The Journal of Chemical Physics

RESEARCH ARTICLE | FEBRUARY 03 2026

Achieving all-atom molecular dynamics accuracy from the Poisson–Boltzmann method through machine learning

Special Collection: Festschrift in honor of Christoph Dellago: Exploring Paths and Barriers in Statistical Mechanics

Ema Slejko ; Amaury Coste ; Tilen Potisk ; Julija Zavadlav ; Matej Praprotnik

J. Chem. Phys. 164, 054107 (2026)
https://doi.org/10.1063/5.0313624

The Journal
of Chemical Physics

ARTICLE

pubs.aip.org/aip/jcp

Achieving all-atom molecular dynamics accuracy from the Poisson–Boltzmann method through machine learning

Cite as: J. Chem. Phys. 164, 054107 (2026); doi: 10.1063/5.0313624
Submitted: 24 November 2025 • Accepted: 20 January 2026 •
Published Online: 3 February 2026

Ema Slejko,[1,2] **Amaury Coste,**[1,3] **Tilen Potisk,**[1,2] **Julija Zavadlav,**[4] **and Matej Praprotnik**[1,2,5,a)]

AFFILIATIONS

[1] Theory Department, National Institute of Chemistry, SI-1001 Ljubljana, Slovenia
[2] Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia
[3] Université Grenoble Alpes, CEA, IRIG-MEM-LSim, FR-38054 Grenoble, France
[4] Professorship of Multiscale Modeling of Fluid Materials, TUM School of Engineering and Design, Technical University of Munich, DE-85748 Garching near Munich, Germany
[5] Universitat de Barcelona Institute of Complex Systems, ES-08028 Barcelona, Spain

Note: This paper is part of the Special Topic Festschrift in Honor of Christoph Dellago: Exploring Paths and Barriers in Statistical Mechanics.

[a)] **Author to whom correspondence should be addressed:** praprot@cmm.ki.si

ABSTRACT

All-atom molecular dynamics (MD) simulations are a standard tool for probing the structural and dynamical properties of biomolecular systems, but their accuracy comes at the cost of high computational demands. To overcome spatial–temporal limitations, implicit models or coarse-graining are often employed, but usually at the expense of reduced accuracy. This limitation is also evident in the Poisson–Boltzmann (PB) mean-field theory, which efficiently captures long-range electrostatics but fails to account for crucial short-range interactions. In this work, we bridge this gap by introducing a graph neural network (GNN) Δ-learning approach trained on the difference between all-atom MD and PB, resulting in DIS-PB (deep implicit solvation model using the PB potential as a prior). DIS-PB, which models solutes and salt ions explicitly by MD while water is coarse-grained out, captures both short-range electrostatic correlations as well as long-range electrostatic interaction tails. Applied to a system of the DNA molecule in 1 mol l^{-1} salt solution, our method reproduces structural properties (NDPs, RDFs, and binding probability patterns) with high fidelity, showing that the GNN-corrected PB can reach the accuracy of all-atom MD at a lower computational cost.

I. INTRODUCTION

Electrostatic interactions play a dominant role in numerous biological processes due to their long-range nature and the high charge density of biomolecules such as proteins and nucleic acids.[1] The interactions between nucleic acids, such as DNA and RNA, and counterions are essential for their proper functioning. Counterions help neutralize the negatively charged phosphate groups along the nucleic acid's backbone. In addition to binding at specific sites and forming long-range interactions, counterions also interact diffusely, creating an "ion atmosphere" around such biomolecules.[2–4] Therefore, accurately modeling their electrostatic properties requires careful consideration of the underlying computational models used to study electrostatic interactions.

Two primary types of models are used in computational electrostatics: explicit solvent models and implicit solvent models. Explicit solvent models account for the full degrees of freedom of solvent molecules (typically water) and ions, which offers a more accurate representation of the system. However, these models come with high computational costs because the computational complexity scales with the number of atoms involved, making them unsuitable for large systems or long timescales. Alternatively, coarse-graining (CG) modeling can be employed to reduce the computational overhead.[5–9] This approach reduces the degrees of freedom in the simulation by either grouping correlated atoms into effective interaction sites or by representing components of the system using a simplified implicit model. Implicit methods treat the solvent as a polarizable continuum and counterions as a mean-field charged "cloud."[10] Implicit models drastically reduce the number of explicit particle interactions, thereby lowering the computational cost, as the solvent usually comprises more than 90% of the system. Some of the commonly employed implicit methods include PB,[11,12] COSMO/polarized continuum model,[13] the generalized Born model,[14–16] or calculating the effective potential between solutes.[17] We will focus on the PB method, which has been extensively used for various applications, such as calculating the electrostatic potentials surrounding DNA and RNA,[18,19] examining salt effects on ligand and protein binding to DNA,[20–22] studying nucleic acid pK_a shifts,[23] and ion distributions around nucleic acids.[24–28] Despite its widespread use, the PB model treats ions as an ideal gas within a constant dielectric medium, which limits its ability to account for ion–ion correlations and restricts its use in capturing specific microscopic interactions.

In contrast to the PB model, all-atom MD simulations can provide a wealth of microscopic and dynamic information about biological systems.[29–43] These simulations rely on force fields (FFs) derived from empirical data and *ab initio* calculations. The quality of employed FF constrains the accuracy of the simulation results. MD simulations have been used to investigate potential correlations between monovalent cations binding to DNA, structural perturbations, nucleic acid stability, and nucleic acid–counterion interactions.[44–48] The most widespread FFs are the so-called classical force fields, which use a combination of fixed-charge Coulomb potential and Lennard-Jones (LJ) interactions to model the intermolecular potential. A more detailed and explicit representation of system particles, including solvent, ions, and solute, results in significant computational complexity and time inefficiency. To address these limitations, multiscale simulations offer a practical solution by combining the accuracy of all-atom MD simulations with the efficiency of faster, less detailed methods such as CG modeling or implicit solvent models. The adaptive resolution simulation scheme[49–52] allows dynamic switching between all-atom, CG,[53] or implicit[54] representations of solvent and ions, providing an efficient way to simulate large systems while maintaining a high level of accuracy in key regions of the system, such as the nucleic acid or protein. This approach has been applied to simulate a DNA molecule[52,54] efficiently and to investigate the mechanisms underlying phase transitions in high-density DNA arrays.[55,56]

In recent years, machine learning (ML) potentials have been developed to bridge the gap between the computational efficiency of force fields and the high accuracy of more complex *ab initio* methods.[57] These models use different ML architectures (such as deep neural networks, graph networks, or kernel methods) to accurately fit the potential energy surface (PES) of the *ab initio* system, with a computational cost comparable to classical FFs.[58–69] Various approaches, including iterative Boltzmann inversion,[70,71] force matching,[72–74] and relative entropy,[75] have been used to approximate the potential of mean-force with high-dimensional energy functions, particularly for modeling non-bonded interactions in systems such as aqueous salt solutions[76,77] and soft matter.[71] ML models represent the PES as a sum of atomic terms, capturing many-body interactions among atoms within a specified cutoff distance. The incorporation of 3-body non-bonded interactions has been demonstrated to influence both structural and dynamic properties in interatomic interactions[78–80] as well as between CG particles.[81] The construction of ML-based PES is computationally efficient, as its complexity scales linearly with system size. However, it involves significant approximation since electrons and nuclei, being charged particles, experience long-range Coulomb interactions. For isotropic systems such as water solutions, the ML potential without an explicit electrostatic term provides accurate results.[82] In contrast, long-range electrostatic interactions are critical for accurately modeling anisotropic systems such as battery materials.[83] Despite this, short-range ML models are trained on *ab initio* methods, which include long-range electrostatics, suggesting that short-range models can approximate certain long-range effects. To better account for long-range interactions, various ML architectures have been developed, such as charge equilibration in the fourth generation potential,[67,84–86] the long-distance equivariant framework LODE,[57,87] Ewald summation for long-range interactions,[88,89] a long-range ML framework for computing electrostatic potentials by long-range descriptors,[90] virtual-node message-passing GNN,[91] etc. A common strategy to explicitly incorporate long-range electrostatics within ML models is to decompose the PES into two components: one representing short-range interactions, modeled using standard ML techniques, and the other capturing long-range interactions through a predefined, physics-based potential. This method is referred to as Δ-learning.[92–94]

ML algorithms can generate both all-atom and CG models[95–97] successfully replicating structural[98–104] and dynamical[105–107] properties. While many studies have focused on pure water solvents or treated ions implicitly, ions are critical in biological processes such as protein binding and nucleic acid folding.[4,108,109] Due to the growing interest in liquid electrolytes for lithium-ion batteries, ML force fields are also being applied in the MD of such systems.[110–112] However, ML potentials that explicitly include ion interactions often rely on all-atom models, which also represent solvent molecules in full atomic detail.[113–117]

Unlike previously described ML methods that refine all-atom MD results to *ab initio* levels,[118] our study aims to bridge the accuracy–efficiency trade-off between the PB model and the all-atom MD. Our reference data consist of all-atom MD simulations of a DNA molecule in aqueous salt solution. As a baseline, we construct a prior potential from the PB electrostatic field of a CG DNA representation. While the PB approach efficiently captures long-range electrostatics, it neglects essential short-range interactions. To address this limitation, we introduce an ML potential trained via Δ-learning on the difference between all-atom and PB forces for a system containing the CG DNA and ions (but no water). The ML component records short-range physics, including ion–ion correlation, hydration effects, and collective behavior. In the final DIS-PB simulations, the system contains only the CG DNA and ions, with water omitted, and combines the ML and PB contributions into an efficient yet accurate alternative to all-atom MD. The workflow of the model is depicted in Fig. 1. Our workflow builds directly on the previously introduced DIS framework.[119] In the present work, we retain the same overall learning strategy but modify the formulation of the prior model: whereas DIS employed a Lennard-Jones–Coulomb

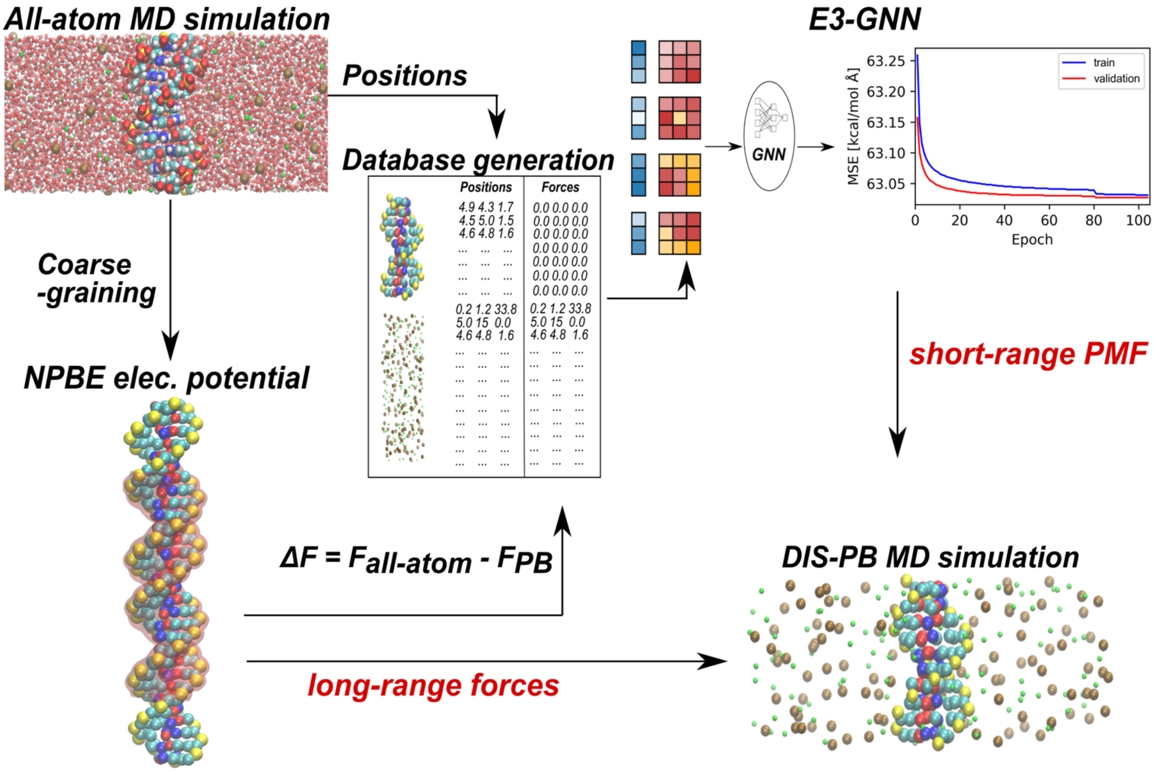

FIG. 1. DIS-PB model workflow. The all-atom MD simulation box and DIS-PB MD simulation box show the cross sections of a periodic DNA molecule in a 1 mol l^{-1} sodium chloride salt solution, with sodium ions shown in green and chloride ions in ochre. The all-atom model includes water molecules and DNA with carbon, nitrogen, phosphate, oxygen, and hydrogen atoms depicted in cyan, blue, yellow, red, and white, respectively. The DIS-PB model contains only ions and the CG DNA molecule (types and colors of atoms match the all-atom representation). The PB electrostatic potential is calculated with APBS for an elongated CG DNA molecule. For database generation, ionic positions are obtained from an all-atom MD simulation, while forces are defined as the difference between the all-atom MD forces and the PB forces. These data are the input for the GNN Allegro, which is trained with a force-matching method. In the DIS-PB model, MD simulations combine a short-range ML potential with long-range contributions calculated as an additional ionic force from the PB potential.

reference potential, DIS-PB replaces this prior with a PB electrostatic potential. This change is motivated by limitations identified in the original DIS method and is intended to assess whether a PB-based prior can improve the efficiency–accuracy balance. To evaluate the accuracy of our approach, we compare several structural and statistical properties, including ion–DNA normalized density profiles (NDPs), ion–ion radial distribution functions (RDFs), binding pattern probabilities, cumulative distribution functions, and excess ion numbers, across all-atom MD, DIS-PB simulations, and simulations with the standard (non-Δ) ML potential.

II. METHODS

A. All-atom MD simulation/database generation

For the all-atom MD simulation, we simulate a 10 base-pair DNA duplex with the sequence d(CTCTCGAGAG)$_2$ using a periodic box (89.7 × 80.7 × 33.99 Å^3) containing TIP3P[120] water and an appropriate number of monovalent cations (Na^+) and anions (Cl^-), inserted by replacing water molecules to construct a 1 mol l^{-1} salt solution. The AMBER 03[121] force field is used for the DNA molecule, and the Joung and Cheatham parameters[122] are applied for ions, with additional corrections for ion–phosphate interactions.[123] MD simulations are performed using LAMMPS,[124] integrating Newton's equations of motion with the velocity Verlet algorithm and a 1 fs time step.[125,126] In this work, we employ periodic boundary conditions along the helical (z) axis, where corresponding intramolecular DNA interactions—defined by bond, angle, and dihedral potentials—are used to connect each strand to its periodic image. This setup effectively mimics an infinitely long DNA molecule while maintaining structural continuity across periodic boundaries. Throughout the simulations, the DNA molecule is kept rigid. After a 2 ns equilibration in the NPT ensemble with the Nosé–Hoover temperature thermostat[127] (a coupling constant of 0.1 ps) and Nosé–Hoover pressure barostat (a coupling constant of 0.1 ps), the system is heated for 10 ns in the NVT ensemble at the temperature of 300 K. The production run is carried out in an NVT ensemble at a temperature of 300 K. Non-bonded van der Waals and electrostatic interactions between all particles are treated with a hybrid potential composed of LJ potential and Coulomb potential within cutoff distances of 0.9 and 1.2 nm, respectively. The electrostatic interactions beyond the cutoff are corrected with the PPPM solver.[128]

B. Poisson–Boltzmann potential/prior

1. Solving the nonlinear Poisson–Boltzmann equation

The PB equation is a second-order elliptic partial differential equation that models the electrostatic potential around a fixed charge distribution in an ionic solution. To accurately compute the electrostatic potential, three components must be considered: the solute molecule, the solvent, and the solvated ions. The Poisson's equation,

$$-\nabla\cdot[\epsilon(\boldsymbol{r})\nabla\Phi(\boldsymbol{r})]=\rho(\boldsymbol{r}) \quad \text{for } \boldsymbol{r}\in\Omega, \quad \text{where} \quad \Phi(\boldsymbol{r})=b(\boldsymbol{r}) \quad \text{for } \boldsymbol{r}\in\partial\Omega, \tag{1}$$

describes the electrostatic potential $\Phi(\boldsymbol{r})$ generated by charge distribution $\rho(\boldsymbol{r})$ in a polarizable continuum. Here, $\epsilon(\boldsymbol{r})=\epsilon_0\epsilon_r(\boldsymbol{r})$ is a spatially dependent permitivity, written in terms of the vacuum permitivity ε_0 and relative permitivity $\epsilon_r(\boldsymbol{r})$. Equation (1) is generally solved on some finite domain Ω, with the potential specified as $b(\boldsymbol{r})$ on the domain boundary $\partial\Omega$. Typically, Dirichlet boundary conditions are employed to calculate $b(\boldsymbol{r})$, as it is an analytic and asymptotically correct form of the potential (Coulomb's law or Debye–Hückel).

In a biomolecular system, it is useful to consider two types of distributions. First, the biomolecule (solute), characterized by partial atomic charges, is typically represented as a fixed charge distribution,

$$\rho_f(\boldsymbol{r})=\sum_{i=1}^{M}Q_i\delta(\boldsymbol{r}-\boldsymbol{r}_i), \tag{2}$$

where each of the M atomic partial charges $\{Q_i\}$ is represented by a delta function $\delta(\boldsymbol{r}-\boldsymbol{r}_i)$ centered at the corresponding atomic position $\{\boldsymbol{r}_i\}$. Second, the contributions of the counterions are modeled in a continuous (mean-field) approximation by a Boltzmann distribution, which defines the mobile charge distribution,

$$\rho_m(\boldsymbol{r})=\sum_{j}^{m}c_jq_j\exp\left[-\beta(q_j\Phi(\boldsymbol{r})+V_j(\boldsymbol{r}))\right], \tag{3}$$

for m counterion species with charges $\{q_j\}$, bulk concentrations $\{c_j\}$, and steric potentials $\{V_j\}$ (i.e., potentials that prevent biomolecule–counterion overlap). In Eq. (3), $\beta=1/\mathrm{kT}$ is the inverse thermal energy, with k being the Boltzmann constant and T being the temperature. In the case of a one-to-one electrolyte, Eq. (3) reduces to

$$\rho_m(\boldsymbol{r})=\overline{\kappa}^2(\boldsymbol{r})\sinh\left[\Phi(\boldsymbol{r})\right], \tag{4}$$

where coefficient $\overline{\kappa}^2(\boldsymbol{x})$ describes both ion accessibility (indirectly via $\exp[-\beta V(\boldsymbol{x})]$) and bulk ionic strength. Combining the expressions for the fixed biomolecule [Eq. (2)] and mobile counterion [Eq. (4)] distributions with Poisson's equation [Eq. (1)] yields the nonlinear PB equation (NPBE) for the one-to-one electrolyte,

$$-\nabla\cdot[\epsilon(\boldsymbol{r})\nabla\Phi(\boldsymbol{r})]+\overline{\kappa}^2(\boldsymbol{r})\sinh\left[\Phi(\boldsymbol{r})\right]=\sum_i q_i\delta(\boldsymbol{r}-\boldsymbol{r}_i) \quad \text{for } \boldsymbol{r}\in\Omega, \quad \text{where} \quad \Phi(\boldsymbol{r})=b(\boldsymbol{r}) \quad \text{for } \boldsymbol{r}\in\partial\Omega, \tag{5}$$

and the corresponding electric field,

$$\boldsymbol{E}(\boldsymbol{r})=-\nabla\Phi(\boldsymbol{r}). \tag{6}$$

The coefficients of the PB equation—the relative permittivity $\epsilon(\boldsymbol{r})$ and ion accessibility function $\overline{\kappa}^2(\boldsymbol{r})$—depend on the structure of the biomolecule (atomic coordinates and radii), as well as on the ionic radius, and are constructed using standard surface-based definitions as implemented in the APBS software. Fixed charge distribution is represented by atomic charges and atom locations that appear in the delta functions as the set of points $\{\boldsymbol{r}_i\}$.

In particular, $\epsilon_r(\boldsymbol{r})$ is discontinuous along the biomolecular surface and assumes low solute dielectric values (between 2 and 20) inside and high bulk solvent values (generally near 80) outside the surface. The ion accessibility function $\overline{\kappa}^2(\boldsymbol{r})$ is defined using the inflated van der Waals surface—the union of spheres centered at atomic positions with radii equal to the atomic van der Waals radius plus the counterion radius. This ion accessibility field can also be used to derive a steric exclusion contribution to the PB potential, which prevents ions from penetrating the biomolecular volume. All details of these constructions are standard and follow the APBS implementation described in Ref. 10.

In this work, we do not modify the underlying definitions of $\epsilon_r(\boldsymbol{r})$ and $\overline{\kappa}^2(\boldsymbol{r})$, but instead examine the sensitivity to selected parameters such as ion radius, grid spacing, and solute dielectric constant that significantly influence their evaluation and, consequently, the overall result of PB potential, as shown in Fig. 4.

The PB equation may only be solved analytically for simplified models, requiring numerical simulation for more complex geometries.[129] To solve the NPBE, the APBS[130–132] software is used, employing the finite-difference method. This method discretizes the PB equation on a non-uniform Cartesian grid, converting continuous functions and operators into discrete counterparts. This transformation yields a system of linear equations that can be efficiently solved via matrix inversion.[129]

We use the extended CG DNA model as the input structure for the APBS calculations. The CG mapping, based on the CG DNA model by Kovalevá *et al.*,[133] represents each nucleotide with six CG sites. The all-atom DNA model contains 634 atoms, whereas the input to the PB and ML models includes only 120 explicit DNA atoms. Although solving the PB equation with the all-atom DNA model is not substantially slower than with the CG DNA model, we choose to use the CG model for both PB and ML calculations to maintain clarity in the workflow. As shown in Fig. 2, using the all-atom DNA model does not significantly improve the PB solution; therefore, the choice of model has little impact on the results. The PB potential naturally decreases with distance from the DNA molecule in the x and y directions. By choosing sufficiently large lateral dimensions for both the MD and APBS boxes, the potential at their boundaries converges to zero, minimizing artifacts from periodicity. Along the z axis, however, we aim to model an infinitely long DNA strand, so the MD box spans exactly one DNA pitch, as described in Sec. II A. As shown in Fig. 3, the APBS potential does not decay along z. To prevent inconsistencies between the finite APBS solution—which would otherwise truncate the potential at the DNA ends—and the periodic boundary conditions in MD, the CG model in APBS is extended in both the $+z$ and $-z$ directions to include three helical turns of DNA. This setup ensures that the PB potential is properly evaluated at the boundaries of the MD box along the z axis, making it fully consistent with the periodic representation of an infinitely long DNA molecule in the MD simulation. Choosing the optimal computational model for PB calculations with

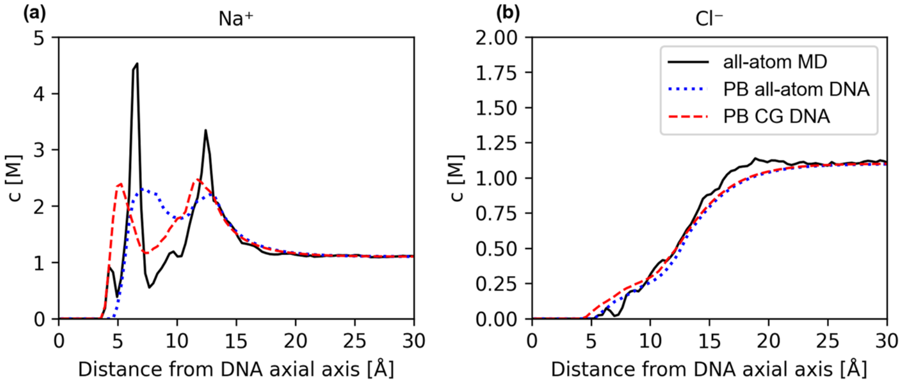

FIG. 2. Cylindrical concentration profiles of Na^+ (a) and Cl^- (b) relative to the DNA axial axis. The PB concentration profile is calculated using $c_i(\boldsymbol{r}) = c_i^0 \exp[-\beta(q_i\Phi(\boldsymbol{r}) + V_i(\boldsymbol{r}))]$, where $c_i^0, q_i, \Phi(\boldsymbol{r}), V_i(\boldsymbol{r})$ are the bulk concentration, charge, PB potential, and solute–ion interaction potential for species i, respectively. Results are shown for all-atom MD simulation (black-solid line) and PB solutions for the all-atom DNA model (blue-dotted line) with r_{ion} of 1.0 Å and PB solution for the CG DNA model with r_{ion} of 2.0 Å (red-dashed line) at 1.0 mol l^{-1}.

APBS requires selecting several parameters, including the molecular surface defining the dielectric boundary and the treatment of the solvent dielectric region. Among the most important are the ion size—treated in some studies as an adjustable parameter fitted to experimental data[26,134] or derived from radial distribution functions in all-atom MD simulations[28]—and the grid spacing, which varies in the literature from 0.5 Å[19] to 1.5 Å.[26] We have carefully optimized various parameters through multiple tests, as indicated in Fig. 4. For solving the PB equation, we select an automatically configured sequential focusing multi-grid. The PB equation is initially solved on a coarse grid box (150 × 150 × 120 $Å^3$), followed by a fine grid box (95 × 95 × 64 $Å^3$), both centered at the center of mass (CoM) of a CG DNA molecule [49, 50, 17 Å]. This grid configuration results in an APBS box larger than the MD simulation box in all three spatial directions. We use 257, 257, and 193 grid points in the x, y, and z directions, respectively, creating a 3D grid with grid spacings of 0.37, 0.37, and 0.33 in the x, y, and z directions, respectively. The PB equation is solved using a single Debye–Hückel boundary condition, with the solute having a dielectric constant of 2, to account for the effect of polarizability,[22,135,136] and the solvent dielectric set to 78.4, consistent with experimental data for water at 300 K. The ionic radius is set to 2 Å and the concentration to 1 mol l^{-1}. The CG DNA molecule defines the dielectric boundary, with atomic radii and partial charges consistent with those in the all-atom simulation. Nucleic acid atomic charges are assigned to nearest-neighbor grid points. The solvent-excluded surface is modeled with a rolling probe of radius 1.4 Å.

2. Electrostatic forces from the Poisson–Boltzmann potential

The PB equation is a mean-field electrostatic model that simplifies various atomic details and all non-Coulombic solvent–polymer interactions.[18] It also neglects direct ion–ion interactions and only

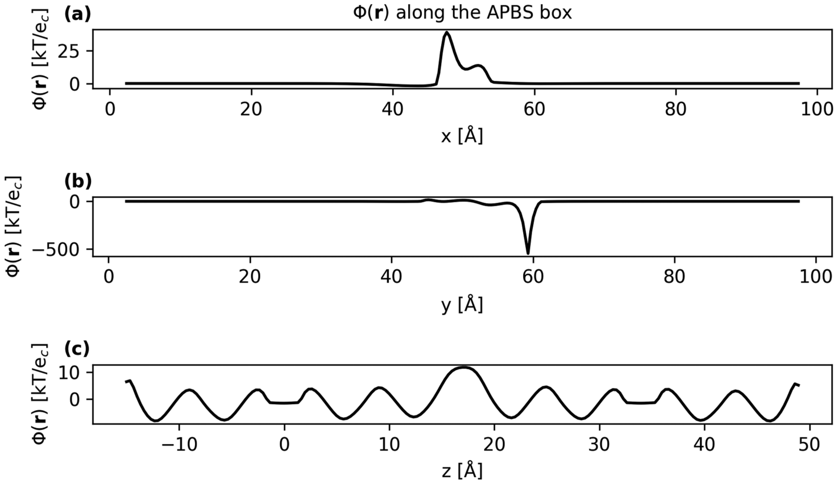

FIG. 3. PB electrostatic potential $[\Phi(\boldsymbol{r})]$ calculated with APBS along the x (at fixed y = 50.027 Å and z = 16.965 Å grid points) (a), y (at fixed x = 49.845 Å and z = 16.965 Å grid points) (b), and z (at fixed x = 49.845 Å and y = 50.027 Å grid points) (c) axis of the APBS box. The units for the PB electrostatic potential are k (in J K^{-1}) for the Boltzmann constant, T (in K) for the system temperature, and e_c for the elementary charge (in C).

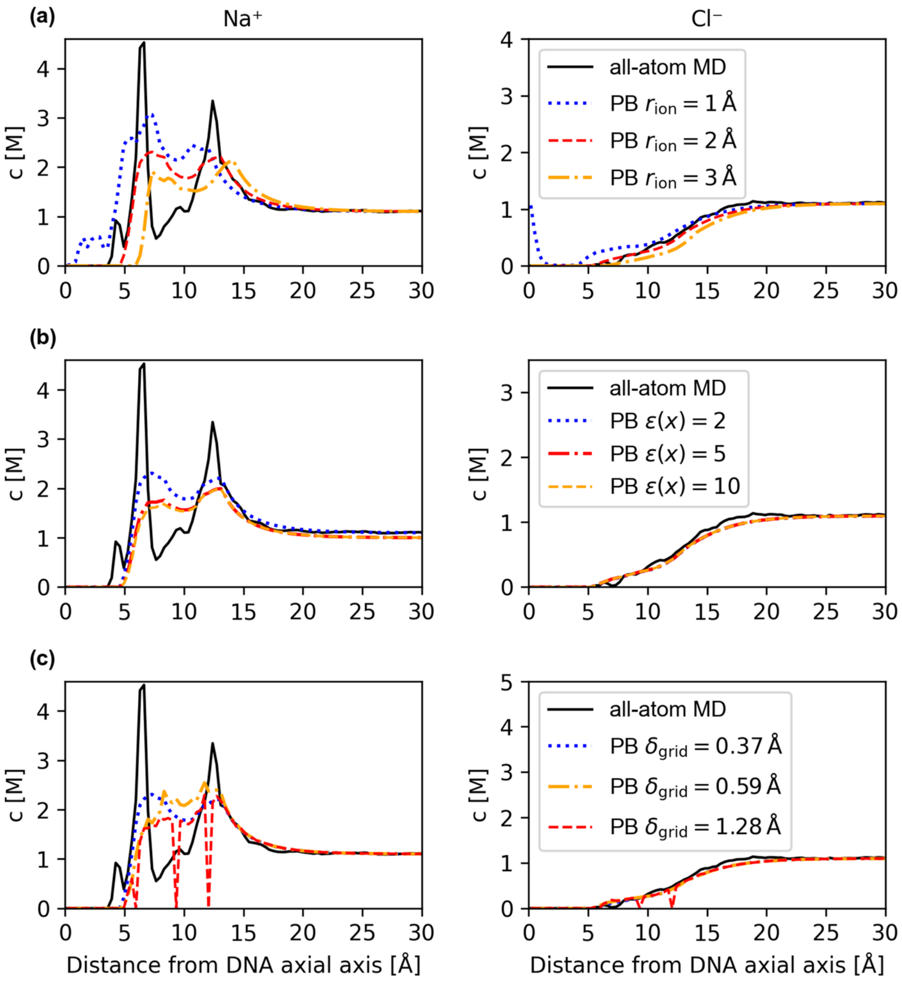

FIG. 4. Cylindrical concentration profiles of Na$^+$ (left) and Cl$^-$ (right) relative to the DNA axial axis. The PB concentration profile is calculated using $c_i(\boldsymbol{r}) = c_i^0 \exp[-\beta(q_i\Phi(\boldsymbol{r}) + V_i(\boldsymbol{r}))]$, where $c_i^0, q_i, \Phi(\boldsymbol{r}), V_i(\boldsymbol{r})$ are the bulk concentration, charge, PB potential, and solute–ion interaction potential for species i, respectively. Panels show all-atom MD (black-solid-line) compared with PB results for varying: (a) ion radii of 1 Å (blue-dotted-line), 2 Å (red-dashed-line), and 3 Å (orange-dashed-dotted-line); (b) inner dielectric constants of 2 (blue-dotted-line), 5 (red-dashed-line), and 10 (orange-dashed-dotted-line); and (c) grid spacings of 0.37 Å (blue-dotted-line), 0.59 Å (orange-dashed-dotted-line), and 1.28 Å (red-dashed-line). All results correspond to a 1.0 mol l^{-1} salt solution.

approximates the ion's size to a certain extent.[28] These simplifications result in non-physical behavior during MD simulations. As a mean-field-based model, the PB equation is most suitable for describing electrostatic interactions at large distances. Deviations from experimental data and detailed all-atom simulations become more pronounced at short-range distances, especially in the presence of a strong electric field.[28] This can be observed in Fig. 2, where the concentration profile calculated directly from the PB solution deviates significantly from the all-atom MD result within the ~12 Å distance from the DNA axial axis but converges well beyond that region. We solve the NPBE with APBS. In our simulations, the CG DNA molecule is kept fixed, and we focus exclusively on the interactions of mobile ions. The full PB potential at position $\boldsymbol{r}_i$ is defined as

$$U^{PB}(\boldsymbol{r}_i) = q_i\Phi(\boldsymbol{r}_i) + V^{steric}(\boldsymbol{r}_i), \tag{7}$$

where q_i is the ion charge (1 and −1 for Na$^+$ and Cl$^-$ ions), $\Phi(\boldsymbol{r})$ is the electrostatic potential, $V^{steric}(\boldsymbol{r})$ is the steric exclusion potential, and $\boldsymbol{r}_i = (x_i, y_i, z_i)$ is the position vector of the i-th ion.

The corresponding PB force acting on the i-th ion is then given by

$$\boldsymbol{F}_i^{PB} = q_i\boldsymbol{E}(\boldsymbol{r}_i) - \nabla_i V^{steric}(\boldsymbol{r}), \tag{8}$$

where $\boldsymbol{E}(\boldsymbol{r}_i) = -\nabla_i\Phi(\boldsymbol{r})$ is the electric field at the ion position $\boldsymbol{r}_i$. The steric component of the PB force strongly repels Na$^+$ ions from the DNA region during the MD simulations, leading to poor sampling of ion–DNA configurations during ML training and, consequently, degraded overall performance of the DIS-PB model. For the construction of the PB prior, we, therefore, retain only the electrostatic contribution,

$$\boldsymbol{F}_i^{PB_{prior}} = q_i\boldsymbol{E}(\boldsymbol{r}_i). \tag{9}$$

The PB prior, Eq. (9), accounts for the mean-field, long-range Coulombic force on the ions. The remaining contribution corresponding to short-range ion–solute interactions is instead learned by the ML algorithm.

As APBS solves the NPBE using FDM, the electrostatic potential is output as values on a three-dimensional grid. During the MD simulation, ion positions do not generally coincide

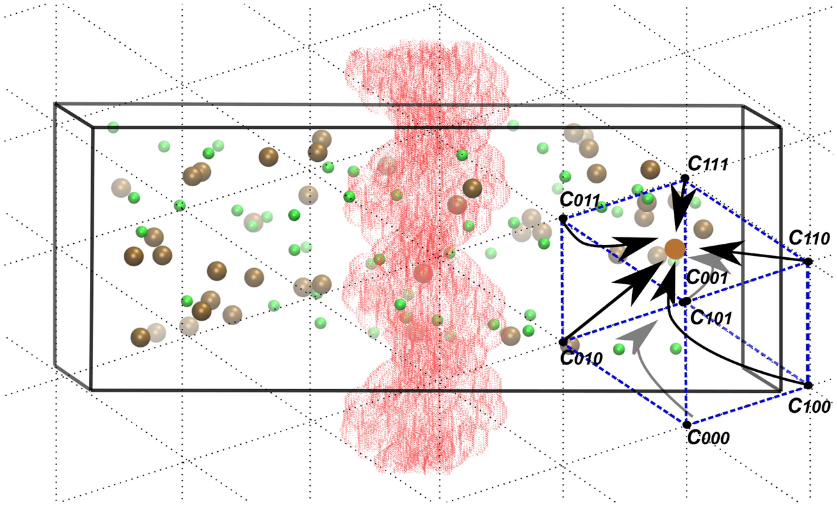

FIG. 5. Schematic representation of the trilinear interpolation of the APBS electric field from the values at the grid points to the positions of the ions during the MD simulation. The central region of the box shows the electric field components in the x, y, and z directions for the elongated CG DNA molecule. For visual clarity, the grid cells are shown much larger than the actual grid spacing used in the APBS calculations.

with the grid points of the mesh. From APBS, we obtain three 3D grids for each component of the electric field, i.e., $\boldsymbol{E}(x,y,z) = (E_x(x,y,z), E_y(x,y,z), E_z(x,y,z))$. During the MD simulation, each electric field is interpolated from the APBS grid values to the ion position using trilinear interpolation, providing the local field needed to compute $\mathbf{F}_i^{\mathrm{PB}}$, as schematically shown in Fig. 5.

First, the relative distances in the x, y, and z directions from the ion position to the nearest grid points are calculated as

$$\Delta x = \frac{x - x^0}{x^1 - x^0}, \quad \Delta y = \frac{y - y^0}{y^1 - y^0}, \quad \Delta z = \frac{z - z^0}{z^1 - z^0}, \tag{10}$$

where (x, y, z) are the ion coordinates, (x_0, y_0, z_0) are the nearest grid points lower than the ion position, and (x_1, y_1, z_1) are the nearest grid points higher than the ion position. The electric field at the ion position is then obtained via trilinear interpolation,

$$\begin{aligned} c = {} & c_{000}(1-\Delta x)(1-\Delta y)(1-\Delta z) + c_{100}\Delta x(1-\Delta y)(1-\Delta z) \\ & + c_{010}(1-\Delta x)\Delta y(1-\Delta z) + c_{110}\Delta x\Delta y(1-\Delta z) \\ & + c_{001}(1-\Delta x)(1-\Delta y)\Delta z + c_{101}\Delta x(1-\Delta y)\Delta z \\ & + c_{011}(1-\Delta x)\Delta y\Delta z + c_{111}\Delta x\Delta y\Delta z, \end{aligned} \tag{11}$$

where c_{000} through c_{111} correspond to the electric field values at the nearest eight grid points, and c is the interpolated value at the ion position (schematically shown in Fig. 5). After interpolating all three components of the electric field, the electrostatic force on the ions is calculated straightforwardly using Eq. (9). This approach is implemented within the LAMMPS[124] framework via a custom fix module.

C. The DIS-PB model

As described in Sec. II B, the PB method is a mean-field theory that efficiently captures long-range electrostatic tails. However, its simplicity and computational efficiency come at the cost of reduced accuracy in short-range interactions. In Sec. I, we outlined how ML approaches have recently emerged as powerful tools for representing complex many-body potentials, with architectures particularly well suited for modeling short-range interactions. To combine the strengths of both approaches, we employ a Δ-learning strategy[104,137,138] in which the ML model is trained on the difference between all-atom MD forces and PB forces, thereby capturing the short-range contributions that PB neglects. A cutoff of 12 Å is used for pair distances between all atom types during training, which is sufficient to capture the relevant short-range interactions and to include an appropriate number of neighbors per atom in the coarse-grained system. As illustrated in Fig. 6, 12 Å corresponds to the approximate minimum distance from the CG DNA axial axis below which short-range PB interactions deviate significantly from the all-atom solution; beyond this distance, the PB and all-atom profiles are in good agreement. It should be noted that the 12 Å cutoff used during training and in the DIS-PB MD simulations with the Allegro force field does not directly correspond to the radial distance from the DNA axis shown in Fig. 6. The cutoff refers to pairwise distances, whereas Fig. 6 shows distances from the CG DNA axial axis, with DNA atoms extending up to ∼10 Å from the axis. Newer versions of Allegro allow different cutoffs for different atom types, enabling more precise control of the overall distance from the CG DNA axis. When such ML potential is combined with the PB potential, the resulting effective potential is capable of closely reproducing all-atom MD results. We evaluate this approach on a DNA system in an aqueous salt solution. The all-atom MD simulations include explicit DNA, ions, and water molecules, whereas the PB potential calculations, ML training, and MD simulations with ML and PB potentials combined are based on a CG DNA representation with explicit ions, while water is omitted completely. As a result, all water–water and water–ion explicit interactions present in all-atom MD are excluded, which, in turn, contributes to the computational speed-up of our approach. We refer to this framework as the DIS-PB model.

Due to its high data efficiency, strict locality, and ability to accurately interpolate to out-of-distribution configurations, we choose the GNN Allegro ML potential based on a many-body equivariant neural network (ENN) architecture.[65,66,68,69,139–142] Collections of nodes and edges represent the atomistic structures. In the graph representation, the nodes correspond to atoms, and edges connect

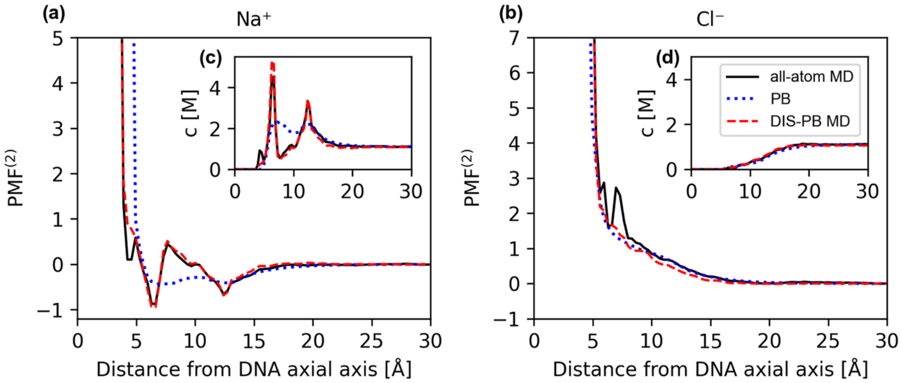

FIG. 6. 2-particle potential of mean force $[\mathrm{PMF}^{(2)}]$ profiles for Na$^+$ (a) and Cl$^-$ (b) relative to the DNA axial axis. Insets show the corresponding cylindrical concentration profiles for Na$^+$ (c) and Cl$^-$ (d). The PB concentration profile is calculated using $c_i(\boldsymbol{r}) = c_i^0 \exp[-\beta(q_i\Phi(\boldsymbol{r}) + V_i(\boldsymbol{r}))]$, where c_i^0, q_i, $\Phi(\boldsymbol{r})$, $V_i(\boldsymbol{r})$ are the bulk concentration, charge, PB potential, and solute–ion interaction potential for species i, respectively. PMF$^{(2)}$ is computed as $\mathrm{PMF}^{(2)} = -RT \ln(c_i(\boldsymbol{r})/c_i^0)$, where R is the gas constant, T is the temperature, and $c_i(\boldsymbol{r})/c_i^0$ is the normalized concentration profile. Results are shown for all-atom MD simulation (black-solid line), PB solution (blue-dotted line), and DIS-PB MD simulation (red-dashed line) at 1.0 mol l^{-1} salt solution.

the atoms within a cutoff-distance sphere. This makes the Allegro model compatible with the spatial decomposition approach used in LAMMPS, leading to excellent computational efficiency and scalability across devices. The parameters for the ML model are trained using a force-matching approach.[143] The training loss is defined as

$$L = \frac{1}{3BN}\sum_{i=1}^{BN} \|\boldsymbol{F}_i^{ML} - \boldsymbol{F}_i^{\Delta}\|^2, \tag{12}$$

where B, N, E^{ML}, $\boldsymbol{F}_i^{ML} = -\nabla_i E^{ML}$, and $\boldsymbol{F}_i^{\Delta}$ denote the batch size, number of ions, ML potential energy of the system,[65,119] ML-predicted force difference, and true (reference) force difference acting on atom (ion) i, respectively. Since we employ the Δ-learning approach, forces on ions had to be stored every 1 ps during the production phase of the all-atom MD simulation. PB forces are calculated for the ions at the positions from the all-atom MD simulation. These stored forces are then subtracted to obtain the true forces used for training,

$$\boldsymbol{F}_i^{\Delta} = \boldsymbol{F}_i^{all-atom} - \boldsymbol{F}_i^{PB_{prior}}, \tag{13}$$

where $\boldsymbol{F}_i^{all-atom}$ are the forces obtained from the all-atom simulation using a classical MD force field (in our case, AMBER 03), and $\boldsymbol{F}_i^{PB_{prior}}$ are the forces calculated by Eq. (9). The ML potential is optimized using the Adam optimizer[144] with the default parameters from PyTorch.[145] The training dataset contains 2.0×10^5 configurations obtained from the reference all-atom MD simulation, which are split sequentially by Allegro into training (80%) and validation (20%) datasets. The accuracy of prediction for both datasets is evaluated with the force matching approach described in Eq. (12). Other training hyperparameters are in large agreement with the DIS model by Coste *et al.*[119]

DIS-PB MD simulations are performed in LAMMPS,[124] integrating Newton's equations of motion with the velocity Verlet algorithm using a 1 fs time step.[125,126] The PB equation is combined with the ML potential via a custom fix that adds the corresponding external forces to the ions. Thus, during DIS-PB MD simulations, the total force on each ion $\boldsymbol{F}_i^{DIS-PB}$ is computed as the sum of the PB prior force $\boldsymbol{F}_i^{PB_{prior}}$ and the ML-predicted force $\boldsymbol{F}_i^{ML}$,

$$\boldsymbol{F}_i^{DIS-PB} = \boldsymbol{F}_i^{PB_{prior}} + \boldsymbol{F}_i^{ML}. \tag{14}$$

The number of ions and the box dimensions are consistent with those used in the all-atom MD simulations (Sec. II A), while the CG DNA model is the same as described above for the PB calculations. Simulations are carried out in the NVT ensemble, with the temperature maintained at 300 K using the Nosé–Hoover thermostat. Periodic boundary conditions effectively render the CG DNA molecule infinite, which is kept rigid throughout the simulation. Each production run lasts 25 ns following 1 ns of equilibration.

Unlike the DIS model of Coste *et al.*,[119] in which ion–ion and ion–DNA interactions are represented explicitly using a pairwise Lennard-Jones potential combined with a Wolf-summed Coulomb term, the prior used in DIS-PB is based on a Poisson–Boltzmann electrostatic field. The LJ–Wolf prior provides a pairwise, real-space approximation to long-range electrostatics, whereas the PB-based prior incorporates solvent screening at the mean-field continuum level and captures non-local electrostatic effects through a spatially extended field. This formulation allows the ML correction in DIS-PB to focus exclusively on local contributions such as ion–ion correlations and near-surface ion–DNA interactions, while the PB prior accounts for the non-local electrostatic component. Due to this conceptual distinction, in which the PB electrostatics enter as an external field rather than explicit pairwise forces, the PB-based prior enables efficient GPU execution of the MD simulations, leading to a substantial computational acceleration compared to DIS.

D. Umbrella sampling and PMF calculation

The potential of mean force (PMF) provides the free-energy profile along a chosen reaction coordinate, reflecting the effective interactions between the two chosen particles. In our case, the collective variable is defined as the distance between the DNA's CoM and the sodium ion, constrained to the straight line extending from the DNA's CoM in the positive x-direction. Its sampled range spans 2.5–20 Å. We obtain the PMF using umbrella sampling implemented via the `colvars` module in LAMMPS.[146] A total of 36 harmonic windows are used, each with a spring constant of 5 kcal mol^{-1} Å^{-2} along the reaction coordinate and a stiff perpendicular restraint of 250 kcal mol^{-1} Å^{-2} to confine the ion to the sampling axis. Each window is simulated for 2 ns with a sampling interval of 10 fs.

We reconstruct the unbiased free-energy profile using the weighted histogram analysis method (WHAM),[147] employing 120 histogram bins and 50 Monte Carlo iterations, as implemented in the WHAM package developed by Grossfield's group.[148] This procedure is carried out for three models: the all-atom MD reference simulation, the DIS-PB MD simulation, and the Allegro-only potential trained directly on all-atom forces. For comparison, we also include the PB electrostatic potential $\Phi(\mathbf{r})$ as well as the full PB potential along the same distance between DNA's CoM and the sodium ion in the positive x-direction.

III. RESULTS AND DISCUSSION

We aim to achieve the all-atom MD accuracy by applying a Δ-learning approach that combines the PB model, which serves as the baseline capturing long-range electrostatics, with an ML potential that recovers the missing short-range interactions. As shown in Fig. 6, the PB solution fails to accurately reproduce the ion distribution at short range (below 12 Å) due to the absence of specific interactions in its underlying theory. To improve the PB solution, we introduce an ML potential trained on the difference between the all-atom and PB forces, thereby capturing the short-range contributions. Both the PB and ML components employ the same CG representation of the DNA molecule, in which only selected DNA atoms are retained and water molecules are omitted. This coarse-grained design reduces the computational cost compared to all-atom MD simulations.

To optimize the performance of our DIS-PB model, we adapt several ML hyperparameters from the original DIS model.[119] In addition, to evaluate what Allegro can learn without Δ-learning, we train it directly on the full all-atom forces using the same hyperparameters as in the DIS-PB model.

To assess the effective ion–DNA interactions across the different models, we first examine the PMF for a Na^+ ion approaching the DNA molecule's CoM along the x-axis (Fig. 7). The PMF is obtained using the umbrella sampling protocol described in Sec. II D. As the ion approaches the DNA, the PB prior and the full PB potential depart markedly from the all-atom reference and fail to reproduce the essential short-range ion–DNA interactions captured in the all-atom PMF, consistent with the discussion in Sec. II B. Adding the ML-based correction in this region produces a substantial improvement: the DIS-PB PMF closely follows the all-atom result. By contrast, the Allegro-only force field—trained directly on all-atom forces—matches the all-atom PMF slightly less accurately than DIS-PB at short separations (up to 8 Å). However, in the intermediate range (8 to 12 Å), it provides a slightly better match than DIS-PB, and it performs comparably in the long-range limit. Beyond 12 Å, the ion–DNA interaction becomes negligible relative to thermal fluctuations, and the PMFs from all force fields fall within approximately kT of each other.

The NDPs of Na^+ ions in Fig. 8(a) show that the Allegro-only model effectively reproduces the structural distribution of Na^+ ions around DNA. However, due to the absence of non-local interactions between Cl^- ions and CG DNA atoms, the model fails to capture the expected electrostatic repulsion. As a result, Cl^- ions penetrate the CG DNA structure and accumulate near carbon, nitrogen, and oxygen atoms in the major groove [Fig. 8(b)],

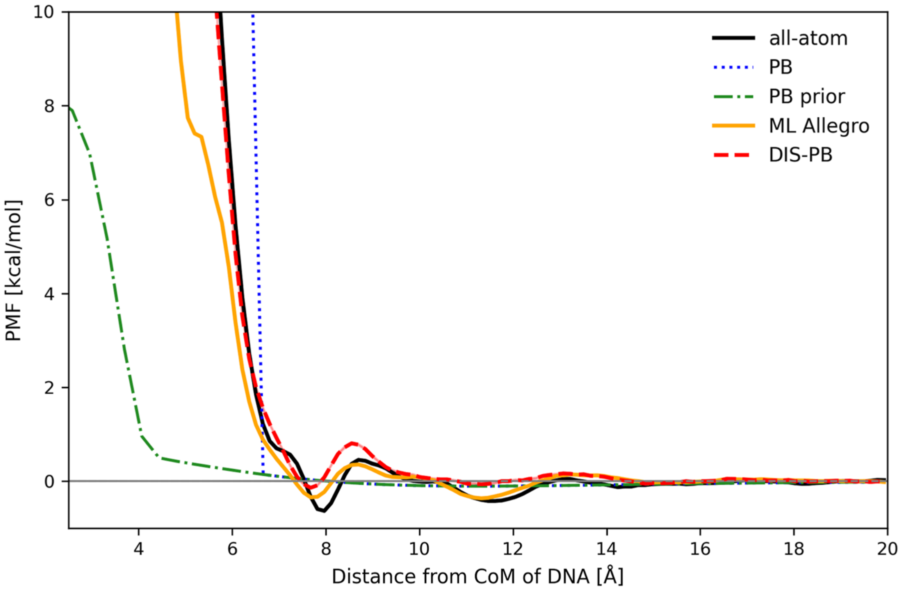

FIG. 7. PMF for Na^+ ion as a function of the distance from the DNA CoM along the x-axis. The PMF is calculated as described in Sec. II D. Results are shown for the all-atom MD simulation (black-solid line), total PB potential Eq. (7) (blue-dotted line), PB prior potential corresponding to the electrostatic contribution, i.e., the first term of Eq. (7) (green dotted-dashed line), DIS-PB MD simulation (red-dashed line), and Allegro-only MD simulation (orange-dashed-dotted line), at 1.0 mol l^{-1} salt solution.

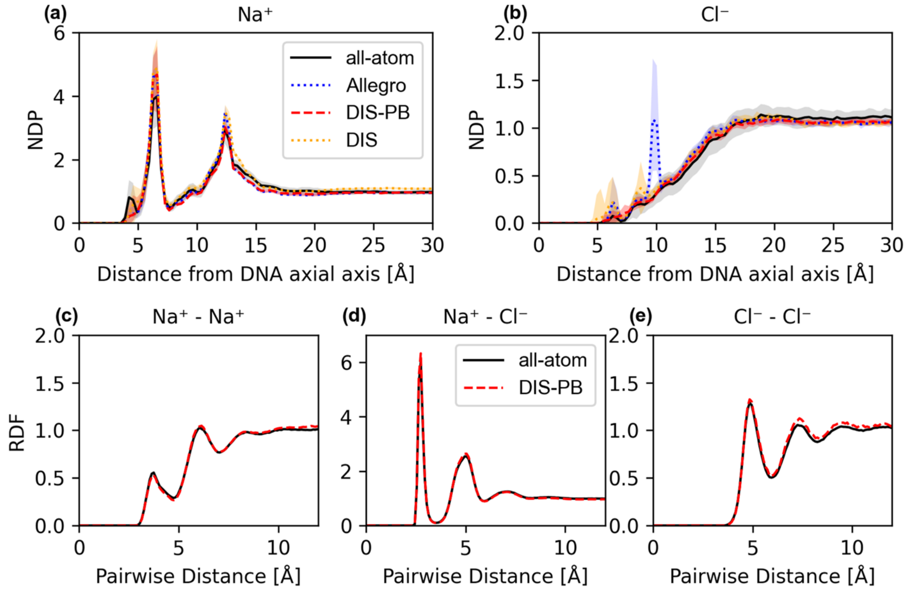

FIG. 8. Cylindrical number density profiles (NDPs) of Na^+ (a) and Cl^- (b) relative to the DNA axial axis. Results are shown for all-atom (black-solid line), Allegro-only (blue-dotted line), DIS-PB (red-dashed line), and DIS (orange-dotted line) simulations at 1.0 mol l^{-1}. Shaded areas denote the standard deviations obtained by block averaging over 1 ns. Radial distribution functions (RDFs) for all-atom (black-solid line) and DIS-PB (red-dashed line) simulations are presented for Na–Na (c), Na–Cl (d), and Cl–Cl (e) pairs.

leading to unstable simulation behavior. In contrast, this unphysical behavior is absent from the DIS-PB model, ensuring simulation stability by consistently repelling Cl^- from all DNA atoms. In addition, the Na^+ NDP profile is in better agreement with all-atom results. This improved behavior highlights the importance of incorporating the PB mean-field prior in combination with a learned short-range potential. In the comparison between the DIS model and DIS-PB, we did not observe any significant improvement in the accuracy of the NDP profiles. To evaluate the representation of ion–ion interactions, which are not included in PB theory, we calculated RDF profiles in the CG DNA salt system. The RDFs generated using the DIS-PB model closely match those from all-atom MD simulations, demonstrating that Allegro's local graph-based neural network architecture effectively captures short-range, many-body interactions [Figs. 8(c)–8(e)].

We further examine the development of the ion atmosphere moving away from the DNA axis by calculating the number of excess Na^+ ions. This is done by integrating the difference between the radial and the bulk concentration across a cylindrical volume. The deviation between the all-atom and DIS-PB models is around 1.5 Na^+ [Fig. 9(a)], indicating a good agreement. To investigate specific Na^+ binding patterns to the CG DNA atom types, we identify ions residing in the first hydration shell and compute the probability distribution of binding configurations.[41] As shown in Fig. 9(b), the most frequently involved atom types in the most probable binding patterns are phosphate (P), then guanine oxygen (GO6), guanine carbon (GC8), adenine carbon (AC8), and thymine oxygen (TO4)—all located in the DNA major groove. Less frequent contributions to the most probable binding patterns include the guanine nitrogen (GN2), thymine oxygen (TO2), and cytosine oxygen

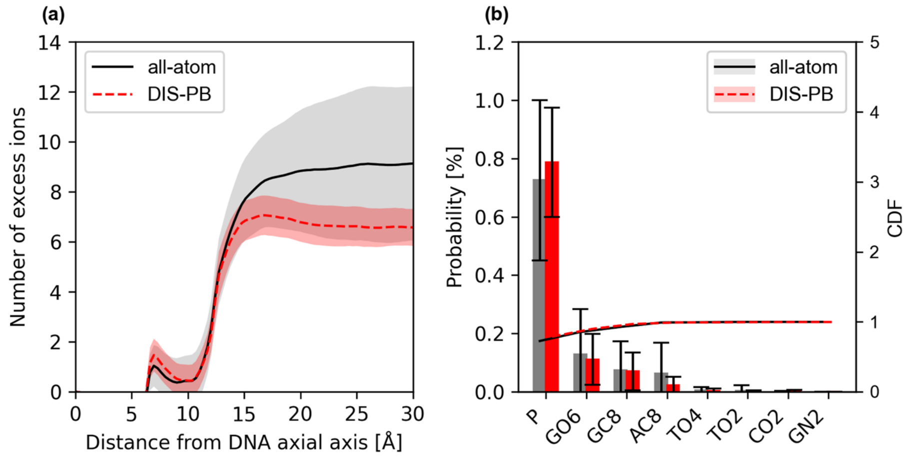

FIG. 9. Number of excess ions as a function of distance from the DNA axial axis (a). Results are shown for all-atom (black-solid line) and DIS-PB (red-dashed line) simulations. Shaded regions represent the standard deviations obtained by block averaging over 1 ns. Probability and cumulative distribution functions (CDFs) of Na^+ binding patterns to DNA (b). The x-axis indicates the nucleic acid atoms forming inner-sphere contacts with the cation. Results are shown for all-atom (black) and DIS-PB (red), with error bars denoting the standard deviations.

(CO2), which are located in the minor groove. DIS-PB and all-atom results agree very well.

A key advantage of Allegro lies in its scalability, enabled by its strictly local architecture. Unlike the DIS model reported by Coste,[119] which employs a hybrid model combining LJ–Wolf and Allegro potentials, and most other methods that integrate short-range ML potential with long-range interactions, our approach uses purely Allegro ML potential supplemented by an external non-local, long-range potential. Since PB electrostatics can be calculated at only a fraction of the computational cost of all-atom MD, and our model employs a highly coarse-grained representation with implicit water, one might expect a substantial speed-up. However, evaluating the ML potential for ions—which accounts for both ion–ion and ion–DNA interactions—is relatively demanding, more expensive than standard pair potentials, and constitutes the majority of the computational cost in our DIS-PB MD simulations. As a result, the speed-up is more modest. We anticipate a much clearer computational advantage for larger systems, where the number of water molecules in an all-atom model would grow significantly while the relative cost of the ML potential remains about the same. Even so, our design efficiently exploits GPU resources and yields considerable performance gains in MD simulations. First, we compare the speed of the original DIS model[119] and our DIS-PB model. Because the original DIS model cannot be deployed on a GPU, we have performed both simulations on a single CPU. The main computational bottleneck for both DIS and DIS-PB methods is the evaluation of the ML potential, which was trained with similar hyperparameters, resulting in comparable complexity. As a result, the simulation rates of the two methods are identical, yielding 1.42 ns per day for both the DIS and DIS-PB approaches, whereas the all-atom MD simulation is considerably slower at 0.6 ns per day. To further examine computation efficiency, we compared the performance of the all-atom MD and DIS-PB MD, both run on a single GPU. The all-atom model produces a simulation rate of 12.736 ns per day, whereas the DIS-PB model achieves a rate of 16.251 ns per day. This speed-up highlights the computational efficiency of the DIS-PB model, enabling longer simulations and larger system sizes. The reported simulation rates refer to the runtime of the MD simulations only; the one-time cost of generating the all-atom reference data for ML training is not included, as this reference simulation is identical to the all-atom MD used for validation and is performed only once.

A key limitation of our approach is that the DNA molecule is kept rigid throughout the MD simulations, which restricts the model to describing only ion behavior without accounting for DNA conformational flexibility. This simplification likely alters ion condensation sites and, consequently, the ion distribution and condensation behavior.[123,149,150] However, at the same time, the rigid DNA setup provides a practical advantage: it allows fast and direct incorporation of the PB potential, which would otherwise vary with every DNA conformational change. To achieve a more realistic description of DNA dynamics, we plan to couple the implicit electrostatic solver (APBS) directly to the MD engine (LAMMPS), thereby allowing DNA to move during the simulation under the combined action of the PB forces on the DNA atoms and the standard MD interaction potentials used in the system. Within this framework, the electrostatic potential will be recomputed every few ps, ensuring that the ionic distribution adjusts self-consistently to the updated DNA conformation. Furthermore, because the model is trained on a single rigid DNA sequence, the learned ML potential is not transferable to other sequences, as it lacks explicit information about DNA structure and sequence dependence. Since our model is trained on a single solvent concentration, it is not directly transferable to other salt concentrations. This limitation could be overcome by making the ML model concentration-dependent, incorporating the solvent concentration as an additional input. Another limitation arises from the Δ-learning approach itself, which couples two complex potentials with many parameters. Any modification in the PB calculation requires new data generation and may necessitate adjustment of ML hyperparameters, making the procedure time-consuming and challenging to systematically approach the accuracy of all-atom MD simulations. A possible improvement would be to incorporate the PB prior directly into the ML training and to couple the APBS software for PB calculations with LAMMPS for MD simulations in an on-the-fly manner so that both contributions evolve consistently during the simulation.

IV. CONCLUSION

In conclusion, we have developed the DIS-PB machine learning model, based on an equivariant GNN and Δ-learning with the PB baseline. Standard ML approaches are particularly well suited for capturing short-range contributions to the potential, owing to their many-body character, flexibility, and, in the case of Allegro, strict locality that ensures scalability. In contrast, the PB method efficiently describes long-range electrostatics but lacks short-range accuracy due to its simplifying assumptions. By combining these complementary strengths, DIS-PB captures both regimes when applied to a simplified system composed only of CG DNA and explicit ions, with water omitted entirely. This coarse-graining substantially reduces the number of particles and interactions the model must consider, which accelerates ML training and makes the resulting MD simulations faster. At the same time, the Δ-learning correction ensures that essential short-range and long-range effects are preserved, allowing the simulations to approach the accuracy of all-atom MD at a lower computational cost. In future work, we will extend the applicability of the DIS-PB model to more complex biomolecular systems, broadening its scope to a wider range of problems in computational biophysics.

ACKNOWLEDGMENTS

The authors thank the HPC RIVR consortium (www.hpc-rivr.si) and EuroHPC JU (eurohpc-ju.europa.eu) for providing the computing resources of the HPC system Vega at the Institute of Information Science (www.izum.si). The authors also acknowledge the financial support of the Slovenian Research and Innovation Agency (Grant No. P1-0002).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Ema Slejko: Investigation (lead); Methodology (equal); Software (lead); Validation (lead); Writing – original draft (lead); Writing – review & editing (equal). **Amaury Coste**: Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing – review & editing (equal). **Tilen Potisk**: Investigation (equal); Methodology (equal); Software (supporting); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Julija Zavadlav**: Conceptualization (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal). **Matej Praprotnik**: Conceptualization (lead); Funding acquisition (lead); Investigation (equal); Methodology (lead); Supervision (lead); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data, necessary software, and scripts that support the findings of this study are openly available at https://github.com/L17-Projects/DIS-PB.git.

REFERENCES

[1] P. Grochowski and J. Trylska, "Continuum molecular electrostatics, salt effects, and counterion binding—A review of the Poisson–Boltzmann theory and its modifications," Biopolymers **89**, 93–113 (2008).

[2] J. K. Frederiksen and J. A. Piccirilli, "Identification of catalytic metal ion ligands in ribozymes," Methods **49**, 148–166 (2009).

[3] E. Freisinger and R. K. O. Sigel, "From nucleotides to ribozymes—A comparison of their metal ion binding properties," Coord. Chem. Rev. **251**, 1834–1851 (2007).

[4] D. E. Draper, D. Grilley, and A. M. Soto, "Ions and RNA folding," Annu. Rev. Biophys. Biomol. Struct. **34**, 221–243 (2005).

[5] W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," J. Chem. Phys. **139**, 090901 (2013).

[6] S. Riniker, J. R. Allison, and W. F. van Gunsteren, "On developing coarse-grained models for biomolecular simulation: A review," Phys. Chem. Chem. Phys. **14**, 12423–12430 (2012).

[7] J. Jin, A. J. Pak, A. E. P. Durumeric, T. D. Loose, and G. A. Voth, "Bottom-up coarse-graining: Principles and perspectives," J. Chem. Theory Comput. **18**, 5759–5791 (2022).

[8] P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. H. Bruininks, T. A. Wassenaar *et al.*, "Martini 3: A general purpose force field for coarse-grained molecular dynamics," Nat. Methods **18**, 382–388 (2021).

[9] G. E. Balatti, M. F. Martini, and M. Pickholz, "Investigating the impact of the glycolipid content on Aurein 1.2 pores in prokaryotic model bilayers: A coarse-grain molecular dynamics simulation study," J. Phys. Chem. B **127**, 5190–5198 (2023).

[10] N. A. Baker, "Biomolecular applications of Poisson–Boltzmann methods," in *Methods in Enzymology* (Elsevier, 2004), Vol. 383, pp. 94–118.

[11] D. D. Nguyen, B. Wang, and G.-W. Wei, "Accurate, robust, and reliable calculations of Poisson–Boltzmann binding energies," J. Comput. Chem. **38**, 941–948 (2017).

[12] S. Ringe, H. Oberhofer, C. Hille, S. Matera, and K. Reuter, "Function-space-based solution scheme for the size-modified Poisson–Boltzmann equation in full-potential DFT," J. Chem. Theory Comput. **12**, 4052–4066 (2016).

[13] F. Lipparini and B. Mennucci, "Perspective: Polarizable continuum models for quantum-mechanical descriptions," J. Chem. Phys. **144**, 160901 (2016).

[14] H. Nguyen, A. Pérez, S. Bermeo, and C. Simmerling, "Refinement of generalized Born implicit solvation parameters for nucleic acids and their complexes with proteins," J. Chem. Theory Comput. **11**, 3714–3728 (2015).

[15] I. S. Tolokh, D. G. Thomas, and A. V. Onufriev, "Explicit ions/implicit water generalized born model for nucleic acids," J. Chem. Phys. **148**, 195101 (2018).

[16] P. Collas, "Coulomb scattering in the Born approximation and the use of generalized functions," Am. J. Phys. **89**, 799–805 (2021).

[17] J. J. Molina, M. Duvail, P. Guilbaud, and J.-F. Dufrêche, "Coarse-grained lanthanoid chloride aqueous solutions," J. Mol. Liq. **153**, 107–111 (2010).

[18] P. Batys, S. Luukkonen, and M. Sammalkorpi, "Ability of the Poisson–Boltzmann equation to capture molecular dynamics predicted ion distribution around polyelectrolytes," Phys. Chem. Chem. Phys. **19**, 24583–24593 (2017).

[19] T. J. Robbins, J. D. Ziebarth, and Y. Wang, "Comparison of monovalent and divalent ion distributions around a DNA duplex with molecular dynamics simulation and a Poisson–Boltzmann approach," Biopolymers **101**, 834–848 (2014).

[20] V. K. Misra, J. L. Hecht, K. A. Sharp, R. A. Friedman, and B. Honig, "Salt effects on protein-DNA interactions. The λcI repressor and *Eco*RI endonuclease," J. Mol. Biol. **238**, 264–280 (1994).

[21] V. K. Misra, K. A. Sharp, R. A. Friedman, and B. Honig, "Salt effects on ligand-DNA binding: Minor groove binding antibiotics," J. Mol. Biol. **238**, 245–263 (1994).

[22] V. K. Misra, J. L. Hecht, A.-S. Yang, and B. Honig, "Electrostatic contributions to the binding free energy of the λcI repressor to DNA," Biophys. J. **75**, 2262–2273 (1998).

[23] C. L. Tang, E. Alexov, A. M. Pyle, and B. Honig, "Calculation of pKas in RNA: On the structural origins and functional roles of protonated nucleotides," J. Mol. Biol. **366**, 1475–1496 (2007).

[24] Y. Bai, M. Greenfeld, K. J. Travers, V. B. Chu, J. Lipfert, S. Doniach, and D. Herschlag, "Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids," J. Am. Chem. Soc. **129**, 14981–14988 (2007).

[25] A. Savelyev and G. A. Papoian, "Electrostatic, steric, and hydration interactions favor Na^+ condensation around DNA compared with K^+," J. Am. Chem. Soc. **128**, 14506–14518 (2006).

[26] V. B. Chu, Y. Bai, J. Lipfert, D. Herschlag, and S. Doniach, "Evaluation of ion binding to DNA duplexes using a size-modified Poisson–Boltzmann theory," Biophys. J. **93**, 3202–3209 (2007).

[27] Y. Bai, V. B. Chu, J. Lipfert, V. S. Pande, D. Herschlag, and S. Doniach, "Critical assessment of nucleic acid electrostatics via experimental and computational investigation of an unfolded state ensemble," J. Am. Chem. Soc. **130**, 12334–12341 (2008).

[28] S. Kirmizialtin, A. J. Silalahi, R. Elber, and M. O. Fenley, "The ionic atmosphere around A-RNA: Poisson–Boltzmann and molecular dynamics simulations," Biophys. J. **102**, 829–838 (2012).

[29] J. Sponer, G. Bussi, M. Krepl, P. Banáš, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete, P. Jurecka *et al.*, "RNA structural dynamics as captured by molecular simulations: A comprehensive overview," Chem. Rev. **118**, 4177–4338 (2018).

[30] N. Calimet, M. Simoes, J.-P. Changeux, M. Karplus, A. Taly, and M. Cecchini, "A gating mechanism of pentameric ligand-gated ion channels," Proc. Natl. Acad. Sci. U. S. A. **110**, E3987–E3996 (2013).

[31] P. Pokorná, M. Krepl, S. Campagne, and J. Šponer, "Conformational heterogeneity of RNA stem-loop hairpins bound to FUS-RNA recognition motif with disordered RGG tail revealed by unbiased molecular dynamics simulations," J. Phys. Chem. B **126**, 9207–9221 (2022).

[32] R. J. Workman, S. Gorle, and B. M. Pettitt, "Effects of conformational constraint on peptide solubility limits," J. Phys. Chem. B **126**, 10510–10518 (2022).

[33] K. Sarthak, D. Winogradoff, Y. Ge, S. Myong, and A. Aksimentiev, "Benchmarking molecular dynamics force fields for all-atom simulations of biological condensates," J. Chem. Theory Comput. **19**, 3721–3740 (2023).

[34] E. P. Barros, J. M. Schiffer, A. Vorobieva, J. Dou, D. Baker, and R. E. Amaro, "Improving the efficiency of ligand-binding protein design with molecular dynamics simulations," J. Chem. Theory Comput. **15**, 5703–5715 (2019).

[35] Y. Hirano, N. Okimoto, S. Fujita, and M. Taiji, "Molecular dynamics study of conformational changes of tankyrase 2 binding subsites upon ligand binding," ACS Omega **6**, 17609–17620 (2021).

[36] Y. Liu, M. B. Prigozhin, K. Schulten, and M. Gruebele, “Observation of complete pressure-jump protein refolding in molecular dynamics simulation and experiment,” J. Am. Chem. Soc. **136**, 4265–4272 (2014).

[37] R. Nassar, E. Brini, S. Parui, C. Liu, G. L. Dignon, and K. A. Dill, “Accelerating protein folding molecular dynamics using inter-residue distances from machine learning servers,” J. Chem. Theory Comput. **18**, 1929–1935 (2022).

[38] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, “Improved side-chain torsion potentials for the Amber ff99SB protein force field,” Proteins: Struct., Funct., Bioinf. **78**, 1950–1958 (2010).

[39] G. Bai, Y. Pan, Y. Zhang, Y. Li, J. Wang, Y. Wang, W. Teng, G. Jin, F. Geng, and J. Cao, “Research advances of molecular docking and molecular dynamic simulation in recognizing interaction between muscle proteins and exogenous additives,” Food Chem. **429**, 136836 (2023).

[40] J. Wang, A. Alekseenko, D. Kozakov, and Y. Miao, “Improved modeling of peptide-protein binding through global docking and accelerated molecular dynamics simulations,” Front. Mol. Biosci. **6**, 112 (2019).

[41] S. Cruz-León and N. Schwierz, “RNA captures more cations than DNA: Insights from molecular dynamics simulations,” J. Phys. Chem. B **126**, 8646–8654 (2022).

[42] M. Karplus and J. A. McCammon, “Molecular dynamics simulations of biomolecules,” Nat. Struct. Biol. **9**, 646–652 (2002).

[43] G. M. Giambaşu, T. Luchko, D. Herschlag, D. M. York, and D. A. Case, “Ion counting from explicit-solvent simulations and 3D-RISM,” Biophys. J. **106**, 883–894 (2014).

[44] M. A. Young, B. Jayaram, and D. L. Beveridge, “Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: Fractional occupancy of electronegative pockets,” J. Am. Chem. Soc. **119**, 59–69 (1997).

[45] M. Feig and B. M. Pettitt, “Sodium and chlorine ions as part of the DNA solvation shell,” Biophys. J. **77**, 1769–1781 (1999).

[46] S. Y. Ponomarev, K. M. Thayer, and D. L. Beveridge, “Ion motions in molecular dynamics simulations on DNA,” Proc. Natl. Acad. Sci. U. S. A. **101**, 14771–14775 (2004).

[47] M. Rueda, E. Cubero, C. A. Laughton, and M. Orozco, “Exploring the counterion atmosphere around DNA: What can be learned from molecular dynamics simulations?,” Biophys. J. **87**, 800–811 (2004).

[48] P. Várnai and K. Zakrzewska, “DNA and its counterions: A molecular dynamics study,” Nucleic Acids Res. **32**, 4269–4280 (2004).

[49] M. Praprotnik, S. Matysiak, L. D. Site, K. Kremer, and C. Clementi, “Adaptive resolution simulation of liquid water,” J. Phys.: Condens. Matter **19**, 292201 (2007).

[50] M. Praprotnik, L. D. Site, and K. Kremer, “Multiscale simulation of soft matter: From scale bridging to adaptive resolution,” Annu. Rev. Phys. Chem. **59**, 545–571 (2008).

[51] S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, and L. Delle Site, “Modeling diffusive dynamics in adaptive resolution simulation of liquid water,” J. Chem. Phys. **128**, 024503 (2008).

[52] J. Zavadlav, R. Podgornik, and M. Praprotnik, “Adaptive resolution simulation of a DNA molecule in salt solution,” J. Chem. Theory Comput. **11**, 5035–5044 (2015).

[53] J. Zavadlav, R. Podgornik, M. N. Melo, S. J. Marrink, and M. Praprotnik, “Adaptive resolution simulation of an atomistic DNA molecule in MARTINI salt solution,” Eur. Phys. J. Spec. Top. **225**, 1595–1607 (2016).

[54] J. Zavadlav, J. Sablić, R. Podgornik, and M. Praprotnik, “Open-boundary molecular dynamics of a DNA molecule in a hybrid explicit/implicit salt solution,” Biophys. J. **114**, 2352–2362 (2018).

[55] J. Zavadlav, R. Podgornik, and M. Praprotnik, “Order and interactions in DNA arrays: Multiscale molecular dynamics simulation,” Sci. Rep. **7**, 4775 (2017).

[56] R. Podgornik, J. Zavadlav, and M. Praprotnik, “Molecular dynamics simulation of high density DNA arrays,” Computation **6**, 3 (2018).

[57] M. Ceriotti, C. Clementi, and O. Anatole von Lilienfeld, “Machine learning meets chemical physics,” J. Chem. Phys. **154**, 160401 (2021).

[58] J. Behler and M. Parrinello, “Generalized neural-network representation of high-dimensional potential-energy surfaces,” Phys. Rev. Lett. **98**, 146401 (2007).

[59] E. Kocer, T. W. Ko, and J. Behler, “Neural network potentials: A concise overview of methods,” Annu. Rev. Phys. Chem. **73**, 163–186 (2022).

[60] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, “Machine learning force fields,” Chem. Rev. **121**, 10142–10186 (2021).

[61] T. Plé, N. Mauger, O. Adjoua, T. J. Inizan, L. Lagardère, S. Huppert, and J.-P. Piquemal, “Routine molecular dynamics simulations including nuclear quantum effects: From force fields to machine learning potentials,” J. Chem. Theory Comput. **19**, 1432–1445 (2023).

[62] F. Noé, A. Tkatchenko, K.-R. Müller, and C. Clementi, “Machine learning for molecular simulation,” Annu. Rev. Phys. Chem. **71**, 361–390 (2020).

[63] P. Reiser, M. Neubert, A. Eberhard, L. Torresi, C. Zhou, C. Shao, H. Metni, C. van Hoesel, H. Schopmans, T. Sommer, and P. Friederich, “Graph neural networks for materials science and chemistry,” Commun. Mater. **3**, 93 (2022).

[64] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, “On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events,” npj Comput. Mater. **6**, 20 (2020).

[65] A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. J. Owen, M. Kornbluth, and B. Kozinsky, “Learning local equivariant representations for large-scale atomistic dynamics,” Nat. Commun. **14**, 579 (2023).

[66] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, “E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials,” Nat. Commun. **13**, 2453 (2022).

[67] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, “A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer,” Nat. Commun. **12**, 398 (2021).

[68] T. Plé, L. Lagardère, and J.-P. Piquemal, “Force-field-enhanced neural network interactions: From local equivariant embedding to atom-in-molecule properties and long-range effects,” Chem. Sci. **14**, 12554–12569 (2023).

[69] O. T. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Sauceda, and K.-R. Müller, “SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects,” Nat. Commun. **12**, 7273 (2021).

[70] D. Reith, M. Pütz, and F. Müller-Plathe, “Deriving effective mesoscale potentials from atomistic simulations,” J. Comput. Chem. **24**, 1624–1636 (2003).

[71] R. Potestio, C. Peter, and K. Kremer, “Computer simulations of soft matter: Linking the scales,” Entropy **16**, 4199–4245 (2014).

[72] S. Izvekov and G. A. Voth, “A multiscale coarse-graining method for biomolecular systems,” J. Phys. Chem. B **109**, 2469–2473 (2005).

[73] Y. Wang, W. G. Noid, P. Liu, and G. A. Voth, “Effective force coarse-graining,” Phys. Chem. Chem. Phys. **11**, 2002–2015 (2009).

[74] L. Lu, J. F. Dama, and G. A. Voth, “Fitting coarse-grained distribution functions through an iterative force-matching method,” J. Chem. Phys. **139**, 121906 (2013).

[75] M. S. Shell, “The relative entropy is fundamental to multiscale and inverse thermodynamic problems,” J. Chem. Phys. **129**, 144108 (2008).

[76] B. Hess, C. Holm, and N. van der Vegt, “Osmotic coefficients of atomistic NaCl (aq) force fields,” J. Chem. Phys. **124**, 164509 (2006).

[77] J.-W. Shen, C. Li, N. F. A. van der Vegt, and C. Peter, “Transferability of coarse-grained potentials: Implicit solvent models for hydrated ions,” J. Chem. Theory Comput. **7**, 1916–1927 (2011).

[78] C. Scherer and D. Andrienko, “Understanding three-body contributions to coarse-grained force fields,” Phys. Chem. Chem. Phys. **20**, 22387–22394 (2018).

[79] P. Gkeka, G. Stoltz, A. Barati Farimani, Z. Belkacemi, M. Ceriotti, J. D. Chodera, A. R. Dinner, A. L. Ferguson, J.-B. Maillet, H. Minoux *et al.*, “Machine learning force fields and coarse-grained variables in molecular dynamics: Application to materials and biological systems,” J. Chem. Theory Comput. **16**, 4757–4775 (2020).

[80] W. G. Noid, “Perspective: Advances, challenges, and insight for predictive coarse-grained models,” J. Phys. Chem. B **127**, 4174–4207 (2023).

[81] Y. Wang, G. Csanyi, and C. Ortner, “Many-body coarse-grained molecular dynamics with the atomic cluster expansion,” arXiv:2502.04661 (2025).

[82] B. Cheng, E. A. Engel, J. Behler, C. Dellago, and M. Ceriotti, “Ab initio thermodynamics of liquid and solid water,” Proc. Natl. Acad. Sci. U. S. A. **116**, 1110–1115 (2019).

[83] C. G. Staacke, H. H. Heenen, C. Scheurer, G. Csányi, K. Reuter, and J. T. Margraf, “On the role of long-range electrostatics in machine-learned interatomic potentials for complex battery materials,” ACS Appl. Energy Mater. **4**, 12562–12569 (2021).

[84] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, “Accurate fourth-generation machine learning potentials by electrostatic embedding,” J. Chem. Theory Comput. **19**, 3567–3579 (2023).

[85] M. Vondrák, K. Reuter, and J. T. Margraf, “q-pac: A Python package for machine learned charge equilibration models,” J. Chem. Phys. **159**, 054109 (2023).

[86] P. Fusch, M. Sanocki, and J. Zavadlav, “Learning non-local molecular interactions via equivariant local representations and charge equilibration,” npj Comput. Mater. **11**, 287 (2025).

[87] A. Grisafi and M. Ceriotti, “Incorporating long-range physics in atomic-scale machine learning,” J. Chem. Phys. **151**, 204105 (2019).

[88] B. Cheng, “Latent Ewald summation for machine learning of long-range interactions,” npj Comput. Mater. **11**, 80 (2025).

[89] A. Kosmala, J. Gasteiger, N. Gao, and S. Günnemann, “Ewald-based long-range message passing for molecular graphs,” in *International Conference on Machine Learning* (PMLR, 2023), pp. 17544–17563.

[90] P. Loche, K. K. Huguenin-Dumittan, M. Honarmand, Q. Xu, E. Rumiantsev, W. B. How, M. F. Langer, and M. Ceriotti, “Fast and flexible long-range models for atomistic machine learning,” J. Chem. Phys. **162**, 142501 (2025).

[91] A. Caruso, J. Venturin, L. Giambagli, E. Rolando, F. Noé, and C. Clementi, “Extending the RANGE of graph neural networks: Relaying attention nodes for global encoding,” arXiv:2502.13797 (2025).

[92] Z. Deng, C. Chen, X.-G. Li, and S. P. Ong, “An electrostatic spectral neighbor analysis potential for lithium nitride,” npj Comput. Mater. **5**, 75 (2019).

[93] S. Wengert, G. Csányi, K. Reuter, and J. T. Margraf, “Data-efficient machine learning for molecular crystal structure prediction,” Chem. Sci. **12**, 4536–4546 (2021).

[94] S. Wengert, G. Csányi, K. Reuter, and J. T. Margraf, “A hybrid machine learning approach for structure stability prediction in molecular co-crystal screenings,” J. Chem. Theory Comput. **18**, 4586–4593 (2022).

[95] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, “DeePCG: Constructing coarse-grained models via deep neural networks,” J. Chem. Phys. **149**, 034101 (2018).

[96] S. Thaler and J. Zavadlav, “Learning neural network potentials from experimental data via differentiable trajectory reweighting,” Nat. Commun. **12**, 6884 (2021).

[97] N. E. Charron, K. Bonneau, A. S. Pasos-Trejo, A. Guljas, Y. Chen, F. Musil, J. Venturin, D. Gusew, I. Zaporozhets, A. Krämer *et al.*, “Navigating protein landscapes with a machine-learned transferable coarse-grained model,” Nat. Chem. **17**, 1284 (2025).

[98] B. E. Husic, N. E. Charron, D. Lemm, J. Wang, A. Pérez, M. Majewski, A. Krämer, Y. Chen, S. Olsson, G. de Fabritiis *et al.*, “Coarse graining molecular dynamics with graph neural networks,” J. Chem. Phys. **153**, 194101 (2020).

[99] Y. Chen, A. Krämer, N. E. Charron, B. E. Husic, C. Clementi, and F. Noé, “Machine learning implicit solvation for molecular dynamics,” J. Chem. Phys. **155**, 084101 (2021).

[100] A. E. P. Durumeric, N. E. Charron, C. Templeton, F. Musil, K. Bonneau, A. S. Pasos-Trejo, Y. Chen, A. Kelkar, F. Noé, and C. Clementi, “Machine learned coarse-grained protein force-fields: Are we there yet?,” Curr. Opin. Struct. Biol. **79**, 102533 (2023).

[101] P. Katzberger and S. Riniker, “Implicit solvent approach based on generalized Born and transferable graph neural networks for molecular dynamics simulations,” J. Chem. Phys. **158**, 204101 (2023).

[102] S. Thaler, M. Stupp, and J. Zavadlav, “Deep coarse-grained potentials via relative entropy minimization,” J. Chem. Phys. **157**, 244103 (2022).

[103] S. Thaler, G. Doehner, and J. Zavadlav, “Scalable Bayesian uncertainty quantification for neural network potentials: Promise and pitfalls,” J. Chem. Theory Comput. **19**, 4520–4532 (2023).

[104] A. Kramer, A. E. Durumeric, N. E. Charron, Y. Chen, C. Clementi, and F. Noé, “Statistically optimal force aggregation for coarse-graining molecular dynamics,” J. Phys. Chem. Lett. **14**, 3970–3979 (2023).

[105] P. R. Vlachas, J. Zavadlav, M. Praprotnik, and P. Koumoutsakos, “Accelerated simulations of molecular systems through learning of effective dynamics,” J. Chem. Theory Comput. **18**, 538–549 (2021).

[106] S. Wang, Z. Ma, and W. Pan, “Data-driven coarse-grained modeling of non-equilibrium systems,” Soft Matter **17**, 6404–6412 (2021).

[107] M. J. Del Razo, D. Crommelin, and P. G. Bolhuis, “Data-driven dynamical coarse-graining for condensed matter systems,” J. Chem. Phys. **160**, 024108 (2024).

[108] B. P. Fingerhut, “The mutual interactions of RNA, counterions and water–quantifying the electrostatics at the phosphate–water interface,” Chem. Commun. **57**, 12880–12897 (2021).

[109] H. T. Nguyen, N. Hori, and D. Thirumalai, “Theory and simulations for RNA folding in mixtures of monovalent and divalent cations,” Proc. Natl. Acad. Sci. U. S. A. **116**, 21022–21030 (2019).

[110] I.-B. Magdău and G. Csányi, “A predictive framework for liquid electrolytes takes root with bamboo: Molecular dynamics,” Nat. Mach. Intell. **7**, 983–984 (2025).

[111] S. Banerjee and A. Tkatchenko, “Non-local interactions determine local structure and lithium diffusion in solid electrolytes,” Nat. Commun. **16**, 1672 (2025).

[112] N. Bergmann, N. Bonnet, N. Marzari, K. Reuter, and N. G. Hörmann, “Machine learning the energetics of electrified solid-liquid interfaces,” Phys. Rev. Lett. **135**, 146201 (2025).

[113] M. Hellström and J. Behler, “Structure of aqueous NaOH solutions: Insights from neural-network-based molecular dynamics simulations,” Phys. Chem. Chem. Phys. **19**, 82–96 (2017).

[114] M. Hellstrom, M. Ceriotti, and J. Behler, “Nuclear quantum effects in sodium hydroxide solutions from neural network molecular dynamics simulations,” J. Phys. Chem. B **122**, 10158–10171 (2018).

[115] C. Schran, F. L. Thiemann, P. Rowe, E. A. Müller, O. Marsalek, and A. Michaelides, “Machine learning potentials for complex aqueous systems made simple,” Proc. Natl. Acad. Sci. U. S. A. **118**, e2110077118 (2021).

[116] Y. Shao, M. Hellström, A. Yllö, J. Mindemark, K. Hermansson, J. Behler, and C. Zhang, “Temperature effects on the ionic conductivity in concentrated alkaline electrolyte solutions,” Phys. Chem. Chem. Phys. **22**, 10426–10430 (2020).

[117] J. Zhang, J. Pagotto, and T. T. Duignan, “Towards predictive design of electrolyte solutions by accelerating ab initio simulation with neural networks,” J. Mater. Chem. A **10**, 19560–19571 (2022).

[118] P. Katzberger, F. Pultar, and S. Riniker, “Transferring knowledge from MM to QM: A graph neural network-based implicit solvent model for small organic molecules,” J. Chem. Theory Comput. **21**, 7450–7459 (2025).

[119] A. Coste, E. Slejko, J. Zavadlav, and M. Praprotnik, “Developing an implicit solvation machine learning model for molecular simulations of ionic media,” J. Chem. Theory Comput. **20**, 411–420 (2023).

[120] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, “Comparison of simple potential functions for simulating liquid water,” J. Chem. Phys. **79**, 926–935 (1983).

[121] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee *et al.*, “A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations,” J. Comput. Chem. **24**, 1999–2012 (2003).

[122] I. S. Joung and T. E. Cheatham III, “Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations,” J. Phys. Chem. B **112**, 9020–9041 (2008).

[123] J. Yoo and A. Aksimentiev, “Improved parametrization of Li^+, Na^+, K^+, and Mg^{2+} ions for all-atom molecular dynamics simulations of nucleic acid systems,” J. Phys. Chem. Lett. **3**, 45–50 (2012).

[124] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. In’t Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen *et al.*, “LAMMPS—A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales,” Comput. Phys. Commun. **271**, 108171 (2022).

[125] L. Verlet and D. Levesque, “On the theory of classical fluids II,” Physica **28**, 1124–1142 (1962).

[126] L. Verlet, “Computer ‘experiments’ on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules,” Phys. Rev. **159**, 98 (1967).

[127] D. J. Evans and B. L. Holian, “The Nose–Hoover thermostat,” J. Chem. Phys. **83**, 4069–4074 (1985).

[128] J. W. Eastwood, R. W. Hockney, and D. N. Lawrence, “P3M3DP-the three-dimensional periodic particle-particle/particle-mesh program,” Comput. Phys. Commun. **19**, 215–261 (1980).

[129] J. R. Nagel *et al.*, "Solving the generalized Poisson equation using the finite-difference method (FDM)," in *Lecture Notes* (Department of Electrical and Computer Engineering, University of Utah, 2011), Vol. 52.

[130] M. Holst and F. Saied, "Multigrid solution of the Poisson–Boltzmann equation," J. Comput. Chem. **14**, 105–113 (1993).

[131] M. J. Holst and F. Saied, "Numerical solution of the nonlinear Poisson–Boltzmann equation: Developing more robust and efficient methods," J. Comput. Chem. **16**, 337–364 (1995).

[132] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, "Electrostatics of nanosystems: Application to microtubules and the ribosome," Proc. Natl. Acad. Sci. U. S. A. **98**, 10037–10041 (2001).

[133] N. A. Kovaleva, I. P. Koroleva, M. A. Mazo, and E. A. Zubova, "The 'sugar' coarse-grained DNA model," J. Mol. Model. **23**, 66 (2017).

[134] S. A. Pabit, X. Qiu, J. S. Lamb, L. Li, S. P. Meisburger, and L. Pollack, "Both helix topology and counterion distribution contribute to the more effective charge screening in dsRNA compared with dsDNA," Nucleic Acids Res. **37**, 3887–3896 (2009).

[135] M. K. Gilson, A. Rashin, R. Fine, and B. Honig, "On the calculation of electrostatic interactions in proteins," J. Mol. Biol. **184**, 503–516 (1985).

[136] B. Honig, A. M. Pyle, K. Chin, and K. A. Sharp, "Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function," Nat. Struct. Biol. **6**, 1055–1061 (1999).

[137] J. D. Morrow, J. L. A. Gardner, and V. L. Deringer, "How to validate machine-learned interatomic potentials," J. Chem. Phys. **158**, 121501 (2023).

[138] S. M. Goodlett, J. M. Turney, and H. F. Schaefer, "Comparison of multifidelity machine learning models for potential energy surfaces," J. Chem. Phys. **159**, 044111 (2023).

[139] I. Batatia, S. Batzner, D. Kovács, A. Musaelian, G. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, "Scaling the leading accuracy of deep equivariant models to biomolecular simulations of realistic size," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (ACM, 2023), pp. 1–12

[140] I. Batatia, S. Batzner, D. P. Kovács, A. Musaelian, G. N. C. Simm, R. Drautz, C. Ortner, B. Kozinsky, and G. Csányi, "The design space of E(3)-equivariant atom-centred interatomic potentials," Nat. Mach. Intell. **7**, 56–67 (2025).

[141] S. Stocker, J. Gasteiger, F. Becker, S. Günnemann, and J. T. Margraf, "How robust are modern graph neural network potentials in long and hot molecular dynamics simulations?," Mach. Learn.: Sci. Technol. **3**, 045010 (2022).

[142] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, "Big data meets quantum chemistry approximations: The δ-machine learning approach," J. Chem. Theory Comput. **11**, 2087–2096 (2015).

[143] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, "The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models," J. Chem. Phys. **128**, 244114 (2008).

[144] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 (2014).

[145] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, "PyTorch: An imperative style, high-performance deep learning library," arXiv:1912.01703 (2019).

[146] G. Fiorin, M. L. Klein, and J. Hénin, "Using collective variables to drive molecular dynamics simulations," Mol. Phys. **111**, 3345–3362 (2013).

[147] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," J. Comput. Chem. **13**, 1011–1021 (1992).

[148] A. Grossfield, "WHAM: The weighted histogram analysis method, version 2.0.11," http://membrane.urmc.rochester.edu/?page_id=126 (accessed November 2025).

[149] H. S. Antila and M. Sammalkorpi, "Polyelectrolyte decomplexation via addition of salt: Charge correlation driven zipper," J. Phys. Chem. B **118**, 3226–3234 (2014).

[150] H. S. Antila, M. Härkönen, and M. Sammalkorpi, "Chemistry specificity of DNA–polycation complex salt response: A simulation study of DNA, polylysine and polyethyleneimine," Phys. Chem. Chem. Phys. **17**, 5279–5289 (2015).