A Framework for Modeling DNA based Molecular Systems*

Sudheer Sahu, Bei Wang, and John H. Reif

Department of Computer Science, Duke University
Box 90129, Durham, NC 27708-0129, USA.
{sudheer,beiwang,reif}@cs.duke.edu

Abstract. Recent successes in building large scale DNA nanostructures and in constructing DNA nanomechanical devices have inspired scientists to design more complex nanoscale systems. The design process can be made considerably more efficient and robust with the help of simulators that can model such systems accurately prior to their experimental implementation. In this paper, we propose a framework for a discrete event simulator for simulating the DNA based nanorobotical systems. It has two major components: a physical model and a kinetic model. The physical model captures the conformational changes of molecules, molecular motions and molecular collisions. The kinetic model governs the modeling of various chemical reactions in a DNA nanorobotical systems including the hybridization, dehybridization and strand displacement. The feasibility of such a framework is demonstrated by some preliminary implementations.

1 Introduction and related work

1.1 Motivation

Recent research has explored DNA as a material for self-assembly of nanoscale objects [21, 54, 65, 85, 106, 110, 111], for performing computation [1, 12, 10, 11, 62, 61, 64, 104, 105, 107], and for the construction of nanomechanical devices [2, 22, 23, 30, 60, 66, 96, 80, 86–89, 99, 98, 112, 116, 117]. One potential application of an autonomous unidirectional DNA device is to perform computation. Recently Yin et al [115, 114] proposed the design of an autonomous universal turing machine and cellular automata. Another potential application beyond computation is the design of a controllable moving device integrated into a DNA lattice for efficient transportation. The major challenges in front of the researchers interested in designing complex DNA based nanodevices, are the time consuming and costly experiments. A lot of times the effect of alterations in only a few parameters need to be tested, and the entire set of experiments need to be repeated from scratch. Accurate computer simulations that capture the essential physical and chemical properties can serve as an effective tool in the design process.

* A preliminary version of this paper was published in DNA12[82]

1.2 Prior Simulators for DNA Computing

Previous simulators for DNA computing include:

- *VNA simulator*[73, 74]: a simulator to aid the protocols for DNA computing. The simulator consists of two main parts, one for finding reactions among existing molecules and generating new ones, and the other for numerically solving differential equations to calculate the concentration of each molecule.
- *Virtual test tubes*[35–37]: a simulator for biochemical reactions based on the kinetics of molecular interactions.
  Neither of these deal with the shapes of nanostructure, and therefore not suitable for simulation of nanorobotics or nanofabrication applications.
- *Hybrisim*[44]: a simulator that deals with the detailed simulation of hybridization only between two strands, and therefore, very limited in use.

Sales-Pardo et. al. [83] modeled a ssDNA as a bead-pin rotational polymer chain and used a modified Monte Carlo simulation to investigate the dynamics of a single-stranded DNA and its associated hybridization events. The geometric constraints of the nucleic chain were handled by a lattice model.

Isambert and Siggia [45] modeled RNA helices as rods and single stranded RNA as Gaussian chains. Kinetic Monte Carlo method was used to sample RNA conformational changes. They also used the short-scale and the large-scale conformation descriptors, i.e. *nets* and *crosslinked gel*, to model geometric constraints related to complex RNA folding conformations.

Bois et. al. [15] investigated the possible effects of topological constraints in DNA hybridization kinetics. Recently Dirks et. al. [28] developed an algorithm aimed at analyzing the thermodynamics of unpseudo-knotted multiple interacting DNA strands in a dilute solution.

1.3 Our Results and Organization of this Paper

In this paper, we describe a comprehensive framework for simulation of DNA based nanorobotic devices.

Our method of simulation is different from the commonly used Gillespi algorithm [38, 49, 39, 100, 34, 79]. In the Gillespi algorithm the concentrations of various reactants are stored as $X_1, \ldots X_n$. And also the rate constants of various possible chemical reactions are also stored as $c_1, \ldots c_m$. Then the calculation of rates of various reactions gives the probabilities for various reactions. The appropriate reaction R_μ is then chosen probabilistically, and after the execution of this reaction the concentrations $X'_1, \ldots X'_n$ of various chemicals are updated appropriately. The algorithm is computationally expensive: more so, in the systems

of our interest where the number of macromolecular interactions are too large. For example, the potentially huge number of possible products from two different DNA single stranded molecules depending upon their alignment with each other, and each product formation will have a different rate constant. Moreover, in nanorobotical and nanofabrication applications, the shapes of various nanostructures involved are as important as the concentrations of the reactants and the reaction rates. Therefore, physical simulations are performed to model the molecular conformations and the chemical reactions are monitored explicitly.

In this paper, we describe a framework for the design of a discrete event simulator, which simulates DNA based nanorobotical devices. Section 2 gives an overview of the system. Section 3 describes the physical simulation of the molecules. Section 4 discusses the event simulation based on the kinetic and thermodynamic studies. Section 5 describes the adaptive time-steps to optimize the physical simulation, and Section 6 describes the analysis of the complete algorithm. Section 7 presents some preliminary results to support such a framework. Discussions and future work is described in Section 8. It should be noted that in this paper, we present the framework for building such a simulator and not the simulator itself. In the subsequent text any reference to simulator is a reference to this framework.

2 Our Discrete Event Simulation

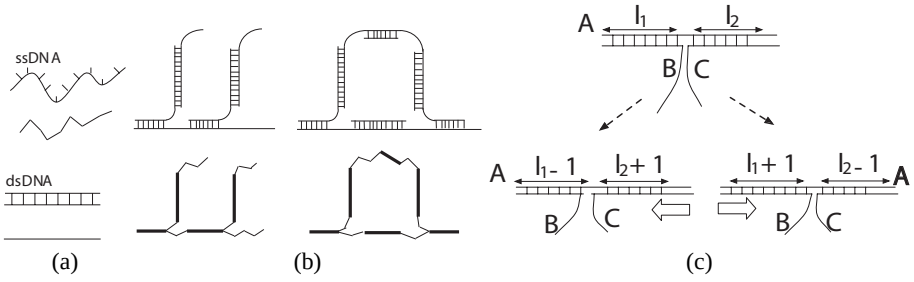

Fig. 1. (a) Schematic view of the molecules in the modeled system. Bold solid lines represent the WLC model used for dsDNA segments while thin solid lines represent the WLC model used for ssDNA segments. (b) Figure shows a complex DNA nanostructure reduced to a collection of WLC segments with different parameters (bold solid line for dsDNA and thin line for ssDNA segments).(c) Strand displacement: molecule B and C compete against each other to hybridize with molecule A.

The simulator performs the molecular-level simulations and provides an useful tool to study DNA based nanomechanical devices. It has two major components. The first component is the physical simulation of the molecule conformations. The second component is the event simulation (hybridization, dehybridization and strand displacement events) which depends on the kinetic and thermodynamic properties of the molecules. Due to the large number of molecules in a given solution, we sample and simulate molecules within a small test volume, assuming the solution is uniform.

The modeled system consists of three types of molecules, single-stranded DNA (ssDNA) molecules, double-stranded DNA (dsDNA) molecules and complex DNA nanostructures with both single-stranded and double stranded segments, as shown in Figure 1.

For the sake of simplicity, we assume no pseudo-knots formation for the complex DNA nanostructures. Therefore, to a first approximation, the complex DNA nanostructure is reducible to a collection of WLC segments with different parameters (i.e. persistence length, elasticity, diffusion coefficients etc). For more complicated DNA nanostructures, we can adapt the geometric descriptors used in [45, 15], as discussed in Section 8.

The secondary structure of a nanostructure can be represented as an undirected graph called connectivity graph. Individual strands are represented as nodes, and hybridization relationships between strands are represented by edges between corresponding nodes. There is an edge between nodes representing two strands, if and only if the two strands hybridize with each other with any alignment. An appropriate data structure for this will be an adjacency list. However, with every node (i.e. double stranded region of any strand) the information about its neighboring unhybridized region also needs to be stored. As shown in Figure 2, this data structure stores individual molecular configurations including sequence and secondary structure.

During the simulation, three types of reaction take place in the solution: the hybridization between a pair of ssDNA segments with complementary base-pairing, the dehybridization of the dsDNA portion of a nanostructure and the strand displacement. The DNA molecules contain potential hybridization sites at their free-ends (sticky ends). During the simulation, when two molecule come into contact (reactive collision), a potential hybridization event is reported. The corresponding free-end base-pairs are investigated to determine the probability of its actual occurrence. Strand displacement is a reaction in which two strands compete against each other to hybridize with a common strand as shown in Figure 1 (c). Strand B

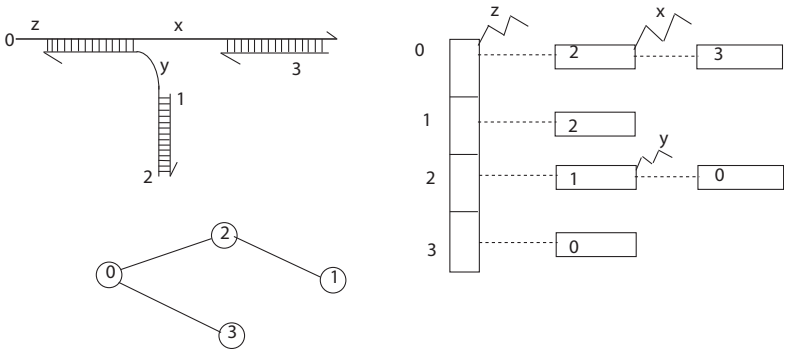

Fig. 2. Suggested data structure for modeling the DNA based complex nanostructures, and connectivity graph for the strands in the nanostructure. It should be noted that the information about the unhybridized sections of the strands is stored at the nodes that represent the neighboring duplex portions as shown in the Figure.

and C compete against each other to hybridize with strand A. At a time instance, B (or C) makes one more bond with A and removes one bond of C (or B).

The required discrete event simulation with Δt as the time-interval is described as follows. Algorithm 1 describes the major steps of the simulation. MQ stores all the nanostructures in the system. T is the total simulation time. Δt is the simulation time per step. *Initialize* is a function that initializes the MQ based on the user input. The detailed algorithms are described in the subsequent Sections.

Algorithm 2 describes steps involved in generating random conformations for all molecules in the system. *Enqueue* and *Dequeue* are standard queuing operations that insert and delete an element in the queue. *MCSimulation(m)* generates new conformations for the molecule m.

Algorithm 1 Discrete Event Simulation

Initialize(MQ)
while $t \leq \mathsf{T}$ **do**
  $t = t + \Delta t$
  {PHYSICAL SIMULATION}
  Physical simulation
  Collision detection
  {EVENT SIMULATION}
  Hybridization
  Dehybridization
  Strand displacement
end while

Algorithm 2 Physical Simulation

for $\forall m_i \in MQ$ **do**
  MCSimulation(m_i)
end for

Algorithm 3 MCSimulation (m)

m^* =RandomConformation(m)
if SelfCollision(m^*) **then**
  continue to next iteration
end if
$\Delta E = E(m^*) - E(m)$
if ($\Delta E > 0$) **then**
  $x \in_{var} [0, 1]$
  if ($x > \exp -\frac{\Delta E}{K_B T}$) **then**
    continue to next iteration
  end if
end if
$m = m^*$

Algorithm 4 Collision Detection

for $\forall m_i, m_j \in MQ, i \neq j$ **do**
  if collide(m_i, m_j) **then**
    e =HEvent(m_i, m_j)
    Enqueue(HQ, e)
  end if
end for

Algorithm 5 Hybridization

while HQ is NOT empty **do**
  e = Dequeue(HQ)
  Hybridize(e)
  Update(MQ)
  if PotentialSD(e) **then**
    Enqueue(SDQ, e)
  end if
end while

Algorithm 6 Dehybridization

for $\forall m_i \in MQ$ **do**
  for $\forall b \in$ bonds of m_i **do**
    if PotentialDehybridization(b) **then**
      Dehybridize(b)
    end if
  end for
  if any dehybridization performed **then**
    DFS on connectivity graph of new m_i, each connected component is a new molecule formed.
    Update(MQ)
  end if
end for

Algorithm 7 Strand Displacement

while SDQ is NOT empty **do**
  e = Dequeue(SDQ)
  e^* = StrandDisplacement($e, \Delta t$)
  if IncompleteSD(e^*) **then**
    Enqueue(SDQ^*, e^*)
  end if
  Update(MQ, e^*)
end while
$SDQ = SDQ^*$

The *MCSimulation(m)* function, described as Algorithm 3, is based on the Metropolis algorithm[41, 71]. $E(m)$ is the energy associated with conformation of molecule m. ΔE, defined as the energy change of the system due to the transition to new conformation, determines the probability that the molecule achieves the new conformation.

Algorithm 4 describes reactive collision detection which leads to potential hybridization events. *Collide(*m_i, m_j*)* returns true if the sticky ends of molecule m_i and m_j collide. e is a data structure that stores an event (hybridization, dehybridization or strand displacement), including all the molecular configurations involved in the event and supplementary information related to the event. For example, in the case of hybridization, it stores the molecular configurations and the information of the hybridization sites. *HEvent(*m_i, m_j*)* creates a potential hybridization event based on a collision between molecules m_i and m_j. HQ stores all potential hybridization events.

Algorithm 5 presents the algorithm involved in hybridization. *Hybridize(e)* probabilistically determines the hybridization product based on the change in free energy as described in Section 4. *PotentialSD(e)* returns true if event e is a potential strand-displacement event. SDQ stores all potential strand-displacement events. When two nanostructures hybridize to form a larger nanostructure, their corresponding connectivity graphs are merged together to form the connectivity graph representing the newly formed nanostructure. *Update(MQ)* updates the configurations of the molecule in the system based on the occurred event e.

Algorithm 6 describes the dehybridization event. *PotentialDehybridization(m)* returns true if molecule m could potentially dehybridize. *Dehybridization(m)* probabilistically dehybridizes molecule m. The corresponding edges are deleted from the connectivity graph representing the nanostructure. Each connected component, thus formed, represents one newly formed nanostructure. *Update(MQ)* updates the configurations of the molecules in the system based on the event e that occurred.

Algorithm 7 shows the steps involved in the strand displacement event. *Strand-Displacement*($e, \Delta t$) probabilistically proceeds with the strand displacement event e within time frame Δt. *IncompleteSD(e)* returns true if the strand displacement event has not completed within the given time frame.

3 Our Physical simulation

The discrete worm-like chain model (WLC) is used to model the polymer-like DNA molecules in solution. Monte Carlo (MC) computer simulations are used to determine their conformations.

3.1 The Discrete Wormlike Chain Model for DNA

The advancement of single molecule dynamics offers experimental validations of various DNA polymer models, among which Gaussian Chain Model, Freely-Jointed Chain (FJC) and Worm-Like Chain (WLC) are widely investigated [77, 52, 46, 91, 75, 31, 51, 5, 109, 90, 55, 16, 17, 53]. The choice of a polymer model depends on the physical property of the DNA chain, affordable computation and molecular-details of interest [29]. Our simulation is constructed using the discrete

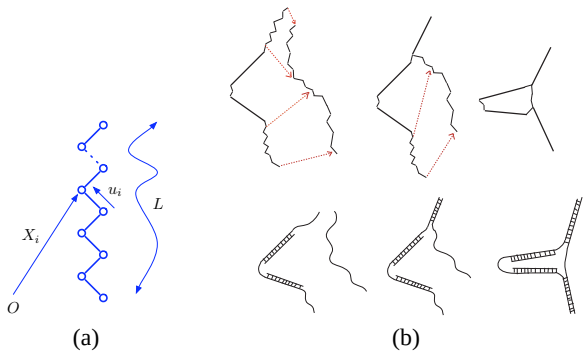

Fig. 3. (a) WLC model (b) Collision detection in 3D (c) Figure illustrates various steps wrt the physical motion of the strands during hybridization.

wormlike chain model. Marko and Siggia [67, 68] used the model to derive the elastic theory suitable for DNA and further completed the model to include bending and twisting elasticity of DNA and the free energy required for deformation. Bustamante et al [17] proposed an interpolation of the Marko-Siggia model for fitting and experimental elasticity curve of single DNA molecules. Klenin et al [50] modeled linear and circular DNA where the DNA polymers are represented by a WLC of stiff segments connected by bending torsion and stretching potentials. Tinnoco et al [97] used WLC as their polymer chain conformation to investigate force effect on thermodynamics and kinetics of single molecule reaction. Larson et al [58, 27] used a similar model to predict the behavior of tethered dsDNA in a constant-velocity flow. Experimental data has shown some reasonably good agreement with the model [72].

The DNA molecule (Figure 3 (a)) is initialized as $N + 1$ beads $(0, 1..N)$ connected by N mass-less extendable segments (springs) of the same length [27, 33, 59]. The contour length of the chain is L. The position of the bead i is denoted

as $\mathbf{x}_i$. The segment vectors are given by

$$\mathbf{u}_i = \mathbf{x}_i - \mathbf{x}_{i-1} \tag{1}$$

Therefore the chain is represented by a set of $N+1$ vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ [20]. We use WLC to model ssDNA, dsDNA and complex DNA nanostructures. Specifically for a complex DNA nanostructure, different parameters as described in Section 3.5 are applied to different segments of the chain depending on whether the segment is double-stranded or single-stranded.

3.2 Monte Carlo Simulation

The molecules are simulated through Monte Carlo simulation for a desired number of time steps using Algorithm 3. According to the Metropolis algorithm used in the simulation, $E(m)$ is the energy associated with conformation of molecule m. The computation of $E(m)$ will be discussed in Section 3.4. ΔE is defined as the energy change of the system due to the new conformation. K_B is the Boltzman constant, and T is the absolute temperature. MQ is the set of all molecules in the simulation. *RandomConformation* is a function that achieves a new conformation of the molecule through random walk in three dimension. *SelfCollision* detects and excludes the self-crossing conformations. The detailed algorithm is shown in Algorithm 3. Similar methods have been used in [118, 7, 63].

3.3 Random Conformation

The random conformation of the DNA molecule is generated by a random walk in three dimensions. Based on [6],

$$\Delta\mathbf{x}_i = \mathbf{R}_i \tag{2}$$

where $\Delta\mathbf{x}_i$ is the change of $\mathbf{x}_i$ in time step Δt, and $\mathbf{R}_i$ is the random displacement. Let D be the diffusion coefficient. We assume $\mathbf{R}_i$ as a Gaussian random variable which is distributed according to

$$W(\mathbf{R}_i) = (4A\pi)^{-3/2} \exp(-\mathbf{R}_i/4A) \tag{3}$$

where $A = D\Delta t$. The diffusion coefficient D of a macromolecule in an ideal dilute solution is computed according to $D = K_B T/f$, where f is the hydrodynamic frictional coefficient of the macromolecule [93]. For a rigid, rod-like molecule f can be written as $f = 3\pi\eta L/(\ln\rho + \gamma)$, where η is the viscosity of the solution, L is the length of the DNA molecule, ρ is the axial ratio and γ is a correction for end effects [93].

3.4 Energy

Now we describe how we calculate $E(m)$ as stated in Algorithm 3. Our current simplified model neglects the following energies (though a more accurate model should take them into consideration [118, 26]): pairing potential between complementary bases, stacking energy from the vertical interactions between neighboring base pairs and hydrodynamic interaction energy with the solvent. We shall consider the torsional rigidity in the forms of bending torque and twisting torque for the DNA molecules in a more sophisticated model. The total energy of a DNA conformation is given as the sum of stretching, bending, twisting and electrostatic interaction energy among negatively charged phosphate groups along the chain [50, 118, 56], which are denoted as E^s, E^b, E^t and E^e, respectively.

$$E^{total} = E^s + E^b + E^t + E^e \tag{4}$$

Stretching Energy. The stretching energy is defined as

$$E^s = \frac{1}{2} Y \sum_{i=1}^{N} (u_i - l_0)^2 \tag{5}$$

where l_0 is the segment equilibrium length, Y is the stiffness parameter defined previously [118].

Bending Energy. The bending elastic energy of the coarse-gained bead-rod model is

$$E^b = -\frac{\kappa}{l_0} \sum_{j=2}^{N} \frac{\mathbf{u}_j \cdot \mathbf{u}_{j-1}}{|\mathbf{u}_j||\mathbf{u}_{j-1}|} \tag{6}$$

where l_0 is the length of the connecting rod and $\mathbf{u}_j/|\mathbf{u}_j|$ is the unit vector directed from bead j-1 to j [29, 101]. The bending rigidity κ is related to the persistence length P by

$$\kappa = K_B T P \tag{7}$$

Twisting Energy. The twisting energy is defined as

$$E^t = \frac{C}{2l_0} \sum_{i=2}^{N-1} (\tau_i)^2 \tag{8}$$

where l_0 is the segment equilibrium length, C is the torsional rigidity constant and τ_i is the twist angle between the $(i-1)$th and ith segments [50]. The computation of the twist energy can be found in [50].

Electrostatic Energy. The DNA intra-chain electrostatic repulsion/attraction can be described by the Debye-Hükel approximation as the electrostatic potential between any two non-adjacent segments i and j [118, 56, 50],

$$E^e_{i,j} = \frac{\nu^2}{D} \int d\lambda_i \int d\lambda_j \frac{\exp\left(-\kappa r_{i,j}\right)}{r_{i,j}} \tag{9}$$

where $r_{i,j}$ is the distance between two charges at arc length parameters $d\lambda_i$ and $d\lambda_j$ along the chain. κ is the inverse of the Debye length and is given as $\kappa = 8\pi e^2 I / K_B T D$, where I is the ionic strength, e is the proton charge, and D is the dielectric constant of water. ν is the linear charge density, which is $\nu = -2e/\Delta$, where Δ is the distance between base pairs [50, 56].

3.5 Parameters

We use the WLC model for both ssDNA and dsDNA for modeling consistency. It is important to notice that there are different sets of parameters used for each of them.

Parameters for ssDNA. Let L be the contour length of the ssDNA, $L = l_{bp} N_{bp} = l_0 N$. Here l_{bp} is the length of the ssDNA per base pair. N_{bp} is the number of bases. N is the number of beads (monomer) in our WLC model. l_0 is the length per segment. The average length of ssDNA in the system is approximately $25 - 30\ bp$. According to [113], $l_{bp} = 0.7\ nm$. Many groups have obtained the force/extension data for ssDNA in different salt environments[118, 91, 81, 8, 18]. Parameters used in our model are obtained from [118], where $l_0 = 1.5\ nm$ and $Y = 120\ K_B T / nm^2$. The persistence length P is $0.7\ nm$ [91]. The diffusion coefficient D of ssDNA is obtained from [93] as approximately $1.52 \times 10^{-6}\ cm^2 s^{-1}$ for a $20\ bp$ strand. The diameter of the ssDNA backbone is $1\ nm$ [25].

Parameters for dsDNA. For dsDNA, the parameters associated with the equations are different, i.e. $l_0 = 100\ nm$ [50, 24, 69], $P = 50\ nm$, $Y = 3K_B T/2P$ [24, 94], $l_{bp} = 0.34\ nm$ [113], and $D = 1.07 \times 10^{-6}\ cm^2 s^{-1}$ [93]. For a short dsDNA segment ($20\ bp$), the WLC model can be simplified to the straight, rigid cylinder model with reasonable adequacy [3, 69]. WLC models are used for simulation consistency.

3.6 Motion of the complex nanostructure.

The MC simulation described previously can be applied to a complex nanostructure. Such a nanostructure is reducible to a collection of ssDNA and dsDNA WLC

segments as shown in Figure 1 b). Perturbations to each segment are done independently. The total energy is computed as a summation of the energies associated with individual segments. For a more accurate model, loop energy [14, 9] of DNA strands can also be considered in the DNA nanostructures that contain the loops in the systems of our interest.

3.7 Physical model for hybridization

Though extensive research has been done for RNA folding simulation [32, 108], to the best of our knowledge there is no empirical results that describe:

1. The motions of each individual strand during the hybridizations.
2. The actual physical location of the hybridized products relative to other molecules in the system.

Therefore we make the following hypotheses:

1. Upon collision that leads to potential hybridization, two strands immediately align their bases involved in the formation of duplex with the right orientation.
2. During the hybridization process, the displacement of the two strands is inversely proportional to their masses (or number of bases in the structure).

The model can be subsequently improved as the empirical evidence becomes available. Figure 3 c) illustrates one schematic to depict our hypotheses.

3.8 Discussions of other physical models

To calculate the mesoscale dynamics of the DNA molecules, Brownian Dynamics (BD) simulation techniques can be applied to replace explicit solvent molecules with a stochastic force [29, 56, 50, 47, 42, 19, 92, 57, 43]. To obtain more accurate molecular details of a particular DNA nanostructure, we may improve our MC physical model with the BD simulation. However, BD simulation becomes infeasible as the number of molecules increases per test volume. For descriptions of the BD algorithm and explicit force computations, refer to [4, 50, 70, 56].

4 Event Simulation

In the event simulation module we use thermodynamics and kinetics principles to calculate the probabilities of various events. Possible events in systems of our interest are hybridization, dehybridization (melting/dissociation) and strand displacement.

4.1 Hybridization

The nearest-neighbor (NN) model is used to model the hybridization event[48]. The model assumes that the stability of a given base-pair depends on the identity and orientation of neighboring base pairs [48]. Empirical data is used to determine parameters for all possible alignments of base pairs. The model has been shown to describe the thermodynamics of DNA structures that involve mismatches and neighboring base pairs beyond the Watson-Crick pairs [78, 84].

Let ΔG° be the standard free energy released as heat by a single hybridization event. It can be calculated from the standard enthalpy and entropy of the reaction [48]:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{10}$$

For reactions taking place in commonly used buffers, the standard enthalpy and entropy can be reliably estimated from the oligonucleotide sequence according to a nearest neighbor stacking model [40]:

$$\Delta H^\circ = \Delta H^\circ_{ends} + \Delta H^\circ_{init} + \sum_{k\in\{stacks\}} \Delta H^\circ_k \tag{11}$$

$$\Delta S^\circ = \Delta S^\circ_{ends} + \Delta S^\circ_{init} + \sum_{k\in\{stacks\}} \Delta S^\circ_k \tag{12}$$

A coarser approximation for DNA of length s can be used[105], so that $\Delta H^\circ \sim -8s\ kcal\ mol^{-1}$ and $\Delta S^\circ \sim (-22s-6)\ cal\ mol^{-1}\ K^{-1}$. *BIND* [40], the thermodynamic simulator for DNA hybridization can be used to calculate the ΔH°, ΔS° and ΔG° for the reaction between two DNA molecules.

When a potential hybridization event that involves molecules m_1 and m_2 is detected due to a collision, the simulator examines all possible alignments of m_1 and m_2. For hybridization according to alignment i, its free energy ΔG°_i is computed using the NN model. Let $m_1{m_2}^i$ be its hybridization product. Let p_i be the stability measurement of $m_1{m_2}^i$. Then it is known that $p_i \propto \exp(-\Delta G^\circ_i/RT)$. Let P^i_h be the probability of hybridization with alignment i. For all j such that p_j is below a given threshold, we reset $p_j = 0$, and only retain p_j values above that threshold. The hybridization product $m_1{m_2}^i$ is formed with probability P^i_h, where

$$P^i_h = \frac{p_i}{\sum_j p_j} \tag{13}$$

This is represented by the formation of the connectivity graph of $m_1{m_2}^i$ by joining the individual connectivity graphs of the molecules m_1 and m_2.

4.2 Dehybridization

Let k_f be the forward reaction rate constant for the hybridization and k_r be the reverse reaction rate constant (rate constant for the dehybridization). For very short DNA, the forward reaction has a diffusion-controlled rate-determining step approximately independent of its length and sequence, so

$$k_f = A_f e^{-E_f/RT} \tag{14}$$

where $k_f \approx 6 \times 10^5\ mol^{-1}s^{-1}$ and $A_f = 5 \times 10^8\ mol^{-1}s^{-1}$. The activation energy for the event is $E_f = 4\ kcal\ mol^{-1}$. A more accurate model can consider the effect of the DNA length, the sequence and the salt concentration on k_f [103], which is shown as,

$$k_f = \frac{k'_N \sqrt{L_s}}{N} \tag{15}$$

where L_s is the length of the shortest strand participating in duplex formation, N is the total number of base pairs present in non-repeating sequence, k'_N is the nucleation rate constant. For $0.2 \leq [\text{Na}^+] \leq 4.0$, k'_N is estimated as $\{4.35 \log_{10}[\text{Na}^+] + 3.5\} \times 10^5$.

The reverse reaction rate k_r is very sensitive to the DNA length and sequence:

$$k_r = k_f e^{\Delta G^\circ/RT} \tag{16}$$

where ΔG is the change in free energy during the hybridization (forward reaction).

Consider a molecule m_1m_2 with concentration $[m_1m_2]$. Let k_r be the rate constant for dehybridization of m_1m_2. Assuming that R_r is the reaction rate of dehybridization, we have

$$R_r = k_r[m_1m_2] \tag{17}$$

Thus, the number of molecules dehybridized in time Δt is $R_t \Delta t$. Therefore the probability P_d that the molecule m_1m_2 dehybridizes in Δt can be approximated as

$$P_d = \frac{k_r[m_1m_2]\Delta t}{[m_1m_2]} = k_r \Delta t \tag{18}$$

Thus, the probability of dehybridization of a double stranded section of a nanostructure is dependent on the value of k_r for that section of nanostructure.

For every molecule m_i in the system, the probability of dehybridization is evaluated for each of double stranded sections in it, and the dehybridizations of these sections is carried out probabilistically. In terms of connectivity graphs, it

means the deletion of edges corresponding to the double stranded sections that dehybridized. In case more than one sections were dehybridized, *depth first search* is performed on new connectivity graph of m_i to identify individual connected components that represent the connectivity graphs of the products of dehybridization event on molecule m_i.

4.3 Strand Displacement

Strand displacement is modeled as a random walk in which the direction of migration of the branching point (junction) along the DNA is chosen probabilistically and is independent of its previous movements.

It has been shown that the branch migration and strand displacement is a biased random walk due to mismatches [13]. In other words, migration probability towards the direction with mismatches is substantially decreased. Consider the DNA nanostructure involving molecule A, B and C in Figure 1(c) (top part). We refer to it as molecule ABC, and the nanostructures after 1 base pair left migration and 1 base pair right migration are referred to as $lABC$ and $rABC$, respectively. Let G°_{ABC}, G°_{rABC} and G°_{lABC} be the free energies of the molecules ABC, $lABC$ and $rABC$, respectively. Let $\Delta G^\circ_r = G^\circ_{rABC} - G^\circ_{ABC}$ and $\Delta G^\circ_l = G^\circ_{lABC} - G^\circ_{ABC}$. Let p_r be the probability of the right-directional migration and p_l be the probability of the left-directional migration. It has been shown in [13] that $p_r \propto \exp(-\Delta G^\circ_r/RT)$, similarly $p_l \propto \exp(-\Delta G^\circ_l/RT)$, where the change of free energies can be computed by the NN model[48].

Let τ be the migration (strand displacement) time per base pair. Let N be the number of nucleotide pair migrations during a time frame of Δt. Let κ be the migration rate constant, which is the number of base pair migrated per second, $\kappa = \mathsf{N}/\Delta t$. Therefore $\tau = 1/\kappa$. At 37°, $\kappa = 6 \pm 2\ Kbp\ sec^{-1}$ and $\tau = 170 \pm 50\mu sec$ [95]. The dependence of κ on salt concentration is discussed in [76].

Thus, the strand displacement event can be modeled as a random walk with each time step equal to τ, and probability of migration in either direction calculated as described above.

5 Optimizations in time stepping

The simulation captures various processes that takes place at different time-scales. Ideally, the smallest time unit should be chosen as the time step ($\delta t \sim 1\mu$ sec) to resolve the conformations and trajectory of each individual molecule using the WLC model and MC simulation. But this would make the overall simulation extremely slow. We attempt to overcome the limitations of such a short time-scale

approach. Inspired by ideas in the kinetic Monte Carlo method [102], long-time system dynamics of the system consists of diffusive jumps from state to state. In general, there can be series of simulation steps where no collisions take place between the strands as they remain far apart. In the case a DNA strand is reasonably far apart (say 10-15 times its length) from all other molecules, it is treated as a rigid body, conformational changes within it can be ignored, and only its movements as a rigid body need to be considered.

Another important optimization is to increase the length of the time step itself. If all the strands are far apart, we can guarantee that within a particular time-interval δT, there will not be any collisions. In that case, a large time step δT can be taken by the simulation to evaluate the next state of the system. When the distance reaches a given threshold where the changes in the conformations of strands can no longer being ignored, we change to the original smaller scale time step δt. The implementation of this computational efficient technique of adaptive time step, requires the distance between the closest pair of potential reactive molecules be stored and updated appropriately.

6 Algorithm analysis

We present an average case analysis of the simulation algorithm presented earlier. Consider that the system consists of m nanostructures each consisting of n distinct sections (single-stranded and double-stranded) when decomposed into the WLC model. For a WLC simulation of a nanostructure, since each nanostructure consists of n segments, in every run of the MCSimulation loop, the time taken is $O(n)$. Assume that on an average, the MCSimulation loop needs to run $f(n)$ times before finding a good configuration. Therefore, the time for each step of physical simulation is $O(mnf(n))$. Naive implementation of collision detection takes $O(m^2n^2)$ time. In the event simulation part, assume that the number of collisions detected is c. Since for each collision all the alignments between two reacting strands are tested, if the average length of each single-stranded section in a molecule is l, it takes $O(cl)$. Each double stranded section is tested for a possibility of dehybridization reaction. If the average number of double stranded regions per molecule is b, then it takes $O(bm)$. For every dehybridization event, *DFS* is performed to evaluate new connected components in $O(b^2m)$. Thus, combining the physical and event simulation the total time taken in each step is $O(m^2n^2 + mnf(n) + cl + b^2m)$. The dominating terms are the first two and therefore, it can be reduced to $O(m^2n^2 + mnf(n))$. It can be concluded that the major portion of the time taken by the algorithm is in the physical simulation. Thus, it is important to optimize the time-complexity of the physical simulation

of the molecules in the system as described in previous section. $f(n)$ can be extremely large, in cases where the molecule is stuck in a low-energy conformation. Better collision detection methods can be used to improve the first term. For the strands that are a reasonable distance away from other strands, the strand can be treated as a rigid unit and the term $f(n)$ disappears causing great reduction in time-complexity. In case, all strands are far apart from each other, we may even take the advantage of using larger time steps, and fast forwarding the simulation through the uninteresting states of the system.

7 Preliminary Results

Our preliminary results demonstrate the feasibility of such a framework in modeling DNA based molecular systems.

7.1 Physical Simulation

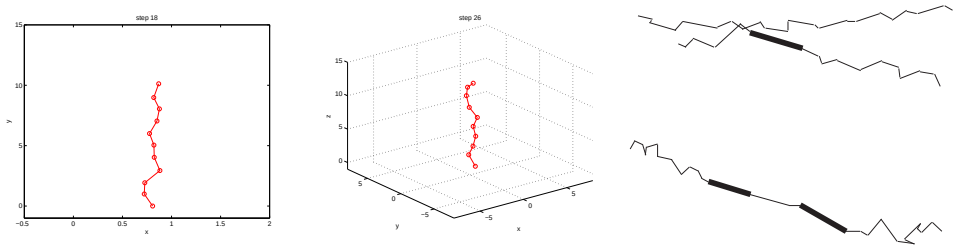

Fig. 4. (a) 2D and 3D snapshots of the simulation for a single tethered DNA (b) Simulation of a hybridization event

The results presented here are obtained using the less computer-intensive Monte Carlo simulation of a discrete WLC model. The physical simulation module is demonstrated through the simulation of a tethered ssDNA. The same module applies to the modeling of other DNA molecules in the system. For demonstration purposes, we neglect twisting energy and focus primarily on the stretching energy and optional bending energy of the tethered DNA. Ideally, relatively long runs are carried out to generate initial conditions for simulations of the tethered-DNA chains, allowing the chains to reach their equilibrium configurations [59]. These configurations are then saved for the actual simulation. The figures shown here

are snapshots of a simulation during different time steps, from both 2D and 3D perspective (Figure 7.1 (a)). The scales for the x-axis and the y-axis are enlarged to show the details of the conformational changes relative to the horizontal plane. The simulations are preliminary but promising.

7.2 Event Simulation

We present here a snapshot of a hybridization event in simulation based on our framework in Figure 7.1 (b). Bold black lines represent the double stranded DNA regions, while the thinner lines are single-stranded. When two molecules come in vicinity of each other (Figure 7.1 (b) top), they combine to form the nanostructure shown (Figure 7.1 (b) bottom). The ssDNA we display in the above snapshots are $20 - 30\ bp$.

8 Discussion and Future work

We presented a comprehensive framework for building a software tool for simulating a DNA based molecular system, and not the actual software tool itself.

We believe that the methods presented here make a good framework for designing the simulator for DNA based molecular systems. We have described how to capture geometric constraints of the molecules with the polymer theory and MC simulation. The preliminary results in the paper support the feasibility of the approach. We also described the approximations and limitations in this framework and the ways of improving them.

It is important to note that, as a framework, the physical simulation component and event simulation component can be decoupled as we improve each component individually.

Various improvements to simulation model can be made to improve the accuracy of the physical simulation:

- To reflect topological constraints by modeling more complicated DNA nanostructures such as pseudo-knots [45, 15].
- To provide more biophysical sound behavior of DNA strands by considering stacking energy and electrostatic energy .
- To achieve the molecular details by replacing the MC simulation with a BD simulation once computational resources are available.
- To validate its correctness against polymer theory and experimental data (for example, the average radius of gyration and the diffusion constant) and to update the physical simulation component to result in more realistic simulation.

A further extension to our framework would be to consider more complicated interactions, i.e. the enzyme restriction event and the hairpin formation. Another extension is to incorporate sequence design capabilities. We would like to design and optimize sequences based on the given nanostructure conformations. Furthermore, a conformation change of a nanodevice can be decomposed into units of local deformations to ease the sequence design.

Acknowledgment

The work is supported by NSF EMT Grants CCF-0523555 and CCF-0432038.

References

1. L. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266:1021–1024, 1994.
2. P. Alberti and J.L. Mergny. DNA duplex-quadruplex exchange as the basis for a nanomolecular machine. *Proc. Natl. Acad. Sci. USA*, 100:1569–1573, 2003.
3. S. A. Allison and S. Mazur. Modeling the free solution electrophoretic mobility of short dna fragments. *Biopolymers.*, 46:359–373., 1998.
4. S.A. Allison and J.A. McCammon. Multistep brownian dynamics: application to short wormlike chains. *Biopolymers*, 23:363–375, 1984.
5. S. R. Aragon and R. Pecora. Dynamics of wormlike chains. *Macromolecules*, 18(10):1868–1875, 1985.
6. R.G.C. Arridge. *An introduction to polymer mechanics*. 1985.
7. G.A. Arteca, T. Edvinsson, and C. Elvingson. Compaction of grafted wormlike chains under variable confinement. *Phys. Chem. Chem. Phys.*, 3:3737–3741, 2001.
8. B. Maier B, D. Bensimon, and V. Croquette. Replication by a single dna polymerase of a stretched single-stranded dna. *Proc. Natl. Acad. Sci. U.S.A.*, 97(22):12002–7, October 2000.
9. W R Bauer, R A Lund, and J H White. Twist and writhe of a dna loop containing intrinsic bends. *Proc Natl Acad Sci U S A*, 90:833–837, 1993.
10. Y. Benenson, R. Adar, T. Paz-Elizur, Z. Livneh, and E. Shapiro. DNA molecule provides a computing machine with both data and fuel. *Proc. Natl. Acad. Sci. USA*, 100:2191–2196, 2003.
11. Y. Benenson, B. Gil, U. Ben-Dor, R. Adar, and E. Shapiro. An autonomous molecular computer for logical control of gene expression. *Nature*, 429:423–429, 2004.
12. Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh, and E. Shapiro. Programmable and autonomous computing machine made of biomolecules. *Nature*, 414:430–434, 2001.
13. I. Biswas, A. Yamamoto, and P. Hsieh. Branch migration through dna sequence heterology. *J. Mol. Bio*, 1998.
14. R.D. Blake and S. G. Delcourt. Loop energy in dna. *Biopolymers*, 26:2009–26, 1987.
15. Justin S. Bois, Suvir Venkataraman, Harry M. T. Choi, Andrew J. Spakowitz, Zhen-Gang Wang, and Niles A. Pierce. Topological constraints in nucleic acid hybridization kinetics. *Nucleic Acids Research*, 33(13):4090–4095, 2005.

16. C. Bouchiat, M.D. Wang, J. Allemand, T. Strick, S. M. Block, and V. Croquette. Estimating the persistence length of a worm-like chain molecules from force-extension measurements. *Biophys. J.*, 76:409–413, January 1999.
17. C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith. Entropic elasticity of lambda-phage dna mechanics. *Science*, 265:1599, 1994.
18. C. Bustamante, S. Smith, J. Liphardt, and D. Smith. Single-molecule studies of dna mechanics. *Current Opinion in Structural Biology*, 10:279–285, 2000.
19. J.E. Butler and E.S.G. Shaqfeh. Brownian dynamics simulations of a flexible polymer chain which includes continuous resistance and multi-body hydrodynamic interaction. *Journal of Chemical Physics*, 122:14901, 2005.
20. G. A. Carri and M. Marucho. Statistical mechanics of worm-like polymers from a new generating function. *J. Chem. Phys.*, 121(12):6064–6077, 2004.
21. N. Chelyapov, Y. Brun, M. Gopalkrishnan, D. Reishus, B. Shaw, and L. Adleman. DNA triangles and self-assembled hexagonal tilings. *J. Am. Chem. Soc.*, 126:13924–13925, 2004.
22. Y. Chen and C. Mao. Putting a brake on an autonomous DNA nanomotor. *J. Am. Chem. Soc.*, 126:8626–8627, 2004.
23. Y. Chen, M. Wang, and C. Mao. An autonomous DNA nanomotor powered by a DNA enzyme. *Angew. Chem. Int. Ed.*, 43:3554–3557, 2004.
24. S. Cocco, J. F. Marko, and R. Monasson. Theoretical models for single-molecule dna and rna experiments: from elasticity to unzipping. *C. R. Physique*, 3:569–584, 2002.
25. C. Desruisseaux, D. Long, G. Drouin, and G. W. Slater. Electrophoresis of composite molecular objects. 1. relation between friction, charge and ionic strength in free solution. *Macromolecules*, 34:44–59, 2001.
26. M. N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette. Stretching single stranded dna, a model polyelectrolyte. *Phys. Rev. Lett.*, 89:248102, 2002.
27. P. Dimitrakopoulos. Stress and configuration relaxation of an initially straight flexible polymer. *J. Fluid Mech.*, 513:265–286, 2004.
28. R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree, and N. A. Pierce. Thermodynamic analysis of interacting nucleic acid strands. *SIAM Rev*, 49 (1):65–88, 2007.
29. P.S. Doyle and P.T. Underhill. Brownian dynamics simulations of polymers and soft matter. In S. Yip, editor, *Handbook of Materials Modeling*, pages 2619–2630. Springer, 2005.
30. L. Feng, S.H. Park, J.H. Reif, and H. Yan. A two-state DNA lattice switched by DNA nanoactuator. *Angew. Chem. Int. Ed.*, 42:4342–4346, 2003.
31. M. Fixman and J. Kovac. Polymer conformation statistics iii: Modified gaussian models of the stiff chains. *J. Chem. Phys.*, 58:1564–1568, 1973.
32. C. Flamm, W. Fontana, I. L. Hofacker, and P. Schuster. Rna folding at elementary step resolution. *RNA*, 6(3):325–38, 2000.
33. J.B. Fournier. Wormlike chain or tense string? a question of resolution. *Continuum Mechanical Thermodynamics*, 14:241, 2002.
34. M.D. Frank-Kamenetskii. Biophysics of dna molecule. *Phys. Rep.*, 288:13 – 60, 1997.
35. M. Garzon, E. Drumwright, R. Deaton, and D. Renault. Virtual test tubes: a new methodology for computing. In *In Proceedings of 7th International Symposium on String Processing and Information Retrieval*, pages 116–121. IEEE Computer Society Pzzress, 2000.
36. M. Garzon and C. Oehmen. Biomolecular computation in virtual test tubes. *Lecture Notes in Computer Science LNCS*, 2340:117–128, 2002.
37. M. H. Garzon, D. R. Blain, and A. J. Neel. Virtual test tubes. *Natural Computing*, 3:461–477, December 2004.

38. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.
39. D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115:1716–1733, 2001.
40. A. J. Hartemink and D. K. Gifford. Thermodynamics simulation of deoxyoligonucleotide hybridization for dna computation. 1997.
41. W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
42. P. J. Heath, J. A. Gebe, S. A. Allison, and J. M. Schurr. Comparison of analytical theory with brownian dynamics simulations for small linear and circular dnas. *Macromolecules*, 29:3583, 1996.
43. J. S. Hur and E. S. G. Shaqfeh. Brownian dynamics simulations of single dna molecule in shear flow. *J. Rheol.*, 44(4):713–742, July-August 2000.
44. N. Ichinose. Hybrisim: Dna hybridization simulator. http://www.genome.ist.i.kyoto-u.ac.jp/ ichinose/bio/hybrisim/.
45. H. Isambert and E. D. Siggia. Modeling rna folding paths with pseudoknots: application to hepatitis delta virus ribozyme. *Proc Natl Acad Sci U S A.*, 97(12):6515–20, 2000.
46. H. M. James and E. Guth. Theory of the elastic properties of rubber. *Journal of Chemical Physics*, 10:455–481, 1943.
47. R. M. Jendrejack, J.J. Pablo, and M. D. Graham. Stochastic simulations of dna in flow: Dynamics and the effects of hydrodynamic interactions. *Journal of Chemical Physics*, 116(17):7752, 2002.
48. J. Santalucia Jr. A unified view of polymer, dumbbell and oligonucleotide dna nearest-neighbor thermodynamics. *PNAS*, 95:1460–1465, 1998.
49. A. M. Kierzek. Stocks: Stochastic kinetic simulations of biochemical systems with gillespie algorithm. *Bioinformatics*, 18:470–481, 2002.
50. K. Klenin, H. Merlitz, and J. Langowski. A brownian dynamics program for the simulation of linear and circular dna and other wormlike chain polyelectrolytes. *Biophys J*, 74(2):780–788, February 1998.
51. J. Kovac and C. Crabb. Modified gaussian model for rubber elasticity. 2. the wormlike chain. *Macromolecules*, 15(2):537, 1982.
52. M. Kuhn and F. Grun. Relationships between elastic constants and stretching double refraction of highly elastic substances. *Kolloid-Z*, 101:294, 1942.
53. S. Kutter. *Elasticity of polymers with internal topological constraints*. PhD thesis, August 2002.
54. T.H. LaBean, H. Yan, J. Kopatsch, F. Liu, E. Winfree, J.H. Reif, and N.C. Seeman. The construction, analysis, ligation and self-assembly of DNA triple crossover complexes. *J. Am. Chem. Soc.*, 122:1848–1860, 2000.
55. B. Ladoux, J. P. Quivy, P. S. Doyle, G. Almouzni, and J. L. Viovy. Direct imaging of single-molecules: from dynamics of a single dna chain to the study of complex dna-protein interactions. *Sci. Prog.*, 84:267, 2001.
56. J. Langowski. Polymer chain models of dna and chromatin. *The European Physical Journal E*, 19:241–249, March 2006.
57. R. G. Larson, H. Hu, D. E. Smith, and S. Chu. Brownian dynamics simulation of a dna molecule in an extensional flow field. *J. Rheol.*, 43(2):267–304, March-April 1999.
58. R.G. Larson, T. Perkins, D. Smith, and S. Chu. Hydrodynamics of a dna molecule in a flow field. *Phys. Rev. E.*, 55:1794–1797, 1997.

59. R.G. Larson, T.T. Perkins, D.E. Smith, and S. Chu. Brownian dynamics simulations of a dna molecule in an extensional flow field. *J. Rheol.*, 43:267, 1999.
60. J. Li and W. Tan. A single DNA molecule nanomotor. *Nano Lett.*, 2:315–318, 2002.
61. D. Liu, M. Wang, Z. Deng, R. Walulu, and C. Mao. Tensegrity: Construction of rigid DNA triangles with flexible four-arm dna junctions. *J. Am. Chem. Soc.*, 126:2324–2325, 2004.
62. Q. Liu, L. Wang, A.G. Frutos, A.E. Condon, R.M. Corn, and L.M. Smith. DNA computing on surfaces. *Nature*, 403:175–179, 2000.
63. A. Malevanets and J. M. Yoemans. Dynamics of short polymer chains in solution. *Europhysics Letters*, 52(2):231, 2000.
64. C. Mao, T.H. LaBean, J.H. Reif, and N.C. Seeman. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature*, 407:493–496, 2000.
65. C. Mao, W. Sun, and N.C. Seeman. Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy. *J. Am. Chem. Soc.*, 121:5437–5443, 1999.
66. C. Mao, W. Sun, Z. Shen, and N.C. Seeman. A DNA nanomechanical device based on the B-Z transition. *Nature*, 397:144–146, 1999.
67. J. Marko and E. D. Siggia. Bending and twisting elasticity of dna. *Macromolecules*, 27:981, 1994.
68. J. F. Marko and E. D. Siggia. Stretching dna. *Macromolecules*, 28:8759, 1995.
69. R. J. Meagher, J. Won, L. C McCormick, S. Nedelcu, M. M. Bertrand, J. L. Bertarm, G. Drouin, A. E. Barron, and G. W. Slaters. End-labeled free-solution electrophoresis of dna. *Electrophoresis*, 26:331–350, 2005.
70. J. Mercier and G. W. Slater. Solid phase dna amplification: a brownian dynamics study of crowding effects. *Biophysical Journal*, 89:32–42, July 2005.
71. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.
72. M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha. Probing single-stranded dna conformation flexibility using fluorescence spectroscopy. *Biophysical Journal*, 86:2530–2537, April 2004.
73. A. Nishikawa, M. Hagiya, and M. Yamamura. Virtual dna simulator and protocol design by ga. In *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO'99*, volume 2, pages 1810–1816, 1999.
74. A. Nishikawa, M. Yamamura, and M. Hagiya. Dna computation simulator based on abstract bases. *Soft Computing*, 5:25–38, 2001.
75. T. Odijk. Stiff chains and filaments under tension. *Macromolecule*, 28:7016–7018, 1995.
76. I.G. Panyutin and P. Hsieh. The kinetics of spontaneous dna branch migration. *Proc Natl Acad Sci U S A.*, 91(6):2021–5, 1994 Mar 15.
77. J.S. Pedersen, M. Laso, and P. Schurtenberger. Monte carlo study of excluded volume effects in wormlike micelles and semiflexible polymers. *Phys Rev E.*, 54(6):5917–5920, December 1996.
78. N. Peyret, P. A. Seneviratne, H. T. Allawi, and J. Santalucia. Nearest-neighbor thermodynamics and nmr of dna sequences with internal aa,cc,gg and tt mismatches. *Biochemistry*, 38:3468, 1999.
79. C. Rao and A. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm,. *J. of Chem. Phys.*, 118:4999–5010, 2003.
80. J.H. Reif. The design of autonomous DNA nanomechanical devices: Walking and rolling DNA. *The 8th International Meeting on DNA Based Computers (DNA 8)*, 2002.

81. M. Rief, H. Clausen-Schaumann, and H. E. Gaub. Sequence-dependent mechanics of single dna molecules. *Nature Structural Biology*, 6:346 – 349, 1999.
82. S. Sahu, B. Wang, and J. H. Reif. A framework for modeling dna based molecular systems. *LNCS*, 4287:250–265, 2006.
83. M. Sales-Pardo, R. Guimera, A. A. Moreira, J. Widom, and L. A. Amaral. Mesoscopic modeling for nucleic acid chain dynamics. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 71:051902, 2005.
84. J. Santalucia and D Hicks. The thermodynamics of dna structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415, 2004.
85. R. Sha, R. Liu, D.P. Millar, and N.C. Seeman. Atomic force microscopy of parallel DNA branched junction arrays. *Chemistry and Biology*, 7:743–751, 2000.
86. W.B. Sherman and N.C. Seeman. A precisely controlled DNA biped walking device. *Nano Lett.*, 4:1203–1207, 2004.
87. J.S. Shin and N.A. Pierce. A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.*, 126:10834–10835, 2004.
88. F.C. Simmel and B. Yurke. Using DNA to construct and power a nanoactuator. *Phys. Rev. E*, 63:041913, 2001.
89. F.C. Simmel and B. Yurke. A DNA-based molecular device switchable between three distinct mechanical states. *Appl. Phys. Lett.*, 80:883–885, 2002.
90. S. B. Smith, L Finzi, and B. Bustamante. Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science*, 258:1122, 1992.
91. S.B. Smith, Y. Cui, and C. Bustamante. Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules. *Science*, 271:795–799, Feb 1996.
92. M. Somasi, B. Khomami, N. J. Woo, J. S. Hur, and E. S. G. Shaqfeh. Brownian dynamics simulations of bead-rod and bead-spring chains: numerical algorithms and coarse-graining issues. *J. Non-Newtonian Fluid Mech.*, 108:227–255, 2002.
93. E. Stellwagen and N. C. Stellwagen. Determining the electrophoretic mobility and translational diffusion coefficients of dna molecules in free solution. *Electrophoresis*, 23(16):2794–2803, 2002.
94. C. Storm and P. C. Nelson. Theory of high-force dna stretching and overstretching. *Physical Review E.*, 67:051906, 2003.
95. B. J. Thompson, M. N. Camien, and R.C.Warner. Kinetics of branch migration in double-stranded dna. *Proc Natl Acad Sci U S A*, 73(7):2299–303, 1976 Jul.
96. Y. Tian, Y. He, Y. Chen, P. Yin, and C. Mao. Molecular devices - a DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. Intl. Ed.*, 44:4355–4358, 2005.
97. I. Tinoco and C. Bustamante. The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem.*, 101-102(1):513–33, 2002.
98. A.J. Turberfield, J.C. Mitchell, B. Yurke, Jr. A.P. Mills, M.I. Blakey, and F.C. Simmel. DNA fuel for free-running nanomachines. *Phys. Rev. Lett.*, 90:118102, 2003.
99. A.J. Turberfield, B. Yurke, and Jr. A.P. Mills. DNA hybridization catalysts and molecular tweezers. *DNA5*, 2000.
100. T. E. Turner, S. Schnell, and K. Burrage. Stochastic approaches for modelling in vivo reactions. *Computational Biology and Chemistry*, 2004.
101. A.V. Vologodskii. Monte carlo simulation of dna topological properties. In *Topology in Molecular Biology*, Biological and Medical Physics, Biomedical Engineering, pages 23–41. Springer Berlin Heidelberg, 2006.

102. A.F. Voter. *Introduction to kinetic monte carlo method*. Springer, NATO publishing unit, 2005.
103. J. G. Wetmur and N. Davidson. Kinetics of renaturation of dna. *J. Mol. Biol.*, 31:349–370, 1968.
104. E. Winfree. Complexity of restricted and unrestricted models of molecular computation. In R. J. Lipton and E.B. Baum, editors, *DNA Based Computers 1*, volume 27 of *DIMACS*, pages 187–198. American Mathematical Society, 1996.
105. E. Winfree. Simulation of computing by self-assembly. Technical Report 1998.22, Caltech, 1998.
106. E. Winfree, F. Liu, L.A. Wenzler, and N.C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–544, 1998.
107. E. Winfree, X. Yang, and N.C. Seeman. Universal computation via self-assembly of DNA: Some theory and experiments. In L.F. Landweber and E.B. Baum, editors, *DNA Based Computers II*, volume 44 of *DIMACS*, pages 191–213. American Mathematical Society, 1999.
108. M. T. Wolfinger, W. A. Svrcek-Seiler, C. Flamm, I. L. Hofacker, and P. F. Stadler. Exact folding dynamics of rna secondary structures. *J.Phys.A: Math.Gen.*, 37:4731–4741, 2004.
109. H. Yamakawa and T. Yoshizaki. Dynamics of helical wormlike chains. i. dynamic model and diffusion equation. *Journal of Chemical Physics*, 75(2):1016–1030, July 1981.
110. H. Yan, T.H. LaBean, L. Feng, and J.H. Reif. Directed nucleation assembly of DNA tile complexes for barcode patterned DNA lattices. *Proc. Natl. Acad. Sci. USA*, 100(14):8103–8108, 2003.
111. H. Yan, S.H. Park, G. Finkelstein, J.H. Reif, and T.H. LaBean. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science*, 301(5641):1882–1884, 2003.
112. H. Yan, X. Zhang, Z. Shen, and N.C. Seeman. A robust DNA mechanical device controlled by hybridization topology. *Nature*, 415:62–65, 2002.
113. J. Yan and J. F. Marko. Localized single-stranded bubble mechanism for cyclization of short double helix dna. *Phys. Rev. Lett.*, 93(10):108108, September 2004.
114. P. Yin, S. Sahu, A.J. Turberfield, and J.H. Reif. Design of autonomous DNA cellular automata. In *Proc. 11th International Meeting on DNA Computing*, pages 376–387, 2005.
115. P. Yin, A.J. Turberfield, S. Sahu, and J.H. Reif. Design of an autonomous DNA nanomechanical device capable of universal computation and universal translational motion. In *Proc. 10th International Meeting on DNA Computing*, pages 344–356, 2004.
116. B. Yurke, A.P. Mills, and A.J. Turberfield. A molecular machine made of and powdered by DNA. *Biophysics*, 78:2629, 2000.
117. B. Yurke, A.J. Turberfield, Jr. A.P. Mills, F.C. Simmel, and J.L. Neumann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406:605–608, 2000.
118. Y. Zhang, H. Zhou, and Z. Ou-Yang. Stretching single-stranded dna: Interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys J.*, 81:1133–1143, August 2001.