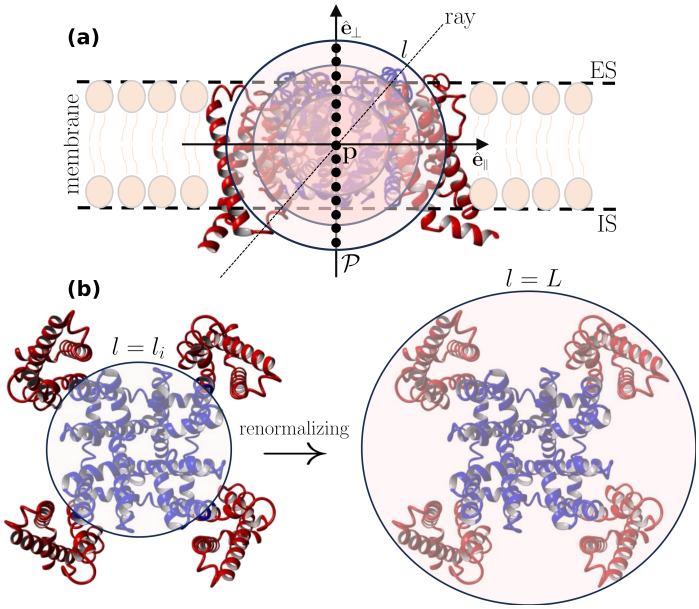
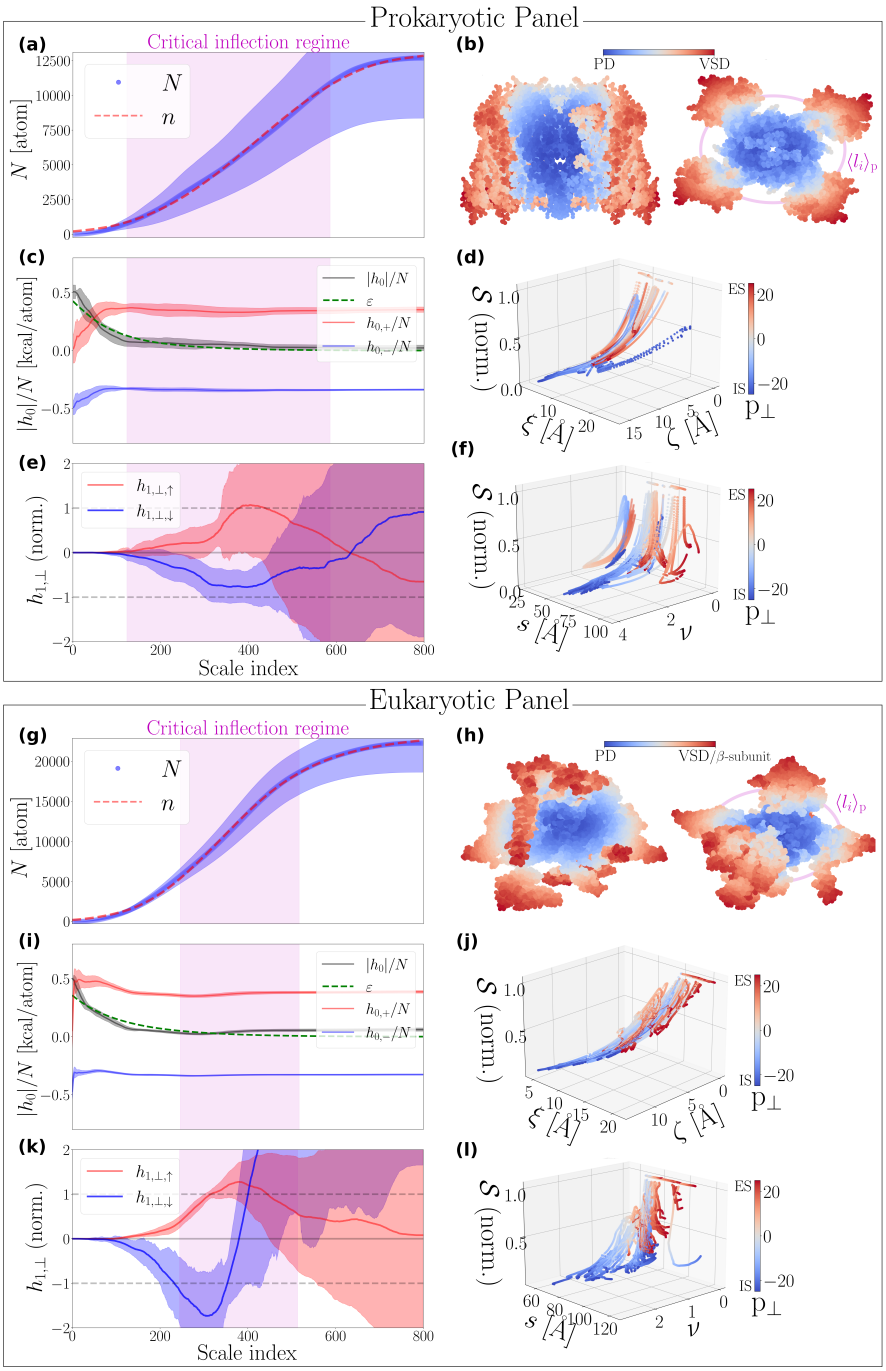


Fig. 1 Geometry and renormalizability of the prototype NaVAb channel. **(a)**, The illustrated molecular side view corresponds to a pre-open NaVAb molecule (PDB code: 3rvy). The pore domain (PD) and voltage sensor domains (VSDs) are illustrated in blue and red, respectively. $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ are the membrane-parallel and membrane-perpendicular unit vectors, respectively. We introduce consecutive pore points, $\mathbf{p} \in \mathcal{P}$ (SI Eq. (S1)), forming a path through the NaVAb pore. Each pore point serves as the center of an ensemble of nested spherical subsystems, each of them characterized by a radius l [Å] (Eq. (1)). An infinite number of paths (rays) can pass through a pore point, radiating in all possible directions. Analyzing the properties of the atomic environment along individual radial paths is impractical due to their infinite number. The renormalization group procedure solves this problem by collapsing radial paths into l, rendering atomic environment properties inherently dependent on l. This is equivalent to *coarse-graining* [59], and enables the computation of relevant scaling exponent. For example, substituting $\ln(n)$ for $\mathcal{O}$ in Eq. (30), returns via (31) the order parameter scaling exponent, $\Lambda = \pm \frac{\xi d_f}{l}$ (SI Eq. (S15)). **(b)**, Top view, from the extracellular side (ES). l_i [Å] represents the characteristic size of the PD, also marking the inflection point of the cumulative atom number (Eqs. (5),(6)). L [Å] is the radius of the smallest sphere that contains the entire molecule (SI Eq. (S4)). IS stands for intracellular side.

Fig. 2 Spatial organization features of the NaVCh superfamily. Statistical summary of the spatial organization features of 71 prokaryotic (subtypes: NaChBac, NaVAb, NaVMs, NaVAe, NaVRh) and 50 eukaryotic (subtypes: NaVPas, NaVEe1, NaV1-8, NaVEh) NaVCh atomic structures (SI Tab. S1), for pore points, $\mathbf{p} \in \mathcal{P}$, and scale indices, $\alpha = 1, 2, \ldots, 800$. **(a)**,**(g)**, Collapsed traces of the empirical, N [atom], and best-fitted theoretical, n [atom], cumulative atom numbers. **(b)** and **(h)** illustrate the prototype NaVAb channel (PDB code: 3rvy) and the human NaV1.7 channel (PDB code: 7w9k), respectively. Atoms are colored according to their PD/VSD ordering score (SI S1.8.1.2), and projected onto a plane perpendicular (left side) and parallel (right side) to the membrane. $\langle l_i \rangle_{\mathrm{p}}$ is the mean value (computed along $\mathcal{P}$) of the characteristic size of the pore domain (PD). VSD stands for voltage sensor domain. **(c)**,**(i)**, Collapsed traces of the empirical, $\frac{|h_0|}{n}$ [kcal/atom], and theoretical, ε [kcal/atom], absolute atomic hydropathic energies. $\frac{h_{0,+}}{n}$ is the hydrophobic component, and $\frac{h_{0,-}}{n}$ is the hydrophilic component of $\frac{h_0}{n}$, respectively. **(d)**,**(j)**, The interplay between the normalized statistical mechanical entropy, $\mathcal{S}$, and the pair of interaction ranges, $\{\xi, \zeta\}$. **(e)**,**(k)**, Collapsed traces of 'pointing extracellularly (ES)' and 'pointing intracellularly (IS)' instances of the normalized membrane-vertical HDF amplitude, $h_{1,\perp}$ [kcal×Å /atom]. $h_{1,\perp}$ is normalized over $h_{1,\perp}|_{l=l_i}$. If $h_{1,\perp}|_{l=l_i} > 0$, $h_{1,\perp}$ is labeled as a 'pointing ES' (↑) instance; otherwise, 'pointing IS' (↓). **(f)**,**(l)**, Nonextensivity of $\mathcal{S}$ [kcal/(Θ×atom)], as revealed from its weak and strong dependence on the molecular radial size, s [Å], and the interaction range ratio, $\nu = \frac{\xi}{\zeta}$, respectively. Note that $\mathcal{S}$ is normalized over $\mathcal{S}_{\max}$ (Eq. (18)). The critical inflection regime marks the interval from which all l_i values are drawn. Trace collapse procedures are described in SI S1.8.1.1.



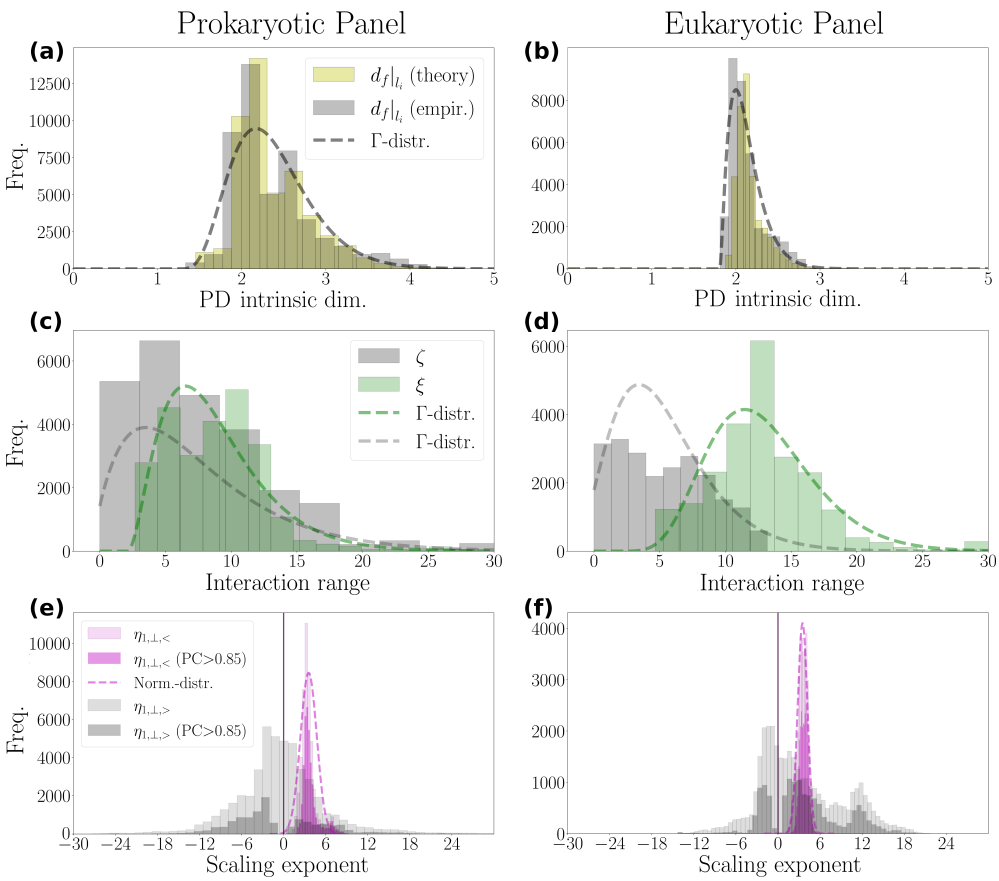


Fig. 3 Intrinsic dimension of a pore domain and attractive-vs.-repulsive interaction range statistics. Statistical compilation of the pore domain (PD) intrinsic dimension and interaction range characteristics for 71 prokaryotic (subtypes: NaChBac, NaVAb, NaVMs, NaVAe, NaVRh) and 50 eukaryotic (subtypes: NaVPas, NaVEe1, NaV1-8, NaVEh) NaVCh atomic structures (SI Tab. S1). **(a)**,**(b)**, Histograms of the empirical and theoretical measures of the PD intrinsic dimension. The Γ-distribution mean value of the empirical intrinsic dimension measure is 2.4 and 2.1, with corresponding standard deviations of 0.5 and 0.2, for prokaryotic and eukaryotic PDs, respectively. **(c)**,**(d)**, Histograms of the attractive and repulsive interaction ranges given by ζ [Å] and ξ [Å], respectively. The Γ-distribution mean value of ξ is 8.8 Å for prokaryotes and 12.9 Å for eukaryotes, while the mean value of ζ is 7.5 Å for prokaryotes and 5.2 Å for eukaryotes. **(e)**,**(f)**, Histograms of the scaling exponents of the membrane-perpendicular dipole field component, $h_{1,\perp}$. The exponents $\eta_{1,\perp,<}$ and $\eta_{1,\perp,>}$ account for the scaling behavior of $h_{1,\perp}$ over l-intervals covering the PD and the VSDs, respectively (Eq. (23)). They are computed according to procedures described in SI S1.8.3. PC stands for Pearson coefficient.

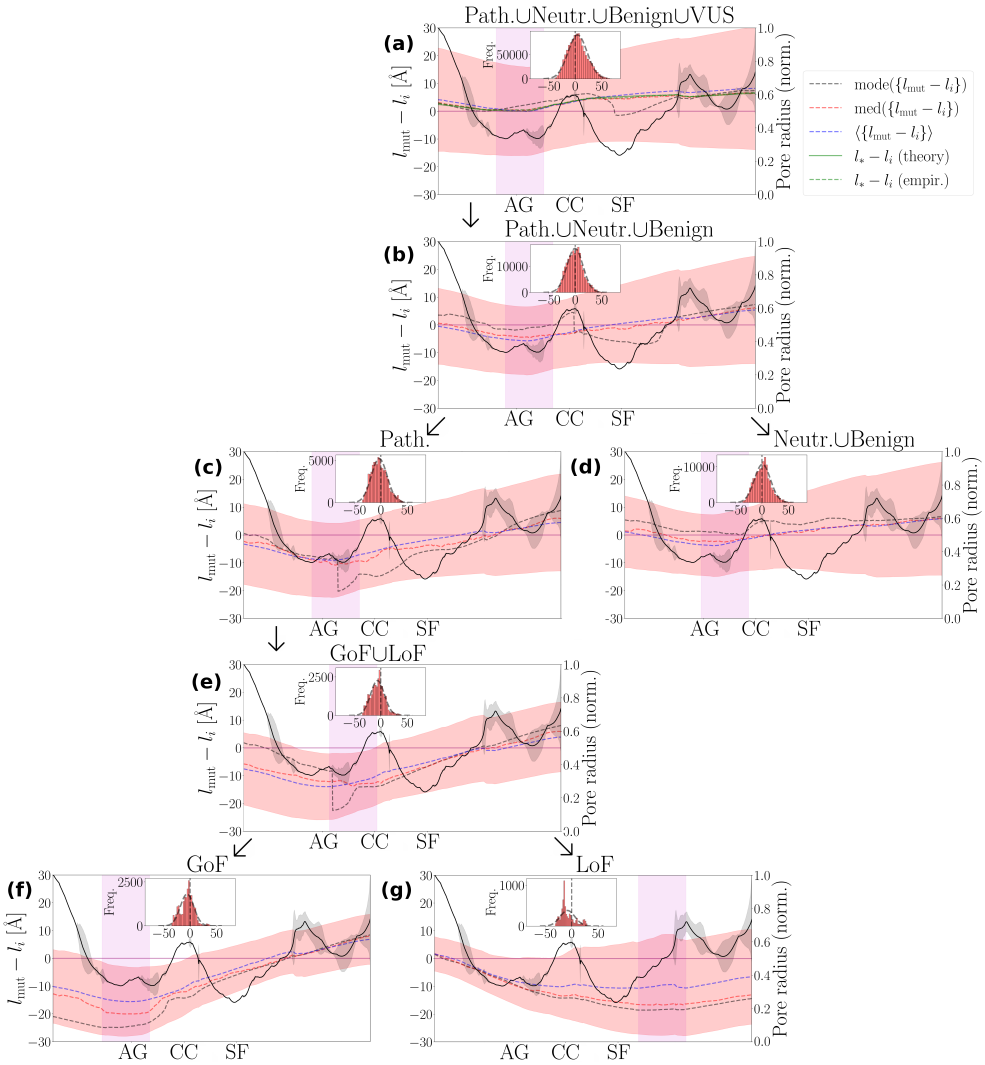


Fig. 4 Distributional characteristics of mutations inside a NaV1.7 structure. We summarize the statistical properties of the $\{l_{\mathrm{mut}} - l_i\}$ dataset in terms of the median med$(\cdot)$ and the mean, $\langle\cdot\rangle$, for pore points, $\mathbf{p} \in \mathcal{P}$. l_{mut} represents the Euclidean distance between $\mathbf{p}$ and the geometric center of the residue being mutated. We break down $\{l_{\mathrm{mut}} - l_i\}$ to different subsets each of them corresponds to a different mutation set. Insets visualize the histogram of the $\{l_{\mathrm{mut}} - l_i\}$ dataset, i.e., the collapsed distribution incorporating contributions from all pore points. In **(a)**, we show the $\{l_{\mathrm{mut}} - l_i\}$-distribution of the parent set of all (i.e., Path.∪Benign∪Neutr.∪VUS) *SCN9A*-gene mutations. l_* identifies the structural location where the coin-flipping entropy is maximized. Note that if $n_L \approx K$, l_* can be approximated with $l_{i,\mathrm{b}}$ (Eq. (8)). The empirical and theoretical instance of l_* is determined by the l-value for which $N/N_{\mathrm{total}} = 0.5$ and $n/n_L = 0.5$ is satisfied, respectively. The parent dataset is broken down to the following: **(b)**, pathogenic and non-pathogenic (Path.∪Benign∪Neutr.). **(c)**, pathogenic (Path). **(d)**, non-pathogenic (Benign∪Neutr.). **(e)**, Gain-of-function (GoF) and loss-of-function (LoF) (GoF∪LoF). **(f)**, GoF. **(g)**, LoF. Note that a negative and positive sign of $l_{\mathrm{mut}} - l_i$ indicate that the residue being mutated is likely residing in the PD and the VSDs, respectively. The areas highlighted in light magenta represent the pore regions where the median is minimized.

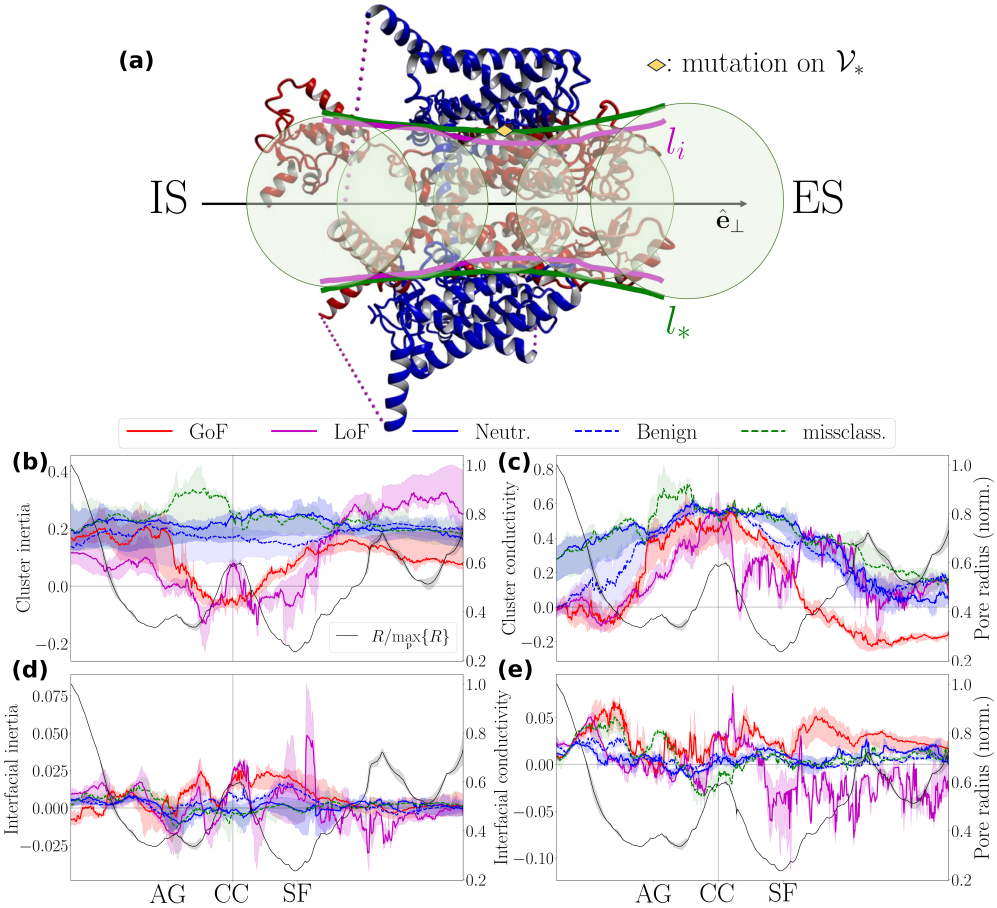


Fig. 5 Inertia and conductivity profile of structural locations attracting pain-disease-associated mutations. **(a)**, Side view of a human NaV1.7 channel (PDB code: 7w9k). The PD and VSDs are illustrated in red and blue, respectively. For clarity, the β subunits are not shown. The dense arrangement of $\mathcal{B}_*$-spheres along $\mathcal{P}$ creates the smooth cylindrical-like surface (in green), denoted as $\mathcal{V}_* = \partial\bigcup_{\mathbf{p}}\mathcal{B}_*$. For comparison, the characteristic size of the PD, l_i, is also illustrated (in magenta). Note that the inequality $l_* > l_i$ arises from the atom packing condition $\nu < 1$ (SI Fig. S23). At the AG pore region, where $\nu \nearrow 1$, $l_* \approx l_i$ implies $l_* \approx l_{i,\mathrm{b}}$ with $n_L \approx K$ (see Eq. (8)). In **(b)** and **(c)** we statistically summarize the n-cluster inertias, $\hat{\phi}_{2k}$, and the n-cluster conductivities, $\hat{\phi}_{2k+1,\perp}$, characterizing structural locations where the gain-of-function (GoF), loss-of-function (LoF), Neutral (Neutr.) and Bening mutations appear for $k = 0, 1, \ldots 5$ **(d)** and **(e)** provide analogous information at an interface level: we statistically summarize the ∂n-shell inertias, $\hat{\mathcal{I}}_{2k}$, and the ∂n-shell conductivities, $\hat{\mathcal{I}}_{2k+1,\perp}$. Clouds around the index traces represent the min-max range of the underlying data points. The misclassified (missclass.) subset contains pain-disease-associated mutations systematically misclassified by the machine-learning algorithm (see Fig. 6(b)). The conductivity index derived from the first-order hydropathic dipole field is separately shown in SI Fig. S4, as the pair $\{h_{1,+,\perp}, h_{1,-,\perp}\}$ satisfies (24) only partially (SI S2.3.1). The construction of the statistical summary indices $\hat{\phi}_j$ and $\hat{\mathcal{I}}_j$ is described in SI S1.8.2. Molecular illustrations were generated using the Yasara software [60].

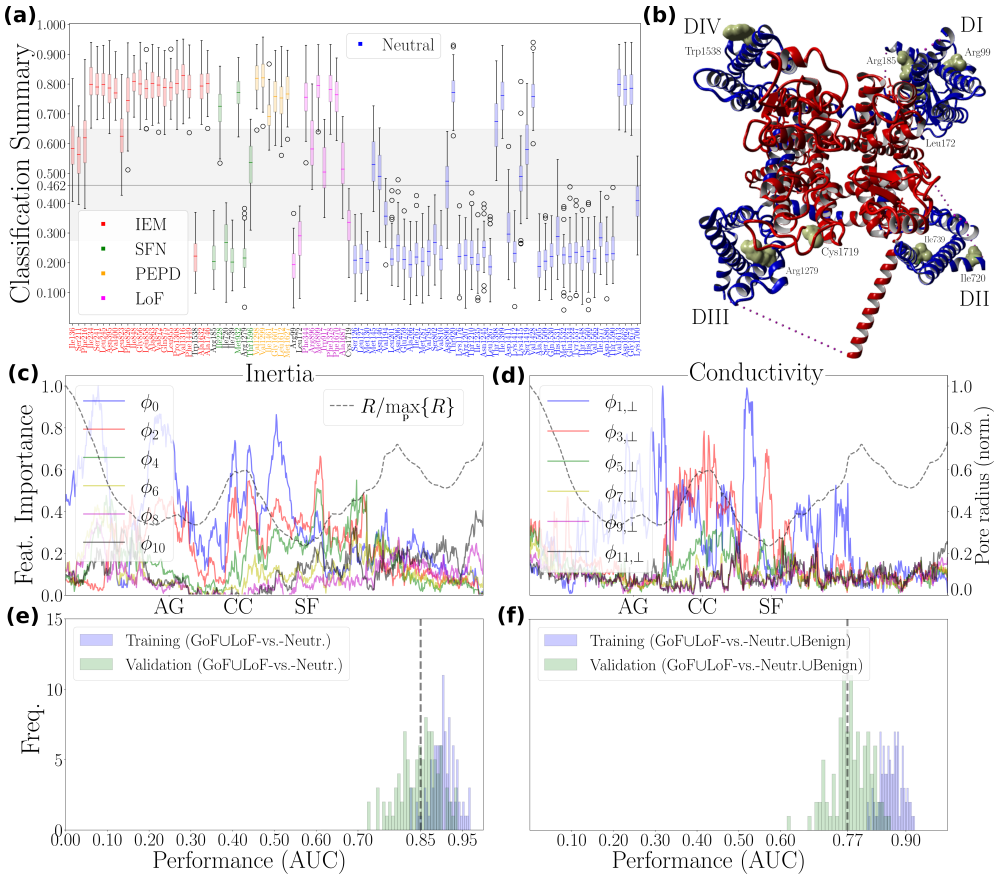


Fig. 6 Machine-learning-assisted verification: pain-disease-associated mutation hotspots exhibit substantially distinct perturbation potential profile. **(a)**, Summary of a machine learning experiment evaluating the distinctiveness of the pain-disease-associated class (`class_0`: GoF∪LoF) in relation to the neutral class (`class_1`: Neutr.). Gain-of-function (GoF) subclass incorporates structural locations whose mutability is phenotypically associated with inherited erythromelalgia (IEM), small fiber neuropathy (SFN), and paroxysmal extreme pain disorder (PEPD) (SI Tab. S3). Loss-of-function (LoF) subclass incorporates structural locations whose mutability is phenotypically associated with insensitivity to pain (SI Tab. S3). Neutrals are structural locations whose mutability is not likely to be phenotypically associated with human pain disease, sourced from [37, 58]. The optimal threshold at 0.462 corresponds to the median of the best thresholds obtained through bootstrapping; each bootstrap sample of `class_0` probabilities yields a threshold that maximizes the f1-score. X-axis ticks indicate the location of a mutation event in the NaV1.7 structure. Note that x-axis ticks of misclassified pain-disease-associated mutations (i.e., those that do not pass the linear threshold) are highlighted in black. **(b)**, Top-view illustration of the human NaV1.7 channel (PDB code: 7w9k). For clarity, the β-subunits are not shown. The van der Waals surface of misclassified structural locations is highlighted in yellow. The PD and the VSDs are shown in red and blue color, respectively. In **(c)** and **(d)** we plot the importance score for the features ϕ_{2k} and ϕ_{2k+1}, $k=0,1,\ldots,5$, accounting for inertia and conductivity constraints, respectively. In **(e)** and **(f)** we illustrate the distribution of the final area under the curve (AUC) scores obtained from the final classification round for two different machine learning experiments, namely, the one summarized in (a) and the another one with `class_0`: GoF∪LoF vs. `class_1`: Neutr.∪Benign, respectively. Benign are structural locations whose mutability is not likely to be associated with disease (SI 1.9). Details concerning machine learning algorithm design and parameter selection can be found in SI 1.10. The Yasara software [60] was used for molecular illustrations.

Table 1 Machine-learning experiments summary. The medians of the AUC and f1 scores were obtained during the final classification round (SI 1.10.2). The first and second number of each $(\cdot,\cdot)$-pair are median values obtained during training and testing the final classification model, respectively. Inertia (iner.) and conductivity (cond.) constraints are probed by the feature inputs $\{\phi_{2k}, \mathcal{I}_{2k}\}$ and $\{\phi_{2k+1,\perp}, \mathcal{I}_{2k+1,\perp}\}$, respectively. *SCN9A*-gene mutation datasets I, II, and III contain the classes {`class_0`: GoF∪LoF, `class_1`: Neutr.}, {`class_0`: GoF∪LoF, `class_1`: Neutr.∪Benign}, and {`class_0`: Path., `class_1`: Neutr.∪Benign}, respectively. GoF and LoF stand for gain-of-function and loss-of-function, respectively, and represent mutations associated with increased or diminished pain sensation (listed in SI Tab. S3). Neutrals (neutr.) are carefully selected human pain disease negative controls, sourced from [37, 58]. Benign are generally not expected to be associated with a disease phenotype. Note that the pathogenic (path.) class contains both pain-disease-associated and pain-disease-unrelated pathogenic mutations.

	PDB: 7w9k, res.: 2.2 Å			PDB: 6j8j, res.: 3.2 Å		
Pert. modes	iner./cond.	iner.	cond.	iner./cond.	iner.	cond.
AUC (dataset I)	$(0.90, 0.85)$	$(0.89, 0.83)$	$(0.87, 0.84)$	$(0.90, 0.83)$	$(0.87, 0.82)$	$(0.87, 0.83)$
f1 (dataset I)	$(0.81, 0.78)$	$(0.78, 0.76)$	$(0.81, 0.80)$	$(0.80, 0.78)$	$(0.79, 0.76)$	$(0.80, 0.78)$
AUC (dataset II)	$(0.87, 0.77)$	$(0.80, 0.74)$	$(0.85, 0.77)$	$(0.84, 0.72)$	$(0.79, 0.70)$	$(0.79, 0.72)$
f1 (dataset II)	$(0.75, 0.72)$	$(0.76, 0.74)$	$(0.74, 0.72)$	$(0.79, 0.73)$	$(0.80, 0.78)$	$(0.74, 0.70)$
AUC (dataset III)	$(0.77, 0.63)$	$(0.71, 0.60)$	$(0.71, 0.64)$	$(0.79, 0.65)$	$(0.74, 0.64)$	$(0.71, 0.61)$
f1 (dataset III)	$(0.71, 0.62)$	$(0.67, 0.61)$	$(0.68, 0.64)$	$(0.72, 0.62)$	$(0.68, 0.61)$	$(0.67, 0.59)$