Ion-Pot and Ion-Curr: Ion-Aware Benchmarks for Electrostatic Field and Observable Prediction in Complex Geometries

Jingqian Liu [*1] **Pinhao Gu** [*2] **Aleksei Aksimentiev** [3]

Abstract

Electrostatic interactions and ionic transport are central to biomolecular function, but continuum solvers for these problems remain costly at scale. Recent learning-based surrogates have shown promise, yet most existing settings focus on fixed ionic conditions and pay limited attention to the role of mobile ions. We introduce two complementary benchmarks for ion-aware biomolecular surrogate modeling: **Ion-Pot**, for electrostatic potential prediction around biomolecules under varying ionic conditions, and **Ion-Curr**, for ionic current prediction in complex nanopore-based biomolecular systems. Together, these tasks span two related prediction regimes: high-dimensional field reconstruction and observable-level transport prediction. We further develop voxel-based convolutional models tailored to each setting and show that they provide accurate and efficient surrogates for both tasks. For Ion-Pot, explicit conditioning on ion concentration improves predictive performance over no-ion-input baselines, highlighting the importance of ion-aware learning for electrostatic field prediction. For Ion-Curr, a lightweight 39K-parameter architecture accurately predicts solver-derived current labels from complex nanopore geometries and analyte configurations, suggesting that surrogate complexity may depend strongly on the dimensionality and structure of the prediction target. These results establish Ion-Pot and Ion-Curr as a paired testbed for biomolecular electrostatics under varying ionic conditions and geometry-dependent ion-transport surrogate modeling. The implementations and instructions for downloading the Ion-Curr dataset are available at https://github.com/jingqian-liu/Ion-Pot_Ion-Curr/tree/main.

[1]Siebel School of Computing and Data Science, University of Illinois Urbana-Champaign, IL, US [2]Department of Chemistry, University of Illinois Urbana-Champaign, IL, US [3]Department of Physics, University of Illinois Urbana-Champaign, IL, US. Correspondence to: Aleksei Aksimentiev <aksiment@illinois.edu>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

1. Introduction

Electrostatics near biomolecules play an important role in shaping how their surfaces interact with their environment and with other molecules, influencing processes such as solvation, binding, and selective recognition (Ren et al., 2012; Zhou & Pang, 2018). In practice, continuum biomolecular electrostatics has long relied on Poisson–Boltzmann solvers such as APBS (Baker et al., 2001; Jurrus et al., 2018) and DelPhi (Li et al., 2013; 2019), as well as interface-accurate methods such as MIBPB (Chen et al., 2011). Although these solvers are physically grounded and widely used, they remain computationally expensive at large scale. This cost has motivated growing interest in machine-learning surrogates for biomolecular electrostatics, including graph-based models such as PBGNN (Wu et al., 2025) and solver-trained surrogates such as PBML (Chen et al., 2024), with the promise of orders-of-magnitude faster inference.

Despite this progress, existing surrogate models have largely focused on solvation free energies or electrostatic potentials in pure solvent, leaving an important aspect of biomolecular electrostatics underexplored in current learning settings: the role of mobile ions in solution. Ionic conditions are not merely a minor correction to vacuum electrostatics; they fundamentally reshape the electrostatic environment through screening and strongly influence biomolecular behavior under realistic solvent conditions (Ren et al., 2012). For example, salt concentration modulates the range and strength of electrostatic interactions at protein surfaces, and changes in ionic strength can alter ligand association, protein–protein binding, and other charge-mediated interactions (Sharp & Honig, 1990; Zhou & Pang, 2018). In specialized biomolecular systems such as ion channels and biological nanopores, which are nano-scale pores often formed by proteins that allow ions and molecules to pass through a confined channel, ions play an even more direct role, since ionic rearrangement and flux determine the measurable current signal itself (Hille, 2001; Ayub & Bayley, 2013). Surrogates that

explicitly account for ions may therefore be worth exploring, as ionic effects constitute a central physical mechanism linking biomolecular structure to electrostatic response and function.

Motivated by this gap, we introduce two complementary benchmark problems for ion-aware biomolecular surrogate modeling. The first considers prediction of the full electrostatic potential field around biomolecules under varying ionic conditions. The second considers prediction of ionic current through biomolecular systems with complex geometries. Although these targets differ in form, both are derived from continuum descriptions involving mobile ions: the former requires recovering a high-dimensional spatial field, whereas the latter requires predicting a single observable induced by biomolecular geometry. Based on this design, we construct two datasets. The first, **Ion-Pot**, is a benchmark for ion-dependent electrostatic field prediction, containing 3,875 proteins evaluated at nine ion concentrations using APBS. The second, **Ion-Curr**, contains approximately 80K samples that map complex nanopore-based biomolecular geometries to ionic current under a reference ionic condition and applied electric bias. The Ion-Curr current labels are generated using a FEniCS-based solver built on the Steric Exclusion Method (SEM), whose predictions have been shown to closely match ionic transport estimated from molecular dynamics simulations (Wilson et al., 2019).

To benchmark these problems, we adopt voxel-based convolutional neural network (CNN) models. This choice is motivated by the spatial structure of both tasks: full-field electrostatic prediction requires learning local and multi-scale three-dimensional patterns, while current prediction in ion channels and nanopores depends strongly on confined geometric features such as pore shape, constrictions, and analyte configurations. CNNs therefore provide a natural inductive bias for structured 3D grids and a practical alternative to graph-based models. In particular, for nanopore current prediction, voxel representations can focus computation on the geometry most relevant to transport, whereas graph-based representations may introduce many nodes for atoms buried inside the biomolecule or surrounding membrane whose contribution to the observable is only indirect. We found that the choice leads to favorable results, which are summarized in the following contributions.

Our main contributions are:

- We introduce **Ion-Pot** and **Ion-Curr**, two benchmark datasets for ion-aware biomolecular electrostatic surrogate modeling, covering ion-dependent electrostatic potential prediction and ionic current prediction, respectively.
- We present two CNN-based models tailored to these tasks: a U-Net-based architecture with approximately 2M parameters for Ion-Pot, achieving $R^2 > 0.98$ while providing $128\times$ speedup over APBS, and a lightweight 39K-parameter model for Ion-Curr, achieving $R^2 > 0.99$ and a speedup of approximately $5700\times$ over the numerical solver.
- For Ion-Pot, we demonstrate that explicitly conditioning on ion concentration improves predictive performance relative to no-ion-input baselines, underscoring the importance of ion-aware learning for electrostatic field prediction.
- For the ionic current prediction task, we find that lightweight models are sufficient to achieve very high accuracy. This suggests that, for certain structured observable-level prediction tasks under physically constrained regimes, high accuracy does not necessarily require high model complexity.

2. Related work

Learning-based surrogate models for PDEs. Learning surrogate models for partial differential equations (PDEs) has attracted significant attention in recent years. Early approaches include physics-informed neural networks (PINNs) (Raissi et al., 2019), which incorporate PDE residuals into the training objective to enforce physical consistency. While PINNs provide a flexible framework for solving both forward and inverse problems, they typically require per-instance optimization and may struggle with high-dimensional or geometrically complex domains (Raissi et al., 2019; Karniadakis et al., 2021).

To enable fast amortized inference across families of PDE problems, more recent work has shifted toward learning the solution operator itself. Neural operator methods learn mappings between function spaces, allowing prediction across varying inputs without retraining. The Fourier Neural Operator (FNO) (Li et al., 2021) parameterizes the solution operator in spectral space, while DeepONet (Lu et al., 2021) uses a branch–trunk architecture to approximate operators from finite-dimensional input encodings. Extensions such as geometry-aware neural operators, including Geo-FNO (Li et al., 2023a), GINO (Li et al., 2023b), and CORAL (Serrano et al., 2023), further extend this paradigm to irregular domains by incorporating learned deformations, point-cloud representations, or coordinate-based neural fields, making them more suitable than earlier PINN-style approaches for complex geometric settings.

Beyond neural operators, a growing line of work develops learning-based PDE surrogates that operate more directly on complex geometries. Graph neural network (GNN) approaches work on unstructured meshes and have shown strong results for fluid dynamics around airfoils (Pfaff et al., 2021), vehicle aerodynamics (Bonnet et al., 2022), and elas-

ticity on irregular domains (Lino et al., 2022). CNN have also been used for PDE surrogate modeling on geometry-varying problems, often through geometry-adaptive mappings or grid-based encodings of irregular domains, as in PhyGeoNet (Gao et al., 2021) and CNN surrogates for aerodynamic flow prediction (Guo et al., 2016; Huang et al., 2020). Point-cloud methods and geometry-aware transformers (Li et al., 2023c; Hao et al., 2023) further generalize to varying domain topologies. Together, these methods reflect a broader trend in PDE surrogate modeling: moving from instance-specific solvers toward learned models that can generalize across both varying inputs and increasingly complex geometric domains.

Biomolecular electrostatics and ion-transport surrogates. Beyond general-purpose PDE surrogate modeling, learning-based approaches have also begun to emerge in biomolecular systems, especially for Poisson–Boltzmann (PB) prediction of biomolecular electrostatics. Benchmark efforts such as SPACIER (Wu & Luo, 2025) provide standardized PB-based datasets spanning proteins, DNA, and small molecules. Building on this setting, the same line of work also introduced the GNN-based model PBGNN (Wu et al., 2025), which achieves high accuracy on surface field and PB-related energy prediction tasks. Other data-driven approaches have explored alternative feature representations: for example, Sliheet et al. (Sliheet et al., 2026) proposed a feature-engineered DNN for predicting PB-related scalar quantities, including reaction-field, Coulombic, and electrostatic solvation energies. Complementing these supervised surrogates, Huang et al. (Huang et al., 2026) showed that dielectric-boundary PB electrostatics can also be solved directly with neural-network trial functions via variational optimization, yielding a physics-informed neural solver rather than an amortized data-driven surrogate.

For ionic current in biomolecular systems, continuum descriptions more commonly arise from ion-transport models such as Poisson–Nernst–Planck (PNP) and Poisson–Nernst–Planck–Stokes (PNPS) equations rather than PB alone (Zhang & Lu, 2021; Mitscha-Baude et al., 2017). Recent machine-learning work in this direction has mainly taken the form of physics-informed neural solvers rather than supervised surrogates: for example, Sun et al. (Sun et al., 2024) developed a PINN framework for coupled electrokinetic microfluidic problems, and Huang et al. (Huang et al., 2025) introduced enriched PINNs (EPINNs) for dynamic PNP systems with adaptive loss weighting and resampling. However, geometry-to-current surrogates over structurally diverse biomolecular systems, particularly ion channels and biological nanopores with highly diverse internal geometries, remain much less explored.

Positioning of this work. Our work sits at the intersection of biomolecular electrostatics and ion-transport surrogate modeling, and emphasizes ion-aware learning across distinct prediction tasks. **Ion-Pot** extends prior Poisson–Boltzmann learning benchmarks by moving from fixed-solvent or local-output settings to full-field electrostatic prediction under varying ionic conditions, with explicit evaluation of generalization across the concentration axis. **Ion-Curr**, in contrast, targets an ion-transport observable rather than a full spatial field, framing ionic current prediction in biologically realistic nanopore and ion-channel-like geometries as a supervised geometry-to-current surrogate problem.

3. Dataset

3.1. Ion-Pot: Ion-aware Protein Electrostatic Potential Dataset

Ion-Pot is a protein electrostatics benchmark derived from AlphaFoldDB (Varadi et al., 2022), designed to evaluate learned surrogates of the Poisson–Boltzmann equation under varying ionic conditions. Starting from up to 10,000 monomeric proteins (molecular mass $\leq 50\,\mathrm{kDa}$), we apply quality filters based on AlphaFold confidence (pLDDT), geometric fit to the computational grid, and solver validity, retaining **3,875** structures (details in Appendix A.1). The distribution of molecular mass is shown in Figure 1.

The electrostatic potential $\phi(\mathbf{r})$ is governed by the nonlinear Poisson–Boltzmann equation (NPBE). In this work, we model the ionic environment as a 1:1 monovalent potassium chloride electrolyte, a common salt condition in biomolecular systems:

$$-\nabla \cdot \big[\varepsilon(\mathbf{r})\,\nabla\phi(\mathbf{r})\big] + \bar{\kappa}^2(\mathbf{r})\,\sinh\big[\phi(\mathbf{r})\big] = \rho_f(\mathbf{r}), \quad (1)$$

where $\varepsilon(\mathbf{r})$ is the position-dependent dielectric ($\varepsilon_{\mathrm{in}} = 2$ inside the protein, $\varepsilon_{\mathrm{out}} \approx 78.5$ in solvent), $\rho_f(\mathbf{r})$ is the fixed charge density, and $\bar{\kappa}^2(\mathbf{r}) = \lambda(\mathbf{r})\,\kappa^2$ is the spatially modulated Debye–Hückel screening parameter. Here $\bar{\kappa}^2 \propto c$ is the bulk inverse-Debye-length-squared at salt concentration c, and $\lambda(\mathbf{r})$ is the ion accessibility function that excludes mobile ions from the protein interior and Stern layer. At $c = 0.00$ M, the ionic term vanishes (Poisson limit); as c increases, the Debye length $\lambda_D = 1/\kappa$ shrinks and the potential is screened more rapidly.

We solve Eq. (1) with APBS at five training concentrations $c \in \{0.00, 0.05, 0.15, 0.45, 0.50\}$ M at 0.5 Å grid resolution, yielding $3{,}875 \times 5 = 19{,}375$ samples. Each NPBE solve costs on the order of 30 s per protein on a single CPU core (29.8 ± 3.5 s; Appendix A), so generating the full multi-concentration label set is a substantial offline cost on the order of 10^2 CPU-hours. This solver cost is what motivates a learned surrogate: once trained, the U-Net predicts a full field in ≈ 0.23 s per protein, a $128\times$ speedup over APBS at matched grid resolution. Each sample is a tuple (x, c, ϕ): a two-channel voxel input $x \in \mathbb{R}^{2\times193\times193\times193}$ (Van der

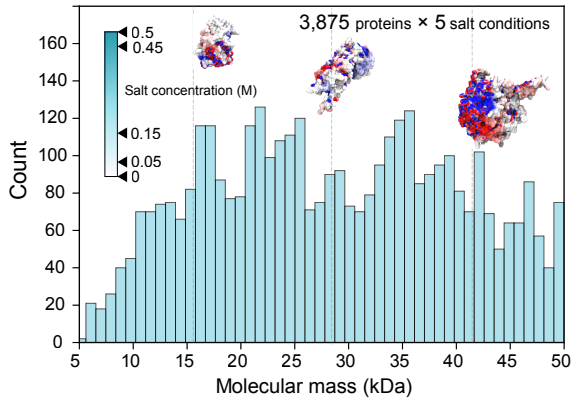

Figure 1. Ion-Pot dataset

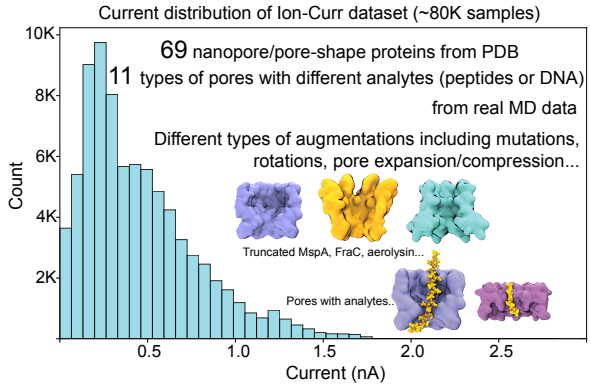

Figure 2. Ion-Curr dataset

Waals accessibility and charge density), the scalar concentration c, and the target potential $\phi \in \mathbb{R}^{1\times193\times193\times193}$.

All three backbones consume this same voxel representation. The protein enters the network only through the two physically derived input channels (van der Waals accessibility and charge density) together with the scalar concentration c; no explicit amino-acid sequence or residue-identity feature is supplied to any architecture. Sequence information is therefore available to the model only implicitly, through its effect on the partial charges assigned by `pdb2pqr` (PARSE force field (Sitkoff et al., 1994), pH 7.4; Appendix A.1) and on the molecular geometry, rather than as a separate sequence input.

Structures are clustered with Foldseek (Van Kempen et al., 2024) by structural similarity, with no sequence-identity cutoff so that even low-identity remote homologs fall in the same cluster; whole clusters are then assigned to train / val / test at 70 / 10 / 20. No structurally similar protein is therefore shared across splits, preventing structural data leakage (**2,712 / 387 / 776** proteins; **13,560 / 1,935 / 3,880** samples; clustering parameters in Appendix A.1).

To assess generalization along the ionic-strength axis, we additionally evaluate on held-out concentrations: *interpolation* at $c \in \{0.10, 0.30\}$ M (within the training range) and *extrapolation* at $c \in \{0.70, 1.00\}$ M.

We provide a detailed description of the Ion-Pot dataset preparation procedure in Appendix A.1. The full dataset is not included with the submission due to its large storage requirement.

3.2. Ion-Curr nanopore-based current dataset

Ionic transport in biomolecular systems is governed by complex electrostatic interactions within heterogeneous three-dimensional environments. Among these systems, biological nanopores and ion channels provide a well-controlled and experimentally relevant setting, where ion transport is strongly modulated by pore geometry, internal structure, and confined analytes (Mayer et al., 2022). Due to their structural diversity and measurable current responses, these systems serve as an effective testbed for benchmarking geometry-to-ion transport mappings.

Ion-Curr is a dataset for learning a direct mapping from 3D voxelized nanopore geometry to ionic current. The ground-truth current is computed using an SEM–FEniCS pipeline, which provides an efficient alternative to fully atomistic molecular dynamics (MD) simulations. While MD-based current estimation typically requires long trajectories for convergence, SEM (Wilson et al., 2019) (i) approximates local ionic conductivity from the distance to nearby biomolecular atoms using a stepwise function derived from simulations that measure conductivity near biomolecular heavy atoms; and (ii) given this spatially varying conductivity field, solves the steady-state conduction equation (Eq. (2)), where σ is the local conductivity and ϕ is the electric potential, using FEniCS to obtain the ionic flux density and resulting current. This approach has been shown to closely match converged MD-derived currents across multiple system types while being orders of magnitude faster. By using an empirical relation between local geometry and conductivity derived from MD data, this framework avoids solving the more complex coupled and nonlinear PNP equations without sacrificing agreement with MD-derived currents.

$$\nabla \cdot \big(\sigma(\mathbf{r})\nabla\phi(\mathbf{r})\big) = 0 \tag{2}$$

The dataset consists of approximately 80K samples constructed from a diverse set of biological nanopores. It includes eight commonly studied nanopores (MspA, CsgG, aerolysin, α-hemolysin, FhuA, CytK, FraC and ClyA), along with 60+ additional pore-shaped proteins from the PDB and their variants, Figure 2. Among these, 11 pore types include analytes (DNA or peptides) threaded through the pore, with configurations derived from dedicated MD simulations.

To increase diversity, we apply several geometry-level aug-

mentations, including rotations around the pore axis, mutations near constriction regions, partial truncation (chopping) of analytes, and controlled expansion or compression of pore geometry. Each augmented sample is assigned the label of its parent nanopore system, so that all samples derived from the same parent system are treated as a single group. We use these parent-system labels when constructing the training, validation, and test splits, ensuring that augmented variants of the same nanopore system do not appear across different splits.

Each sample is represented as voxelized structural features, including distance-to-atom maps, paired with the corresponding ionic current computed by the FEniCS-SEM solver. Ion-Curr labels are generated at a reference condition of 0.2 V bias and 1 M KCl, producing current values spanning approximately 0–2.5 nA and covering the typical operating current range of biological nanopore systems (details in Appendix B.1). Because the SEM formulation models current through an effective conductance, moderate changes in applied voltage or bulk conductivity can be represented by linear rescaling under the corresponding linear-response approximation. In this work, we therefore use Ion-Curr as a geometry- and analyte-to-current surrogate benchmark at a fixed reference condition, rather than as a fully ion-conditioned current benchmark. Large changes in ionic conditions, such as moving to substantially lower salt concentrations (e.g., 0.1 M), may introduce nonlinear electrokinetic effects that require a more detailed solver, such as PNP. Extending Ion-Curr in this direction is left for future work.

This dataset provides a controlled benchmark for evaluating how well machine learning models can map complex 3D biomolecular geometries to ion-transport properties.

4. Experiments and Results

4.1. Ion-Pot

Three backbone models, U-Net (Ronneberger et al., 2015), ResNet (He et al., 2016; Kwon et al., 2021), and FNO (Li et al., 2021), are adopted to evaluate voxel-based prediction performance on the Ion-Pot task. We consider two conditioning mechanisms, additive conditioning (Add) and feature-wise linear modulation (FiLM) (Perez et al., 2018), yielding six conditioned model variants plus three unconditioned baselines (w/o ion).

Specifically, the scalar concentration c is mapped to a feature vector via a multi-frequency sinusoidal embedding followed by a two-layer MLP. Concretely, $c \mapsto \mathrm{SinEmb}_\theta(\alpha c) \to \mathrm{Linear} \to \mathrm{GELU} \to \mathrm{Linear}$, with pre-scaling $\alpha = 10$ and frequency-spread parameter $\theta = 10^4$ (following the Transformer/DDPM convention (Vaswani et al., 2017; Ho et al., 2020). The geometrically-spaced frequency basis maps nearby concentrations to nearby embeddings while exposing multiple frequency bands, which encourages smooth interpolation across the training concentrations. The resulting embedding is applied to the 32-channel intermediate 3D feature map through one of two conditioning operators: *Add*, which broadcasts the embedding spatially and adds it element-wise; or *FiLM*, which splits the embedding into per-channel scale γ and shift β parameters and applies $x \mapsto (1+\gamma)\,x + \beta$.

All models are evaluated using mean absolute error (MAE), root mean squared error (RMSE), weighted absolute percentage error (WAPE), and coefficient of determination (R^2). We adopt WAPE—the ratio of total absolute error to total absolute target magnitude—rather than per-voxel relative error metrics such as mean absolute percentage error (MAPE), because the ground-truth potential becomes near-zero over much of the grid as the Debye length shrinks at high ionic strength, which makes per-voxel relative metrics unstable while leaving WAPE well-defined.

Conditioning ablation. Table 1 reports performance at the unscreened Poisson limit ($c = 0.00\,\mathrm{M}$). Across all three backbones, ionic conditioning yields lower RMSE and higher R^2 than the unconditioned baseline, while the unconditioned baseline shows marginally lower MAE and WAPE. Because $c = 0.00\,\mathrm{M}$ is the unique concentration at which the screening term vanishes, this pattern is consistent with conditioning improving *robustness*: it suppresses larger errors near the protein–solvent interface, where the dielectric boundary is sharpest and where the Debye length $\lambda_D \propto 1/\sqrt{c}$ varies most rapidly with c, without changing average voxel error. Without access to c, the model must learn a single field that averages over all screening regimes. The preferred conditioning operator is backbone-dependent: Add gives the best RMSE and R^2 for U-Net and ResNet, while FiLM is best for FNO. This benefit becomes unambiguous at the held-out non-zero concentrations, where c is informative and conditioning improves every metric over the unconditioned baseline (discussed below).

Interpolation and extrapolation. Table 2 confirms strong interpolation and extrapolation performance: U-Net+Add and ResNet+Add both achieve $R^2 \geq 0.976$ in both regimes, with WAPE in the 15–25% range. For U-Net and ResNet, additive conditioning is uniformly better than FiLM across MAE, RMSE, WAPE, and R^2 in both regimes. For FNO, FiLM slightly outperforms Add on interpolation, while the two are essentially tied on extrapolation. Across backbones, performance is stable from interpolation to extrapolation; for U-Net+Add, WAPE actually improves out of distribution (17.5% interp vs. 15.1% extrap), consistent with the observation that screening makes the prediction problem strictly easier away from $c = 0.00\,\mathrm{M}$. Figure 3 makes this trend

Table 1. Conditioning ablation on the 776 test proteins at $c = 0.00\,\mathrm{M}$ (unscreened Poisson limit). Best result per backbone is shown in **bold**.

Backbone	Cond.	MAE	RMSE	WAPE (%)	R^2
UNet	w/o ion	**1.129**	2.477	**51.651**	0.947
	Add	1.196	**1.975**	54.707	**0.966**
	FiLM	1.163	1.988	53.183	0.966
ResNet	w/o ion	**1.185**	2.497	**54.182**	0.946
	Add	1.270	**2.329**	59.018	**0.952**
	FiLM	1.303	2.675	59.605	0.938
FNO	w/o ion	**0.975**	5.189	**45.587**	0.764
	Add	1.357	4.065	62.053	0.856
	FiLM	1.271	**3.719**	58.109	**0.880**

explicit, showing WAPE and R^2 across the in-distribution ($c = 0$), interpolation, and extrapolation regimes for both convolutional backbones.

Effect of ion conditioning at held-out concentrations. With the unconditioned baseline now evaluated at the held-out concentrations (Table 2), the benefit of ion conditioning is unambiguous away from $c = 0$: the best conditioned variant improves over the no-ion baseline on every metric, for all three backbones and in both regimes. The gains are large and persist under extrapolation; for example, interpolation WAPE drops from 32.7% to 17.5% for U-Net and from 116.5% to 37.3% for FNO, while R^2 rises from 0.964 to 0.989 (U-Net) and from 0.747 to 0.891 (FNO). Additive conditioning beats the baseline on every metric without exception; the only conditioned setting that ever underperforms the baseline is ResNet+FiLM, which improves MAE and WAPE but is marginally worse in RMSE and R^2. This contrasts with the unscreened $c = 0$ limit (Table 1), where the baseline can match conditioning on MAE and WAPE because the screening term vanishes and c carries no information, and confirms that explicit concentration conditioning is what allows a single model to track Debye screening across unseen ionic strengths.

Why $c = 0.00\,\mathrm{M}$ is the hardest regime. The in-distribution result at $c = 0.00\,\mathrm{M}$ ($R^2 \approx 0.966$ for U-Net+Add; Table 1) is lower than the held-out interpolation/extrapolation values ($R^2 \geq 0.985$; Table 2). This is not an evaluation artifact: it reflects that the unscreened Poisson limit is genuinely the hardest regime in the dataset. At $c = 0\,\mathrm{M}$ the potential carries an unscreened long-range $\sim 1/r$ Coulombic tail that spans essentially the entire grid, so the surrogate has to fit a non-trivial signal in every voxel, including the bulk far from the protein. As c increases, the Debye length $\lambda_D \propto 1/\sqrt{c}$ shrinks ($\sim$3 Å at $c = 1\,\mathrm{M}$), so the potential is exponentially screened and is approximately zero over the large majority of the grid; only a thin shell around the molecular surface carries non-trivial signal. The absolute error decreases accordingly: from $c = 0.00\,\mathrm{M}$ to extrapolation at $c = 0.70$–$1.00\,\mathrm{M}$, U-Net+Add MSE drops from $\sim$3.9 to $\sim$1.1, and WAPE drops from $\sim$55% to $\sim$15%. Both magnitude-sensitive metrics agree that screening makes the prediction problem strictly easier. We use WAPE alongside R^2 as the primary metrics for cross-concentration comparisons because WAPE normalizes by the total target magnitude rather than per-voxel relative error, and stays well-behaved across the full concentration range.

Table 2. Interpolation ($c \in \{0.10, 0.30\}$ M) and extrapolation ($c \in \{0.70, 1.00\}$ M, pooled) on the 776 test proteins. `w/o ion` is the unconditioned baseline (no concentration input), reported here at the held-out concentrations. Best per backbone–regime pair among the conditioned variants is shown in **bold**.

Backbone	Eval.	Cond.	MAE	RMSE	WAPE (%)	R^2
UNet	Interp.	w/o ion	0.384	1.982	32.728	0.964
		Add	**0.205**	**1.079**	**17.482**	**0.989**
		FiLM	0.220	1.228	18.743	0.986
	Extrap.	w/o ion	0.403	1.995	36.161	0.964
		Add	**0.169**	**1.059**	**15.143**	**0.990**
		FiLM	0.215	1.267	19.286	0.985
ResNet	Interp.	w/o ion	0.419	2.031	35.674	0.962
		Add	**0.283**	**1.615**	**24.061**	**0.976**
		FiLM	0.376	2.184	32.005	0.957
	Extrap.	w/o ion	0.433	2.040	38.820	0.962
		Add	**0.240**	**1.598**	**21.528**	**0.977**
		FiLM	0.379	2.198	34.018	0.956
FNO	Interp.	w/o ion	1.369	5.270	116.544	0.747
		Add	0.460	3.787	39.156	0.869
		FiLM	**0.439**	**3.464**	**37.348**	**0.891**
	Extrap.	w/o ion	1.412	5.294	126.611	0.744
		Add	0.440	**3.761**	39.473	**0.871**
		FiLM	**0.431**	3.773	**38.620**	0.870

Calibration against naive baselines. To confirm that the high accuracy at large salt concentration reflects genuine learning rather than a trivially easy target, we calibrate against a zero-field predictor $\hat{\phi}(\mathbf{r}) \equiv 0$. By construction this baseline attains $\mathrm{WAPE} = 100\%$ at every concentration (since $\sum_{\mathbf{r}} |\phi - 0| / \sum_{\mathbf{r}} |\phi| = 1$) and $R^2 \leq 0$ (a constant-zero field cannot outperform predicting the mean), independent of the screening regime. Our conditioned surrogates improve over this reference by a wide and roughly constant margin across the full concentration range (U-Net+Add: WAPE from 100% to 15–55%, and R^2 from ≤ 0 to ≥ 0.96), which indicates that the strong high-c numbers are not an artifact of a near-zero ground truth. A stronger, structure-aware reference is the unconditioned (concentration-blind) model itself, now reported at the held-out concentrations

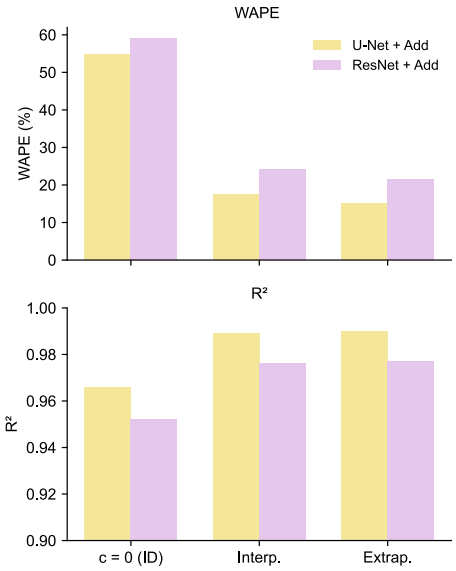

Figure 3. Generalization across the ionic-strength axis for the two convolutional backbones with additive conditioning. WAPE (left) and R^2 (right) are shown for the in-distribution unscreened limit ($c = 0$), held-out interpolation ($c \in \{0.10, 0.30\}$ M), and held-out extrapolation ($c \in \{0.70, 1.00\}$ M). Error is flat or smaller out of distribution: Debye screening makes the target potential decay rapidly and reduces its variance at higher salt. Values are taken from Tables 1 and 2.

in Table 2; ion conditioning improves over it on essentially every metric. Indeed, the unconditioned FNO exceeds 100% WAPE at the held-out concentrations (116.5% interpolation, 126.6% extrapolation), a larger total absolute error than the target magnitude itself and worse in WAPE than the zero-field bound, underscoring that without access to c it cannot adapt to the unseen screening regime. A per-concentration mean-field predictor or an analytic Debye–Hückel screened field would fall between these references and could tighten the calibration further.

Multiscale coupling and the FNO gap. The NPBE around a biomolecule couples three physically distinct scales: a long-range Coulombic tail set by source-charge geometry, a concentration-dependent Debye-screened decay near the protein–solvent interface (λ_D varies from ∞ at $c = 0.00$ M to $\sim$3 Å at $c = 1.00$ M), and fine-grained residue-level charge features on the molecular surface. A useful surrogate must reconstruct all three simultaneously, which makes Ion-Pot a stress test for multiscale inductive bias rather than for raw capacity. FNO achieves $R^2 \geq 0.86$ but consistently underperforms the convolutional models despite being the largest architecture ($\sim$28.3 M complex parameters vs. $\sim$5.8 M for ResNet and $\sim$2 M for U-Net), indicating that the gap is driven by inductive bias rather than model capacity.

Three structural mismatches account for the gap. First, FNO was developed for smooth, spatially homogeneous PDEs on periodic domains (Li et al., 2021); the dielectric jump from $\varepsilon = 2$ (protein interior) to $\varepsilon \approx 78.5$ (solvent) at an irregular molecular surface creates sharp potential gradients that a truncated Fourier basis (12 of 96 unique modes per axis) represents poorly, producing Gibbs-like oscillations near the boundary. Second, the flat single-resolution FNO architecture cannot decompose the three-scale structure described above, whereas U-Net's encoder–decoder hierarchy with skip connections naturally separates coarse and fine scales. Third, the non-periodic boundary conditions (multipole expansion at the grid boundary) and singular point-charge sources are structurally mismatched with the periodic Fourier basis, forcing the representation to waste capacity on edge and singularity artifacts. Together, these observations suggest that for ion-aware electrostatic surrogates with discontinuous coefficients, hierarchical convolutional architectures provide stronger error control than spectral parameterizations of equivalent or larger size.

4.2. Ion-Curr

Ion-Curr is a lightweight 3D CNN model with approximately 39K parameters that maps single-channel voxelized biomolecular geometries to scalar FEniCS-SEM ionic current labels. The architecture is a dilated residual regression network: residual connections are used to support stable feature learning, while dilated convolutions enlarge the receptive field with minimal additional parameters. We use dilation rates of 4 and 6 in the main convolutional blocks, enabling the model to aggregate geometric context across an extended pore region rather than relying only on local voxel neighborhoods.

The model progressively extracts volumetric features, compresses them into a global representation, and predicts the final ionic current with a regression head. This design reflects the physical structure of the task, where current depends on the global conductive pathway, including constrictions, pore continuity, and excluded-volume geometry. Additional architectural details are provided in Appendix B.2. Training is performed using 5-fold GroupKFold cross-validation over nanopore identities.

Table 3. Ablation results for Ion-Curr and its variants.

	MAE (nA) ↓	MAPE (%) ↓	WAPE (%) ↓	R^2 ↑
Ion-Curr	$\mathbf{0.020^{\pm 0.003}}$	$\mathbf{8.2^{\pm 0.8}}$	$\mathbf{4.1^{\pm 0.6}}$	$\mathbf{0.991^{\pm 0.004}}$
No Dilation	$0.037^{\pm 0.008}$	$13.3^{\pm 1.6}$	$7.8^{\pm 1.7}$	$0.978^{\pm 0.010}$

Using the same evaluation metrics as the Ion-Pot model, the performance of Ion-Curr is summarized in Table 3. The model demonstrates strong generalization to held-out test systems, including CytK and several pore-like proteins:

2m6x, 3lj5, 4p1x, 5ngd, 6rb9, 6rhv, and 7t4e. The test set also includes thousands of mutants derived from these seven pores, and some test systems include analytes. Together, these held-out systems cover a current range of 0–2.5 nA and account for approximately 10% of the dataset, comprising 7,839 samples from 8 of the 69 system types. Five-fold cross-validation confirms robustness across training and validation splits, with the mean performance across folds achieving $R^2 > 0.99$. The average MAE is 0.02 nA over the 0–2.5 nA prediction range. MAPE is approximately 8.2% across the full test set, while WAPE decreases to approximately 4% after reducing the disproportionate influence of extreme low-current cases.

Generalization and acceleration of the Ion-Curr model. To further evaluate generalization, we consider two challenging scenarios, as shown in Figure 4. First, we insert an amino acid, either tryptophan or glycine, into CytK (Versloot et al., 2022), a pore included in the test set, and pull it through the pore lumen. The model accurately predicts the residual current, achieving a MAPE of 2.7%, which demonstrates robustness to held-out pores with confined analytes. Second, we perform an MD simulation in a FraC nanopore (Huang et al., 2019), which appears in the training/validation set but is not observed with analytes during training. A poly-aspartic acid peptide with two substitutions, glycine or serine, is placed inside the pore; these two substituted systems differ by only a few atoms. We simulate 70 ns, sample multiple conformations, and compute FEniCS-SEM currents as ground-truth labels. Ion-Curr achieves a MAPE of 4.1% in this setting. These results are interpreted as generalization to numerical solver outputs rather than as a direct validation against physical currents. Direct comparison to MD-derived ionic currents would require substantially longer trajectories, typically on the microsecond scale, because currents estimated from hundreds-of-nanoseconds MD simulations can be highly noisy. This motivates the use of SEM calculations as a stable intermediate target and Ion-Curr as an efficient surrogate for this solver. Prior work has shown consistency between SEM currents and MD currents using microsecond-long trajectories (Wilson et al., 2019); several such systems are included in our dataset.

Together with its lightweight architecture, this accuracy translates into substantial computational acceleration. Ion-Curr provides an approximately 5700× speedup for mapping distance fields to current values compared with the FEniCS-based solver. Runtime for Ion-Curr is measured on an A100 GPU, whereas the solver runtime is measured on an AMD EPYC 7502 CPU. Additional benchmarking details are provided in Appendix B.3.

Ablation study on the Ion-Curr architecture. We conduct an ablation study to evaluate the role of dilated convolutions in the Ion-Curr architecture (Table 3). Removing dilation substantially degrades performance: MAE increases by approximately 85%, while R^2 decreases by about 1.3% relative to the full model. Similar trends are observed for the relative error metrics, with MAPE increasing by approximately 62% and WAPE increasing by approximately 90%. These results indicate that an enlarged receptive field is important for accurate ionic current prediction. This improvement is physically reasonable because dilated convolutions allow the model to aggregate information over a wider spatial context, without substantially increasing the number of parameters or reducing voxel resolution too early. This enables Ion-Curr to capture long-range geometric context that is important for modeling ionic flux through confined biomolecular pores.

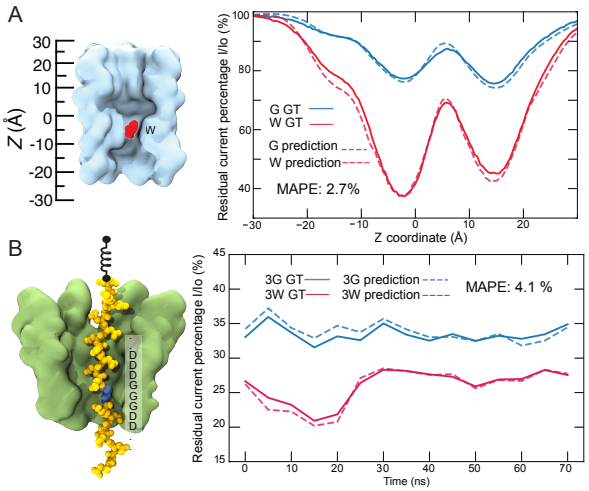

Figure 4. Out-of-box test of Ion-Curr

Efficiency of lightweight models for current-flux prediction. Compared with the models used for full-field reconstruction, Ion-Curr requires only 39K parameters while still achieving strong predictive performance. This suggests a broader hypothesis: the model size needed for a surrogate may scale with the complexity and dimensionality of the target. In particular, predicting a single observable may require far less capacity than reconstructing a full spatial field. In our setting, the target is an integrated current value, so the model does not need to reproduce the entire solution landscape, but only the subset of geometric information most relevant to transport. The relatively structured nature of the problem, including fixed boundary conditions and the conductivity-to-current formulation, likely further reduces the required model complexity.

5. Conclusion

We introduced Ion-Pot and Ion-Curr, two complementary benchmarks for ion-aware surrogate modeling in biomolec-

ular electrostatics. Ion-Pot targets full-field electrostatic potential prediction under varying ionic conditions, while Ion-Curr targets scalar ionic current prediction from complex nanopore geometries. On Ion-Pot, U-Net and ResNet baselines achieve $R^2 > 0.97$ with strong interpolation and extrapolation along the concentration axis, and we find that explicitly conditioning on ion concentration improves accuracy over unconditioned baselines across the held-out ionic strengths, with the largest gains away from the unscreened $c = 0$ limit. On Ion-Curr, a lightweight 39K-parameter CNN achieves $R^2 > 0.99$ and generalizes to unseen pore types and analyte configurations, suggesting that structured observable-level prediction can be potentially accomplished with surprisingly low model complexity. Both models provide orders-of-magnitude speedups over their respective numerical solvers.

Several limitations may need to be addressed in future work. First, Ion-Pot is evaluated mainly with voxel-wise field reconstruction metrics. Although ion conditioning improves prediction over unconditioned baselines, it remains unclear whether better field-level accuracy translates to better downstream physical quantities, such as electrostatic solvation free energies. Evaluating such derived quantities would clarify the practical value of full-field electrostatic surrogates beyond voxel-wise reconstruction. Second, both Ion-Pot and Ion-Curr use fixed-size voxel domains. This design supports controlled benchmarking and efficient training, but limits the representation of larger complexes and spatial features outside the predefined simulation box. Future work could explore adaptive grids, multi-resolution voxel representations, or hierarchical patch-based inference to support larger systems while preserving resolution near molecular surfaces or other key regions. Third, the strong performance of the lightweight Ion-Curr model suggests that observable-level prediction may not always require high model complexity, at least for the SEM-based current formulation studied here. However, this conclusion should be interpreted cautiously. More complex transport settings, such as coupled PNP equations, concentration polarization, or multivalent ions, may require richer architectures or more explicit physical constraints. Testing Ion-Curr-style models under these more demanding regimes would help clarify when lightweight observable predictors are sufficient.

Acknowledgements

We thank Dr. Kumar Sarthak and Dr. Monika Kumari for sharing the nanopore conformational data used in this study. This work was supported by grants from the National Institutes of Health (R21-HG011741 and R01-HG012553). This work also used supercomputing resources provided through ACCESS allocation MCA05S028 and the NCSA DeltaAI system under project XRAS-NCSA-1125.

References

Ayub, M. and Bayley, H. Nanopore-based identification of individual nucleotides for direct rna sequencing. *Nano Letters*, 13(12):6144–6150, 2013.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, 98(18):10037–10041, 2001.

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., and MacKerell, A. D. J. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles. *Journal of Chemical Theory and Computation*, 8(9):3257–3273, 2012.

Bonnet, F., Mazari, J. A., Cinnella, P., and Gallinari, P. AirfRANS: High fidelity computational fluid dynamics dataset for approximating reynolds-averaged navier–stokes solutions. In *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2022.

Chen, D., Chen, Z., Chen, C., Geng, W., and Wei, G.-W. Mibpb: A software package for electrostatic analysis. *Journal of Computational Chemistry*, 32(4):756–770, 2011.

Chen, J., Xu, Y., Yang, X., Cang, Z., Geng, W., and Wei, G.-W. Poisson-Boltzmann-based machine learning model for electrostatic analysis. *Biophysical Journal*, 123(17): 2807–2814, 2024.

Gao, H., Sun, L., and Wang, J.-X. Phygeonet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain. *Journal of Computational Physics*, 428: 110079, 2021.

Guo, X., Li, W., and Iorio, F. Convolutional neural networks for steady flow approximation. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 481–490, 2016.

Hao, Z., Ying, C., Wang, Z., Su, H., Dong, Y., Liu, S., Cheng, Z., Zhu, J., and Song, J. Gnot: A general neural operator transformer for operator learning. *arXiv preprint arXiv:2302.14376*, 2023.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Hille, B. *Ion Channels of Excitable Membranes*. Sinauer Associates, Sunderland, MA, 3 edition, 2001. ISBN 9780878933211.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Huang, G., Willems, K., Soskine, M., Wloka, C., and Maglia, G. Frac nanopores with adjustable diameter identify the mass of opposite-charge peptides with 44 dalton resolution. *Nature Communications*, 10:835, 2019.

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., and Wu, J. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 1055–1059. Ieee, 2020.

Huang, X., Wang, F., Zhang, B., and Liu, H. Enriched physics-informed neural networks for dynamic poisson–nernst–planck systems. *Mathematics and Computers in Simulation*, 237:231–246, 2025.

Huang, Z., Li, B., Wang, Z., and Zhang, Z. Neural network poisson–boltzmann electrostatics for biomolecular interactions. *Journal of Computational Physics*, 544:114446, 2026.

Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, J., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G.-W., Holst, M. J., McCammon, J. A., and Baker, N. A. Improvements to the apbs biomolecular solvation software suite. *Protein Science*, 27(1):112–128, 2018.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

Kwon, I., Jo, G., and Shin, K.-S. A deep neural network based on resnet for predicting solutions of Poisson–Boltzmann equation. *Electronics*, 10(21):2627, 2021.

Li, C., Jia, Z., Chakravorty, A., Pahari, S., Peng, Y., Basu, S., Koirala, M., Panday, S. K., Petukh, M., Li, L., and Alexov, E. Delphi suite: New developments and review of functionalities. *Journal of Computational Chemistry*, 40(28):2502–2508, 2019.

Li, L., Li, C., Zhang, Z., and Alexov, E. On the dielectric "constant" of proteins: smooth dielectric function for macromolecular modeling and its implementation in delphi. *Journal of Chemical Theory and Computation*, 9 (4):2126–2136, 2013.

Li, Z., Kovachki, N. B., Azizzadenesheli, K., liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

Li, Z., Huang, D. Z., Liu, B., and Anandkumar, A. Fourier neural operator with learned deformations for pdes on general geometries. *Journal of Machine Learning Research*, 24(388):1–26, 2023a.

Li, Z., Kovachki, N., Choy, C., Li, B., Kossaifi, J., Otta, S., Nabian, M. A., Stadler, M., Hundt, C., Azizzadenesheli, K., et al. Geometry-informed neural operator for large-scale 3d pdes. *Advances in Neural Information Processing Systems*, 36:35836–35854, 2023b.

Li, Z., Shu, D., and Barati Farimani, A. Scalable transformer for pde surrogate modeling. *Advances in Neural Information Processing Systems*, 36:28010–28039, 2023c.

Lino, M., Fotiadis, S., Bharath, A. A., and Cantwell, C. D. Multi-scale rotation-equivariant graph neural networks for unsteady eulerian fluid dynamics. *Physics of Fluids*, 34(8), 2022.

Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.

Mayer, S. F., Cao, C. J., and Dekker, C. Biological nanopores for single-molecule sensing. *iScience*, 25(4): 103930, 2022.

Mitscha-Baude, G., Buttinger-Kreuzhuber, A., Tulzer, G., and Heitzinger, C. Adaptive and iterative methods for simulations of nanopores with the pnp–stokes equations. *Journal of Computational Physics*, 347:114–133, 2017.

Perez, E., Strub, F., De Vries, H., Dumoulin, V., and Courville, A. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., and Battaglia, P. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*, 2021.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

Ren, P., Chun, J., Thomas, D. G., Schnieders, M. J., Marucho, M., Zhang, J., and Baker, N. A. Biomolecular electrostatics and solvation: A computational perspective. *Quarterly Reviews of Biophysics*, 45(4):427–491, 2012.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

Serrano, L., Le Boudec, L., Kassaï Koupaï, A., Wang, T. X., Yin, Y., Vittaut, J.-N., and Gallinari, P. Operator learning with neural fields: Tackling pdes on general geometries. *Advances in Neural Information Processing Systems*, 36: 70581–70611, 2023.

Sharp, K. A. and Honig, B. Electrostatic interactions in macromolecules: Theory and applications. *Annual Review of Biophysics and Biophysical Chemistry*, 19:301–332, 1990.

Sitkoff, D., Sharp, K. A., and Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *The Journal of Physical Chemistry*, 98(7):1978–1988, 1994.

Sliheet, E., Talha, M. A., and Geng, W. A dnn biophysics model with topological and electrostatic features. *Journal of Chemical Information and Modeling*, 66:3660–3676, 2026.

Sun, R., Jeong, H., Zhao, J., Gou, Y., Sauret, E., Li, Z., and Gu, Y. A physics-informed neural network framework for multi-physics coupling microfluidic problems. *Computers & Fluids*, 284:106421, 2024.

Van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L., Söding, J., and Steinegger, M. Fast and accurate protein structure search with foldseek. *Nature biotechnology*, 42(2):243–246, 2024.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. Alphafold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*, 50(D1):D439–D444, 2022.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Versloot, R. C. A. et al. Beta-barrel nanopores with an acidic–aromatic sensing region identify proteinogenic peptides at low ph. *ACS Nano*, 16(5):7258–7268, 2022.

Wilson, J., Sarthak, K., Si, W., Gao, L., and Aksimentiev, A. Rapid and accurate determination of nanopore ionic current using a steric exclusion model. *ACS Sensors*, 4 (3):634–644, 2019.

Wu, Y. and Luo, R. Spacier: A dataset for modeling electrostatic poisson–boltzmann atomic solvation potentials. In *NeurIPS 2025 Workshop on Machine Learning and the Physical Sciences*, 2025. Workshop paper.

Wu, Y., Zhu, Q., and Luo, R. End-to-end modeling of reaction field energy using data-driven geometric graph neural networks. *Journal of chemical theory and computation*, 21(19):9710–9725, 2025.

Zhang, Q. and Lu, B. An inverse averaging finite element method for solving three-dimensional poisson–nernst–planck model in nanopore system. *The Journal of Chemical Physics*, 155(19):194106, 2021.

Zhou, H.-X. and Pang, X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chemical Reviews*, 118(4):1691–1741, 2018.

A. Ion-Pot: Additional Details

A.1. Dataset Construction

Source and filtering. We query AlphaFoldDB v4 for monomeric proteins with molecular mass ≤ 50 kDa, retrieving up to 10,000 candidate structures. Three successive filters are applied: (i) Confidence: mean per-residue pLDDT ≥ 75 (from the PDB B-factor field), with at least 70% of residues exceeding pLDDT 70 and a minimum length of 30 residues. (ii) Geometric fit: the Van der Waals bounding box must fit within 50 Å per axis to be fully contained in the APBS grid. (iii) Solver validity: after charge assignment with `pdb2pqr` (PARSE force field, pH 7.4), APBS is run and proteins producing non-finite or physically unreasonable output are discarded. After all filters, 3,875 proteins are retained.

APBS parameters. Protein dielectric $\varepsilon_{\text{in}} = 2.0$, solvent dielectric $\varepsilon_{\text{out}} = 78.54$, temperature $T = 298$ K, solvent probe radius 1.4 Å, cation radius 1.76 Å, anion radius 2.27 Å. Ion radii were chosen to match the CHARMM-style Lennard-Jones $R_{\min}/2$ size parameters for K^+ and Cl^- in the CHARMM topology/parameter files (Best et al., 2012).

The computational domain is a cubic grid of 193^3 points (96 Å per side at 0.5 Å spacing) with multiple Debye–Hückel boundary conditions, charge discretization via cubic B-spline interpolation (`chgm=spl0`), and a smoothed molecular surface (`srfm=smol`, surface density 10.0).

Foldseek clustering. The 3,875 structures are clustered with Foldseek (Van Kempen et al., 2024) (`foldseek cluster`; `--cov-mode 1 -c 0.5 -e 1e-3 --min-seq-id 0.0`): two structures are placed in the same cluster when their structural alignment is significant ($E \leq 10^{-3}$) and covers at least 50% of the target structure, with no minimum sequence identity so that structurally similar but low-identity remote homologs are also grouped. Clusters are assigned to train / val / test via greedy packing at 70 / 10 / 20, ensuring no cluster is shared across splits and hence no structural leakage into validation or test.

A.2. Model Architectures and Training

U-Net. We use the MONAI BasicUNet (Ronneberger et al., 2015) with 3D spatial convolutions, 5 encoder–decoder levels, and feature channels $(32, 32, 64, 64, 128)$ across five encoder–decoder levels, followed by a 32-channel output block and a $5 \times 5 \times 5$ convolution to the scalar potential. Skip connections at each level carry high-frequency detail directly from the contracting to the expanding path ($\sim$2 M parameters).

ResNet. We use a 3D residual network (He et al., 2016) with reflection-padded $3 \times 3 \times 3$ convolutions, instance normalization and ReLU: a stem maps the 32-channel input embedding to width 64, a stride-2 convolution downsamples to width 128, six residual blocks operate at the bottleneck, and a stride-2 transposed convolution followed by a final $3 \times 3 \times 3$ projection recovers a single-channel output ($\sim$5.8 M parameters).

FNO. We use a 3D Fourier Neural Operator (Li et al., 2021) with width 32, depth 4, and 12 Fourier modes retained per spatial axis. Each FNO layer computes $\text{GELU}(\mathcal{F}^{-1}[R \cdot \mathcal{F}(x)] + Wx)$, where R is a learned complex weight tensor in Fourier space (retaining modes ≤ 12 per axis) and W is a pointwise $1 \times 1 \times 1$ convolution. Normalized 3D grid coordinates $(x, y, z) \in [0, 1]^3$ are appended as additional input channels. Zero-padding of 6 voxels is applied before the spectral layers and removed afterward to mitigate edge artifacts from the non-periodic domain. The 4 spectral layers each store four complex weight tensors of shape $(32 \times 32 \times 12 \times 12 \times 12)$, yielding $\sim$28.3 M complex parameters—roughly $10\times$ more than ResNet and $25\times$ more than U-Net.

Training. All models are trained with Smooth L1 loss ($\beta = 0.001$), AdamW (lr $= 5 \times 10^{-4}$, cosine-annealed to 10^{-6}), batch size 4, for 24,000 steps.

A.3. Benchmarking speed of the model and solver

We benchmark inference throughput on a single NVIDIA A100-SXM4-40GB GPU using 193^3 input grids (96 Å at 0.5 Å spacing) with 5 warmup and 50 timed batches. The APBS numerical solver (NPBE, mdh boundary conditions) was benchmarked on 50 representative proteins using an AMD EPYC 7502 CPU. Ion-Pot-UNet achieves a 128× speedup over APBS (232 ms vs 29.8 ± 3.5 s per protein) with only 2 M parameters and 9.0 GB peak memory. All three architectures—UNet, FNO (28.3M), and ResNet (5.8M)—achieve two orders of magnitude acceleration while maintaining high accuracy. At

batch size 1, UNet is the fastest (232 ms), followed by FNO (300 ms) and ResNet (335 ms). Memory scales linearly with batch size;

B. Ion-Curr: Additional Details

B.1. Ion-Curr dataset preparation

The Ion-Curr dataset contains eight typical biological nanopores and pore-shaped proteins (full list in Table 4). For each structure, we first truncate regions far from the pore constriction to obtain a consistent system size that fits within a $60 \times 60 \times 90$ Å^3 voxel grid. Such truncation is commonly used to reduce system size and has been shown to have minimal impact on electrostatic potential near the constriction and the resulting ionic current.

All pores are embedded in a POPC lipid membrane. For open-pore systems (without analytes), we apply three types of geometry-level augmentations (see Figure 5):

(1) Mutation. Random mutations are introduced near the constriction region using amino acids (excluding proline, cysteine, and selenocysteine), generating hundreds of variants per pore.

(2) Expansion/Compression. Approximately one quarter of the mutated samples are selected for geometric deformation. For each selected sample, pore atoms are spatially scaled along the X and Y axes using factors ranging from 0.6 to 1.3. Different scaling factors may be applied along the two axes, allowing the augmented structures to mimic expansion, compression, or anisotropic deformation of the nanopore cross-section.

(3) Mixed deformation. For another subset of mutated samples, the pore is partitioned into two or three segments along the z-axis, and each segment is independently expanded/compressed/deformed using scaling factors in the range 0.6–1.3.

Among all pores, 11 types include analytes (DNA or peptides) threaded through the pore. These systems are derived from real MD simulations. For example, MspA–DNA and aerolysin systems are obtained from published nanopore MD trajectories, while the remaining systems are generated using a consistent protocol. Specifically, each pore is embedded in a POPC membrane, solvated with water and ions, and equilibrated for 20 ns. A stretched peptide (10–20 residues, randomly generated excluding proline, cysteine, and selenocysteine) is then inserted into the pore, followed by an additional 10 ns simulation to allow free movement and generate diverse current blockage configurations. Only the final frame of each trajectory is included in the dataset.

For pore–analyte systems, we further apply additional augmentations (see Figure 5):

(4) Rotation. Structures are rotated around the z-axis with random angles between 10° and 30°.

(5) Chop. The analyte is truncated to shorter segments (3–6 residues), with random starting positions along the sequence.

These procedures result in a diverse dataset capturing variations in pore geometry, local structure near constrictions, and analyte configurations.

B.2. Ion-Curr architecture and training details

Architecture. Ion-Curr uses a lightweight three-dimensional convolutional regression network that maps a single-channel volumetric input to a scalar current prediction. The model consists of two dilated residual stages followed by global feature aggregation and a small fully connected regression head.

The input volume first passes through a dilated residual block that maps the input from 1 channel to 12 channels. This block contains two sequential $3 \times 3 \times 3$ convolutions. The first convolution uses dilation 4 with padding 4, followed by batch normalization and a ReLU activation. The second convolution uses dilation 2 with padding 2, followed by batch normalization. The residual shortcut uses a $1 \times 1 \times 1$ convolution followed by batch normalization to match the channel dimension. The shortcut output is then added to the main branch, followed by a final ReLU activation. A $2 \times 2 \times 2$ max-pooling layer is applied after this block for downsampling.

The second dilated residual block follows the same structure, but maps channels $12 \rightarrow 24$ and uses dilation rates 6 and 3 for the two $3 \times 3 \times 3$ convolutions, respectively, and followed by a similar shortcut. The block output is passed through a final ReLU activation after adding the shortcut.

After the two residual stages, adaptive average pooling reduces the feature map to $2 \times 2 \times 2$. A final $2 \times 2 \times 2$ convolution

Table 4. Nanopore and pore-like proteins included in Ion-Curr. Underlined entries indicate that the type includes systems with analytes inside.

Biological nanopores			
ClyA	FhuA	FraC	Aerolysin
α-hemolysin	CsgG	CytK	MspA
Pore-like proteins / PDB-derived systems			
1dsx	1hxt	1ldf	1oed
1tlz	2a9h	2j58	2jqy
2lhf	2m6i	2m6x	2r4p
2rb9	2wmz	2x00	3b07
3lj5	3o44	3szd	3w9t
4bum	4gq1	4h56	4khr
4plx	4mto	4twl	4uoc
5ec5	5gaq	5ngd	5svu
5zug	6bbg	6dg7	6g8z
6h2f	6hjy	6m6z	6oqh
6rb9	6rhv	6uls	6v03
7bgh	7jn0	7olq	7q9y
7qn6	7re5	7rqw	7t4e
7ux1	7xqh	8j4i	8s90
8uzl	8zar	9eym	9ezc
9fdg			

maps the 24 channels to 48 channels and collapses the spatial dimensions to $1 \times 1 \times 1$, followed by ReLU and flattening into a 48-dimensional vector. The regression head then maps this vector through a two-layer MLP ($48 \rightarrow 24 \rightarrow 1$), with ReLU after the first linear layer and no activation at the output.

Training. Models were trained for 90 epochs using the Adam optimizer with an initial learning rate of 1×10^{-3}. The learning rate was linearly decayed by 9×10^{-6} per epoch. Training was performed with a batch size of 64 and optimized using the MAE loss. Mixed-precision training was used throughout.

To prevent data leakage, samples were split by nanopore identity. Specifically, 10% of pore labels were held out as a fixed test set, and the remaining 90% were partitioned into 5 folds using GroupKFold cross-validation. Each fold was trained independently using a same random seed. For each fold, the checkpoint with the lowest validation loss was selected and used for final evaluation on the test set.

B.3. Benchmarking speed of the model and solver

The inference throughput of the Ion-Curr model was benchmarked on a single NVIDIA A100-SXM4-80GB GPU. Evaluation across three batch sizes (64, 128, and 256) utilized a protocol of 5 warmup and 50 timed batches with CUDA-synchronized wall-clock timing. This configuration yielded throughputs of 570.5, 555.7, and 527.4 samples/s, with peak GPU memory consumption of 5,063 MB, 10,119 MB, and 20,239 MB, respectively. In contrast, the equivalent SEM calculation, which calculates current from distance maps, requires approximately 10s per sample on an AMD EPYC 7502 CPU. This FEniCS-based finite-element solver, which computes the Laplace/Poisson equations on a 3D conductivity field, demonstrated highly consistent CPU-bound performance, with latency ranging from 8.17s to 11.57s over 113 runs.

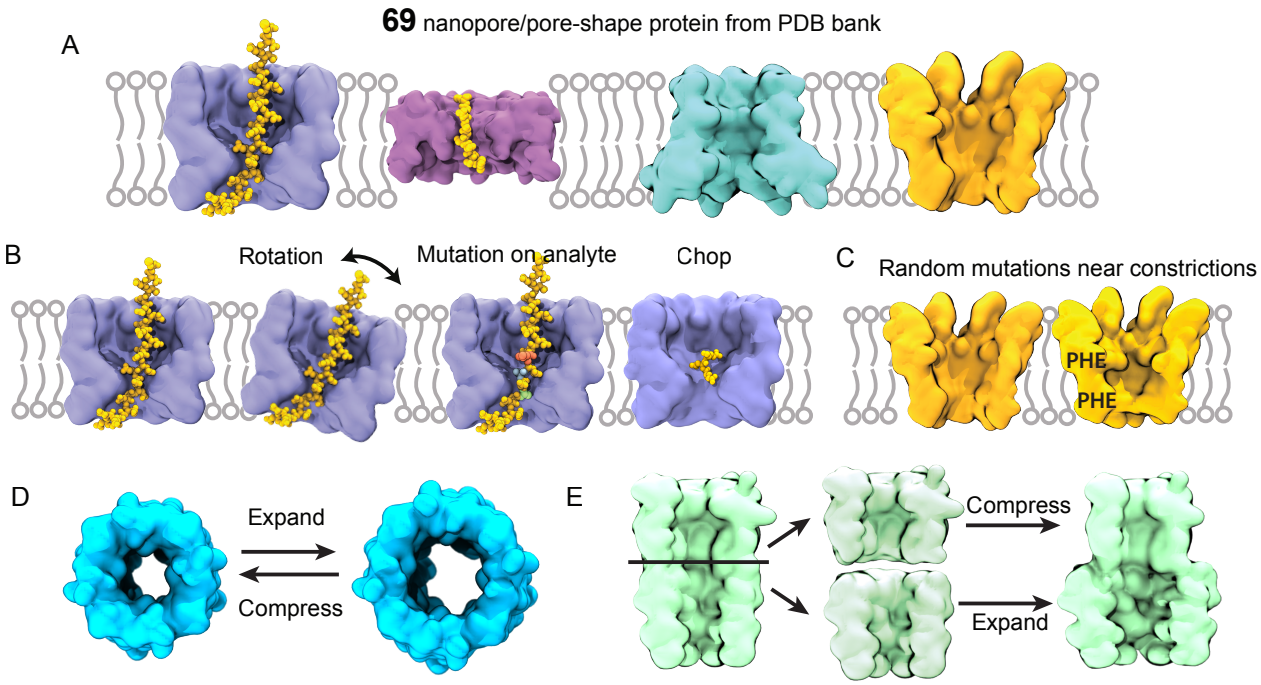

Figure 5. Data augmentation approaches for Ion-Curr