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. In this paper, we propose a framework for a discrete event simulator for simulating the DNA based nano-robotical systems. We describe a physical model that captures the conformational changes of the solute molecules. We also present methods to simulate various chemical reactions due to the molecular collisions, including hybridization, dehybridization and strand displacement. The feasibility of such a framework is demonstrated by some preliminary results.

1 Introduction and Related Work

Recent research has explored DNA as a material for self-assembly of nanoscale objects [19,47,58,75,96,100,101], for performing computation [1,11,9,10,55,54,57,94,95,97], and for the construction of nano-mechanical devices [2, 20, 21, 28, 53, 59, 86, 70, 76, 77, 78, 79, 89, 88, 102, 106, 107]. One key application of an autonomous unidirectional DNA device is to perform computation. Recently Yin proposed the design of an autonomous universal turing machine and cellular automata [105, 104]. One potential application beyond computation is the design of a controllable moving device which can be integrated into a DNA lattice for efficient transportation. One major challenge in the design of DNA based devices is the cost and time required for the experiments. Computer simulations can be performed to capture the essential physical and chemical properties, and serve as an effective tool in the design process.

Our method of simulation is different from the commonly used Gillespi algorithm [33,42,34,90,32,69]. In the system of our interests, the geometry of the nano-structures plays an important role apart from the concentrations of the reactants and the reaction rates. Physical simulations are performed to model the molecular conformations and the chemical reactions are monitored explicitly.

Sales-Pardo et. al. 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 [72]. The geometric constraints of the nucleic chain was handled by a lattice model [72]. Isambert and Siggia modeled RNA helices as rods and single stranded RNA as Gaussian chains [38]. Kinetic Monte Carlo method was used to sample RNA conformational changes [38]. They also used the short-scale and the large-scale conformation descriptors, i.e. *nets* and *crosslinked gel*, to model geometric constrains related to complex RNA folding conformations. Bois

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

C. Mao and T. Yokomori (Eds.): DNA12, LNCS 4287, pp. 250–265, 2006.

et. al. investigated the possible effects of topological constraints in DNA hybridization kinetics [13]. Recently Dirks et. al. developed an algorithm aiming at analyzing the thermodynamics of unpseudo-knotted multiple interacting DNA strands in a dilute solution [27].

In this paper, we describe a framework for the design of a discrete event simulator, which simulates DNA based nano-robotical 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 framewok. 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.

2 Discrete Event Simulation

The simulator performs the molecular-level simulations and provides an useful tool to study DNA based nano-mechanical 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 kinetics, thermodynamics and geometry of the molecules. Due to the large number of molecules in a given solution, we sample and simulate molecules within a small cell volume, assuming the solution is well mixed.

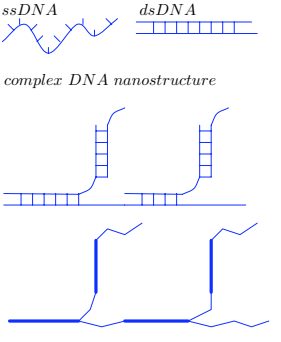

Fig. 1. Schematic view of the molecules in the modeled system. Bold solid lines represent the WLC model used for dsDNA segments while thin solid lines represent the WLC model used for ssDNA segments.

The modeled system consists of three types of molecules, single-stranded DNA (ssDNA), double-stranded DNA (dsDNA) and complex DNA nano-structure with both single-stranded and double stranded segments, as shown in Figure 1. We assume no self-hybridization and no pseudo-knots formation for the complex DNA nano-structures. Therefore, to the first approximation, the complex DNA nano-structure is reducible to a collections of WLC segments with different parameters, i. e. persistence length. For more complicated DNA nano-structures, we can adapt the geometric descriptors used in [38, 13], as discussed in section 8.

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 nano-structure and the strand displacement. The DNA molecule contains potential hybridization sites at its free-end (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 2. Strand B 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. m_i is a data structure that stores individual molecular configurations, including sequence and secondary structure. MQ stores all m_i in the system. T is the total simulation time. Δt is the simulation time per step. $Initialize$ is a function that initialize the MQ based on the user input. The detailed algorithms are described in the subsequent sections.

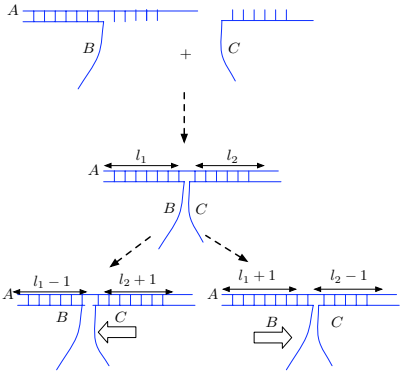

Fig. 2. Strand displacement: molecule B and C compete against each other to hybridize with molecule A

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

Algorithm 3 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 inter-molecule relations. 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 colliding molecule m_i and m_j. HQ stores all potential hybridization events.

Algorithm 4 presents 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. $Update$ (MQ, e) updates the configurations of the molecule in the system based on the occurred event e.

Algorithm 5 describes dehybridization event. $PotentialD(m)$ returns true if molecule m could potentially dehybridize. $Dehybridization(m)$ probabilistically dehybridizes molecule m.

Algorithm 6 shows the steps involved in the strand displacement event. $StrandDisplacement(e, \Delta t)$ 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.

Algorithm 1. Discrete Event Simulation

```
Initialize(MQ)
while t ≤ T do
    t = t + Δt
    {PHYSICAL SIMULATION}
    Physical simulation
    Collision detection
    {EVENT SIMULATION}
    Hybridization
    Dehybridization
    Strand displacement
end while
```

Algorithm 2. Physical Simulation

```
for ∀m_i ∈ MQ do
    MCSimulation(m_i)
end for
```

Algorithm 3. Collision Detection

```
for ∀m_i, m_j ∈ MQ, i ≠ j do
    if collide(m_i, m_j) then
        e =HEvent(m_i, m_j)
        Enqueue(HQ, e)
    end if
end for
```

Algorithm 4. Hybridization

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

Algorithm 5. Dehybridization

```
for ∀m_i ∈ MQ do
    if PotentialD(m_i) then
        e = Dehybridization(m_i)
        Update(MQ, e)
    end if
end for
```

Algorithm 6. Strand Displacement

```
while SDQ is NOT empty do
    e = Dequeue(SDQ)
    e* = StrandDisplacement(e, Δt)
    if IncompleteSD(e*) then
        Enqueue(SDQ*, e*)
    end if
    Update(MQ, e*)
end while
SDQ = SDQ*
```

Algorithm 7. MCSimulation (m), $m \in MQ$

```
m* =RandomConformation(m)
if SelfCollision(m*) then
    continue to next iteration
end if
ΔE = E(m*) − E(m)
if (ΔE > 0) then
    x ∈_var [0, 1]
    if (x > exp −ΔE/(K_B T)) then
        continue to next iteration
    end if
end if
m = m*
```

3 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 Discrete Wormlike Chain Model

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 [67,45,39,81,65,29,44,5, 99, 80, 48, 14, 15, 46]. The choice of a polymer model depends on the physical property of the DNA chain, affordable computation and molecular-details of interest [26].

Our simulation is constructed using the discrete wormlike chain model. Marko and Siggia 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 [60, 61]. Bustamante el al. proposed an interpolation of the Marko-Siggia model for fitting and experimental elasticity curve of single DNA molecules [15]. Klenin et al. modeled linear and circular DNA where the DNA polymers are represented by a WLC of stiff segments connected by bending torsion and stretching potentials [43]. Tinnoco et. al. used WLC as their polymer chain conformation to investigate force effect on thermodynamics and kinetics of single molecule reaction [87]. Larson et al used a similar model to predict the behavior of tethered dsDNA in a constant-velocity flow [51, 25]. Experimental data has shown some reasonably good agreement with the model [64].

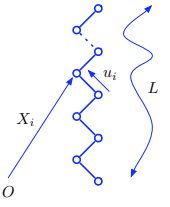

Fig. 3. WLC model

The DNA molecule (Figure 3) is initialized as $N+1$ beads $(0, 1..N)$ connected by N mass-less extendable segments (springs) of the same length [25, 31, 52]. 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 $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ [18]. We use WLC to model ssDNA, dsDNA and complex DNA nanostructure. Specifically for complex DNA nanostructure, different parameters 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 as Algorithm 7. 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 a later section. 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 detail algorithm is shown in Algorithm 7. Similar methods have been used in [108, 7, 56]. To achieve random conformation of the molecules, more permutations can be used such as random rotation of an interval chain or bond-length change, which are described in [108].

3.3 Collision Detection: Cylinder Model

To simulate the motion of the molecule, each segment occupies a finite volume. Instead of using repulsive forces (weak and short-ranged) to maintain the excluded volume of the individual segments of the chain [17], for two disjoint segments, we assume a minimum distance D between them in three-dimension. In other words, we assume each segment is a cylinder with a certain radius R, when two cylinders contact ($2R \geq D$), a collision occurs (Figure 4). If the two cylinders belong to the same chain, the self-avoiding criteria is violated. If the two cylinders belong to neighboring DNA molecules, a potential hybridization event occurs.

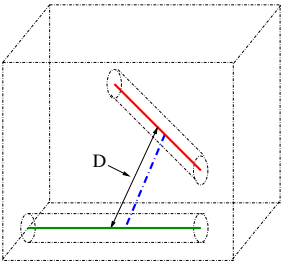

Fig. 4. Collision detection in 3D

3.4 Random Conformation

The random conformation of the DNA molecule is generated by a random walk in three dimension. 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, $\mathbf{R}_i$ 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 [83]. f of a rigid, rod-like molecule 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 [83].

3.5 Energy

Now we describe how we calculate $E(m)$ as stated in Algorithm 7. Our current simplified model neglects the following energies though more accurate model should take them into consideration [108, 24]: 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 [43, 108, 49], 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 [108].

Please refer to [74] for description of the *bending energy*, *twisting energy*, *electrostatic energy*, and other physical models.

3.6 Parameters

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

Parameters for ssDNA. Let L be the contour length of the ssDNA, $L = l_{bp} N_{bp} = l_0 N$. 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 [103], $l_{bp} = 0.7\ nm$. Many groups have obtained the force/extension data for ssDNA in different salt environment [108, 81, 71, 8, 16]. Parameters used in our model is obtained from [108], where $l_0 = 1.5\ nm$ and $Y = 120\ K_B T/nm^2$. The persistence length $P = 0.7\ nm$ [81]. The diffusion coefficient D of ssDNA is obtained from [83] as approximately $1.52 \times 10^{-6}\ cm^2 s^{-1}$ for a 20 bp strand. The diameter of the ssDNA backbone is $1\ nm$ [23].

Parameters for dsDNA. For dsDNA, the parameters associated with the equations are difference, i.e. $l_0 = 100\ nm$ [43, 22, 62], $P = 50\ nm$, $Y = 3K_B T/2P$ [22, 84], $l_{bp} = 0.34\ nm$ [103], and $D = 1.07 \times 10^{-6}\ cm^2 s^{-1}$ [83]. For short dsDNA segment (20 bp), WLC model can be simplified as the straight, rigid cylinder model with reasonable adequacy [3, 62]. WLC models are used for simulation consistency.

3.7 Motion of the Complex Nano-structure

The MC simulation described previously can can applied to the the complex nano-structure. Since it is reducible to a collection of ssDNA and dsDNA WLC segements, perturbations of each segment is done independently. The total energy is computed as a summation of the energies associated with individual segments. For more accurate model, loop energy and the energy associated with each branching point should also be considered.

3.8 Physical Model for Hybridization

Though extensive research has been done for RNA folding simulation [30, 98], to the best of our knowledge, there is no empirical results that describe: 1) the location of contact that initializes the hybridization; 2) the motions of each individual strands during the hybridizations; nor 3) the actual physical location of the hybridized products relative to other molecules in the system.

We make the following hypothesis: 1) location of contact is not explicitly modeled in the simulation; 2) upon collision that leads to potential hybridization, two strands immediately align their bases involved in the formation of duplex with the right orientation; 3) during the hybridization process, the displacement of the two strands is inversely proportional to their mass (or number of bases in the structure). The model can be subsequently improved as the empirical evidence become available. Figure 5 illustrates one schematic to depict our hypothesis.

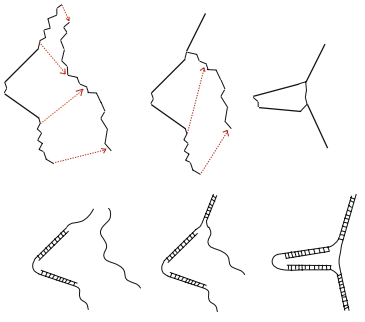

Fig. 5. Figure illustrates various steps wrt the physical motion of the strands during hybridization

4 Event Simulation

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

4.1 Hybridization

The nearest-neighbor (NN) model is used to model the hybridization event [41]. The model assumes that the stability of a given base-pair depends on the identity and orientation of neighboring base pairs [41]. Empirical data are used to determine parameters for all possible alignments of base pairs. The model has been shown to describe the thermodynamics of DNA structure that involves mismatches and neighboring base pairs beyond the Watson-Crick pairs [68,73]. 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 $m_1{m_2}^i$, $p_i \propto \exp(-\Delta G_i^\circ / RT)$. Let P_h^i be the probability of hybridization according to alignment i, for all p_j that exceeds a given threshold, we have

$$P_h^i = \frac{p_i}{\sum_j p_j} \tag{6}$$

4.2 Dehybridization

Let P_d be the dehybridization probability of a molecule m_1m_2 and $[m_1m_2]$ be its concentration. Let k_r be the reverse rate constant, R_r be the reverse rate, where $R_r = k_r[m_1m_2]$. The number of molecules dehybridized in time Δt is $R_t\Delta t$. Therefore the probability that the molecule m_1m_2 dehybridizes in Δt can be approximated as

$$P_d = \frac{k_r[m_1m_2]\Delta t}{[m_1m_2]} = k_r\Delta t \tag{7}$$

Refer to [74] for more details.

4.3 Strand Displacement

Strand displacement is modeled as a random walk in which the direction of movement of the branch point 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 [12]. In other words, migration probability towards the direction with mismatches are substantially decreased. Based on Figure 2, molecule ABC is denoted as the DNA nanostructure involving molecule A, B and C before the strand displacement. Let G°_{ABC} be its free energy. Denote G°_{rABC} and G°_{lABC} as the free energy of ABC after 1 base pair migration towards right, and left, 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 [12] 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 described previously.

5 Adaptive Time Step

We use adaptive time steps in our simulation. The simulation captures various processes at different time-scales. Ideally, the smallest time unit should be chosen as the time step $\delta t \sim 10^{-6}$ to resolve the conformations and trajectory of each individual molecule using the WLC model and MC simulation. Inspired by ideas in the kinetic Monte Carlo method [92], long-time system dynamics of the system consists of diffusive jumps from state to state. There are series of simulation steps where no collisions take place and molecules remain far apart. We attempt to overcome the limitations of such a short time-scale approach. In other words, we differ the time intervals between long-ranged molecules and short-ranged molecules.

If all the strands are far apart, we can guarantee that within a particular time-interval δT there will not be any collisions. We treat each molecule in the system as a unit of rigid body and assign a random momentum to each unit. We apply this large-scale time step δT as the simulation step at that instance. δT moves the entire system from state to state, which is computational efficient. We store the distance between the closest pair of potential reactive molecules. As the distance reaches a given threshold where the conformations of molecules can no longer being ignored, we change to a smaller-scale time step δt.

6 Algorithm Analysis

The major portion of the time taken by the algorithm is in the physical simulation, so it suffices to analyze the time-complexity of the physical simulation of the molecules in the system. The discrete WLC model is used to replace the continuos WLC for computational simplification. As the (discrete) WLC consists of N segments, it is an approximation that improves as N increase. For a WLC simulation of a single chain (dsDNA or ssDNA), due to the self-collision detections, it runs in $O(N^2)$ time for a single simulation step. Similar analysis applies to complex DNA nano-structures where N is equal in this case to the total number of double-stranded or single-stranded segments in the structure. Let M be the total number of the molecules in the small cell volume. For a long-ranged simulation period, each molecule in the system is treated as a rigid unit. Therefore the complexity per simulation step is $O(M)$. If M' is the number of molecule pairs that reach the short-range simulation threshold, then the simulation time per step is $O(M'N^2)$.

7 Preliminary Results

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

7.1 Physical Simulation

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 purpose, 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 [52]. Then these configurations are saved for the actual simulation. The figures shown here are snapshots of a simulation during different time steps, from both 2D and 3D (Figure 6) perspective, visualized by Matlab. 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.

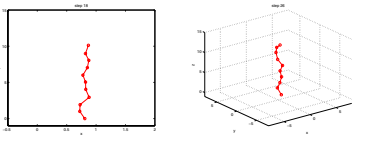

Fig. 6. 2D and 3D snapshots of the simulation for a single tethered DNA

7.2 Event Simulation

We present here a snapshot of a hybridization event in simulation based on our framework in Figure 7. Bold black lines represent the double stranded DNA regions, while the thinner lines are single-stranded. The ssDNA we display in the above snapshots are $20 - 30\ bp$.

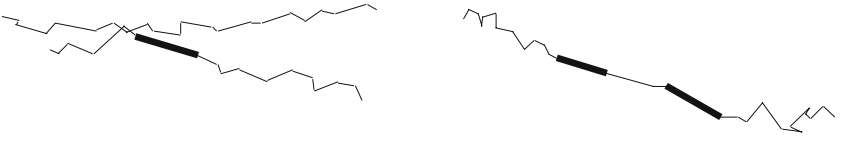

Fig. 7. Simulation of a hybridization event

8 Discussion and Future Work

We presented a comprehensive framework for building a software tool for simulating DNA based molecular system, and not the actual software tool itself. 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. As we improve the accuracy of the physical simulation, i.e. to reflect topological constraints by modeling more complicated DNA nano-structures such as pseudo-knots [38, 13]; to provide more biophysical sound behavior of DNA strands by considering stacking energy and electrostatic energy; or to achieve the molecular details by replacing the MC simulation with a BD simulation once computational resources are available, we can validate its correctness against polymer theory and experimental data, i.e. average radius of gyration and the diffusion constant. We can constantly update the physical simulation component to result in more realistic simulation.

During the physical simulation the random perturbations often lead to a configuration that can be achieved only with a low probability. Can we optimize the simulation so that we sample a larger space of configurations to avoid these with low probability, therefore making the simulation more computational efficient?

The first extension to our framework is to consider more complicated interactions, i.e. the enzyme restriction event and the hairpin formation. The second extension is to incorporate sequence design capabilities. We would like to design and optimize sequences based on the given nano-structure conformations. Furthermore, a conformation change of a nano-device can be decomposed into units of local deformations to ease the sequence design.

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

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*, 18:1868, 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. 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.
10. 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.
11. 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.
12. I. Biswas, A. Yamamoto, and P. Hsieh. Branch migration through dna sequence heterology. *J. Mol. Bio*, 1998.
13. Justin S. Bois, Suvir Venkataraman1, Harry M. T. Choi1, Andrew J. Spakowitz, Zhen-Gang Wang and Niles A. Pierce1,2,* Topological constraints in nucleic acid hybridization kinetics. *Nucleic Acids Research*, 33(13):4090–4095, 2005.
14. 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, January 1999.
15. C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith. Entropic elasticity of lambda-phage dna mechanics. *Science*, 265:1599, 1994.
16. C. Bustamante, S. Smith, J. Liphardt, and D. Smith. Single-molecule studies of dna mechanics. *Current Opinion in Structural Biology*, 10:279, 2000.
17. 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(014901), 2005.
18. 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.
19. 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.
20. Y. Chen and C. Mao. Putting a brake on an autonomous DNA nanomotor. *J. Am. Chem. Soc.*, 126:8626–8627, 2004.
21. Y. Chen, M. Wang, and C. Mao. An autonomous DNA nanomotor powered by a DNA enzyme. *Angew. Chem. Int. Ed.*, 43:3554–3557, 2004.
22. S. Cocco, J. F. Marko, and R. Monasson. Theoretical models for single-molucule dna and rna experiments: from elasticity to unzipping. *to appear in CRAS, special issue dedicated to Single Molecule Experiments*, 2002.
23. 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.
24. 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.
25. P. Dimitrakopoulos. Stress and configuration relaxation of an initially straight flexible polymer. *J. Fluid Mech.*, 513:265–286, 2004.

26. P.S. Doyle and P.T. Underhill. Brownian dynamics simulations of polymers and soft matter. *S. Yip, (ed.), Handbook of Materials Modeling*, pages 2619–2630, 2005.
27. R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree and N. A. Pierce. Thermodynamic analysis of interacting nucleic acid strands. *SIAM Rev*, in press.
28. 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.
29. M. Fixman and J. Kovac. Polymer conformation statistics iii: Modified gaussian models of the stiff chains. *J. Chem. Phys.*, 58:1564–1568, 1973.
30. C. Flamm, W. Fontana, I. L. Hofacker, and P. Schuster. RNA folding at elementary step resolution. *RNA*, 6(3):325-38, 2000.
31. J.B. Fournier. Wormlike chain or tense string? a question of resolution. *Continuum Mechanical Thermodynamics*, 14:241, 2002.
32. M.D. Frank-Kamenetskii. Biophysics of dna molecule. *Phys. Rep.*, 288:13 – 60, 1997.
33. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.
34. D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115:1716–1733, 2001.
35. A. J. Hartemink and D. K. Gifford. Thermodynamics simulation of deoxyoligonucleotide hybridization for dna computation. 1997.
36. 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.
37. 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.
38. 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.
39. H. M. James and E. Guth. Theory of the elastic properties of rubber. *Journal of Chemical Physics*, 10:455–481, 1943.
40. 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.
41. J. Santalucia Jr. A unified view of polymer, dumbbell and oligonucleotide dna nearest-neighbor thermodynamics. *PNAS*, 95:1460–1465, 1998.
42. A. M. Kierzek. Stocks: Stochastic kinetic simulations of biochemical systems with gillespie algorithm. *Bioinformatics*, 18:470–481, 2002.
43. 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.
44. J. Kovac and C. Crabb. Modified gaussian model for rubber elasticity. 2. the wormlike chain. *Macromolecules*, 15(2):537, 1982.
45. M. Kuhn and F. Grun. Relationships between elastic constants and stretching double refraction of highly elastic substances. *Kolloid-Z*, 101:294, 1942.
46. S. Kutter. Elasticity of polymers with internal topological constraints. *PhD Thesis*, August 2002.
47. 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.
48. 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.

49. J. Langowski. Polymer chain models of dna and chromatin. *Manuscript*, 2006.
50. 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.
51. 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.
52. 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.
53. J. Li and W. Tan. A single DNA molecule nanomotor. *Nano Lett.*, 2:315–318, 2002.
54. 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.
55. 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.
56. A. Malevanets and J. M. Yoemans. Dynamics of short polymer chains in solution. *Europhysics Letters*, 52(2):231, 2000.
57. 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.
58. 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.
59. 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.
60. J. Marko and E. D. Siggia. Bending and twisting elasticity of dna. *Macromolecules*, 27:981, 1994.
61. J. F. Marko and E. D. Siggia. Stretching dna. *Macromolecules*, 28:8759, 1995.
62. 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.
63. J. Mercier and G. W. Slater. Solid phase dna amplification: a brownian dynamics study of crowding effects. *Biophysical Journal*, 89:32–42, July 2005.
64. 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.
65. T. Odijk. Stiff chains and filaments under tension. *Macromolecule*, 28:7016–7018, 1995.
66. 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.
67. 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.
68. 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.
69. 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.
70. 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.
71. M. Rief, H. Clausen-Schaumann, and H. E. Gaub. Sequence-dependent mechanics of single dna molecules. *Nature Structural Biology*, 6:346 – 349, 1999.
72. 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.
73. J. Santalucia and D Hicks. The thermodynamics of dna structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415, 2004.

74. S. Sahu, B. Wang, and J. H. Reif. A Framework for Modeling DNA Based Molecular Systems. *Technical Report*, Duke University, 2006.
75. 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.
76. W.B. Sherman and N.C. Seeman. A precisely controlled DNA biped walking device. *Nano Lett.*, 4:1203–1207, 2004.
77. J.S. Shin and N.A. Pierce. A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.*, 126:10834–10835, 2004.
78. F.C. Simmel and B. Yurke. Using DNA to construct and power a nanoactuator. *Phys. Rev. E*, 63:041913, 2001.
79. F.C. Simmel and B. Yurke. A DNA-based molecular device switchable between three distinct mechanical states. *Appl. Phys. Lett.*, 80:883–885, 2002.
80. 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.
81. 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.
82. 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.
83. 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.
84. C. Storm and P. C. Nelson. Theory of high-force dna stretching and overstretching. *Physical Review E.*, 67:051906, 2003.
85. 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.
86. 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.
87. I. Tinoco and C. Bustamante. The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys Chem.*, 101-102:513, December 2002.
88. 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.
89. A.J. Turberfield, B. Yurke, and Jr. A.P. Mills. DNA hybridization catalysts and molecular tweezers. *DNA5*, 2000.
90. T. E. Turner, S. Schnell, and K. Burrage. Stochastic approaches for modelling in vivo reactions. *Computational Biology and Chemistry*, 2004.
91. A.V. Vologodskii. Monte carlo simulation of dna topological properties. *Preprint*, 2004.
92. A.F. Voter. Introduction to kinetic monte carlo method. *Springer*, NATO publishing unit, 2005.
93. J. G. Wetmur and N. Davidson. Kinetics of renaturation of dna. *J. Mol. Biol.*, 31:349–370, 1968.
94. 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.
95. E. Winfree. Simulation of computing by self-assembly. Technical Report 1998.22, Caltech, 1998.
96. 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.

97. 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.
98. 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.
99. H. Yamakawa and T. Yoshizaki. Dynamics of helical wormlike chains. i. dynamic model and diffusion equation. *Journal of Chemical Physics*, 75(2):1016, July 1981.
100. 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.
101. 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.
102. H. Yan, X. Zhang, Z. Shen, and N.C. Seeman. A robust DNA mechanical device controlled by hybridization topology. *Nature*, 415:62–65, 2002.
103. 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.
104. 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.
105. 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.
106. B. Yurke, A.P. Mills, and A.J. Turberfield. A molecular machine made of and powdered by DNA. *Biophysics*, 78:2629, 2000.
107. 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.
108. 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(2):1133–1143, August 2001.