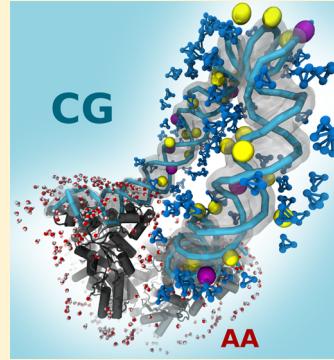


Exploring LacI–DNA Dynamics by Multiscale Simulations Using the SIRAH Force Field

Matias R. Machado* and Sergio Pantano*

Biomolecular Simulations Group, Institut Pasteur de Montevideo, Montevideo, Uruguay, 11400

ABSTRACT: The lac repressor protein (LacI) together with its target regulatory sequence are a common model for studying DNA looping and its implications on transcriptional control in bacteria. Owing to the molecular size of this system, standard all-atom (AA) simulations are prohibitive for achieving relevant biological time scales. As an alternative, multiscale models, which combine AA descriptions at particular regions with coarse-grained (CG) representations of the remaining components, were used to address this computational challenge while preserving the relevant details of the system. In this work, we implement a new multiscale approach based on the SIRAH force field to gain deeper insights into the dynamics of the LacI–DNA system. Our methodology allows for a dual resolution treatment of the solute and solvent, explicitly representing the protein, DNA, and solvent environment without compromising the AA region. Starting from the P1 loop configuration in an undertwisted conformation, we were able to observe the transition to the more stable overtwisted state. Additionally, a detailed characterization of the conformational space sampled by the DNA loop was done. In agreement with experimental and theoretical evidence, we observed the transient formation of kinks at the loop, which were stabilized by the presence of counterions at the minor groove. We also show that the loop's intrinsic flexibility can account for reported FRET measurements and bent conformations required to bind the CAP transcription factor.



INTRODUCTION

The lactose operon (*lac*) of *Escherichia coli* is a paradigmatic protein–DNA system used by Jacob and Monod to understand gene regulation in bacteria.¹ Within this system, the *lac* repressor protein (LacI) is one of the master regulators that controls gene expression. To accomplish this, LacI binds to two out of three specific DNA sequences along the *lac* operon, named operators, and brings them together in space, forcing the formation of a DNA loop that prevents transcription from proceeding because of steric hindrance.

LacI is a homotetramer composed of a dimer of homodimers associated by a four-helix bundle at the C-terminal region.² This complex adopts a V-shaped configuration that can open like a hinge. The N-terminal region of each homodimer, named the headpiece, is responsible for binding DNA (Figure 1A). The palindromic nature of the operators' sequence allows for the formation of different loop topologies, referred to as parallel (P1 or P2) when the 3' end of the upstream operator points toward the 5' end of the other and *vice versa* or antiparallel (A1 or A2) when the 3' and 5' ends point toward each other.³ The intrinsic flexibility of DNA may also impact the loop's conformation, so different theoretical models were proposed in which undertwisted (U), overtwisted (O), or wrapped configurations were considered.^{4–6} In addition, if the V-shaped LacI opens to a more linear form, then an extended (E) loop can be achieved.^{3,7,8} All of these aspects impart a tremendous amount of structural diversity for the LacI–DNA interaction and represent a challenge to any structural study of this complex. Moreover, the distribution of observed states depends

on the length and composition of the sequence between the looped operators as well as the experimental conditions.^{9,10}

The main structural information on the LacI–DNA complex comes from experimental determinations by X-ray diffraction of headpieces, homodimers, and homotetramer.^{2,11–13} That structural information, obtained at different resolutions, was integrated into an atomistic model of the complete LacI–DNA complex lacking the looping DNA segment (PDB ID: 1Z04).¹⁴ However, because the flexibility of the LacI–DNA complex is crucial for its function, a dynamical description of its mechanistic details may reveal new insights into the structural features that affect gene repression. The gold standard experimental technique used to explore this interaction under noncrystalline conditions is Förster resonance energy transfer (FRET), which can provide geometrical information with a resolution on the nanometer scale. Integration of this data with modeling and molecular dynamics (MD) simulations may offer atomic-level information on the structure, flexibility, and interactions in this complex molecular system. However, simulating the LacI–DNA complex and its solvent environment using an all-atom (AA) description may be too computationally expensive for achieving experimentally or biologically relevant time scales. As an alternative, a reasonable trade off between computational cost and system complexity can be acquired by using coarse-grained (CG) models of biomolecules.¹⁵ Several examples of CG descriptions have been applied to the study of LacI–DNA.^{3,6,14,16–21} In general, CG

Received: June 18, 2015

Published: August 26, 2015

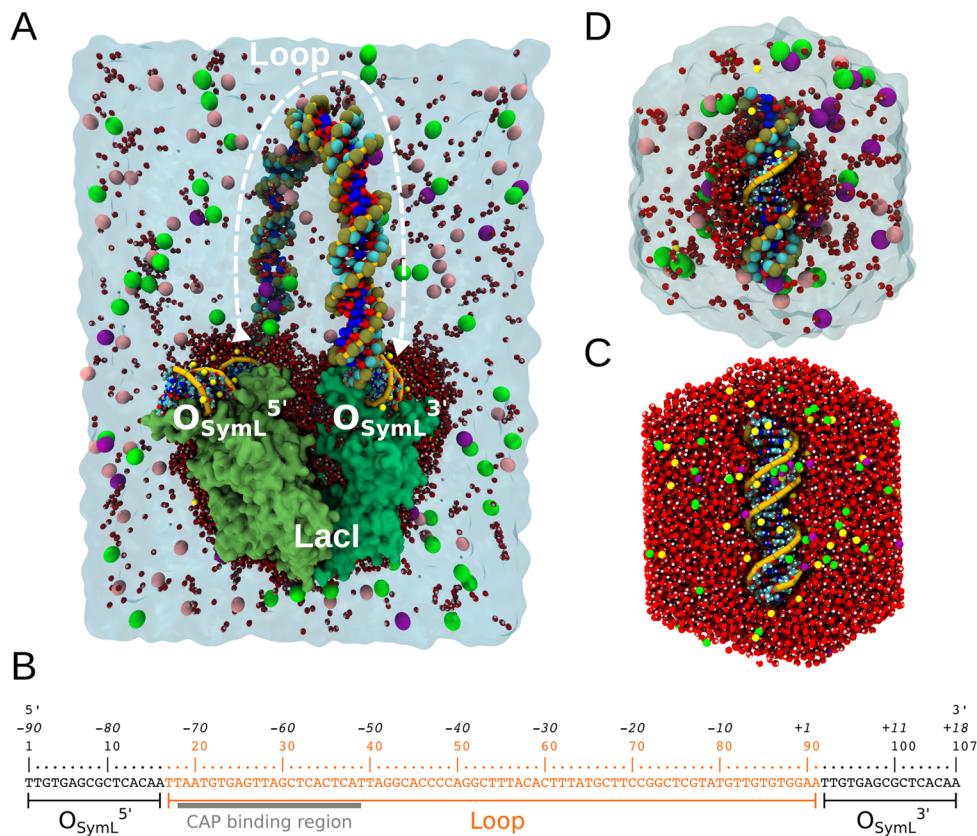


Figure 1. Molecular systems studied. (A) Representation of the multiscale LacI–DNA system. The subunits of the LacI protein are shown as dark and light green surfaces, and operator sequences ($O_{\text{SymL}}^{5'}$ and $O_{\text{SymL}}^{3'}$) and the DNA loop are indicated as van der Waals spheres. Atomistic water and Na^+ are in red and yellow, respectively, and CG ions Na^+ , K^+ , and Cl^- are in pink, purple, and green, respectively. The CG water (WT4) embeds the system. (B) Sequence and base pair numbers of *lac* corresponding to the DNA in panel A. Positions relative to the transcription start site are indicated in italic as a reference. (C, D) Molecular representation of systems 20-mer[AA] and 20-mer[AA/CG], respectively, following the same style as that in panel A.

models for this system use an elastic rod model to represent the DNA loop and rigid conformations of the LacI protein to mimic spatial topological restraints. However, depending on the particular CG simplifications, some relevant details, like electrostatics, DNA sequence identity, protein flexibility, and so on, may be lost. Nevertheless, when some fine details are mandatory for explaining a given phenomena, AA and CG approaches may be combined into multiscale (AA/CG) representations.^{15,22,23} This kind of approach was used by Villa et al.²⁴ to recover relevant information on the dynamics of the headpieces in the presence of DNA. Although the DNA loop was described as an elastic rod in their model, the protein was treated fully atomistically within a solvation box containing water and ions. That dual resolution description allowed for the protein's flexibility to be considered, without which it would not be possible to unveil its relevant role in the interaction with the DNA fiber.

Major drawbacks in many CG or AA/CG approaches are their transferability and intricate implementation, which may hinder their extensive usage by the scientific community. In this work, we developed a new multiscale approach to explore the flexibility of the DNA loop created by LacI when it binds to two distant symmetrical operator sequences, $O_{\text{SymL}}^{5'}$ and $O_{\text{SymL}}^{3'}$, in the *lac* promoter of *E. coli* (Figure 1B). The simulations were performed in the context of a fully explicit water and electrolyte environment, which allowed the protein's flexibility and the loop's dynamics to be recovered. The LacI–DNA complex was

described using a dual resolution approach in which the protein, the two operators, and a 1 nm thick solvation shell of water and ions were considered at the AA level and the loop region of DNA and its surrounding aqueous environment were treated at the CG level (Figure 1A). It is worth noting that the DNA filament is represented as a double-stranded helix in which both resolution levels coexist and interactions are treated within the same standard classical Hamiltonian implemented in widely used simulation packages such as AMBER and GROMACS. This feature avoids the use of external forces in different parts of the system. Furthermore, this new approach does not modify the existent parameters in the SIRAH force field,^{25–28} which are available at <http://www.sirahff.com>.

The methodology was validated by comparative simulations in a 20-mer double-stranded DNA (Figure 1C,D) that showed good agreement between AA and AA/CG descriptions on several structural and dynamical features of both solute and solvent. By applying this approach to the LacI–DNA system, we were able to reproduce experimental and theoretical observations previously reported in the literature. Furthermore, our results gave a new perspective on the LacI–DNA system by exposing the relevant role of loop flexibility and solvent environment in the binding to catabolite activator protein (CAP).

METHODOLOGY

Molecular Systems. Three systems were simulated:

- (i) Atomistic 20-mer double-stranded DNA (20-mer[AA]): 5'-d(C₁A₂T₃G₄C₅A₆T₇G₈C₉A₁₀T₁₁G₁₂C₁₃A₁₄T₁₅G₁₆-C₁₇A₁₈T₁₉G₂₀)-3' (Figure 1C).
- (ii) Multiscale representation of system 20-mer[AA] (20-mer[AA/CG]), in which the central region (base pairs G₈ to C₁₃) was described atomistically and its flanking sequences were depicted at the CG level (Figure 1D).
- (iii) Multiscale representation of the LacI–DNA complex starting from the P1 loop configuration in the under-twisted (U) state. The protein and the operators' sequences (base pairs 1–16 and 92–107) were modeled at atomistic resolution, and the DNA loop was represented at the CG level (Figures 1A and B).

In all cases, the model building procedure started from the Cartesian coordinates of structures containing all of the atoms. Systems 20-mer[AA] and 20-mer[AA/CG] were built in the canonical B-form of DNA²⁹ using the NAB utility of AMBER.³⁰ Residues belonging to the CG region were mapped according to the SIRAH scheme for nucleotides²⁶ using the tools provided in SIRAH's package distribution. Residues at the AA region remain unchanged. The solute was placed in an octahedral simulation box of 1.2 nm. The system 20-mer[AA] was solvated by atomistic water, K⁺ counterions, and 0.15 M NaCl. In the case of system 20-mer[AA/CG], a 0.5 nm shell of atomistic water molecules with 9 Na⁺ ions was added around the AA base pairs. The ion's concentration within the AA region was set to neutralize 70% of the polyelectrolyte charge according to the counterion condensation theory.³¹ The bulk solvent was represented by SIRAH CG water (named WT4); CG K⁺ ions were used to neutralize the remaining charge of the DNA. Additionally, 0.15 M of CG NaCl was added to mimic physiological conditions.

The structure of the LacI–DNA complex in the P1 configuration and U conformation was taken from ref 24, which is available at http://www.ks.uiuc.edu/Research/vmd/script_library/scripts/dnademos. The building protocol was essentially the same as that described before for system 20-mer[AA/CG]. After mapping the CG residues, the solute was centered in a square simulation box of 16 × 20 × 26 nm³. A solvent shell of 1 nm around the AA region was modeled at atomistic resolution including 48 Na⁺ ions. The bulk solvent was represented at the CG level by WT4, the system's charge was neutralized by CG K⁺ ions, and 0.15 M of CG NaCl was added (Figure 1A).

Interaction Parameters. Amino acids, nucleotides, and ions at the AA level were described by the ff99SB force field of AMBER,³² as implemented in GROMACS through ffamber ports.³³ The atomistic water molecules were described by the TIP3P model.³⁴ The SIRAH force field was used to represent the CG DNA fragments and the bulk solvent.^{25,26} The mechanical coupling at the AA/CG DNA interface was taken from our previous publication.²⁷ On the other hand, the dual resolution solvent approach originally presented by Darré et al.³⁵ and Gonzalez et al.²⁸ was used to treat the solvent AA/CG interface. Specific Lennard-Jones (LJ) parameters out of the Lorentz–Berthelot combination rule were set among atomistic water, ions, and the CG DNA to avoid spurious interactions in the proximities of the AA/CG interface. Such behavior is a consequence of the different van der Waals resolutions, which place charged groups of the CG solute and the AA solvent too close.

The LJ potential is defined as $V_{\text{LJ}} = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, where ϵ is the depth of the potential well, σ is the finite distance at which the interparticle potential is zero, and r is the distance between the particles. As in the implementation of our AA/CG model for solvation,³⁵ we assigned the LJ interaction parameters of the water oxygen to any bead of the CG DNA equal to the interaction of a WT4 bead with the centroid representing the C5' atom in our model (i.e., $\sigma = 0.4245$ nm, $\epsilon = 0.50$ kJ mol⁻¹). In a similar fashion, each AA ion was set to interact the same with phosphate beads of the CG DNA as its respective CG ion (i.e., Na⁺: $\sigma = 0.5216$ nm, $\epsilon = 0.68$ kJ mol⁻¹; K⁺: $\sigma = 0.5541$ nm, $\epsilon = 0.68$ kJ mol⁻¹; Cl⁻: $\sigma = 0.5716$ nm, $\epsilon = 0.68$ kJ mol⁻¹). These corrections prevented the system from establishing unbalanced interactions between the AA solvent and CG solute while maintaining the integrity of the AA/CG interface.

Simulation Protocol. All calculations were performed with GROMACS 4.5.5 (<http://www.gromacs.org>). Initially, the whole system was relaxed by energy minimization, and then MD simulations were performed in the NPT ensemble. A reference temperature of 300 K was set by coupling the groups solute and AA and CG solvents separately to the Nosé–Hoover thermostat^{36,37} with a coupling time of 1.2 ps. The pressure was kept at 1 bar by means of the Parrinello–Rahman barostat^{38,39} with a coupling time of 6 ps. A cutoff for nonbonded interactions of 1.2 nm was used, and long-range electrostatics were evaluated using particle mesh Ewald (PME).^{40,41} A time step of 2 fs was used, and all bonds involving hydrogen atoms within the AA region were restrained using the LINCS algorithm.^{42,43} To avoid possible fraying of the ends of the helix frequently observed in long MD simulations,⁴⁴ conservative distance restraints of 1255 kJ mol⁻¹ nm⁻² were added to preserve the Watson–Crick hydrogen bonds at the capping base pairs. Equilibration was carried out for 100 ps, imposing harmonic positional restraints of 1000 kJ mol⁻¹ nm⁻² on the solute. Production runs of 500 ns were performed, and snapshots were recorded every 5 ps for analysis.

Calculated Properties. Comparison of Sampling Spaces. Trajectories of systems 20-mer[AA] and 20-mer[AA/CG] were fitted to their corresponding canonical B form DNA in order to eliminate rotational and translational movements. Covariance matrix calculation and principal component analysis were performed using the GROMACS utilities G_COVAR and G_ANAEIG. The phosphate positions were used in the analysis. The shape likeness in sampling space between both simulation schemes was estimated by computing the following similarity index (SI)

$$\text{SI} = 1 - \sqrt{\left[\text{tr} \left(\sqrt{\frac{M1}{\text{tr}(M1)}} - \sqrt{\frac{M2}{\text{tr}(M2)}} \right) \right]^2} \quad (1)$$

where $M1$ and $M2$ are the two covariance matrices and $\text{tr}()$ is the trace of the matrix. The SI ranges from 1 for identical matrices to 0 when the sampled subspaces are orthogonal.

Solvation Structure around the DNA. The distribution of the solvent molecules around the DNA was characterized by computing the average number of water molecules within 0.5 nm from phosphates in each base pair. In systems 20-mer[AA] and 20-mer[AA/CG], radial distribution functions (RDF) were calculated for water oxygens, WT4 beads, or atomistic and CG Na⁺ ions around phosphates at the central region of the DNA (base pairs 8–13).

Conformational Description of LacI. The relative orientation of the homodimers in the homotetramer of LacI was characterized by measuring the cleft angle α , which was defined as the angle between lines connecting the geometric centers of each operator to the geometric center of the four-helix bundle of LacI.

Sampling Surface of the DNA Loop. The conformational space explored by the DNA loop in the LacI–DNA complex was described by tilt (β) and screw (ϕ) angles. The angle β was defined by the geometric centers of the four-helix bundle in LacI protein, both operators, and the base pair at position 54 of the DNA loop, whereas ϕ corresponded to a torsional angle defined by the geometric centers of base pairs at positions 41, 25, 83, and 67 of the DNA loop (Figure 1B).

Local Bending at DNA Grooves. The local bending was measured from the helical axis of the DNA, which was defined as the line connecting the geometric centers of each couple of phosphate centroids in a base pair (Figure 2A). This axis definition is a reasonable approximation in the case of B form DNA (Figure 2B). Local bending was evaluated at the central base pair (i) of overlapping 5 bp long groove windows by means of the θ angle, which represents the deviation between the helical axis of the 2 bp before and after the reference i bp (vectors $i - 1 \rightarrow i - 2$ and $i + 1 \rightarrow i + 2$ in Figure 2C). Notice that a DNA kink dramatically changes the value of θ .

Ion Occupancy at DNA Grooves. Grooves were defined as overlapping 5 bp long sequence windows. The number density of cations (Na^+ and K^+) within the major and minor grooves of the DNA was calculated within a sphere placed at the geometric center of phosphate atoms defining each groove (Figure 2D). The radius of the sphere was set to 0.8 or 0.5 nm for the major and minor grooves, respectively. The number density at groove i (ρ_i) was estimated as

$$\rho_i = \frac{3}{4N\pi r^3} \sum_{j=1}^N C_{ij} \quad (2)$$

where C_{ij} is the count of cations at groove i and frame j . The sum runs over all N frames in the trajectory.

Comparison Against FRET Experiments. To compare the simulation of the LacI–DNA system against FRET experiments of Edelman et al.⁴⁵ and Morgan et al.,⁴⁶ we measured the distance R between phosphate atoms of residue 20 and the complementary base of residue 90 (Figure 1B), at which the fluorescent dyes Cy3 and Cy5 were attached. Taking the Förster radius (R_0) for the fluorescent pair as 5.6 nm, the FRET efficiency (E) can be estimated as $E = 1/[1 + (R/R_0)^6]$.

RESULTS AND DISCUSSION

Testing the Multiscale Simulation Scheme. The dual resolution scheme presented here implies two particular AA/CG frontiers: (i) along the DNA filament, AA and CG residues are connected with specific bonded interactions defined between atoms and beads at the AA/CG border²⁷ (none of the bonded or nonbonded parameters of either AMBER or SIRAH force fields were modified); (ii) the solvent contains atomistic and CG water with their respective ions. These species interact among themselves and with the solute only via nonbonded interactions tuned to experience only limited mixing.^{28,35}

To test the robustness of this multiscale approach, we evaluated the structure and dynamics of a dual resolution DNA filament and the correctness in the solvation environment.

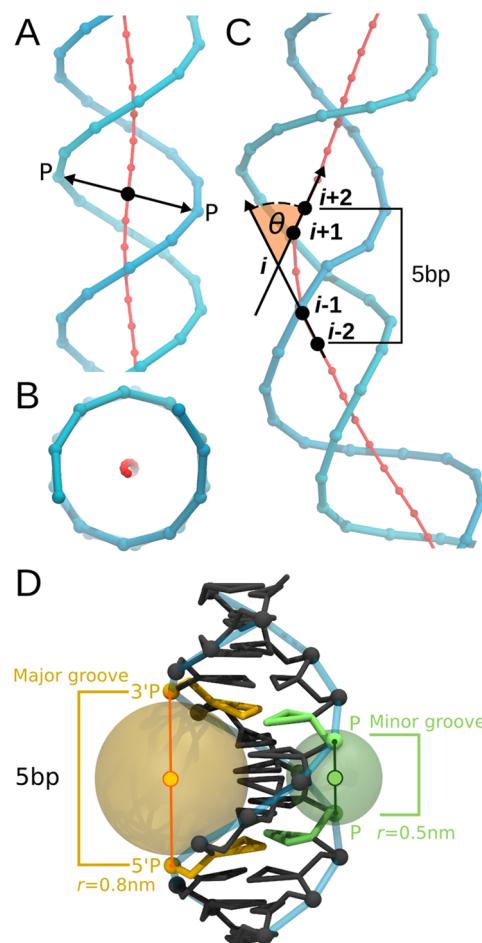


Figure 2. Definition of local bending and ion occupancy at DNA grooves. (A) The backbone is represented with blue tubes, and phosphate positions are shown as blue spheres. The geometric center of phosphate atoms sharing the same base pair is represented by red spheres (black arrows and dots emphasize an example), and the helical axis is shown as a red line that goes through them. (B) Axial view of panel A. (C) Example of a kinked DNA segment taken from simulations of ref 47 that shows the measurement of the bending angle θ at the central base pair i . (D) The relative location of bases defining the mayor (yellow) and minor (green) grooves is shown. Semi-transparent spheres indicate the volume considered to calculate the number density.

Systems 20-mer[AA] and 20-mer[AA/CG] were chosen as reference test cases to validate the presented methodology because we had exhaustively analyzed and compared them previously in an implicit solvent context.²⁷

During the simulations, the DNA's backbone showed average root-mean-square deviation (RMSD) values of 0.6 (SD 0.1) nm and 0.6 (SD 0.2) nm for the 20-mer[AA] and 20-mer[AA/CG] systems, respectively. The structural dynamics of the solute was compared in terms of a similarity index (SI) between the principal components of the covariance matrix that describe the sampling space (see *Methodology*). A comparison of the SI calculated between both halves of the same trajectory resulted in similar values for the AA and AA/CG systems (Table 1), suggesting that both simulation schemes provide a similarly high degree of sampling. This holds true for the whole system or when considering only the central region. Furthermore, calculation of the SI between the 20-mer[AA] and 20-mer[AA/CG] trajectories resulted in sizable overlaps between both

Table 1. Similarity Indices (SI) Calculated on Systems 20-mer[AA] and 20-mer[AA/CG]^a

20-mer	[AA] 250–500 ns	[AA/CG] 250–500 ns	[AA] 0–500 ns
[AA] 0–250 ns	0.708 (100%)		
	0.830 (100%)		
[AA/CG] 0–250 ns		0.743	
		0.719	
[AA/CG] 0–500 ns			0.532 (75%)
			0.683 (82%)

^aValues on top and bottom correspond to SI calculated on all phosphates or only the central ones (base pairs 8–13), respectively. Values in brackets indicate the percentage relative to the maximum expected SI (i.e., 20-mer[AA]).

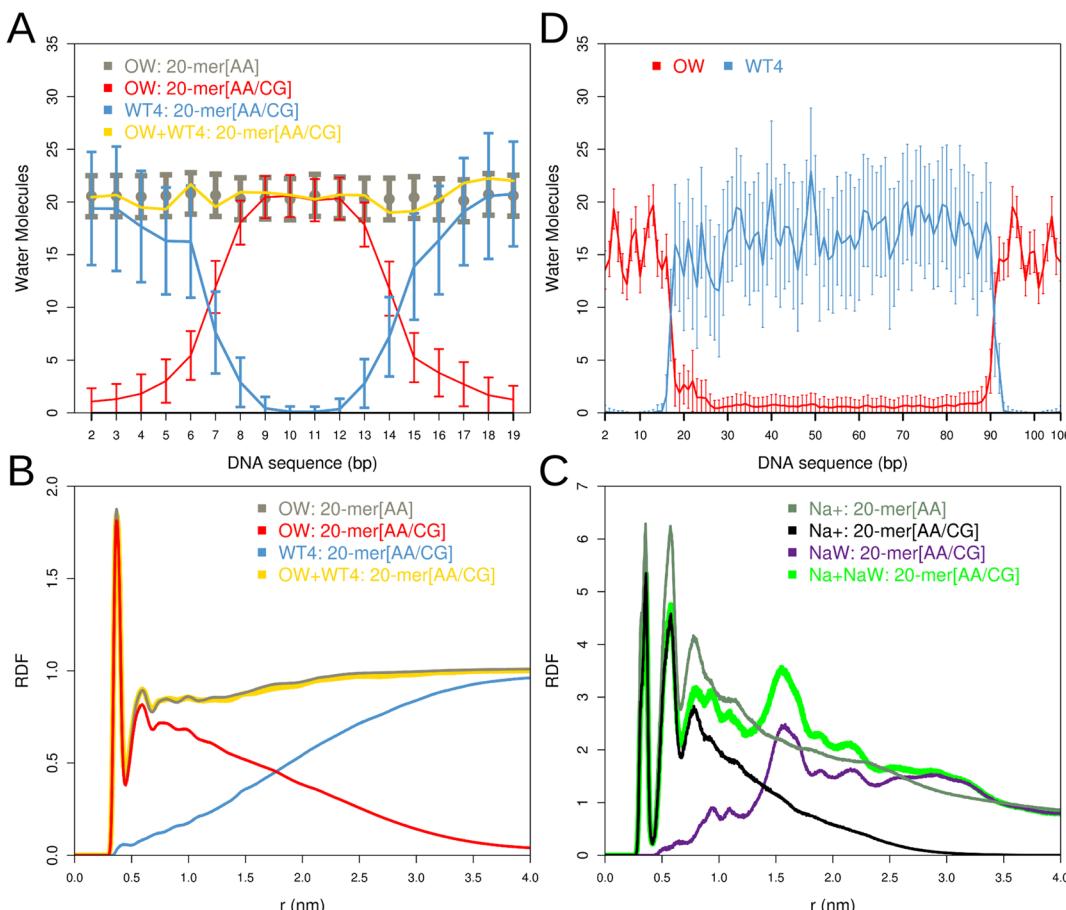


Figure 3. Solvent distribution around DNA in systems 20-mer[AA] and 20-mer[AA/CG]. (A) Average number of water oxygens (OW) or WT4 beads closer than 0.5 nm to either phosphate atom of each base pair. The number of CG beads was normalized by a factor of 2.8 to account for the number of water molecules effectively represented by each WT4 bead. In the case of the AA/CG system, the sum of the AA and CG distributions is also indicated. (B) Radial distribution function (RDF) of OW atoms or WT4 beads around phosphate atoms at the central base pairs of the DNA. (C) Same as panel B for atomistic Na⁺ and its counterpart CG ion (NaW). (D) Same as panel A for the LacI–DNA system.

sampled subspaces (Table 1), especially if the SI calculated on the AA trajectory is considered to be the maximum attainable within this time scale. These results, which are in line with those in our previous reports,^{27,48} suggest a good correspondence between the conformational spaces sampled by the AA and AA/CG representations of the DNA in explicit solvent. Moreover, they indicate that, within our dual resolution approach, the CG region does not have an influence on the AA region.

Besides the structural dynamics, it is also worth ensuring that the AA/CG environment reproduces good solvation with homogeneous density around the solute. The correct solvation of the systems implies a good reproduction of the water structure and ions distribution around the solute. Due to the

dual resolution nature of the solute in system 20-mer[AA/CG], it is desirable for the AA region to remain solvated by atomistic molecules. The same is valid for CG regions. In previous publications, we have shown that a 1 nm shell of atomistic water is enough to avoid spurious effects on the structure of the atomistic solute. Moreover, this method also has been shown to be transferable to different water models, in particular, TIP3P.²⁸ A similar criterion was used here. It is worth noting that, owing to differences in the free energy of solvation of AA water in CG solvent and *vice versa*, both molecular species experience only limited mixing.³⁵ This process happens in hundreds of picoseconds, resulting in an AA solute surrounded by AA water embedded in a CG environment (Figure 1A,D). For this goal, we calculated the number of water molecules of both

species around each base pair. It is important to recall that each WT4 bead accounts for 2.8 water molecules.²⁵ By computing the number of WT4 beads along the DNA fiber, we found high solvation around CG nucleotides, which decreased to negligible values in the neighborhood of the atomistic region, whereas the opposite behavior was found for atomistic solvent. It is noteworthy that adding both AA and CG contributions resulted in good agreement with the fully atomistic situation (Figure 3A). Similarly, a good match was found in the radial distribution function (RDF) of solvent around phosphates. In this case, the sum of the AA and CG solvent molecules resulted in remarkably similar distributions with a good reproduction of the solvation peaks (Figure 3B).

Another very important aspect for polyelectrolytes is to achieve a correct description of the distribution of ions, which help to stabilize the DNA structure. To test this aspect, we calculated and compared the RDF of Na^+ ions at the central region of the DNA for both systems (Figure 3C). In close analogy with the water distribution, the reproduction of the ion structure in the neighborhood close to DNA was remarkable. Noticeable differences appear only beyond 0.7 nm and can be ascribed to interactions of CG ions with CG nucleotides. However, the number of Na^+ ions added to the AA region of system 20-mer[AA/CG] matched the values predicted by the counterion condensation theory,³¹ and the differences in ionic distribution start after or coincide with the Debye length for physiological salt concentration.⁴⁹

Taken together, we conclude that the AA/CG approach resulted in a good reproduction of the global dynamics and solvation of the system.

In the next sections, we applied the presented multiscale representation to study the dynamics of the LacI–DNA system.

Structural Dynamics of LacI Protein. The LacI–DNA system, as shown in Figure 1A, was simulated for 500 ns using the multiscale approach introduced in the previous paragraph. The dual resolution water environment generated a homogeneous solvation around the AA/CG DNA filament. A comparison between Figures 3A and D indicates no dependence of the solvation structure with the geometry of the system.

We first studied and compared the dynamics of the repressor protein, which was represented at the atomistic level. The conformational dynamics of LacI was characterized by the RMSD of $\text{C}\alpha$ atoms from the initial conformation. During the simulation, each protein homodimer preserved the structure with an average RMSD value of 0.3 nm (Figure 4A). This value is comparable with results obtained from simulations of the same system by Villa et al.²⁴ At 25 ns of simulation, the homotetramer suffered a conformational change, which shifted its RMSD to 0.75 nm, and it remained in this state until the end of the simulation. This conformational change was mainly related to a movement in the relative orientation of the two homodimers, which implied the approaching of the headpiece and core domains. As a consequence, the cleft angle α diminishes from $\sim 55^\circ$ to $\sim 40^\circ$ (Figure 4B). Close inspection of the trajectory suggested that residues at either side of the complex attracted each other to form salt bridges, polar contacts, or hydrophobic patches, favoring the interaction of the homodimers (Figure 4C). Such conformational change was not reported in previous simulations shown by Villa et al.,²⁴ most likely because of the use of large external fluctuating forces of 80 pN assumed to mimic DNA tensions, which may have prevented the described conformational change. These forces also increased the flexibility of the headpiece domains.

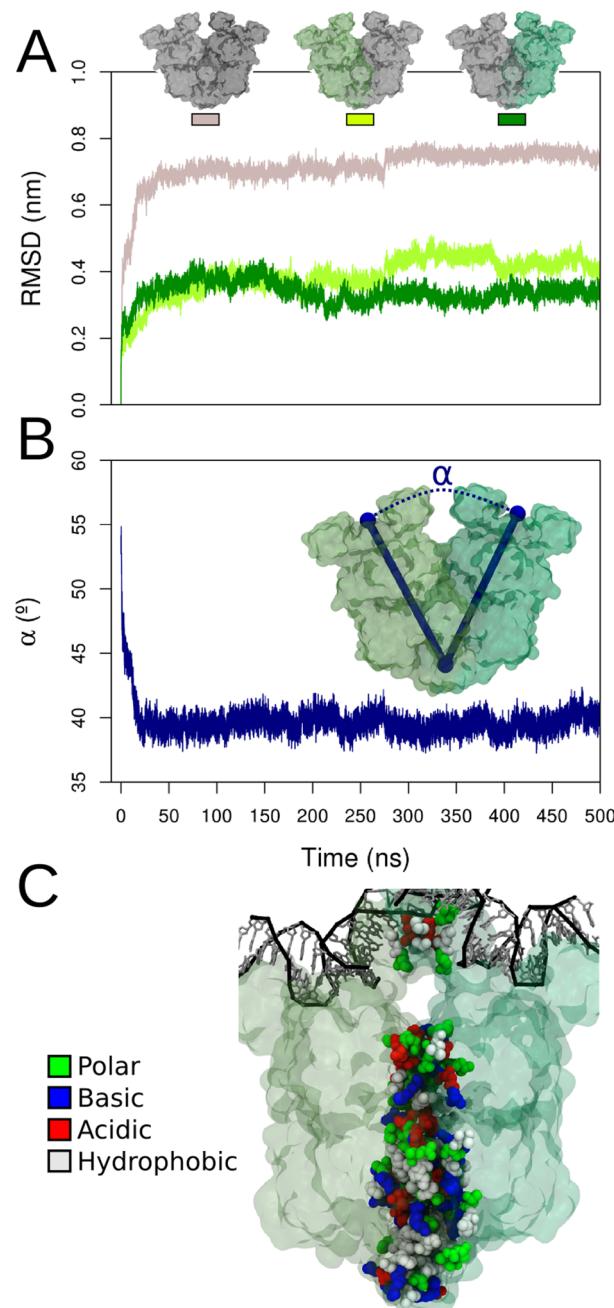


Figure 4. Structural dynamics of LacI. (A) RMSD values for the homotetramer (gray) and each homodimer (dark and light green) along the trajectory. (B) Temporal evolution of the cleft angle α . The inset graphically shows the measured α angle. (C) Representative snapshot at 500 ns of simulation. Homodimers are represented as transparent surfaces. Residues at the interface are represented by the van der Waals radii and colored by residue type.

The absence of external forces in our simulation resulted in modest fluctuations at the headpiece domains (average root-mean-square fluctuations of 0.22 nm in $\text{C}\alpha$ atoms).

Loop Dynamics: Global Conformation. The accuracy of our CG model of DNA⁴⁸ grants the opportunity to explore the loop's dynamics in the context of the whole system and to evaluate its impact on the loop's conformation.

The DNA loop has been described to adopt different topologies in the LacI–DNA complex. All of them depend on the orientation of the operators.³ In this work, we focused on

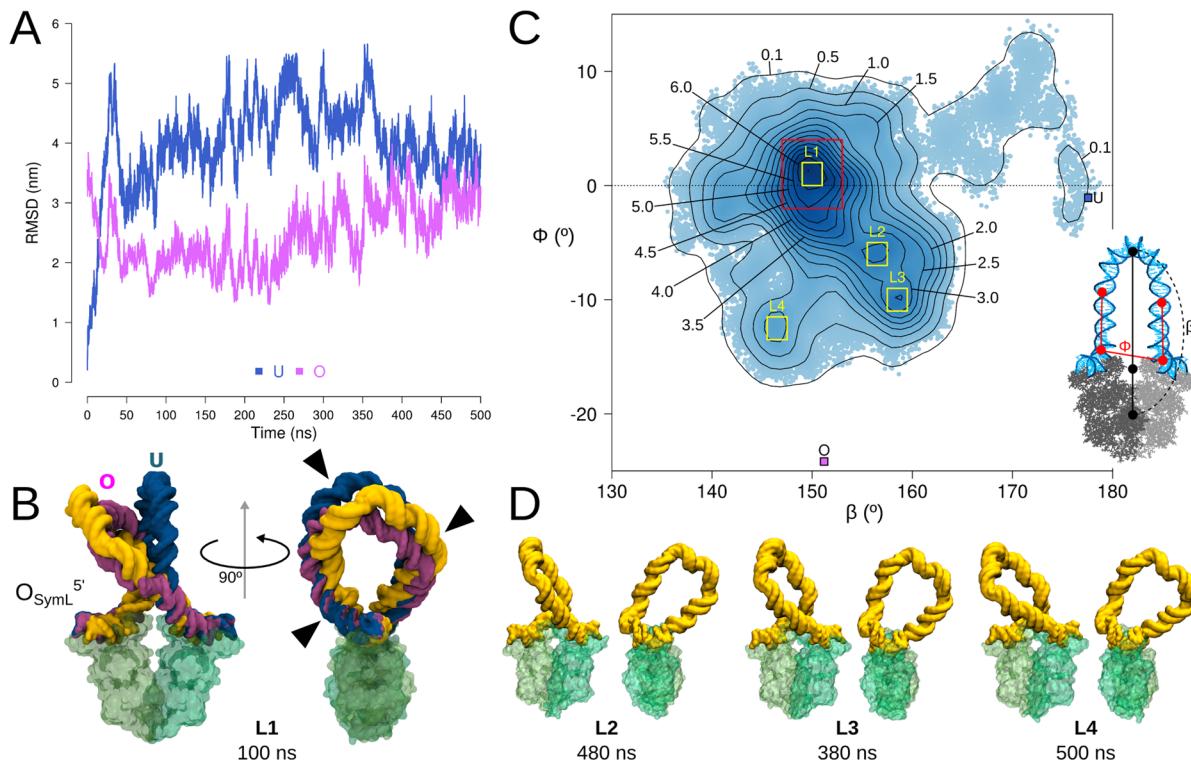


Figure 5. DNA loop conformations. (A) RMSD values of the DNA loop compared against undertwisted (U, blue) or overtwisted (O, purple) conformations of Villa et al.²⁴ (B) Structural comparison between U (blue) and O (purple) solutions to the elastic rod model with a snapshot at 100 ns from our trajectory (yellow). Systems were aligned to the protein coordinates. Black arrows indicate the presence of kinks in the DNA loop of our model. The location of operator $O_{\text{SymL}}^{5'}$ in our model is indicated. (C) Conformational surface of the loop described by β and ϕ angles. The surface is colored according to the density of points, from low (light) to high (obscure) values. Isocontour densities are shown in units of $10^{-3}(\text{deg})^{-2}$. Areas covering conformers L1–L4 are indicated by yellow squares. The red square covers 20% of the total sampled conformations. The positions of the U and O conformations are indicated. (D) Representative structures of loop conformers L2–L4.

the topology of the parallel P1 loop. The conformational space sampled by the DNA loop in our simulation was compared against results from a previous multiscale approach developed by Schulten and co-workers, in which this region was modeled as an elastic rod.^{4,50} The elastic rod model assumes that the DNA is B form with anisotropic sequence bendability and an electrostatic phosphate contribution but without the capacity to form kinks. Such a model described two energy minima corresponding to undertwisted (U) and overtwisted (O) conformations, the latter being the most stable. We started our multiscale simulation from the U conformation of the loop. RMSD values of phosphate atoms with respect to the U and O conformations were calculated along our trajectory (Figure 5A). The initial structure of the loop rapidly shifted from the U conformation to sample shapes more similar to O. Indeed, visual inspection of the loop geometry at the first 100 ns of simulation revealed very close similarity to the O conformation (Figure 5B). The main differences involve regions of strong bending (kinks, Figure 5B, right). Although the structural features were the same, the side toward which each loop tilts was different. In our simulation, the loop got closer to the $O_{\text{SymL}}^{5'}$ operator, whereas in the O conformation reported by Villa et al.,²⁴ it is closer to $O_{\text{SymL}}^{3'}$. This might be consistent with the fact that although the LacI–DNA complex is symmetric in terms of structure it is not from a sequence point of view. So, the tilt of the loop to either operator may have implications on the function of LacI–DNA or its

interactions with other proteins, such as catabolite activator protein (CAP).^{14,51}

To better characterize the conformational landscape of the loop, we decomposed the movements of the DNA into two angular variables that describe tilt (β) and screw (ϕ) (Figure 5C). The U and O conformations have different values of ϕ and β , which allow them to be separated very well along the surface. On these two coordinates, the simulation explored a broad and shallow surface, suggesting high flexibility and a low energy minima. The formation of kinks that release the screw tension from the DNA fiber may be the reason for this observation (see below). Using this analysis, it was, nevertheless, possible to identify four basins representing the most sampled loop conformations, which were denoted L1–L4. Region L1, with an average tilt of 150° and screw of 1° , was the most visited conformation and represented 3% of the total surface. A representative structure within this zone was previously introduced and is shown in Figure 5B. About 20% of the sampled conformations was in the neighborhood of region of L1 (Figure 5C, red rectangle), whereas the remaining percentage was spread along the surface. The other three populated regions, L2–L4, represented 4% of the sampled surface. In agreement with previously reported nonhomogeneous rod models of the LacI–DNA system,¹⁸ representative structures of regions L1–L4 show that strong bending events have profound effects on the resulting geometry of the loop (Figure 5B,D).

Our results on the loop's dynamics were compatible with FRET experiments on the binding of LacI to a similar DNA segment. First, FRET measurements based on donor quenching of a Cy3–Cy5 pair placed close to the operators' sequence obtained an efficiency (E) of ~ 0.74 .⁴⁵ Later, single-molecule FRET experiments on the same system reported an E of ~ 0.90 .⁴⁶ Along our simulation, the loop was able to explore conformations that accounted for both reported FRET efficiencies (Figure 6, top). Indeed, along the time course, we

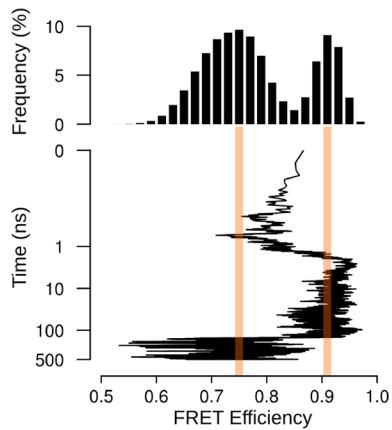


Figure 6. Comparison of FRET experiments from Edelman et al.⁴⁵ and Morgan et al.⁴⁶ with our multiscale simulation of the LacI–DNA system. Top: Histogram of FRET efficiencies from the simulation. Bottom: Time evolution on a logarithmic scale. The experimental values are indicated as transparent lines. See *Methodology* for more details of the measurements.

observed that the system was able to shift repeatedly from one state into the other (Figure 6, bottom). Previous multiscale simulations on the LacI–DNA system pointed to the mobility

of the headpiece domains as being a very important factor in FRET measurements.²⁴ Here, we show that the DNA loop's flexibility can equally explain those FRET observations.

Loop Dynamics: Local Bending and Cation Localization. As already mentioned, our CG model of DNA naturally incorporates sequence-dependent effects and important mechanical aspects like kinking. In addition, the electrostatics is accounted at long-range, taking into consideration all of the components of the system (which means not only phosphate atoms but also base pairs, protein, water, and ions). This level of description opens the possibility of exploring the fine details that link the mechanical and external factors governing the dynamics of the loop. In addition to the long-range deformations of the DNA fiber that affect the global shape of the loop, we also explored the bending plasticity on the scale of 5 bp, i.e., a length scale that is within a groove. The bending of each segment was described by the angle θ (see *Methodology*). Figure 7 introduces the local bending fluctuations of the DNA loop along the sequence and time space. As expected for a short fiber in a looped conformation, the segments comprising it were bent most of the time ($\theta > 10^\circ$). Regions with $\theta > 20^\circ$ were present in particular sections of the filament, although their specific localization varied during time. For example, the region spanning base pairs 60–70 experienced a strong bending event ($\theta > 40^\circ$) at 150–200 ns of simulation; it then went back to a low bending state, and it recovered the strong bending at 450 ns. Spontaneous fluctuations are indicative of low-energy processes on the order of kT . The local curvature pattern indicates the alternation of bent and nearly straight regions. From this analysis, the strong bending events can be correlated with the kinked conformations presented in Figure 5B,D. In agreement with experimental reports,⁵² AT- or GC-rich sequences were able to bend the DNA grooves to the same extent ($\theta \sim 20^\circ$).

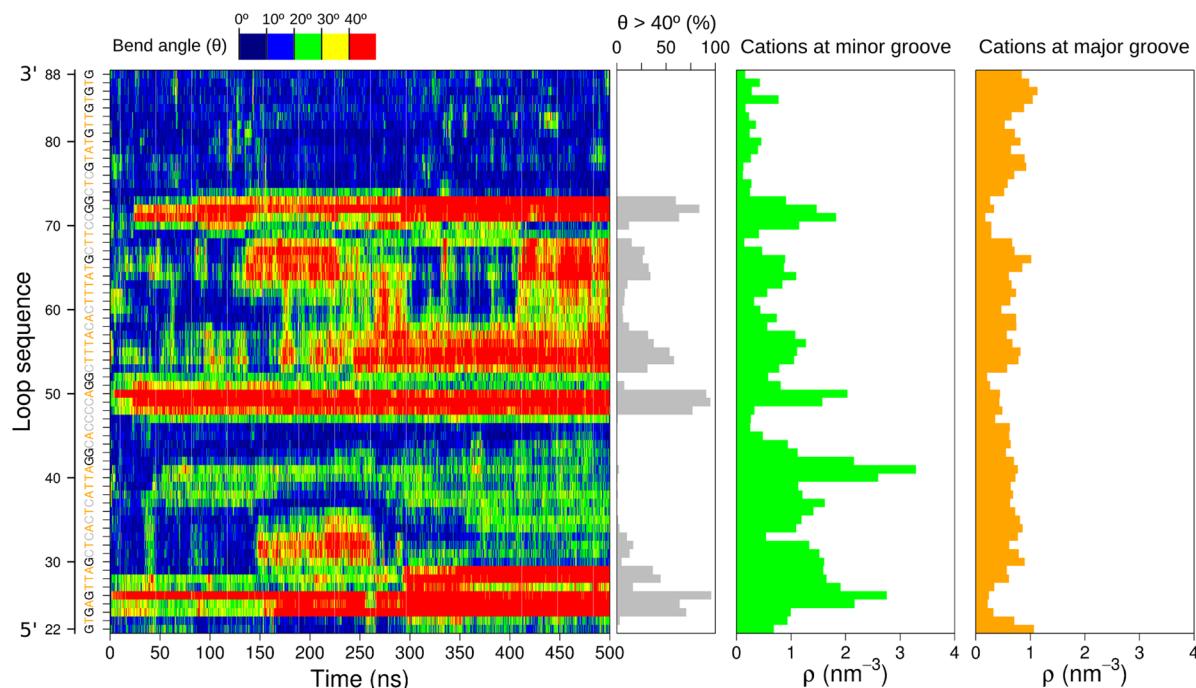


Figure 7. Local bending and cation localization along the DNA sequence. From left to right: Time evolution of the local bending, percentage of bent conformations with $\theta > 40^\circ$, number density of cations at minor and major grooves.

Single-molecule FRET experiments, theoretical analysis, and atomistic simulations showed that a short segment (<100 bp) of double-stranded DNA must form kinks to close a circular conformation.^{53–55} Indeed, electron microscopy and footprinting experiments on the LacI–DNA system confirmed the presence of such structural motifs.⁵⁶ In this regard, the patterns of strong bending observed in our simulation ($\theta > 40^\circ$) are consistent with the alternation of sites with high and low DNase I sensitivities (Figure 7). The presence of several events may be related to already described cooperative effects that are mechanically transmitted at distant sites through the stressed DNA filament.⁵⁷ Moreover, the bending is in agreement with experimental observations of a kink at region 31–35 with sequence TTTAT (Figure 1B) when the system is under supercoiling conditions.⁵⁸ The structural features of the kinks that we observed are the same as those described by us in a previous work,⁴⁷ which implies a strong change in the inter base pair roll, with a consequent widening of the major groove and narrowing of the minor groove. Such a conformational change in the polyelectrolyte is influenced by the solvent environment and ion distribution, which is an aspect that can be captured by the presented methodology.

It is known that the concentration of ions plays an important role in this protein–DNA interaction. In particular, experimental evidence points to the relevance of ions in modulating the LacI–DNA binding energy.⁵⁹ According to theoretical predictions and experimental evidence, ions may help by either stabilizing bent conformations or reducing the bending energy of the DNA by neutralizing phosphate charges at the bent face.^{60,61} In that sense, we observed preferential localization of counterions in the minor groove at zones of strong bending (i.e., around positions 25, 50, 55, 65, and 70 of the sequence) or with persistent flexibility (i.e., around position 40; Figure 7). On the other hand, the ions' distribution along the major groove was more homogeneous and constituted a less obvious pattern. Both observations are in agreement with previously described findings and point to a role of cations in either stabilizing or facilitating local bending conformations in the DNA loop.

Loop Dynamics: CAP Binding Region. Homodimeric catabolite activator protein (CAP) is a bacterial transcription regulator involved in the activation of the transcription of several operons.⁶² Cyclic adenosine monophosphate (cAMP) binds to CAP, causing a cascade of structural rearrangements that allow for selective binding to the consensus DNA sequence 5'-AAATGTGATCT-3'. This, in turn, induces DNA bending and a rearrangement of CAP, which is followed by the binding of RNA polymerase.⁶³

CAP binds the lac sequence at base pairs 18–39 (Figure 1B) and kinks the double-stranded DNA.⁶⁴ Experimental evidence showed that both LacI and CAP can be bound at the same time.^{65,66} The high flexibility that we observed in the loop at the CAP binding region raises a question about its contribution to protein–DNA recognition. Asymmetric phosphate neutralization experiments along the CAP binding sequence showed that the DNA alone could achieve bent conformations close to the experimental structure of the complex.⁶⁷ In agreement with that observation, the presence of strong bending events in the CAP binding region along our simulation (Figure 7) suggests that the DNA flexibility may play a role in an indirect readout mechanism of CAP–DNA recognition. We investigated if the loop was able to assume conformations close to the experimental structure of the CAP–DNA complex. Structural

alignment of the CAP binding region rendered a poor RMSD fit, which, in the best case, was ~0.9 nm (Figure 8A). However,

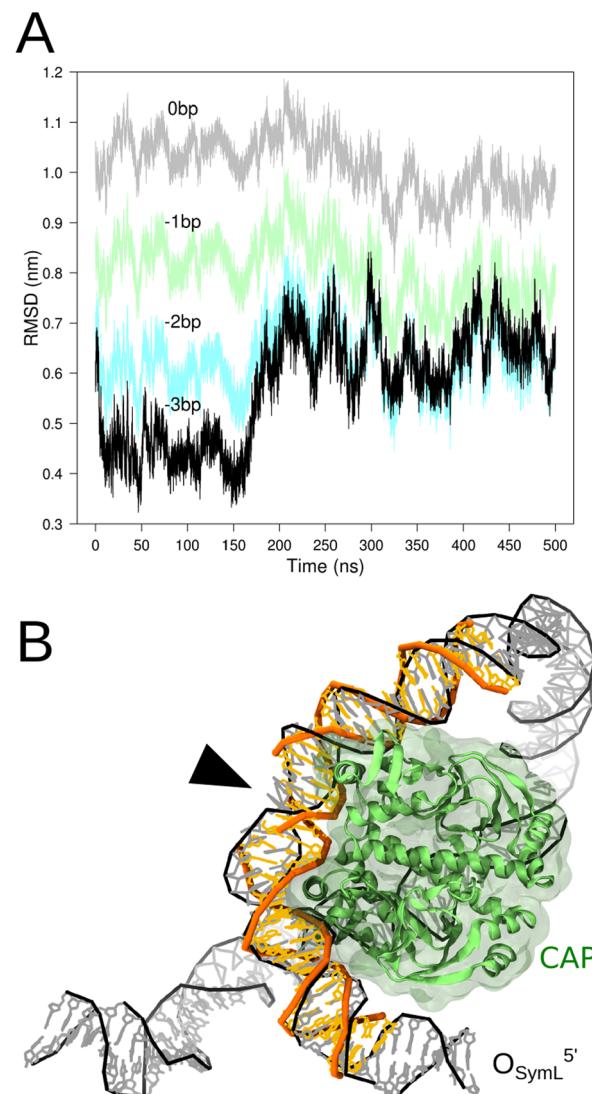


Figure 8. Interaction of CAP with the loop sequence of the LacI–DNA system. (A) RMSD of the CAP binding region at the loop with respect to its conformation in the experimental structure of the CAP–DNA complex (PDB ID: 1ZRC).⁶⁴ C1' atoms were used in the structural alignment to avoid discontinuities in the DNA backbone of the experimental structure. Results shifting the binding region toward the 5' end by –1, –2, and –3 bp are compared with those of the wild-type position (0 bp) (B) Structural representation of the RMSD fit at position –3 bp and 100 ns of simulation. The CAP protein is represented as a green cartoon, and its excluded volume is shown as a transparent surface. The DNA in the experimental structure is represented as orange thick tubes, and the loop in the LacI–DNA system is shown as black thin lines. The black arrow points to a kink at residue 26 of the loop. The LacI protein is omitted for clarity.

if the center of the CAP binding region was shifted –1, –2, or –3 bp toward the 5' end of the DNA, then the similarity increased to RMSD values below 0.4 nm. A shift of –3 bp puts the center of the CAP binding region at residue 26, which exhibited a strong bending event (Figure 7). As an example, Figure 8B shows the structural alignment of the CAP–DNA complex with the LacI–DNA system at 100 ns of simulation. From this picture, it is apparent that the spontaneously arising

kinks shape the DNA to a similar extent as that needed by CAP for binding. Considering the high flexibility of the loop and the transient nature of the kink events, we expect that the CAP binding region should be able to explore conformations very close to the experimental structure of the CAP–DNA complex. So, the loop's dynamics itself may play an important role in the CAP–DNA recognition process. The structural distortions induced by DNA kinks or bends generate a local environment of phosphate crowding that is suitable for protein cationic interactions.⁶⁸ These observations suggests a model in which CAP may recognize the curvature of DNA at medium distance via electrostatic interactions, whereas the final shaping would be mediated by close contacts and desolvation, which is in agreement with previous atomistic MD simulations.⁶⁹

CONCLUSIONS

A new multiscale methodology incorporating fully atomistic and coarse-grained resolution at the level of both solute and solvent was introduced and tested on a highly nontrivial case. This approach relies on the combination of AMBER and SIRAH force fields to obtain an AA/CG description of the system with explicit representation of the solvent. A comparison between reference systems demonstrated a good reproduction of structural and dynamic properties with regard to the solute and its solvent environment. In particular, the characteristics of the AA region were nicely preserved. Some outstanding features of this multiscale approach include the ease with which it can be implemented or extended to any other system configuration without the need for patching or writing new MD codes. Moreover, fast algorithms, such as PME and domain decomposition,⁷⁰ can be applied, as can upcoming techniques like multitime step,^{71,72} which will certainly increase its computational benefit. In addition, there is also a speed gain from the system's size reduction that depends on each case.²⁸ For example, the simulated LacI–DNA system (170 000 particles) is four times smaller than its fully atomistic representation. More importantly, however, is that the model provided a compatible description with other approaches like the continuum and nonhomogeneous elastic rod models of DNA, but it did so with the advantage of naturally including fine details like long-range electrostatics, accurate DNA dynamics, and ion/water distribution.

This work also presented the longest MD simulation performed to date in the study of the LacI–DNA system. The MD introduced the fundamental role of DNA mechanics and the surrounding ions' sphere in releasing tension from the complex by forming kinks at the loop sequence, an important aspect difficult to account for with other methodologies.^{24,46} However, we have reported that the kink's life times and ion binding events are on the order of several microseconds.^{47,48} The presented simulation is far below such sampling times. Furthermore, we considered only the wild-type sequence, whereas the impact of different sequence lengths and compositions have been studied. Despite these limitations, the results proved to be in very good agreement with previous simulations and experimental data, pointing to the robustness of the methodology.

AUTHOR INFORMATION

Corresponding Authors

*(M.R.M.) Tel/Fax: +598-2522 0910. E-mail: mmachado@pasteur.edu.uy

*(S.P.) Tel/Fax: +598-2522 0910. E-mail: spantano@pasteur.edu.uy.

Funding

This work was partially funded by FOCEM (MERCOSUR Structural Convergence Fund), COF 03/11. M.R.M. and S.P. are researchers from the National Scientific Program of ANII (SNI).

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Lewis, M. The Lac Repressor. *C. R. Biol.* **2005**, *328*, 521–548.
- (2) Friedman, A. M.; Fischmann, T. O.; Steitz, T. A. Crystal Structure of Lac Repressor Core Tetramer and Its Implications for DNA Looping. *Science* **1995**, *268*, 1721–1727.
- (3) Swigon, D.; Coleman, B. D.; Olson, W. K. Modeling the Lac Repressor-Operator Assembly: The Influence of DNA Looping on Lac Repressor Conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 9879–9884.
- (4) Villa, E.; Balaeff, A.; Mahadevan, L.; Schulten, K. Multiscale Method for Simulating Protein-DNA Complexes. *Multiscale Model. Simul.* **2004**, *2*, 527–553.
- (5) Tsodikov, O. V.; Saecker, R. M.; Melcher, S. E.; Levandoski, M. M.; Frank, D. E.; Capp, M. W.; Record, M. T., Jr. Wrapping of Flanking Non-Operator DNA in Lac Repressor-Operator Complexes: Implications for DNA looping. *J. Mol. Biol.* **1999**, *294*, 639–655.
- (6) La Penna, G.; Perico, A. Wrapped-Around Models for the Lac Operon Complex. *Biophys. J.* **2010**, *98*, 2964–2973.
- (7) Wong, O. K.; Guthold, M.; Erie, D. A.; Gelles, J. Interconvertible Lac Repressor-DNA Loops Revealed by Single-Molecule Experiments. *PLoS Biol.* **2008**, *6*, e232.
- (8) Hirsh, A. D.; Lillian, T. D.; Lionberger, T. A.; Perkins, N. C. DNA Modeling Reveals an Extended Lac Repressor Conformation in Classic In Vitro Binding Assays. *Biophys. J.* **2011**, *101*, 718–726.
- (9) Han, L.; Garcia, H. G.; Blumberg, S.; Towles, K. B.; Beausang, J. F.; Nelson, P. C.; Phillips, R. Concentration and Length Dependence of DNA Looping in Transcriptional Regulation. *PLoS One* **2009**, *4*, e5621.
- (10) Haeusler, A. R.; Goodson, K. A.; Lillian, T. D.; Wang, X.; Goyal, S.; Perkins, N. C.; Kahn, J. D. FRET Studies of a Landscape of Lac Repressor-Mediated DNA Loops. *Nucleic Acids Res.* **2012**, *40*, 4432–4445.
- (11) Spronk, C. A.; Bonvin, A. M.; Radha, P. K.; Melacini, G.; Boelens, R.; Kaptein, R. The Solution Structure of Lac Repressor Headpiece 62 Complexed to a Symmetrical Lac Operator. *Structure* **1999**, *7*, 1483–1492.
- (12) Lewis, M.; Chang, G.; Horton, N. C.; Kercher, M. A.; Pace, H. C.; Schumacher, M. A.; Brennan, R. G.; Lu, P. Crystal Structure of the Lactose Operon Repressor and Its Complexes with DNA and Inducer. *Science* **1996**, *271*, 1247–1254.
- (13) Bell, C. E.; Lewis, M. A Closer View of the Conformation of the Lac Repressor Bound to Operator. *Nat. Struct. Biol.* **2000**, *7*, 209–214.
- (14) Balaeff, A.; Mahadevan, L.; Schulten, K. Structural Basis for Cooperative DNA Binding by CAP and Lac Repressor. *Structure* **2004**, *12*, 123–132.
- (15) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The Power of Coarse Graining in Biomolecular Simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 225–248.
- (16) Kahn, J. D.; Cheong, R.; Mehta, R. A.; Edelman, L. M.; Morgan, M. A. Flexibility and Control of Protein–DNA Loops. *Biophys. Rev. Lett.* **2006**, *01*, 327–341.
- (17) Zhang, Y.; McEwen, A. E.; Crothers, D. M.; Levene, S. D. Statistical-Mechanical Theory of DNA Looping. *Biophys. J.* **2006**, *90*, 1903–1912.

- (18) Goyal, S.; Perkins, N. C. Looping Mechanics of Rods and DNA with Non-Homogeneous and Discontinuous Stiffness. *Int. J. Non-Linear Mech.* **2008**, *43*, 1121–1129.
- (19) Cherstvy, A. G. Looping Charged Elastic Rods: Applications to Protein-Induced DNA Loop Formation. *Eur. Biophys. J.* **2011**, *40*, 69–80.
- (20) Perez, P. J.; Clauvelin, N.; Grosner, M. A.; Colasanti, A. V.; Olson, W. K. What Controls DNA Looping? *Int. J. Mol. Sci.* **2014**, *15*, 15090–15108.
- (21) Biton, Y. Y.; Kumar, S.; Dunlap, D.; Swigon, D. Lac Repressor Mediated DNA Looping: Monte Carlo Simulation of Constrained DNA Molecules Complemented with Current Experimental Results. *PLoS One* **2014**, *9*, e92475.
- (22) Kar, P.; Feig, M. Recent Advances in Transferable Coarse-Grained Modeling of Proteins. In *Advances in Protein Chemistry and Structural Biology*; Karabencheva-Christova, T., Ed.; Academic Press: Oxford, 2014; Vol. 96, Chapter 5, pp 143–180.
- (23) Fogarty, A. C.; Potestio, R.; Kremer, K. Adaptive Resolution Simulation of a Biomolecule and Its Hydration Shell: Structural and Dynamical Properties. *J. Chem. Phys.* **2015**, *142*, 195101.
- (24) Villa, E.; Balaeff, A.; Schulten, K. Structural Dynamics of the Lac repressor–DNA Complex Revealed by a Multiscale Simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6783–6788.
- (25) Darré, L.; Machado, M. R.; Dans, P. D.; Herrera, F. E.; Pantano, S. Another Coarse Grain Model for Aqueous Solvation: WAT FOUR? *J. Chem. Theory Comput.* **2010**, *6*, 3793–3807.
- (26) Dans, P. D.; Zeida, A.; Machado, M. R.; Pantano, S. A Coarse Grained Model for Atomic-Detailed DNA Simulations with Explicit Electrostatics. *J. Chem. Theory Comput.* **2010**, *6*, 1711–1725.
- (27) Machado, M. R.; Dans, P. D.; Pantano, S. A Hybrid All-Atom/coarse Grain Model for Multiscale Simulations of DNA. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18134–18144.
- (28) Gonzalez, H. C.; Darré, L.; Pantano, S. Transferable Mixing of Atomistic and Coarse-Grained Water Models. *J. Phys. Chem. B* **2013**, *117*, 14438–14448.
- (29) Arnott, S.; Campbell-Smith, P.; Chandrasekaran, R. Atomic Coordinates and Molecular Conformations for DNA–DNA, RNA–RNA, and DNA–RNA Helices. In *CRC Handbook of Biochemistry and Molecular Biology*; Fasman, G. D., Ed.; CRC Press: Cleveland, OH, 1976; Vol. 2, pp 411–422.
- (30) Macke, T. J.; Case, D. A. Modeling Unusual Nucleic Acid Structures. In *Molecular Modeling of Nucleic Acids*; Leontis, N. B., SantaLucia, J., Jr., Eds.; American Chemical Society: Washington, DC, 1997; Vol. 682, pp 379–393.
- (31) Manning, G. S. Counterion Binding in Polyelectrolyte Theory. *Acc. Chem. Res.* **1979**, *12*, 443–449.
- (32) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712–725.
- (33) Sorin, E. J.; Pande, V. S. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophys. J.* **2005**, *88*, 2472–2493.
- (34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (35) Darré, L.; Tek, A.; Baaden, M.; Pantano, S. Mixing Atomistic and Coarse Grain Solvation Models for MD Simulations: Let WT4 Handle the Bulk. *J. Chem. Theory Comput.* **2012**, *8*, 3880–3894.
- (36) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695–1697.
- (37) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (38) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (39) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (41) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (42) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.
- (43) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (44) Cheatham, T. E., III; Case, D. A. Using Amber to Simulate DNA and RNA. In *Computational Studies of RNA and DNA*; Šponer, J., Lankáš, F., Eds.; Springer: Dordrecht, The Netherlands, 2006; pp 45–71.
- (45) Edelman, L. M.; Cheong, R.; Kahn, J. D. Fluorescence Resonance Energy Transfer over Approximately 130 Basepairs in Hyperstable Lac Repressor-DNA Loops. *Biophys. J.* **2003**, *84*, 1131–1145.
- (46) Morgan, M. A.; Okamoto, K.; Kahn, J. D.; English, D. S. Single-Molecule Spectroscopic Determination of Lac Repressor-DNA Loop Conformation. *Biophys. J.* **2005**, *89*, 2588–2596.
- (47) Zeida, A.; Machado, M. R.; Dans, P. D.; Pantano, S. Breathing, Bubbling, and Bending: DNA Flexibility from Multimicrosecond Simulations. *Phys. Rev. E* **2012**, *86*, 021903.
- (48) Dans, P. D.; Darré, L.; Machado, M. R.; Zeida, A.; Brandner, A. F.; Pantano, S. Assessing the Accuracy of the SIRAH Force Field To Model DNA at Coarse Grain Level. In *Advances in Bioinformatics and Computational Biology*; Setubal, J. C., Almeida, N. F., Eds.; Springer International Publishing: New York, 2013; pp 71–81.
- (49) Kalcher, I.; Horinek, D.; Netz, R. R.; Dzubiella, J. Ion Specific Correlations in Bulk and at Biointerfaces. *J. Phys.: Condens. Matter* **2009**, *21*, 424108.
- (50) Balaeff, A.; Mahadevan, L.; Schulten, K. Modeling DNA Loops Using the Theory of Elasticity. *Phys. Rev. E* **2006**, *73*, 031919.
- (51) Swigon, D.; Olson, W. K. Mesoscale Modeling of Multi-Protein-DNA Assemblies: The Role of the Catabolic Activator Protein in Lac-Repressor-Mediated Looping. *Int. J. Non-Linear Mech.* **2008**, *43*, 1082–1093.
- (52) Strauss-Soukup, J. K.; Rodrigues, P. D.; Maher, L. J. Effect of Base Composition on DNA Bending by Phosphate Neutralization. *Biophys. Chem.* **1998**, *72*, 297–306.
- (53) Lankas, F.; Lavery, R.; Maddocks, J. H. Kinking Occurs during Molecular Dynamics Simulations of Small DNA Minicircles. *Structure* **2006**, *14*, 1527–1534.
- (54) Le, T. T.; Kim, H. D. Probing the Elastic Limit of DNA Bending. *Nucleic Acids Res.* **2014**, *42*, 10786–10794.
- (55) Vologodskii, A.; Frank-Kamenetskii, M. D. Strong Bending of the DNA Double Helix. *Nucleic Acids Res.* **2013**, *41*, 6785–6792.
- (56) Krämer, H.; Niemöller, M.; Amouyal, M.; Revet, B.; von Wilcken-Bergmann, B.; Müller-Hill, B. Lac Repressor Forms Loops with Linear DNA Carrying Two Suitably Spaced Lac Operators. *EMBO J.* **1987**, *6*, 1481–1491.
- (57) Lionberger, T. A.; Demurtas, D.; Witz, G.; Dorier, J.; Lillian, T.; Meyhöfer, E.; Stasiak, A. Cooperative Kinking at Distant Sites in Mechanically Stressed DNA. *Nucleic Acids Res.* **2011**, *39*, 9820–9832.
- (58) Borowiec, J. A.; Zhang, L.; Sasse-Dwight, S.; Gralla, J. D. DNA Supercoiling Promotes Formation of a Bent Repression Loop in Lac DNA. *J. Mol. Biol.* **1987**, *196*, 101–111.
- (59) Barkley, M. D.; Lewis, P. A.; Sullivan, G. E. Ion Effects on the Lac Repressor-Operator Equilibrium. *Biochemistry* **1981**, *20*, 3842–3851.
- (60) Strauss, J. K.; Maher, L. J. DNA Bending by Asymmetric Phosphate Neutralization. *Science* **1994**, *266*, 1829–1834.
- (61) Range, K.; Mayaan, E.; Maher, L. J.; York, D. M. The Contribution of Phosphate-Phosphate Repulsions to the Free Energy of DNA Bending. *Nucleic Acids Res.* **2005**, *33*, 1257–1268.

- (62) Kolb, A.; Busby, S.; Buc, H.; Garges, S.; Adhya, S. Transcriptional Regulation by cAMP and Its Receptor Protein. *Annu. Rev. Biochem.* **1993**, *62*, 749–795.
- (63) Zhou, Y.; Zhang, X.; Ebright, R. H. Identification of the Activating Region of Catabolite Gene Activator Protein (CAP): Isolation and Characterization of Mutants of CAP Specifically Defective in Transcription Activation. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6081–6085.
- (64) Napoli, A. A.; Lawson, C. L.; Ebright, R. H.; Berman, H. M. Indirect Readout of DNA Sequence at the Primary-Kink Site in the CAP-DNA Complex: Recognition of Pyrimidine-Purine and Purine-Purine Steps. *J. Mol. Biol.* **2006**, *357*, 173–183.
- (65) Hudson, J. M.; Fried, M. G. Co-Operative Interactions between the Catabolite Gene Activator Protein and the Lac Repressor at the Lactose Promoter. *J. Mol. Biol.* **1990**, *214*, 381–396.
- (66) Vossen, K. M.; Stickle, D. F.; Fried, M. G. The Mechanism of CAP-Lac Repressor Binding Cooperativity at the E. Coli Lactose Promoter. *J. Mol. Biol.* **1996**, *255*, 44–54.
- (67) Hardwidge, P. R.; Zimmerman, J. M.; Maher, L. J. Charge Neutralization and DNA Bending by the Escherichia Coli Catabolite Activator Protein. *Nucleic Acids Res.* **2002**, *30*, 1879–1885.
- (68) Grant, B. N.; Dourlain, E. M.; Araneda, J. N.; Throneberry, M. L.; McFail-Isom, L. A. DNA Phosphate Crowding Correlates with Protein Cationic Side Chain Density and Helical Curvature in protein/DNA Crystal Structures. *Nucleic Acids Res.* **2013**, *41*, 7547–7555.
- (69) Berrera, M.; Pantano, S.; Carloni, P. Catabolite Activator Protein in Aqueous Solution: A Molecular Simulation Study. *J. Phys. Chem. B* **2007**, *111*, 1496–1501.
- (70) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (71) Ferrarotti, M. J.; Bottaro, S.; Pérez-Villa, A.; Bussi, G. Accurate Multiple Time Step in Biased Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 139–146.
- (72) Di Pasquale, N.; Gowers, R. J.; Carbone, P. A Multiple Time Step Scheme for Multiresolved Models of Macromolecules. *J. Comput. Chem.* **2014**, *35*, 1199–1207.