LA-UR-04- 6051

Approved for public release; distribution is unlimited.

Title: MODELING DNA BUBBLE FORMATION AT THE ATOMIC SCALE

Author(s): Violeta Beleva
Kim Rasmussen
Angel E. Garcia

Submitted: Proceedings of the Second Multiscale/Materials Modeling Conference
UCLA Oct. 11-15, 2004

Group Address Theoretical Biology and Biophysics Group Los Alamos National Laboratory Los Alamos, New Mexico 87545

Los Alamos
NATIONAL LABORATORY

Modeling DNA Bubble Formation at the Atomic Scale

Violeta Beleva[1], Kim Rasmussen1[2], and Angel E. Garcia[1,2]
[1]Center for Non Linear Studies (CNLS) and [2]Theoretical Division, T10 MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545
(axg@lanl.gov)

ABSTRACT

We describe the fluctuations of double stranded DNA molecules using a minimalist Go model over a wide range of temperatures. Minimalist models allow us to describe, at the atomic level, the opening and formation of bubbles in DNA double helices. This model includes all the geometrical constraints in helix melting imposed by the 3D structure of the molecule. The DNA forms melted bubbles within double helices. These bubbles form and break as a function of time. The equilibrium average number of broken base pairs shows a sharp change as a function of T. We observe a temperature profile of sequence dependent bubble formation similar to those measured by Zeng et al. (*1*).

1. Introduction

Long nuclei acid molecules melt partially through the formations of bubbles. It is known that CG rich sequences melt at higher temperatures than AT rich sequences. The melting temperature, however, is not solely determined by the CG content, but by the sequence through base stacking and solvent interactions (*2*). Recently, models that incorporate the sequence and nonlinear dynamics of DNA double strands have shown that DNA exhibits a very rich dynamics (*3*). Recent extensions of the Bishop-Peyrard model show that fluctuations in the DNA structure lead to opening in localized regions, and that these regions in the DNA are associated with transcription initiation sites(*4*). 1D and 2D models of DNA may contain enough information about stacking and base pairing interactions, but lack the coupling between twisting, bending and base pair opening imposed by the double helical structure of DNA that all atom models easily describe. However, the complexity of the energy function used in all atom simulations (including solvent, ions, etc) does not allow for the description of DNA folding/unfolding events that occur in the microsecond time scale.

2. Methods

We have developed an all atom model of DNA that contain these couplings, but with a simplified set of interactions, similar to the Go models used for protein folding (*5, 6*). The Go model defines a minimally frustrated, funnel-like, energy landscape(*7*). We use a minimalist representation of the interaction potential which includes base pairing, screened Coulomb, and stacking interactions. Based on the secondary structure of the native structure (i.e., the folded state), we classify atomic interactions as native, and non native. Native interactions are stacking and hydrogen bond interactions that are present in the folded state. Non native interactions are modeled as excluded volume interactions between all other pair of atoms not interacting in the native state. All atoms in the nucleic acid bases interact with neighboring bases in the sequence, and with the atoms in the neighboring bases of their base pair partners. For example, in the double helix $d[CGCG]_2$

(shown below) we define as native interactions the Watson-Crick hydrogen bond interactions between C1 and G8, G2 and C7, C3 and G6, and G5 and C5. Other base pairs are not considered native. We also define as native stacking interactions between all atoms in the G2 base with atoms in C1, C3, G6, and G8—but not with G4 and C5.

C1 G2 C3 G4
G8 C7 G6 C5

All atoms in the backbone interact via the non native potential, regardless of base pairing. In addition, all pairs of P atoms in the phosphates interact via a screened Coulomb potential, with a screening length, L_D, determined by the Debye-Huckel theory—assuming a homogeneous monovalent salt solution in water.
Stacking interactions are modeled by a Lennard Jones potential ($\varepsilon_s[(\sigma/r)^{12} - 2(\sigma/r)^6]$ with σ_s = 3.5 A), hydrogen bonding are modeled by a 10-12 potential ($\varepsilon_{HB}[5(\sigma/r)^{12} - 6(\sigma/r)^{10}]$ with σ_{HB} = 3.0 A), and non native interactions are modeled by a repulsive $\varepsilon_{NN}(\sigma/r)^{12}$, with σ_{NN} = 2.0 A. The native hydrogen bonding, stacking, and non-native parameters are taken to be of order 1, 0.1, and 0.01, respectively, with ε_{HB}/RT_m =1 at 350 K. All bonding interactions are modeled using the Amber united atom force field (*8*). The overall ratio of energy over T is scaled such that we observe a sharp dependence in the opening/closing of base pairs at a fixed T. All calculations are done with a screening length L_D = 10 A. All bonding interactions are modeled using the Amber united atom force field (*8*).

3. Results

To test for the simplified Hamiltonian shown above we model the base pair opening and closing of a 60 base pairs molecule studied by Zeng et al. (*1*). First we determined a set of energy parameters that will show a melting transition at a fixed temperature, T_m= 350 K (Fig. 1). In Fig. 2 we show the dynamics of base pair opening and closing at temperatures below and above the transition T (defined as the T at which half the bases are opened). DNA bubbles are formed at all temperatures, but larger bubbles are formed at higher T.

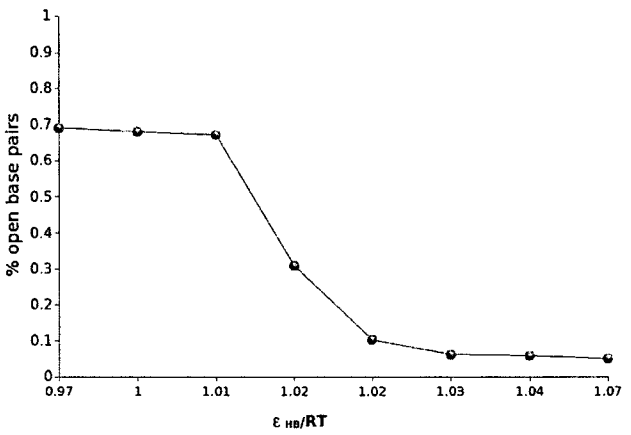

Fig. 1. Fraction of open base pairs as a function of the non-bonding interaction energy ε_{HB}/RT, at a fixed temperature, T=350 K. The curves are calculated from 20 ns trajectories.

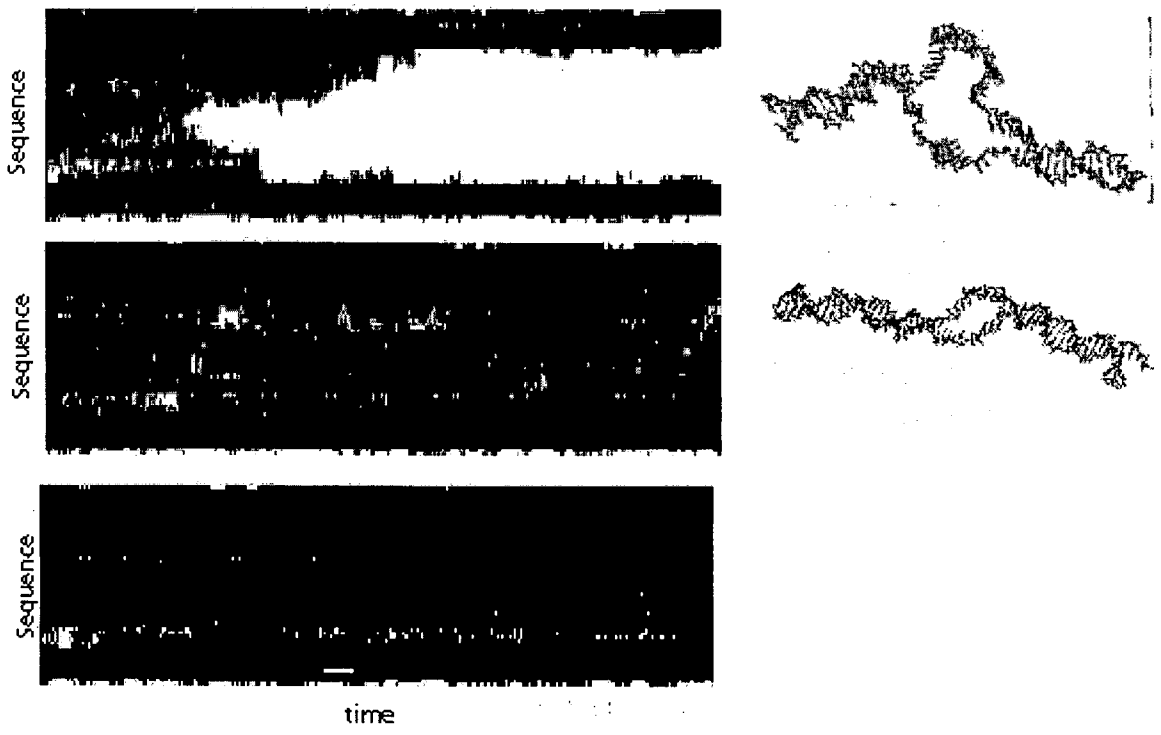

Fig. 2. Time series of base opening (white) and closing (black) during three different 20 ns trajectories at 330 K (bottom), 350 K (center), and 355 K (top). The right hand side plot shows DNA double helical configurations showing bubbles of different sizes. The CG rich sequence at the ends melt at much higher T, while the AT rich region open at $T < T_m$.

4. Conclusions

We have shown that simple minimalist models with atomic detail can reproduce observed bubble formation fluctuations in DNA. Bubbles form at temperatures well below the transition T. Small bubbles are 5 – 10 base pairs long, while large bubbles are as large as the AT rich segment of the DNA molecule.

1. Y. Zeng, A. Montrichok, G. Zocchi, *Phys. Rev. Lett.* **91**, 148101 (2003).
2. K. J. Breslauer, R. Frank, H. Blocker, L. A. Marky, *Proc. Nat. Acad. Sci. (USA)* **83**, 3746-3750 (1986).
3. M. Peyrard, A. R. Bishop, *Phys. Rev. Lett.* **62**, 2755-8 (1989).
4. C. H. Choi *et al.*, *Nuc. Acids Res.* **32**, 1584-1590 (2004).
5. N. Go, *Ann. Rev. Biophys. & Bioeng.* **12**, 183-210 (1983).
6. C. Clementi, A. E. Garcia, J. N. Onuchic, *J. Mol. Bio.* **326**, 933-954 (2003).
7. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, *Proteins Struct. Funct. & Gen.* **21**, 167-195 (1995).
8. S. J. Weiner *et al.*, *J. Amer. Chem. Soc.* **106**, 765-84 (1984).