ENHANCING NUCLEIC ACID SIMULATIONS USING ASYMMETRIC PERIODIC BOUNDARY CONDITIONS

G VENKANNA[1], K CHAKRAVARTHY[2], M SWATHI[3], E SUDHEER[4]

[1,2,3,4] Assistant Professor, Department of Science and Humanities, AnuBose Institute of Technologyfor Women's, KSP Road, New Paloncha, Bhadradri Kothagudem District, Telangana (TS), 507115

Submitted: 02-06-2024 Accepted: 11-07-2024 Published: 18-07-2024

ABSTRACT

Molecular dynamics and coarse-grained simulations are essential tools for understanding the structural dynamics and functional behavior of nucleic acids at multiple length and time scales. However, conventional symmetric periodic boundary conditions (PBCs) often introduce artificial constraints that can affect accuracy, especially in elongated or directionally biased biomolecular systems such as DNA and RNA. This study presents an advanced simulation framework based on asymmetric periodic boundary conditions (APBCs) to enhance the realism and efficiency of nucleic acid simulations. The proposed approach allows differential periodicity along selected spatial dimensions, enabling more faithful representation of nucleic acid conformational flexibility while reducing finite-size and boundary artifacts. APBCs are implemented and evaluated in both all-atom molecular dynamics and coarse-grained models, demonstrating improved stability, reduced computational overhead, and better agreement with known structural and dynamical properties. The results highlight the effectiveness of asymmetric boundary treatments in multiscale nucleic acid modeling, offering a robust alternative to traditional PBC methods for high-accuracy biomolecular simulations.

Keywords: Asymmetric periodic boundary conditions; Molecular dynamics; Coarse-grained simulations; Nucleic acids; Multiscalemodeling; Computational biophysics.

I. INTRODUCTION

Molecular dynamics (MD) and coarse-grained (CG) simulations have become indispensable techniques for investigating the structure, dynamics, and interactions of nucleic acids at atomic and mesoscopic scales. These computational approaches provide valuable insights into DNA and RNA behavior that are often difficult to obtain experimentally, including conformational transitions, mechanical properties, and sequence-dependent effects. Accurate simulation of such systems, however, depends critically on the treatment of boundary conditions used to approximate bulk environments while maintaining computational feasibility.

Periodic boundary conditions (PBCs) are widely employed in biomolecular simulations to eliminate surface effects and mimic infinite systems. In their conventional symmetric form, PBCs assume uniform periodicity along all spatial dimensions, which is suitable for roughly isotropic systems. Nucleic acids, however, often exhibit strong anisotropy due to their elongated structures and directional flexibility. Applying symmetric PBCs to these systems can introduce artificial interactions between periodic images, distort long-range correlations, and impose unnecessary constraints on molecular motion.

To overcome these limitations, asymmetric periodic boundary conditions (APBCs) have been proposed as an alternative framework that allows different periodicity or boundary treatments along selected axes. By tailoring boundary conditions to the intrinsic geometry of nucleic acid systems,

APBCs reduce finite-size artifacts and enable more realistic modeling of directional conformational changes. This approach is particularly advantageous in multiscale simulations that combine all-atom MD with coarse-grained representations, where computational efficiency and physical accuracy must be carefully balanced.

This work focuses on the development and application of APBCs in both molecular dynamics and coarse-grained simulations of nucleic acids. By evaluating their impact on structural stability, dynamical behavior, and computational performance, the study demonstrates how asymmetric boundary treatments can improve the fidelity of nucleic acid simulations and support more accurate multiscalemodeling in computational biophysics.

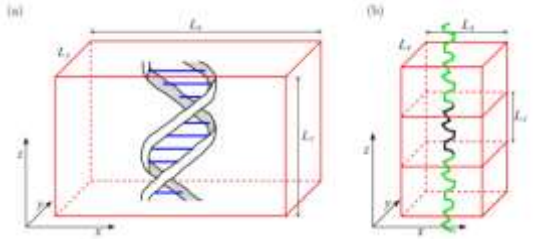

Figure 1: (a) Schematic of the simulation domain (1) for the case with 10 base pairs of DNA, i.e. for n = 1. (b) Discrete worm-like chain segment in an APBC simulation is denoted by the black line. Its periodic image copies by the green line.

in the cuboid computational domain

$$\Omega = [0, L_x] \times [0, L_y] \times [0, L_z]. \qquad (1)$$

The main idea behind the APBC is that DNA is periodic with the period of 10 base pairs, i.e. the APBC will allow us to use 10 n base pairs of DNA in domain **Ω**, where n ∈ N is an integer denoting the number of helical pitches. A schematic of our simulation domain is presented in Figure 1(a) for the case of the simulation with 10 base pairs, i.e. for n = 1. The DNA molecule is positioned parallel to the z-axis and we use periodic boundary conditions in the z-direction. Such a periodic boundary condition in z-direction is less common in allatom MD simulation studies, where the biomolecule of interest is often placed in the middle of the computational domain and it is solvated on all its sides by a layer of water molecules separating the biomolecule from the domain boundary.

Considering the projection into the xy-plane, the DNA molecule is positioned in the middle of the simulated domain. In particular, the DNA molecule is separated by the layer of water molecules from the boundaries of the simulated domain in both x-direction and y-direction. While we use periodic boundary conditions in all three directions, there is an asymmetry (highlighted in our terminology APBC): a modeller has a relative freedom to choose the values of Lx and Ly in the computational domain defined by (1), while the value of Lz is dictated by the properties of the simulated biomolecule. The imposed DNA periodicity fixes the helical twist of the DNA molecule with the simulation box size Lz chosen such that it exactly corresponds to n helical pitches. However, considering simulations at isothermal-isobaric (NpT) ensemble, standard isotropic barostats introduce fluctuations in the domain size leading to changes in Lz as well. To fix Lz, an asymmetric barostat is used in Section 3 of this paper.

The APBC has been used in previous studies5,21,22 to mimic an infinitely long DNA molecule. Except of the asymmetry between the z-direction and x-direction (resp. ydirection), the APBC can lead to a relatively standard all-atom MD set up, with the domain periodic in all three directions, which was previously used to explore the ion atmosphere around the DNA.21,22 However, it is more challenging to use the APBC to study mechanical properties of biopolymers, as we will first illustrate in Section 2 by considering a discrete worm-like chain model. This is followed by all-atom MD

simulations of DNA in Section 3, where we present the use of APBC to investigate mechanical properties of the DNA and the properties of the surrounding solvent.

II. WORM-LIKE CHAIN MODEL

Let us consider the discrete worm-like chain (WLC) model where DNA consists of N segments li ,i = 1, 2, . . . , N, each having the same length, ℓ. Denoting the angle between the adjacent i-th and (i+1)-th segments by θi ,for i = 1, 2, . . . ,(N −1), the chain bending energy is

$$\frac{\mathrm{E}}{k_{\mathrm{B}}T} = \alpha \sum_{i=1}^{N-1} \theta_i^2, \qquad (2)$$

where α is a dimensionless constant. We define the persistence length of the first j-th segments, for j ≤ N, by

$$a_j = \left\langle \frac{\mathbf{l}_1}{\ell} \cdot \sum_{i=1}^{j} \mathbf{l}_i \right\rangle. \qquad (3)$$

That is, aj is the average value of the projection of the vector connecting the end points of the first and the j-th segment on the direction of the first segment. Then the persistence length of the WLC model can be defined as the limit

$$a_{\mathrm{orig}} = \lim_{N\to\infty} a_N, \qquad (4)$$

which effectively is the average value of the projection of the end-to-end vector of a long chain on the direction of the first segment. The average in (3) can be evaluated as

$$a_j = \sum_{i=1}^{j} \frac{\langle \mathbf{l}_1 \cdot \mathbf{l}_i \rangle}{\ell} = \ell \sum_{i=1}^{j} \langle \cos(\theta) \rangle^{i-1} = \ell \frac{1 - \langle \cos(\theta) \rangle^j}{1 - \langle \cos(\theta) \rangle}, \qquad (5)$$

where the average hcos(θ)i is given by

$$\langle \cos(\theta) \rangle = \int_0^\pi \cos(\theta) \sin(\theta) \exp\left[-\alpha\, \theta^2\right] \mathrm{d}\theta \Big/ \int_0^\pi \sin(\theta) \exp\left[-\alpha\, \theta^2\right] \mathrm{d}\theta. \qquad (6)$$

To get formula (6), we note that the distribution of angles between adjacent segments is proportional to sin(θ) exp [−α θ2]. Using (4), (5) and (6), we deduce

$$a_{\mathrm{orig}} = \frac{\ell}{1 - \langle \cos(\theta) \rangle} = \ell \left(2\alpha + \frac{2}{3} + \frac{1}{15\alpha} + \mathcal{O}\left(\frac{1}{\alpha^2}\right) \right), \quad \text{as} \quad \alpha \to \infty, \qquad (7)$$

$$= \ell \left(1 + \frac{\pi^2}{8}\alpha + \mathcal{O}(\alpha^2) \right), \quad \text{as} \quad \alpha \to 0. \qquad (8)$$

In Figure 2(a), we present how the persistence length aorig depends on the stiffness parameter α in interval [0, 3/2], illustrating the accuracy of both expansions (7) and (8). While (7) is derived in the limit α → ∞, it approximates the exact result well for persistence lengths

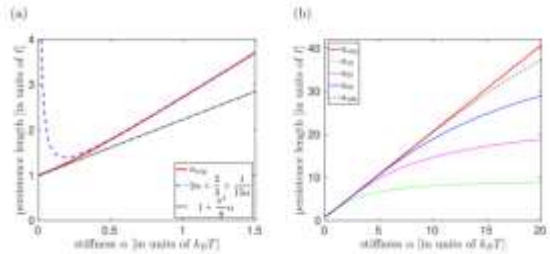

Figure 2: (a) Plot of persistence length aorig, given by (4), as a function of the stiffness parameter α, together with asymptotic results (7) and (8). The dimensionless parameter α can be viewed to express energy in units kBT, while all persistence lengths are plotted in units of the segment lentgh, ℓ. (b) Plot of persistence length aorig, given by (4), and persistence lengths aj , given by (5), for j = 10, 20, 50, 100, as a function of parameter α.

satisfyingaorig> 2ℓ or equivalently for α > 0.62. In Figure 2(b), we plot the dependence of the persistence length aorig on the stiffness parameter α in a larger interval [0, 20] together with the values of aj given by (5). Using the exact result for aorig given on the left hand side of equation (7), we can rewrite (5) as follows

$$a_j = a_{\text{orig}} \left(1 - \langle \cos(\theta) \rangle^j\right). \tag{9}$$

Considering the limit α → ∞ in (6), we have

$$\langle \cos(\theta) \rangle = 1 - \frac{1}{2\alpha} + \frac{1}{6\alpha^2} + \mathcal{O}\left(\frac{1}{\alpha^3}\right), \qquad \text{as} \quad \alpha \to \infty,$$

where the first three terms of the expansion on the right hand side provide an approximation of hcos(θ)i with about 5% relative error for α > 1, and the relative error decreases as we increase α, for example, the relative error is smaller than 1% for α > 2. Substituting this expansion for hcos(θ)i into equation (9), we obtain that for sufficiently large values of α, say for α > 1, we can calculate the persistence length aorig from aj by using the following formula

$$a_{\text{orig}} = \frac{a_j}{1 - \left(1 - \frac{1}{2\alpha} + \frac{1}{6\alpha^2}\right)^j}. \tag{10}$$

2.1 The dependence of persistence length on APBC

Considering that the polymer chain is simulated in the domain (1) with APBC, we have an extra constraint

$$\sum_{i=1}^{N} \mathbf{l}_i = [0, 0, L_z], \tag{11}$$

where N denotes the number of simulated segments along the z-direction. As it is illustrated in Figure 1(b), such a model can be viewed as a model of an (infinitely) long polymer chain by using the periodicity

$$\mathbf{l}_i = \mathbf{l}_{i \bmod N}, \qquad \text{for } i \in \mathbb{N}. \tag{12}$$

However, substituting (11)–(12) into the definition of persistence length (4), we would obtain that aorig = ∞ because the periodic boundary means that the infinitely long filament is effectively straight. Since equation (11) postulates that the vector connecting ends of N segments is fixed, we obtain the most variability in this model by looking at the behaviour of the ⌊N/2⌋ consecutive segments. Due to the symmetry of the problem and condition (4), the average of the vector P⌊N/2⌋i=1 li is equal to [0, 0, Lz/2] for any value of α, but the deviations from this average will depend on α. To illustrate this, we define the average distance of the polymer middle point from the axis of the polymer by

$$\left\langle \left\| [1, 1, 0] \cdot \sum_{i=1}^{\lfloor N/2 \rfloor} \mathbf{l}_i \right\| \right\rangle, \tag{13}$$

that is, we calculate the (Euclidean) norm of the projection of the vector P⌊N/2⌋i=1 li on the x–y–plane. The average (13) is plotted in Figure 3(a) for different values of parameter α and

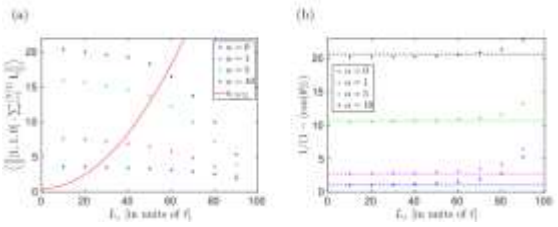

Figure 3: (a) The average distance of the polymer middle point from the axis of the polymer, defined by (13), estimated from Monte Carlo simulations of the WLC model with N = 100 segments for Lz∈ {10, 20, . . . , 90} and α ∈ {0, 1, 5, 10}. The red line shows a⌊N/2⌋ = a50 as a function of the domain length Lz (theoretical result (14) confirmed by simulations for all considered values of α).

(b) The estimate of aorig given by equation (15), where hcos(θ)i is estimated from Monte Carlo simulations of the WLC model with N = 100 segments for Lz∈ {10, 20, . . . , 90} and α ∈ {0, 1, 5, 10}. The theoretical result (without APBC, independent of Lz), given by equation (6), is plotted by dashed lines for each value of α.

domain length Lz. We observe that, for fixed value of Lz, the average (13) increases with the value of the stiffness parameter α. Moreover, Figure 3(a) also shows that the value of the average (13) approaches zero as Lz approaches its maximum possible value, N ℓ. Indeed, if Lz = N ℓ, the polymer is straight and the value of (13) is exactly equal to zero. On the other hand, if Lz is smaller then we obtain a larger value of (13), especially for polymers with larger persistence length (i.e. for large values of α). On the face of it, one possible way to estimate aorig could be to estimate a⌊N/2⌋ from our Monte Carlo simulations and then use formula (10) for j = ⌊N/2⌋. However, formula (10) has been derived for the case of the WLC model in the 3-dimensional physical space R 3 . Considering the APBC, we obtain that a⌊N/2⌋ is independent of α (see Appendix A). We have

$$a_{\lfloor N/2 \rfloor} = \ell \left(\frac{N - \lfloor N/2 \rfloor}{N-1} + \frac{L_z^2(\lfloor N/2 \rfloor - 1)}{N(N-1)\ell^2} \right), \qquad (14)$$

which simplifies to a⌊N/2⌋ ≈ (ℓ/2) + L 2 z /(2ℓN) for large values of N. This result is also visualized in Figure 3(a). In particular, a better strategy to obtain the real persistence length aorig from the APBC simulations is to estimate

$$\langle \mathbf{l}_i \cdot \mathbf{l}_{i+1} \rangle = \langle \cos(\theta) \rangle$$

and then use the exact result for aorig given on the left hand side of equation (7), namely

$$a_{\text{orig}} = \frac{\ell}{1 - \langle \cos(\theta) \rangle}. \qquad (15)$$

The results are presented in Figure 3(b).

III. APBC IN ALL-ATOM MD SIMULATIONS

In this section, we investigate the use of APBC in all-atom MD models of DNA. Our simulations are performed with 10–100 base pairs (bp) of double-stranded DNA (dsDNA). Since we use the APBC, all simulations are effectively simulating (infinitely) long DNA chains. In particular, MD results with the longest simulated chain (100 bp) can be used as the 'ground truth' for the presented APBC simulations with shorter 10–50 bp long DNA chains. We note that the MD simulations of relatively short 50 bp DNA segments without APBC have been previously used in the literature to estimate the DNA persistence length by using a middle section of the simulated DNA segment.

We consider 6 types of (infinitely) long DNA sequences, with repeated nucleotides, namely poly(A), poly(C), poly(AT), poly(CG), poly(AC) and poly(AG), where poly(X) means that the correspoding nucleotide sequence is periodically repeated. We note that these 6 cases correspond to all possible cases of pairs of nucleotides which are repeated infinitely many times. For example, repetitions of dinucleotides AC, CA, TG and GT all correspond to the poly(AC) case, because AC and CA are equivalent due to the periodic boundary conditions along the chain length, and TG is on the complementary strand, with GT being equivalent to TG because of the periodic boundary conditions.

Each infinitely long sequence is modelled in our computational domain (1) with APBC using N = 10 n base pairs of DNA, where n ranges from 1 to 10. The APBC is implemented along the z-direction as detailed in Appendix B.1. First, an (N + 1) bp long dsDNA configuration is constructed in such a way

that the (N + 1)-th base pair is equivalent to the first base pair translated to the z-direction. Then, a nucleotide at the 3'-end of each strand is removed and the bond to the 3'-end (removed) nucleotide is substituted with that to the first base at the 5'-end. The corresponding angles and dihedrals are added to MD structural files as detailed in Table 1 in Appendix B.1. In all MD simulations, we consider domain (1) with Lx = Ly = 200 Å and we vary Lz. In Figures 4, 5 and 7, we choose Lz as a multiple of n (resp. N) with

$$L_z = 3.375N\,\text{Å} = 33.75n\,\text{Å}, \qquad (16)$$

while we study the effect of stretching and shrinking of DNA in Figure 6 by using Lz obtained as the 95%, 100% and 105% of the value given by equation (16). All MD simulations are done in KCl solutions, with K+ ions neutralizing the negatively charged DNA segments. We use the concentration 150 mMKCl in Figures 4, 5 and 6, while we vary the concentration of KCl in Figure 7.

When using APBC with polymer models, there are (locally) two important directions: parallel to the polymer chain and perpendicular to the polymer chain. We consider both of them, in Sections 3.1 and 3.2, respectively. In Section 3.1, we study the effects of APBC on the properties of the DNA chain, where we can make direct analogues to the results obtained for the persistence length of the WLC model in Section 2. This is followed by studying the characteristics of the surrounding solvent in Section 3.2, where we investigate the ion athmosphere around DNA for different concentrations of KCl.

3.1 Mechanical properties along the chain

The persistence length for our (infinitely) long sequences of dinucleotides can be determined by various experimental and theoretical studies23 as summarized in Appendix B. In Figure 4, we present the results of all-atom MD simulations with APBC of N = 10 bp segments using the six cases of repeated dinucleotides. Technical details of these MD simulations are given in Appendix B.

To analyze our MD results, we associate a unit orientation vector hi with each base pair, i.e. i = 1, 2, . . . , N, where N = 10n is the total number of simulated base pairs. Denoting the angle between the i-th and (i + j)-th base pair as φj , we have cos(φj) = hi · hi+j , which we calculate for all i = 1, 2, . . . , N. Averaging the calculated results over all possible values of i, we have

$$\langle \cos(\phi_j) \rangle = \langle \mathbf{h}_i \cdot \mathbf{h}_{i+j} \rangle , \qquad (17)$$

where the accuracy of this average is further improved by calculating it as a time average over long MD time series. More precisely, we calculate three independent time series of length 10 ns and sample our results every 10 ps, disregarding the beginning of each simulation as the time required to equilibrate the system, see Appendix B for more details. Considering N = 10 (i.e. n = 1), we plot the averages (17) in Figure 4(a) for values j = 1, 2, 3, 4, 5. We note that

$$\langle \cos(\phi_j) \rangle = \langle \cos(\phi_{N-j} \rangle = \langle \cos(\phi_{10n-j} \rangle ,$$

because we use APBC. In particular, the values of the averages (17) for j = 6, 7, . . . are already represented in Figure 4(a) by the corresponding values for j = 1, 2, 3, 4, 5. In Figure 4, we observe that the results are clearly sequence dependent for j = 1, with the j = 4 case providing the best match to the j = 1 case. On the other hand, the results are less sequence dependent for j = 2 or j = 5. Given the APBC, there is no variation for j = N = 10 as we have already observed for the WLC model, because of the constraint (11). In Figure 4(b), we present the average separation between the subsequent base pairs for each of the studied

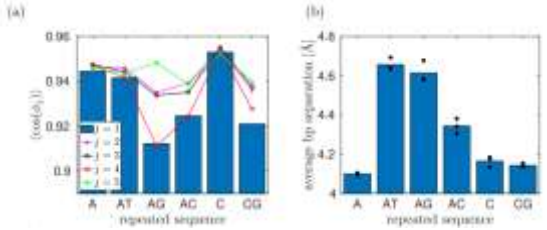

Figure 4: The results of all-atom MD simulations of DNA chains with APBC. (a) The average (17) for each of the 6 considered sequences of repeated nucleotides is calculated using three independent MD simulations. (b) The average separation between the base pairs calculated using three independent MD simulations (blue bars). The results for each individual realization is plotted as a black dot.

case.

Our MD simulations in Figure 4 use the smallest possible value of N (corresponding to n = 1), while one can expect that the results of all-atom MD simulations should be less influenced by the APBC for larger values of n (in theory, the APBC-induced errors should decrease to zero in the limit $n \to \infty$). To investigate this further, we study the dependence of our results on n for the poly(A) case in Figure 5(a). We use three independent MD simulations for n = 1, 2, 3, 4, 5, 10 corresponding to simulations with N ranging from 10bp to 100bp. In each case, we plot the averages (17) for j = 1, 2, . . . , 10. We note that this average is trivially equal to 1 in the case j = 10 for N = 10 bp (because the first and the eleventh base pairs are identical for N = 10 bp), so we omit this artificial value from our plot for 10 bp in Figure 5(a). We observe that the results for n = 1, 2, 3, 4, 5 are matching some trends of the results for 100 bp. In particular, we can make similar conclusions as in Section 2 that the local properties (smaller values of j) are less influenced by using APBC than the averages estimated over the whole simulated polymer length (for j comparable to N).

In Figure 3, we have considered the WLC model with N = 100 segments while varying the

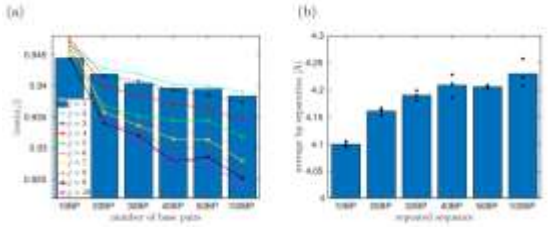

Figure 5: The results of all-atom MD simulations with APBC using the poly(A) DNA chain with N in the range 10–100 bp. (a) The average (17) for each of the 6 values of N considered is calculated using three independent MD simulations. The results are presented for j = 1, 2, . . . , 10 and n = 1, 2, 3, 4, 5, 10. (b) The average separation between the base pairs calculated using three independent MD simulations (blue bars). The results for each individual realization is plotted as a black dot.

domain length Lz. In Figure 6, we present the results of a similar study using all-atom MD simulations with N = 100 bp. The middle bars in Figure 6(a) and Figure 6(b) correspond to the results of the poly(A) case with 100 bp which has already been included in Figure 5(a). Using equation (16), this corresponds to Lz = 337.5 Å. The other simulations correspond to the same set up where we either extend or shrink the value of Lz by 5%, i.e. we use the values of Lz given as

$$L_z = 320.625\,\text{Å}, \qquad L_z = 337.5\,\text{Å}, \qquad \text{and} \qquad L_z = 354.375\,\text{Å}. \tag{18}$$

In Figure 6(b), we observe that the average separation between base pairs increases as we increase Lz. On the other hand, the behaviour of averages (17) is less monotonic as we stretch or shrink the DNA chain, see Figure 6(a). Another way to visualize the results of all atom MD simulations is to consider the average (17) as a function of the distance between the base pairs,20 which is visualized as function hHi in Figure 6(c). To calculate hHi, we

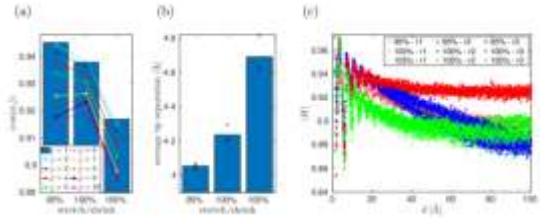

Figure 6: The results of MD simulations of 100 bp poly(A) dsDNA with APBC that use the values of Lz given by equation (18). We average over three independent MD time series for each of the presented case. (a) The average (17) for j = 1, 2, . . . , 10. (b) The average separation of base pairs (blue bars). Dots include the results for individual MD realizations (i.e. we have averaged over the dots to calculate blue bars). (c) The average hHi as a function of distance d. We present results for the 95% (blue), 100% (red) and 105% (green) cases using different colours. Different symbols (circle, square, triangle of the same colour) denote data points calculated by different MD realizations.

averagehhi · hji over all pairs i and j such that the corresponding base pairs are the distance d apart. We present this average, hHi, as a function of the distance d in Figure 6(c). The rate of decay of function hHi with distance d can be used as an alternative way to define and estimate the persistence length from MD simulations.

3.2 Ion atmosphere

The APBC are useful for investigating solvent properties in the direction perpendicular to the polymer chain. In Figure 7, we present the results of such a study, calculating the radial distribution of K+ and Cl− ions. We use three different concentrations of KCl, namely 0.25M, 0.5M and 1M. In each case, we use n = 1, i.e. we use the APBC with 10 bp of poly(A) dsDNA. The results are calculated by averaging over four independent MD time series, each calculated for 10 ns. After the initial transient (of 1 ns) and at equidistant time intervals of 10 ps, we calculate the distance of each ion from the nearest atom of DNA, so our raw data are given in terms of the histograms

$$N_{K^+}(r,\Delta r) = [\text{number of ions with the distance from DNA in interval } (r, r+\Delta r)],$$
$$N_{Cl^-}(r,\Delta r) = [\text{number of ions with the distance from DNA in interval } (r, r+\Delta r)].$$

To get the radial distribution function, these numbers have to be divided by the volume, V (r, Δr), giving the volume of all points which have their distance from the DNA in the interval (r, r + Δr). Then the radial distribution of K+ ions and Cl− ions is defined by

$$g_{K^+}(r) = \lim_{\Delta r\to\infty} \frac{N_{K^+}(r,\Delta r)}{V(r,\Delta r)}, \quad \text{and} \quad g_{Cl^-}(r) = \lim_{\Delta r\to\infty} \frac{N_{Cl^-}(r,\Delta r)}{V(r,\Delta r)}, \qquad (19)$$

where r is the distance from the DNA. To calculate Figure 7, we approximate the limit in equation (19) by choosing (relatively small) value Δr = 1 Å and we approximate the DNA as a straight line (or equivalently as a straight cylinder) in the z-direction, i.e. V (r, Δr) = 2π r Δr Lz, giving

$$g_{K^+}(r) \approx \frac{N_{K^+}(r,\Delta r)}{2\pi\, r\, \Delta r\, L_z}, \quad g_{Cl^-}(r) \approx \frac{N_{K^+}(r,\Delta r)}{2\pi\, r\, \Delta r\, L_z}, \qquad (20)$$

Formulas (20) are visualized in Figure 7 as histograms.

IV. DISCUSSION

Using MD simulations at constant pressure and temperature, we can solvate the DNA with water and ions, fixing the concentration of ions in the bulk. In Section 3.2, we have presented illustrative results of such all-atom MD investigations with APBC. Such simulations can also be used to estimate other solvent properties, for example, the moments of force distributions on ions, which can be used for parametrizing coarse-grained stochastic models of ions used in

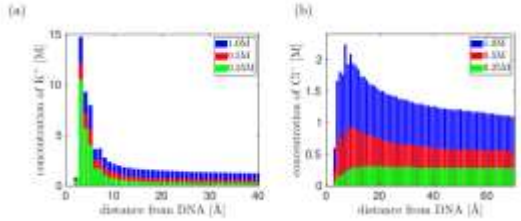

Figure 7: The results of all-atom MD simulations with APBC using the poly(A) DNA chain with N = 10 bp and three different concentrations of KCl in the bulk. (a) The concentration of K+ ions given by (20) as a function of the distance from DNA. (b) The concentration of Cl− ions given by (20) as a function of the distance from DNA.

multiscale and multi-resolution simulations.16,19 The APBC simulations can also be coupled with coarse-grained models of water to design adaptive resolution simulation techniques.5,24,25 In Section 3.2, we have presented the results calculated with APBC using n = 1 helical pitch. In particular, the simulated domain length is around 3.4 nm long and considerably smaller than the DNA's persistence length, which is about 50 nm. To study mechanical properties of DNA, we need to increase the number of helical pitches as we have shown in Section 3.1 with our MD simulation results considering up to n = 10 helical pitches along the z-direction of APBC simulation domain (1).

To get further insight into the correct use of the APBC, we have started our investigation using a discrete worm-like chain (WLC) model in Section 2, where we have observed in Figure 3 that the APBC affect less some local properties of the polymer chains than some global averages. In particular, the persistence length of the polymer chain can be estimated from local properties of relatively short polymer chains, simulated with the help of APBC. The APBC are also applicable to simulations of biopolymers with larger persistence length (for example, actin filaments26,27), when a modeller is interested to understand the properties of the surrounding solvent.

In Appendix B, we provide the technical details of all-atom MD simulations, including the treatment of constant pressure simulations. The barostat used is again asymmetric with no fluctuations of Lz. In the APBC simulations, we have different treatment of the z-direction and all perpendicular directions in the x − y plane. Simulations with 2D periodicity have also been used to study behaviour of a slab of water between two metallic walls,28 which can be treated using three-dimensional Ewald techniques by including the image charges. One advantage of the APBC simulations is that they can be implemented with relatively minor modifications of standard all-atom MD tools29–32 as detailed in Appendix B. Note that, the number of helical turns in the DNA model with APBC is fixed, and thus the model does not allow for over-winding or under-winding of DNA.

V. CONCLUSION

This study demonstrates that asymmetric periodic boundary conditions (APBCs) provide a robust and efficient alternative to conventional symmetric boundary treatments in molecular dynamics and coarse-grained simulations of nucleic acids. By allowing directional periodicity tailored to the intrinsic anisotropic geometry of DNA and RNA systems, APBCs significantly reduce finite-size effects and artificial interactions between periodic images. This results in improved structural stability and more realistic dynamical behavior across different simulation scales.

Furthermore, the successful implementation of APBCs in both all-atom and coarse-grained models highlights their versatility and suitability for multiscale nucleic acid simulations. The approach enhances computational efficiency without compromising physical accuracy, making it particularly valuable for long-time and large-system simulations. Overall, asymmetric periodic boundary conditions offer a promising framework for advancing high-fidelity computational studies of nucleic acids and related biomolecular systems.

REFERENCES

(1) Vologodskii, A. Biophysics of DNA; 2015.
(2) Jacobson, D.; Saleh, O. Counting the ions surrounding nucleic acids. Nucleic Acids Research 2016, 45, 1596–1605.
(3) Mocci, F.; Laaksonen, A. Insight into nucleic acid counterion interactions from inside molecular dynamics simulations is "worth its salt". Soft Matter 2012, 8, 9268–9284.
(4) Várnai, P.; Zakrzewska, K. DNA and its counterions: a molecular dynamics study. Nucleic Acids Research 2004, 32, 4269–4280. (5) Zavadlav, J.; Podgornik, R.; Praprotnik, M. Adaptive resolution simulation of a DNA molecule in salt solution. Journal of Chemical Theory and Computation 2015, 11, 5035–5044. (6) Zavadlav, J.; Sablic, J.; Podgornik, R.; Praprotnik, M. Open-Boundary Molecular Dynamics of a DNA Molecule in a Hybrid Explicit/Implicit Salt Solution. Biophysical Journal 2018, 114, 2352–2362.
(7) Dans, P.; Walther, J.; Gómez, H.; Orozco, M. Multiscale simulation of DNA. Current Opinion in Structural Biology 2016, 37, 29–45.
(8) Korolev, N.; Nordenskiöld, L.; Lyubartsev, A. Multiscale coarse-grained modelling of chromatin components: DNA and the nucleosome. Advances in Colloid and Interface Science 2016, 232, 36–48.
(9) Poppleton, R.; Matthies, M.; Mandal, D.; Romano, F.; Šulc, P.; Rovigatti, L. oxDNA: coarse-grained simulations of nucleic acids made simple. The Journal of Open Source Software 2023, 8, 4693.
(10) Sengar, A.; Ouldridge, T.; Henrich, O.; Rovigatti, L.; Šulc, P. A Primer on the oxDNA Model of DNA: When to Use it, How to Simulate it and How to Interpret the Results. Frontiers in Molecular Biosciences 2021, 8 . (11) Kovaleva, N.; Koroleva, I.; Mazo, M.; Zubova, E. The "sugar" coarse-grained DNA model.Journal of Molecular Modelling 2017, 23, 66. 19
(12) Rolls, E.; Togashi, Y.; Erban, R. Varying the resolution of the Rouse model on temporal and spatial scales: application to multiscale modelling of DNA dynamics. MultiscaleModeling and Simulation 2017, 15, 1672–1693.
(13) Maffeo, C.; Aksimentiev, A. MrDNA: a multi-resolution model for predicting the structure and dynamics of DNA systems. Nucleic Acids Research 2020, 48, 5135–5146. (14) Minhas, V.; Sun, T.; Mirzoev, A.; Korolev, N.; Lyubartsev, A.; Nordenskiöld, L. Modeling DNA Flexibility: Comparison of Force Fields from Atomistic to Multiscale Levels. Journal of Physical Chemistry 2020, 124, 38–49.
(15) Lee, S.; Rasaiah, J. Molecular dynamics simulation of ion mobility. 2. alkali metal and halide ions using the SPC/E model for water at 25◦ C. Journal of Physical Chemistry 1996, 100, 1420–1425.