Multiscale Simulation of Nanobiological Flows

A new multiscale approach for simulating nanobiological flows uses concurrent coupling of constrained molecular dynamics for long biomolecules with a mesoscopic lattice Boltzmann treatment of solvent hydrodynamics. The approach is based on a simple scheme of space–time information exchange between the atomistic and mesoscopic scales.

Modeling biological systems efficiently and reliably calls for innovative computational methods, often requiring sophisticated upgrades and extensions of techniques originally developed for physical and chemical stand-alone applications. Indeed, biological systems exhibit a degree of complexity and diversity that straddles many decades in space–time resolution, to the point that, for many years, biological systems served as a paradigm of the kind of complexity that can only be handled in qualitative or descriptive terms. Advances in computer technology, combined with constant progress and breakthroughs in simulational methods, are closing the gap between quantitative models and actual biological behavior. The main computational challenge raised by biological systems remains the wide and disparate range of spatiotemporal scales involved in their dynamical evolution. Protein folding, morphogenesis, and intra- and extracellular communication are just a few examples.

Various strategies address this challenge. These methods—generally referred to as *multiscale modeling*—are based on composite computational schemes that rely on multiple levels of description of a given biological system, most typically the atomistic and continuum levels. The multiple descriptions are then glued together, through suitable hand-shaking procedures, to produce the final composite multiscale algorithm. To date, mainstream multiscale modeling is based on the coupling of atomistic and continuum models. This choice reflects the historical developments of statistical mechanics and computational physics. In essence, continuum methods reduce the information to a small number of distributed properties (fields), whose space–time evolution is computed by solving a corresponding set of partial differential equations, such as reaction–diffusion–advection equations. Atomistic models, on the other hand, rely on the well-established molecular dynamics (MD) approach, possibly including extensions capable of dealing with a quantum description. This two-stage continuum-to-atomistic representation overlooks an intermediate description level—the *mesoscopic* level—as typically

1521-9615/08/$25.00 © 2008 IEEE
Copublished by the IEEE CS and the AIP

Maria Fyta and Efthimios Kaxiras
Harvard University
Simone Melchionna
University of Rome
Sauro Succi
National Research Council, Italy

represented by Boltzmann kinetic theory and its extensions. Kinetic theory lies between the continuum and atomistic descriptions, so it should provide an appropriate framework for developing robust multiscale methodologies.

Until recently, the common perception of the Boltzmann equation (the central equation of kinetic theory) as computationally nearly as demanding as MD yet of limited use for dense fluids (water being the typical biological medium) because it lacks many-body correlations hindered this approach. But new developments in lattice kinetic theory[1,2] are making this view obsolete. Over the past decade, such developments have shown that suitably discretized forms of minimal kinetic equations, most notably the lattice Boltzmann (LB) equation, are giving rise to efficient algorithms capable of handling complex flowing systems across many scales of motion. Fluid flow behavior is described through minimal forms of the Boltzmann equation, living on a discrete lattice. The lattice dynamics is designed to reflect the basic conservation laws of continuum mechanics and also to host additional (mesoscopic) physics, which continuum models don't easily accommodate. Remarkably, you can achieve both tasks within the same algorithm, which often proves computationally advantageous over the continuum approach based on the Navier-Stokes equations.

Only very recently have researchers begun to incorporate LB advances in a new class of mesoscopic multiscale solvers.[3] In this article, we address these solvers, focusing on the biologically important problem of biopolymer translocation through nanopores. We base our procedure on the assumption that, to capture the translocation process's essential aspects, we don't have to resolve all underlying atomistic details. As a result, we treat both solvent and solute degrees of freedom through appropriate coarse-graining. Because of our methodology's intrinsic coarse-graining nature, the direct mapping to experimental conditions must proceed through the adjustment of appropriate parameters.

Multiscale Coupling Methodology

We couple a mesoscopic fluid solver (the LB method) concurrently to the atomistic scale using explicit atomistic dynamics, which, for simplicity, we call MD in a broad sense. This procedure involves different levels of the statistical description of matter (continuum and atomistic). It handles different scales through the spatial and temporal coupling between the constrained MD for the polymer evolution and the LB treatment of the explicit solvent dynamics. This multiscale framework is well suited to address a class of biologically related problems.

The solvent dynamic doesn't require statistical ensemble averaging because it's represented through a discrete set of pre-averaged probability distribution functions (the single-particle Boltzmann distributions), which are propagated along straight particle trajectories. At variance with Brownian dynamics, the LB approach handles fluid-mediated solvent-to-solvent interactions through an explicit representation of local collisions between the solvent and solute molecules. By leveraging space–time locality, the corresponding algorithm scales linearly with the number of beads, as opposed to Brownian dynamics' superquadratic dependence. This dual field/particle nature greatly facilitates the coupling between the mesoscopic and atomistic levels, both on conceptual and computational grounds. Full details on this scheme appear elsewhere.[4] Researchers have previously coupled LB and MD to investigate single-polymer dynamics,[5] but we extend this to long molecules of biological interest.

We base our procedure on the assumption that, to capture the translocation process's essential aspects, we don't have to resolve all underlying atomistic details.

A word of caution is in order with the Stokes limit, $Re \rightarrow 0$. In this limit, the scale separation between atomistic and hydrodynamic degrees of freedom becomes opaque, as the atomic mean free path becomes comparable with hydrodynamic scales. However, researchers have shown that finite Reynolds corrections to hydrodynamics have negligible effects on the solvent-mediated forces between suspended bodies—for example, Oseen-level hydrodynamics is satisfactorily recovered.[4,6]

We first turn to the atomistic part in our approach and consider a polymer consisting of N monomer units (also referred to as *beads*). We advance in time each polymer bead according to the following set of MD equations:

$$\mathbf{F}_{tot,i} = \mathbf{F}_{c,i} + \mathbf{F}_{drag,i} + \mathbf{F}_{r,i} + \mathbf{F}_{\kappa,i}, \quad (1)$$

where the index i runs over all beads. In this expression, $\mathbf{F}_{c,i}$ is a conservative force describing

bead-to-bead interactions, represented here by a Lennard-Jones potential:

$$V_{LJ}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \quad (2)$$

We truncate this potential at a distance of $r = 2^{1/6}\sigma$ and augment it by an angular harmonic term to account for distortions of the angle between consecutive bonds:

$$V_{ang}(\phi) = \frac{\kappa_\phi \phi^2}{2}, \quad (3)$$

where ϕ is the relative angle between two consecutive bonds, and κ_ϕ is a constant. The present model doesn't include torsional motions, but we can easily incorporate them if needed. The second term in Equation 1, $\mathbf{F}_{drag,i}$, represents the dissipative drag force due to polymer–fluid coupling given by

$$\mathbf{F}_{drag,i} = -m_\gamma(\mathbf{v}_i - \mathbf{u}_i), \quad (4)$$

where $\mathbf{v}_i$, $\mathbf{u}_i$ is the bead and fluid velocity evaluated at the bead position $\mathbf{r}_i$ of bead i with a mass m; and γ is the friction coefficient.

In addition to mechanical drag, the polymer feels the effects of stochastic fluctuations in the fluid environment, which is related to the third term in Equation 1, $\mathbf{F}_{r,i}$, an uncorrelated random force with zero mean acting on bead i. The term $\mathbf{F}_{\kappa,i}$ in Equation 1 is the reaction force resulting from $N - 1$ holonomic constraints for molecules modeled with rigid covalent bonds. Using constraints instead of flexible bond lengths lets us eliminate unimportant high-frequency intramolecular motion that might lead to numerical instabilities. To avoid spurious dissipation, the bead velocities must be strictly orthogonal to the relative displacements. We enforce the constraints on both positions and velocities over positions and momenta separately via the Shake and Rattle algorithms.[7,8]

The LB equation is a minimal form of the Boltzmann kinetic equation in which all molecular motion details are removed except those that are strictly needed to recover hydrodynamic behavior at the macroscopic scale (mass-momentum and energy conservation). The result is an elegant equation for the discrete distribution function $f_p(\mathbf{x}, t)$ describing the probability of finding an LB particle at lattice site $\mathbf{x}$ at time t with a discrete speed $\mathbf{c}_p$. Specifically, we're dealing with nanoscale flows and consider the fluctuating LB equation, which takes the following form:[9]

$$f_p(\mathbf{x} + \mathbf{c}_p\Delta t, t + \Delta t) = f_p(\mathbf{x}, t) - \omega\Delta t(f_p - f_p^{eq})(\mathbf{x}, t) + F_p\Delta t + S_p\Delta t. \quad (5)$$

The particles can only move along the links of a regular lattice defined by discrete speeds, so the synchronous displacements $\Delta\mathbf{x}_p = \mathbf{c}_p\Delta t$ never take the fluid particles away from the lattice. We use the standard 3D 19-speed lattice in this study (see Figure 1 in Sauro Succi and his colleagues' article[3]). The right-hand side of Equation 5 represents the effect of intermolecular solvent-to-solvent collisions, through a relaxation toward local equilibrium, f_p^{eq}, typically a second-order (low-Mach) expansion in the fluid velocity of a local Maxwellian with speed $\mathbf{u}$:

$$f_p^{eq} = w_p\rho\left\{1 + \frac{\mathbf{u}\cdot\mathbf{c}_p}{c_s^2} + \frac{1}{2c_s^4}[uu : (\mathbf{c}_p\mathbf{c}_p - c_s^2\mathbf{I})]\right\}, \quad (6)$$

where c_s is the solvent's sound speed, w_p is a set of weights normalized to unity, $\boldsymbol{I}$ is the unit tensor in configuration space, and ρ is the local density. The relaxation frequency controls the fluid kinematic viscosity υ through the relation $\upsilon = c_s^2(1/\omega - \Delta t/2)$ and Δt is the LB time step.[3] Knowing the discrete distributions f_p lets us calculate the local density ρ, flow speed $\rho\mathbf{u}$, and momentum-flux tensor $\boldsymbol{P}$, using a direct summation on all discrete distributions:

$$\rho(\mathbf{x},t) = \sum_p f_p(\mathbf{x},t) \quad (7)$$

$$\rho\mathbf{u}(\mathbf{x},t) = \sum_p f_p(\mathbf{x},t)\mathbf{c}_p \quad (8)$$

$$\boldsymbol{P}(\mathbf{x},t) = \sum_p f_p(\mathbf{x},t)\mathbf{c}_p\mathbf{c}_p. \quad (9)$$

The momentum-flux tensor's diagonal component gives the fluid pressure, while the off-diagonal terms give the shear stress. Both quantities are available locally and at any point in the simulation. Thermal fluctuations are included through the source term F_p in Equation 5, which is consistent with the fluctuation theorem at all scales. In the same equation, the polymer-fluid back reaction is described through the source term S_p, which represents the momentum input per unit time due to the polymer's reaction on the fluid populations. Equation 10 gives this back reaction:

$$S_p(\mathbf{x},t) = \frac{w_p}{c_s^2}\sum_{i\in D(x)}[\mathbf{F}_{drag,i} + \mathbf{F}_{r,i}]\cdot\mathbf{c}_p, \quad (10)$$

where $D(x)$ denotes the mesh cell to which the ith

bead belongs. All quantities in this equation must reside on the lattice nodes, so we must extrapolate the frictional and random forces from the particle to the grid location.

In the LB solver, free streaming proceeds along straight trajectories, securing exact conservation of the numerical scheme's mass and momentum and facilitating the imposition of geometrically complex boundary conditions. We don't need to solve the computationally expensive Poisson equations because the pressure field is locally available. All interactions are local, rendering the LB scheme ideal for parallel computing. More advanced LB models[10] could be equally well-suited for coupling to atomic-scale dynamics.

The MD solver is marched in time with a stochastic integrator (due to extra nonconservative and random terms),[11] proceeding at a fraction $1/M$ of the LB time step Δt: $\Delta t_{MD} = \Delta t/M$. The time step ratio $M > 1$ controls the scale separation between the solvent and solute timescales and should be chosen as small as possible, consistent with the requirement of providing a realistic description of the polymer dynamics. We repeat the MD cycle M times, freezing the hydrodynamic field at each LB timestamp $t_n = n\Delta t$. At this time step, we transfer spatial information from grid to particle locations (and conversely). For this transfer, we use a simple nearest-grid-point interpolation scheme (see Figure 1) for simplicity. At a time step $t_n = n\Delta t$, the pseudoalgorithm describing a single LB time step will be

1. Velocity interpolation: $\mathbf{u}(\mathbf{x}) \rightarrow \mathbf{u}_i$
2. For $k = 1, M$: Advance the molecular state from t to $t + \Delta t_{MD}$
3. Force extrapolation: $\mathbf{F}_i \rightarrow \mathbf{F}(\mathbf{x})$
4. Advance the Boltzmann populations from t to $t + \Delta t$.

Figure 1 outlines this scheme. In terms of computational efficiency, Δt_{MD} is largely independent of the number of beads N because

- the LB–MD coupling is local,
- the forces are short-ranged, and
- the Shake and Rattle algorithms are empirically known to scale linearly with the number of constraints.

Up to this point, we've described a general scheme that's applicable to any situation in which a long polymer is moving in a solvent. This motion supports a fundamental understanding of polymer dynamics in the solvent's presence and is crucial in relevant biophysical processes.

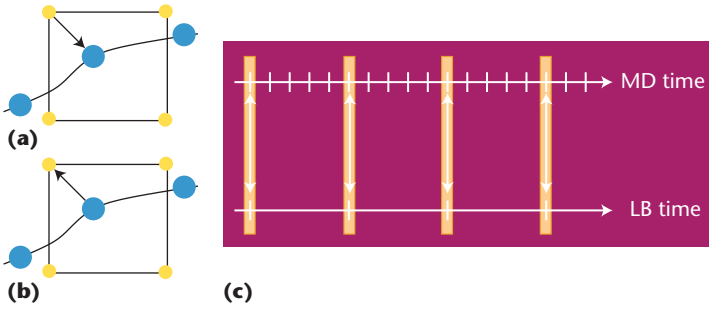

Figure 1. Nearest-grid-point interpolation scheme. We transfer spatial information (a) from grid to particle and (b) from particle to grid. Gray spheres denote beads, and white spheres are lattice sites. (c) The information exchange (lattice Boltzmann [LB]–molecular dynamics [MD] coupling) is sketched through the vertical highlighted regions between the two scales of our multiscale approach. The MD marches in time on a finer timescale (by a factor of five in the sketch and our simulations) than the LB solver.

Biopolymer Translocation through Nanopores

Recent experimental studies have focused on the translocation of biopolymers such as RNA or DNA through nanometer-sized pores. These studies explore in vitro the translocation process through microfabricated channels under the effects of an external electric field or through protein channels across cellular membranes.[12,13] In particular, recent experimental work has focused on the possibility of fast DNA sequencing through electronic means—that is, by reading the DNA sequence while it moves through a nanopore under the effect of a localized electric field.[12] This type of biophysical process is important in phenomena such as viral infection by phages, interbacterial DNA transduction, or gene therapy. Researchers have also analyzed some universal features of DNA translocation theoretically, using nonhydrodynamic coarse-grained or microscopic models.[14,15] However, these complex phenomena involve competition between many-body interactions at the atomic or molecular scale, fluid–atom hydrodynamic coupling, and the interaction of the biopolymer with wall molecules in the pore's region. Resolving these interactions is essential to understanding the physics underlying the translocation process. To this end, we model the dynamics of biopolymer translocation through narrow pores, using the multiscale scheme described previously.

We performed our numerical simulations in a

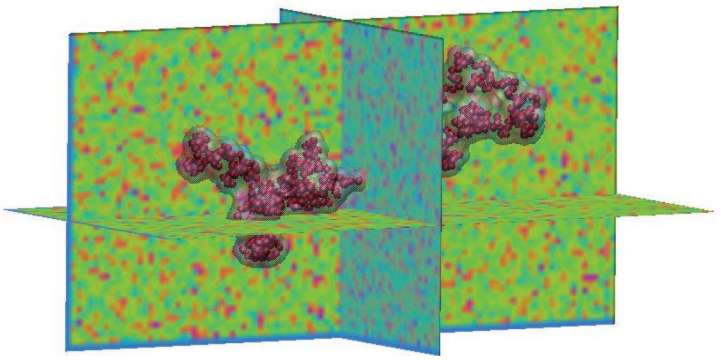

Figure 2. A typical translocation event. The 3D box includes the fluid, the polymer, and a wall (not shown) that separates the box into two chambers. A polymer (red beads) represents double-stranded DNA at an instant when roughly 60 percent of the beads have translocated. The fluid is also shown as a 3D representation of the velocity magnitude at the beads' vicinity and on 2D planes (red denotes high value). A pulling force $\mathbf{F}_{drive,i}$ with a direction from right to left induces translocation.

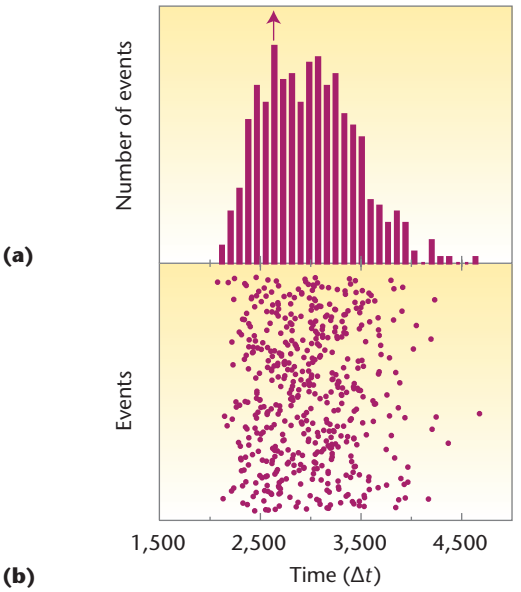

Figure 3. Translocation results for a molecule with 100 beads. (b) The variety of translocation events and (a) the corresponding duration histogram.

3D box of size $N_x \times N_y \times N_z$ in units of the lattice spacing Δx. The box contains both the solvent and the polymer. We take $N_x = 2N_y$, $N_y = N_z$, with a separating wall located in the midsection of the x direction at $x = N_x/2$. For polymers with fewer than 500 beads, we use $N_x = 80$; for larger polymers, we use $N_x = 100$. At $t = 0$, the polymer resides entirely in the right chamber at $x > N_x/2$. This polymer can translocate from one chamber to the other through a square hole of side $d = 3\Delta x$ in the center of the separating wall. Figure 2 shows a 3D representation of a typical translocation event. We also map the fluid speed's magnitude on different planes and as a 3D contour surface surrounding the polymer beads.

Several parameters are involved in the simulation (additional details appear elsewhere[4]). We measure all parameters in units of the LB time step Δt and spacing Δx, which we set equal to 1. The parameters for the Lennard-Jones potential are $\sigma = 1.8$ and $\varepsilon = 2 \times 10^{-3}$, and we set the bond length among the beads to $b = 1.2$. The solvent density and kinematic viscosity are 1 and 0.1, respectively, and the inverse temperature is $\beta = 10^4$. The friction coefficient, $\gamma = 0.1$, is a parameter governing both the polymer's structural relation toward equilibrium and the strength of the coupling with the surrounding fluid. The MD time step is a $1/M$ fraction of the LB time step, as mentioned previously, and we set $M = 5$.

A constant electric force, $\mathbf{F}_{drive,i}$, induces translocation. This force acts along the x direction and is confined in a rectangular channel of size $3\Delta x \times \Delta x \times \Delta x$ along the streamwise (x direction) and cross flow (y, z directions). This driving force, which is an additional term in Equation 1, represents the external field's effect in the experiments. We use $\mathbf{F}_{drive,i} = 0.02$. This choice of parameter values implies that we're describing the fast translocation regime, in which the translocation time is much smaller than the Zimm time, which is the polymer's typical relaxation time toward its native (minimum energy, maximum entropy) configuration. Under these conditions, we can't ignore the polymer dynamics' many-body aspects because the beads along the chain don't move independently.

Here, we model DNA as a polymeric chain of several segments (the beads) and trace its dynamic evolution interacting with a fluid solvent as it passes through a narrow hole that's comparable to the bead size. Each bead maps to several base pairs (bp), ranging from roughly 8 (similar to the hydrated diameter of B-DNA in physiological conditions) to approximately 10^3.[16,17] To estimate this number for our simulations and interpret our results in terms of physical units, we examine the persistence length (l_p) of the semiflexible polymers used in our simulations. We use the formula for the fixed-bond-angle model of a worm-like chain:[18]

$$l_p = \frac{b}{1-\cos\langle\theta\rangle}, \qquad (11)$$

where $\langle\theta\rangle$ complements the average bond angle between adjacent bonds. In lattice units (Δx), we found an average persistence length for the polymers considered to be approximately 12. For λ-phage DNA, l_p ~ 50 nm,[19] which we set equal to l_p for our polymers. Thereby, the lattice spacing is Δx ~ 4 nm, which is also the size of one bead. Given that the bp spacing is ~0.34 nm, one bead maps to approximately 12 bps. With this mapping, the pore size is roughly 12 nm, close to the experimental pores, which are of the order of 10 nm. The polymers presented here with N = 20 – 500 beads correspond to DNA lengths in the range 0.2 – 6 × 10^3 bp. The DNA lengths used in the experiments are larger (up to ~10^5 bp). With appropriate computational resources, our multiscale scheme could handle these lengths.

Having established the quantitative mapping of DNA bp to the simulated beads, we compared the statistical features of the simulated translocation process to experimental studies. We then generated the ensemble of our simulations using different realizations of the initial polymer configuration to account for the process's statistical nature. Finally, we performed extensive simulations of numerous translocation events over 100 to 1,000 initial polymer configurations for each polymer length. Figure 3b depicts the various events for a molecule with 100 beads, which represents double-stranded DNA with 1.2 kbp; Figure 3a shows the projected duration histograms in LB units. We obtained similar distributions for all the polymer lengths considered here by accumulating all events for each length. By choosing lengths that match experimental data, we compared the corresponding experimental duration histograms (see Figure 1c in Arnold Storm and his colleagues' article[20]) to the theoretical ones. This comparison sets the LB time step to Δt, 8 nsec.

Figure 4 shows the time distributions for representative DNA lengths simulated here. In this figure, we use physical units according to the mapping described earlier for direct comparison to similar experimental data.[20] The MD time step for M = 5 will then be Δt_{MD}, 1.6 nsec, indicating that the MD timescale related to the coarse-grained model that handles the DNA molecules is significantly stretched over the physical process. An exact match to all the experimental parameters is, of course, not feasible with coarse-grained simulations. Nevertheless, essential features of DNA translocation are reasonably well reproduced, so we can use the current approach to model similar biophysical processes that involve biopolymers in solution. We can make this process more efficient by a further fine-tuning of the parameters used in the multiscale model.

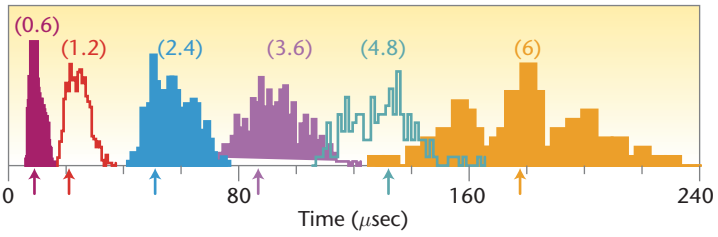

Figure 4. Histograms. Calculated translocation times for various events and DNA lengths are shown in parentheses in kbp (10^3 bp). The arrows point to the most probable time for each polymer size.

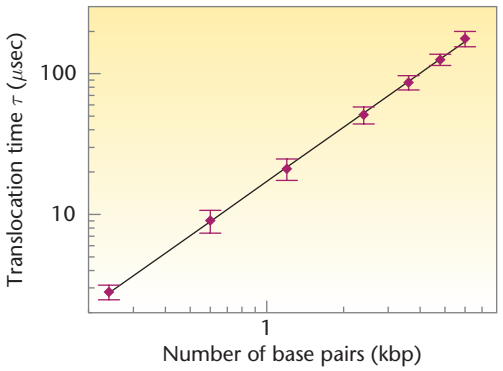

Figure 5. Scaling law of the translocation time with the DNA length. The points correspond to the arrows in Figure 4 and lead to an exponent equal to 1.27 ± 0.03.

The variety of initial polymer realizations produce a scaling-law dependence of the translocation times on length.[15,20] The duration histograms aren't simple Gaussians, but are skewed toward longer times. Accordingly, we use the most probable time (peak of the distribution shown by the arrow in Figure 3a) as the representative translocation time for every distribution. This is also the definition of the translocation time in experiments, to which we compare our results. Calculating the most probable times for each length leads to the nonlinear relation between the translocation time τ and the number of beads N: $\tau(N) \propto N^{\alpha}$, with an exponent α ~ 1.28 ± 0.01. The scaling law reported in Figure 5 is in good agreement with a recent experimental study of double-stranded

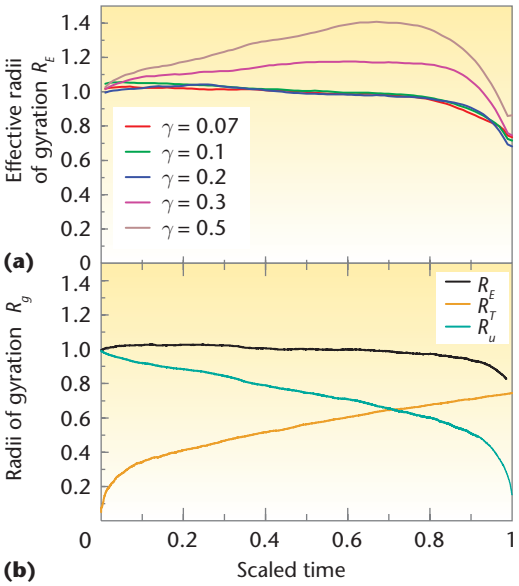

Figure 6. Translocation process dynamics. (b) Radii of gyration with time for a polymer representing 4.8 kbp, and (a) effective radii of gyration for different values of γ for a 2.4 kbp polymer.

DNA translocation, which reported α ~1.27 ± 0.03.[20] In the absence of a solvent, the exponent rises to 1.36 ± 0.03. Such a difference indicates a significant acceleration of the process due to hydrodynamic interactions.

Dynamics of the Translocation Process

Our simulations confirm that the polymer moves through the pore in the form of two almost compact blobs on either side of the wall. One of the blobs (the untranslocated part, denoted by U) is contracting and the other (the translocated part, denoted by T) is expanding. This behavior is visible in Figure 2 for a random event and holds throughout the process, apart from the end points (initiation and completion of the translocation). We assign a radius of gyration $R_I(t)$ (with $I = U, T$) to each of these blobs, following a static scaling law with the number of beads N_I: $R_I(t) \sim N_I^{\mu}(t)$ with $\mu \sim 0.6$ being the Flory exponent for a 3D self-avoiding random walk. Based on the conservation of polymer length, $N_U + N_T = N$, we can define an effective translocation radius as

$$R_E(t) = \left[R_T^{1/\mu}(t) + R_U^{1/\mu}(t)\right]^{\mu}, \quad (12)$$

which should be constant when the static scaling applies. We deduce from our simulations that $R_E(t)$ is approximately constant for all times throughout the process except near the end points. There, the polymer can no longer be represented as two uncorrelated compact blobs, so the static scaling no longer holds.

Figure 6b represents the time evolution of all radii as averages over hundreds of events for a specific polymer length. We scale the time shown so that $t = 1$ denotes an event's total translocation time; we also scale the radii of gyration with respect to $R_U(t = 0)$ for each case. By definition, $R_U(t)$ vanishes at $t = 1$, while R_T increases monotonically from $t = 0$ up to $t = 1$, although it never reaches the value $R_U(t = 0)$; we'll elaborate on this later.

Checking the static scaling's validity with respect to the hydrodynamic field's strength is essential. To this end, we further explore the effective radii of gyration in relation to the parameter γ. We fixed the length to $N = 200$ beads and generated roughly 100 different initial configurations for each value of γ. We present the variation of R_E with this coefficient in Figure 6a, where R_E is clearly almost constant for small γ, but as γ increases, the radii are no longer constant not only at the end points but also throughout the translocation. We interpret large parameter γ values as a strong molecule–fluid coupling. The fluid's influence on the beads, experienced by the back reaction, is large and suppresses the polymer fluctuations such that the translocating biopolymer can no longer be represented as a pair of compact blobs, and the static scaling no longer holds.

Inspection of all the biopolymers at the end of the event reveals that they become more compact after their passage through the pore. This is quantitatively checked through the values of the radii of gyration, which is considerably smaller at the end than it was initially: $R_T(t = 1) < R_U(t = 0)$. The fact that as the polymer passes through the pore it becomes more compact than it was at the event's initial stage might be related to incomplete relaxation, but this remains to be investigated. In Figure 6a, all effective radii of gyration at the translocation's final stage decrease to a value smaller than the initial $R_U(t = 0)$. The ratio $R_T(t = 1)/R_U(t = 0)$ is always smaller than 1 and ranges from 0.72 for $\gamma = 0.1$ to 0.90 for $\gamma = 0.5$.

Throughout its motion, the polymer continuously interacts with the fluid environment. The forces that essentially control the process are the electric drive $\mathbf{F}_{drive,i}$ and the hydrodynamic drag $\mathbf{F}_{drag,i}$ which act on each bead. However, at the end points (initiation and completion of the passage through the pore), entropic forces become important. The fluctuations experienced by the

polymer because of the fluid's presence are correlated to these entropic forces, which, at least close to equilibrium, can be expressed as the gradient of the free energy with respect to the fraction of translocated beads. At the final stage of a translocation event, the radius of the untranslocated part undergoes a visible deceleration (see Figure 6b), and most of the beads have already translocated. It is, therefore, entropically more favorable to complete the passage through the hole rather than revert it—that is, the entropic forces cooperate with the electric field, and the translocation is accelerated.

The entropic forces can also lead to rare events, such as retraction, which occur in our simulations at a rate of less than 2 percent and depend on length, initial polymer configuration, and parameter set. A retraction event is related to a polymer that antitranslocates after having partially passed through the pore. We've visually inspected the retraction events and associate them with the translocated part entering a low-entropy configuration (hairpin-like) subject to a strong entropic pull-back force from the untranslocated part. The polymer's translocated part assumes an elongated conformation, which leads to an increase of the entropic force from the coiled, untranslocated part of the chain. As a result, the translocation is delayed, and the polymer is eventually retracted.

The fact that the entropic forces are related to the number of translocated monomers $N_T(t)$ led us to investigate this quantity's time evolution in more detail. We plotted the number of translocated monomers in Figure 7 for various initial configurations of the polymer, with each curve corresponding to a different completed translocation event. The translocation for a given polymer proceeds along a curve closely related to its initial configuration and its interactions with the fluid. Each polymer follows a distinct trajectory as indicated by the variety of curves. We can't predict the ensuing behavior from the initial polymer configuration in a simple and unique manner.

Energetics of the Translocation Process

As a final step, we study the work performed on the biopolymer throughout its translocation. On general grounds, we expect hydrodynamic interactions to minimize frictional effects and form a cooperative background that assists the polymer's passage through the pore. We investigate the hydrodynamic field's level of cooperation through the synergy factor $S_H(t)$, defined as the work made by the fluid on the polymer per unit time:

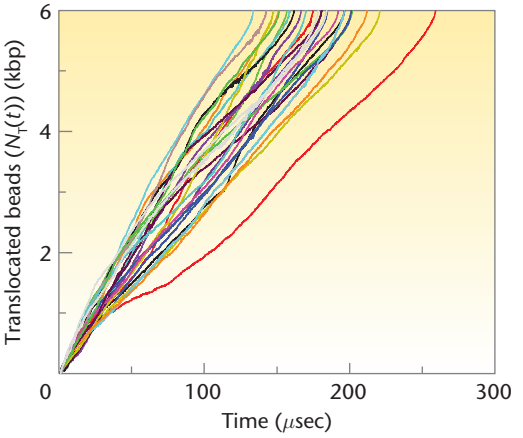

Figure 7. Time evolution of the number of translocated beads $N_T(t)$ for DNA with 6 kbp. Curves correspond to translocation events for various initial configurations.

$$S_H(t) = \frac{dW_H}{dt} = \gamma \left\langle \sum_i^N \mathbf{v}_i(t) \cdot \mathbf{u}_i(t) \right\rangle , \qquad (13)$$

where W_H is the fluid's work on the polymer. Through this definition, positive values of this hydrodynamic work rate indicate a cooperative effect of the solvent, while negative values indicate a competitive effect by the solvent. We can also easily obtain the work done per time step by the electric field ($S_E(t)$) on the polymer through the expression:

$$S_E(t) = \frac{dW_E}{dt} = \left\langle \sum_i^N \mathbf{F}_{drive,i} \cdot \mathbf{v}_i(t) \right\rangle , \qquad (14)$$

where W_E is the electric drive's work on the polymer. The brackets in Equations 13 and 14 denote averages over different realizations of the polymer for the same length. The results for the averages over all realizations are qualitatively similar to the work rates for an individual event of the same length. For all lengths studied here, the total work per time step of the hydrodynamic field ($S_H(t)$) on the whole chain is essentially constant, as Figure 8 shows for an individual event. For the same event, $S_E(t)$ is also constant with time. The figure shows the polymer's kinetic energy K (plotted as γK) for the same event. The kinetic energy is constant with time, as expected, because the temperature in the simulations is held constant, but its fluctuations differ from those of S_E and S_H. We can justify the larger value of K with respect to both S_H and S_E because the bead velocities are larger than the fluid velocity. The hydrodynamic work

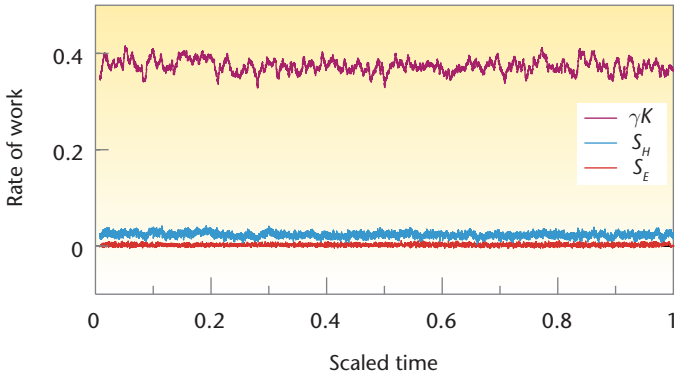

Figure 8. Rates of work (S_H, S_E) and kinetic energy (γK) as functions of time. For one translocation event of a polymer consisting of $N = 300$ beads (3.6 kbp), time is scaled with respect to this event's total time.

per time is also larger than the corresponding electric field work because the latter only acts in the small region around the pore. Moreover, the average of all these quantities over all events for the same polymer length is constant, and shows smaller fluctuations with time than those of any individual event.

In addition to the work rates' variation with time, we analyze their distributions during translocation events. Figure 9 shows that the distribution of $S_H(t)$ lies entirely in the positive range, indicating that hydrodynamics turns the solvent into a cooperative environment (by enhancing the translocation process's speed). The figure also shows the distribution for $S_E(t)$ over all events for the same length. This distribution is mostly positive but has a small negative tail indicating that beads can be found moving against the electric field.

Performance Data

Within our multiscale computational scheme, the total cost of the computation scales roughly to

$$t \sim (t_{LB}V + t_{MD}MN)N_{LB},$$

where

1. t_{LB} is the CPU time required to update a single LB site per time step;
1. t_{MD} is the CPU time to update a single bead per time step (including the overhead of LB–MD coupling);
2. V is the computational domain's volume in lattice units;
3. N is the number of polymer beads;
4. M is the LB–MD time step ratio; and
5. N_{LB} is the number of LB time steps.

Because the LB–MD coupling is local, the forces are short-ranged and the Shake and Rattle algorithms scale linearly with the number of constraints, so t_{MD} is largely independent of N. The LB part scales linearly with the volume occupied by the solvent. Indeed, at constant volume, the simulations' CPU cost scales linearly with the number of beads: the execution times for 50, 100, and 400 beads are 0.433, 0.489, and 0.882 seconds per step, respectively, on a 2-GHz AMD Opteron processor. By excluding hydrodynamics, these numbers become 0.039, 0.075, and 0.318 seconds per step. When we keep the polymer concentration constant, the volume needed to accommodate a polymer of N beads should scale approximately as $N^{1.8}$. A typical translocation event with 500 beads evolves over 30,000 LB steps or 150,000 MD steps. Assuming 250 flops per site per LB step and 2,500 flops per bead per MD step, the previous equation leads to a computational cost of approximately 6 hours. This is comparable with the time observed directly from simulations (approximately 7 hours).

The linear scaling of the CPU time with the molecular size (at constant volume) is the key feature of the LB–MD approach, letting us explore long biomolecules ($N > 1{,}000$) with a relatively modest computational cost. Nevertheless, we must resort to parallel computing, and we expect LB's favorable properties toward parallel implementations to facilitate this task. A parallel implementation will also open the way to simulate systems at least an order of magnitude larger than those considered so far, letting us assign chemical specificity to the biopolymer constituents, rather than the generic nature of the beads that constitute the model polymers in the present study. Work along these lines is in progress.

Acknowledgments

Harvard's Nanoscale Science and Engineering Center, funded by US National Science Foundation award number PHY-0117795, supported Maria Fyta. Simone Melchionna and Sauro Succi thank Harvard's Initiative for Innovative Computing for its hospitality and support.

References

1. D.A. Wolf-Gladrow, *Lattice Gas Cellular Automata and Lattice Boltzmann Models*, Springer Verlag, 2000.
2. R. Benzi, S. Succi, and M. Vergassola, "The Lattice Boltzmann Equation—Theory and Applications," *Physics Reports*, vol. 222, Dec. 1992, pp. 145–197.

3. S. Succi et al., "Applying the Lattice Boltzmann Equation to Multiscale Fluid Problems," *Computing in Science & Eng.*, vol. 3, no. 6, 2001, pp. 26–37.
4. M.G. Fyta et al., "Multiscale Coupling of Molecular Dynamics and Hydrodynamics: Application to DNA Translocation through a Nanopore," *Multiscale Modeling and Simulation*, vol. 5, no. 4, 2006, pp. 1156–1173.
5. P. Ahlrichs and B. Duenweg, "Simulation of a Single Polymer Chain in Solution by Combining Lattice Boltzmann and Molecular Dynamics," *J. Chemical Physics*, vol. 111, no. 17, 1999, pp. 8225–8239.
6. A.J.C. Ladd and R. Verberg. "Lattice-Boltzmann Simulations of Particle-Fluid Suspensions," *J. Statistical Physics*, vol. 104, no. 5-6, 2001, pp. 1191–1251.
7. J.P. Ryckaert, G. Ciccotti, and H.J.C. Berendsen, "Numerical-Integration of Cartesian Equations of Motion of a System with Constraints—Molecular-Dynamics of n-alkanes," *J. Computational Physics*, vol. 23, no. 3, 1977, pp. 327–341.
8. H.C. Andersen, "Rattle—A Velocity Version of the SHAKE Algorithm for Molecular-Dynamics Calculations," *J. Computational Physics*, vol. 52, no. 1, 1983, pp. 24–34.
9. R. Adhikari et al., "Fluctuating Lattice Boltzmann," *Europhysical Letters*, vol. 71, no. 3, 2005, pp. 473–477.
10. S. Ansumali, I.V. Karlin, and H.C. Öttinger, "Minimal Entropic Kinetic Models for Hydrodynamics," *Europhysical Letters*, vol. 63, no. 6, 2003, pp. 798–804.
11. S. Melchionna, "Design of Quasi-Symplectic Propagators for Langevin Dynamics," *J. Chemical Physics*, vol. 127, 2007, article 044108.
12. J.J. Kasianowicz et al., "Characterization of Individual Polynucleotide Molecules Using a Membrane Channel," *Proc. Nat'l Academy of Science*, vol. 93, no. 24, 1996, pp. 13770–13773.
13. J. Li et al., "DNA Molecules and Configurations in a Solid-state Nanopore Microscope," *Nature Materials*, vol. 2, no. 9, 2003, pp. 611–615.
14. S. Matysiak et al., "Dynamics of Polymer Translocation through Nanopores: Theory Meets Experiment," *Physical Rev. Letters*, vol. 96, no. 11, 2006, article 118103.
15. D.K. Lubensky and D.R. Nelson, "Driven Polymer Translocation through a Narrow Pore" *Biophysical J.*, vol. 77, no. 4, 1999, pp. 1824–1838.
16. C. Forey and M. Muthukumar, "Langevin Dynamics Simulations of Genome Packing in Bacteriophage," *Biophysical J.*, vol. 91, no. 1, 2006, pp. 25–41.
17. J.S. Hur, E.S.G. Shaqfeh, and R.G. Larson, "Brownian Dynamics Simulations of Single DNA Molecules in Shear Flow," *J. Rheology*, vol. 44, no. 4, 2000, pp. 713–742.
18. H. Yamakawa, *Modern Theory of Polymer Solutions*, Harper & Row, 1971.
19. P.J. Hagerman, "Flexibility of DNA," *Ann. Rev. Biophysics and Biophysical Chemistry*, vol. 17, 1988, pp. 265–286.
20. A.J. Storm et al., "Fast DNA Translocation through a Solid-State Nanopore," *Nanoletters*, vol. 5, no. 7, 2005, pp. 1193–1197.

Maria Fyta *is a postdoctoral researcher in Harvard University's physics department. Her current research focuses on the multiscale modeling of DNA-related biophysical processes. Fyta has a PhD in physics from the University of Crete in Greece. Contact her at mfyta@physics.harvard.edu.*

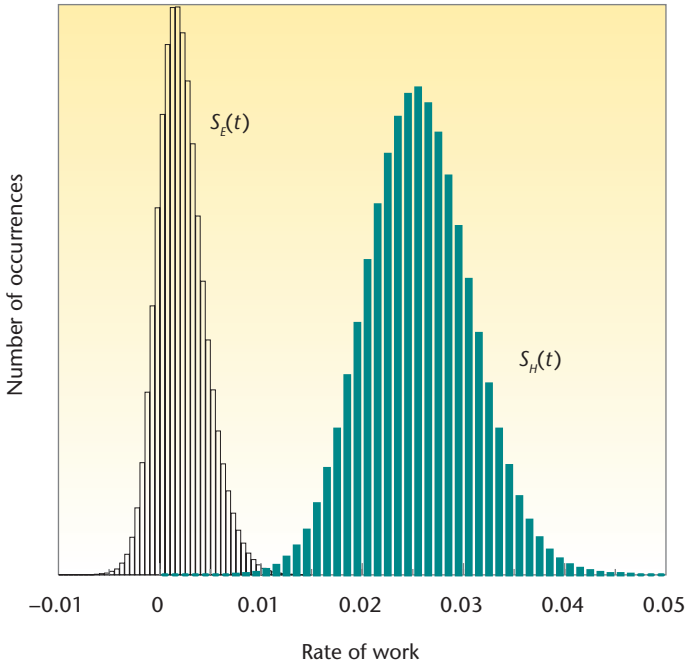

Figure 9. Distribution of the translocation work per unit time averaged over all events for a molecule with 3.6 kbp. The contributions from both the hydrodynamic and electric field are shown ($S_H(t)$ and $S_E(t)$, respectively).

Simone Melchionna *is a researcher at the National Research Council's Institute for the Physics of Matter. His research interests cover complex and biological systems, confined fluids, proteins, and DNA, investigated via computational methods. Melchionna has a PhD in chemistry from the University of Rome–La Sapienza. Contact him at simone.melchionna@roma1.infn.it.*

Efthimios Kaxiras *is a professor of physics and the Gordon McKay Professor of Applied Physics in the Physics Department and the School of Engineering and Applied Sciences at Harvard University. His research interests include the physics of solids, surfaces, clusters, and macromolecules, studied by first-principles electronic structure methods. Kaxiras has a PhD in physics from MIT. Contact him at kaxiras@physics.harvard.edu.*

Sauro Succi *is research director of the National Research Council's Institute of Applied Computing in Rome. His scientific interests cover a range of topics associated with the mathematical and computational modeling of complex system dynamics. Succi has a PhD in plasma physics from the Swiss Polytechnic Institute of Lausanne, Switzerland. Contact him at succi@iac.rm.cnr.it.*