PHYSICAL REVIEW E VOLUME 52, NUMBER 4 OCTOBER 1995

Model simulations of DNA dynamics

Fei Zhang and Michael A. Collins
Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 0200, Australia

(Received 14 March 1995; revised manuscript received 1 June 1995)

A simple physical model is developed to study some characteristics of DNA dynamics. The model uses only a few degrees of freedom so that DNA dynamics can be simulated over *longer times* and for *larger sequences* than is possible with all-atom simulations. The utility of the model is examined by comparing several qualitative and quantitative features of the model DNA dynamics with all-atom simulations. The model helix is stable at low temperatures and gradually unwinds (melts) at high temperatures. The melting transition in this model seems to be fairly sharp in temperature, but the melting kinetics is rather slow in time. Localized fluctuations of the hydrogen bonds are commonly observed, especially in regions enriched with adenine-thymine pairs.

PACS number(s): 87.10.+e, 87.15.By, 63.20.Dj, 63.70.+h

I. INTRODUCTION

Over the past several years, the molecular dynamics of deoxyribonucleic acids (DNA) has attracted increasing attention [1–8]. Many attempts have been made to model DNA dynamics using ideas of energy localization and transport in nonlinear and disordered systems [1]. However, simple phenomenological models, while relatively easy to analyze and simulate, are usually unable to provide detailed dynamical and structural information. At the other extreme, *all-atom* molecular dynamics (MD) simulation [2–8] would appear to be the method of choice by virtue of the fact that it represents the exact classical solution, if the molecular potential energy surface were known. Unfortunately, the computational demand of all-atom MD is so large (there are about 60 atoms per base pair) that only short DNA fragments have been simulated over 1 nsec at most [6], which is much shorter than the characteristic time scale of many physical and biological processes in DNA (e.g., replication, transcription, denaturation, and supercoiling [9]). Therefore, it is useful to develop reduced DNA models. If possible such models should contain the most important degrees of freedom and enough structural details to mimic reality while allowing simulations of large sequences over long time scales. Moreover, an understanding of which reduced models best mimic reality is essential if very simple mathematical models are to be of use.

The objective of the present paper is to explore the development of a molecular model to study some characteristics of DNA dynamics. The simplified model described here proceeds in the spirit of MD simulations, but discards as many degrees of freedom as possible to preserve maximum simplicity. We use the reduced model to simulate several DNA sequences over about 10 nsec at different temperatures ranging from 200 K to 520 K. It is shown that the model is able to reproduce some features of DNA dynamics. In particular, the helix structure of the DNA model is stable at low temperatures, but melts above a certain temperature, where the two strands of DNA begin to unwind and separate from each other, with increasing numbers of disrupted hydrogen bonds. The melting transitions are found to be rather slow in time; it takes longer than 10 nsec to completely denature a DNA sequence of only 100 base pairs.

However, this very simple model allows only strictly planar motion of the DNA base pairs. Hence it does not describe any three-dimensional dynamics, including tilt and vertical shift of the base pairs and bending of the whole DNA helix. Since the model is mathematically equivalent to a two-dimensional lattice model (albeit a relatively complex one), the "melting" transition and fluctuation dynamics should not be expected to accurately model helical DNA.

The paper is organized as follows. In Sec. II, we describe a simplified model constructed from the all-atom DNA double helix, giving all the geometrical and force parameters involved. In Sec. III the molecular dynamics simulation results for several types of DNA sequences are reported in detail. Section IV contains some concluding remarks.

II. DESCRIPTION OF THE MODEL

The simplified DNA model is first built up with a *double helix* structure, whose equilibrium geometry is taken to be that for *B*-form DNA [10]. As is well known, each strand of DNA is composed of a sequence of purine and pyrimidine bases or nucleotides, which are all flat molecules composed of one or two rings of nitrogen and carbon atoms [11]. These molecules are quite stiff with respect to all molecular distortions. In particular, they have no low-frequency conformational distortions so that they may be treated approximately as rigid bodies. Each base is attached to a five-membered deoxyribose ring, which is connected to other such rings by a quite flexible backbone [see Fig. 1(a)]. The deoxyribose ring is not flat and has a number of preferred conformations. While transitions between these preferred conformations can occur at physiological temperatures, the ring is nevertheless significantly more rigid than the deoxyribose—PO_4—deoxyribose backbone linkage. Thus, in order to discard as many freedoms as possible, we ignore distortions of the ribose ring and fix

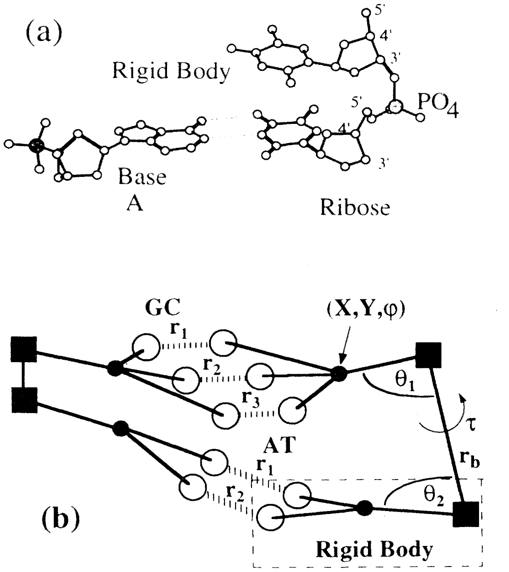

FIG. 1. (a) Schematic representation of a fragment of DNA double helix, showing a base pair and one adjoining nucleotide and phosphate backbone. All hydrogen atoms have been deleted for clarity. (b) Schematic representation of the simplified DNA model constructed from the all-atom model. The center of mass (black circle), hydrogen bonding sites (white circle), and backbone connections (black square) are indicated, with bond lengths, angles, and torsion angles defined.

it rigidly to each base in the standard conformation it takes in idealized B-DNA [10]. The flexible backbone itself is treated as a structureless elastic rod. Each strand of this B-DNA model is then a sequence of *rigid bodies* (base-ribose) *connected by flexible rods.* The bases of each strand are hydrogen bonded to the complementary bases on the other strand, i.e., adenine (A) to thymine (T) and guanine (G) to cytosine (C). In the standard equilibrium B-DNA structure, the bases in a pair are nearly coplanar. In this first simplest model, we only allow these rigid base-ribose bodies to move in this XY plane. When projected onto this plane, the backbone connections at the ribose positions $C3'$ and $C5'$ are quite close to each other. Hence, for simplicity, both the backbone connections are taken to apply at the one location given in Table I. The distance between two adjacent planes is 3.38 Å. The motion of each rigid body can then be completely described in terms of the (X,Y) coordinates of the body's center of mass and an angle ϕ, which determines the orientation of the body with respect to, say, the X axis. This model is depicted schematically in Fig. 1(b), where the internal coordinates are defined.

The model potential energy has the form

$$V=\sum_{n}\sum_{j=1,2}(V_b+V_\theta+V_\tau)+\sum_{n}(V_{\rm LJ}+V_{\rm H}), \qquad (1)$$

which is a sum over all base pairs (n) and over the two strands ($j = 1,2$). Each term in this potential is described as follows. We assume quadratic backbone stretching and bending [the superscript "eq" in the following potential functions refers to the corresponding equilibrium values and the base pair (n) and strand labels (j) are dropped for simplicity]

$$V_b=\frac{1}{2}K_b(r_b-r_b^{\rm eq})^2, \qquad (2)$$

$$V_\theta=\frac{1}{2}K_\theta[\cos(\theta_i)-\cos(\theta_i^{\rm eq})]^2,\quad i=1,2, \qquad (3)$$

and a periodic torsion potential

$$V_\tau=K_\tau[1-\cos(\tau-\tau^{\rm eq})]. \qquad (4)$$

Nonbonded Lennard-Jones interactions between the centers of mass of the base-ribose moieties in a pair prevent unphysical overlap:

$$V_{\rm LJ}=4\epsilon\left[\left(\frac{\sigma}{\bar{R}}\right)^{12}-\left(\frac{\sigma}{\bar{R}}\right)^{6}\right], \qquad (5)$$

where $\bar{R}=R_{\rm c.m.}-R_{\rm c.m.}^{\rm eq}+2^{1/6}\sigma$, $R_{\rm c.m.}$ being the distance between the centers of mass of the rigid bodies in a pair. A modified Morse potential describes each hydrogen bond energy

$$V_H=V_0\{\exp[-\alpha(r_i-r_i^{\rm eq})]-1\}^2$$
$$-\frac{1}{4}V_0\{1+\tanh[\beta(r_i-r^*)]\}, \qquad (6)$$

where r_i refers to the hydrogen bond lengths ($i = 1,2,3$) of Fig. 1, $r_i^{\rm eq}$ is given in Table I, and the term involving tanh mimics the formation of hydrogen bonds with the surrounding solvent (water) once the interbase hydrogen bonds are stretched more than about 2 Å from their equilibrium values. The potential energy minimum for hydrogen bonding to the solvent has been set at half the value for hydrogen bonding to the complementary base.

For a given base-pair sequence, the equilibrium structure of a B-DNA molecule is determined from the parameters of Table I. The appropriate values of $r_b^{\rm eq},\theta_1^{\rm eq},\theta_2^{\rm eq}$, and $\tau^{\rm eq}$ are then determined. The magnitudes of the force parameters in Eqs. (2)–(6) are presented in Table II. In particular, the values of K_b and K_τ are close to the values used by Tan and Harvey [12] in a molecular mechanics study of supercoiled DNA structure. An appropriate value for K_θ is taken to ensure that the three parameters, K_b, K_τ, and K_θ, in combination, are consistent with the observed elastic modulus for twisting of DNA [13]. The hydrogen-bond force parameters are those of Chen and Prohofsky [14].

All the variables (bond lengths, angles, etc.) in the potential V are expressed in terms of the center of mass coordinates $(X_1(n),Y_1(n))$ and $(X_2(n),Y_2(n))$ and orientation angles $\phi_1(n)$ and $\phi_2(n)$ of the two rigid bodies that form the nth base pair between strands 1 and 2. The total Hamiltonian for this model is then $H = T+V$, where V is the potential energy and T is the kinetic energy

$$T=\frac{1}{2}\sum_{n}\left\{\frac{[P_{X_1}(n)]^2+[P_{Y_1}(n)]^2}{M_1(n)}+\frac{[P_{X_2}(n)]^2+[P_{Y_2}(n)]^2}{M_2(n)}+\frac{[P_{\phi_1}(n)]^2}{I_1(n)}+\frac{[P_{\phi_2}(n)]^2}{I_2(n)}\right\}. \tag{7}$$

Here P represents the momentum conjugate to the variable indicated, $M_i(n)$ denotes the total mass of the nth base-ribose rigid body in the ith strand, and $I_i(n)$ denotes the corresponding moment of inertia about the base-ribose center of mass. Using the B-DNA structure data, we have estimated these parameters (see Table I).

Therefore, there are naturally two basic ingredients in our model: strong *nonlinearity* in the torsional potentials (4) and in the interbase hydrogen-bond (H-bond) interactions (6) and *inhomogeneities* in the masses and moments of inertia of the four bases, and in the strength of the H bonds in AT and GC base pairs (an AT pair has two H bonds, while a GC pair has three). It has been postulated that both nonlinearity and inhomogeneity might play important roles in energy localization and transport processes in the DNA thermal denaturation [1].

III. MOLECULAR DYNAMICS SIMULATION RESULTS

We have carried out constant temperature (canonical) simulations of the above DNA model using the Nosé-Hoover thermalization scheme [15], in which the temperature is an adjustable parameter in the modified equations of motion for the dynamical variables. Following the criterion (see Ref. [15]) for choosing an appropriate value of the thermal relaxation time τ_r (or the heat bath mass Q), we have set $\tau_r \approx 1$ psec, which is very close to the characteristic period of hydrogen-bond vibrations between complementary bases. It was determined that the simulation results do not change if different relaxation times (ranging from 0.2 to 5 psec at least) are used. The total linear momentum in the X and Y directions and the angular momentum about the Z axis are initially set to zero and remain unchanged during the simulation. However, the total energy of the DNA molecule is subject to

TABLE I. Equilibrium geometrical parameters for the DNA model. All distances are given in angstroms and angles are given in radians. Orientation angles are measured with respect to the X axis. The $\pm$ sign below is taken as $+$ for bases in strand 1 and $-$ for bases in strand 2. $A_j \equiv (j-1)\pi/5$, $j=1,2,3,\ldots$, is added for the jth base.

Property	Adenosine	Thymine	Guanine	Cytosine
Mass (g/mol)	184	194	170	210
Moment of inertia (g Å^2/mol)	1194.7	1394.6	1010.8	1649.2
Distance of the base center of mass (c.m.) to the helix axis	4.5294	3.9565	4.5896	3.7238
Distance of the base c.m. to each bonding site				
Hydrogen bond 1	3.7438	4.0328	3.9738	3.8947
Hydrogen bond 2	2.1842	3.2872	2.3112	3.0056
Hydrogen bond 3			2.2667	3.7982
Backbone connection	3.5513	3.9588	3.3932	4.1755
Orientation angle of the base c.m. to helix axis vector	$A_j \pm 1.4094$	$A_j \pm 1.3073$	$A_j \pm 1.3471$	$A_j \pm 1.2552$
Dynamical variable $\phi_k(j)$, $k=1,2$	$A_j \mp 2.2042$	$A_j \mp 2.2961$	$A_j \mp 2.4354$	$A_j \mp 2.3370$
Orientation angle at each bonding site				
Hydrogen bond 1	$\phi_k(j)$	$\phi_k(j)$	$\phi_k(j)$	$\phi_k(j)$
Hydrogen bond 2	$\phi_k(j) \pm 0.5863$	$\phi_k(j) \pm 0.6117$	$\phi_k(j) \pm 0.5237$	$\phi_k(j) \pm 0.6227$
Hydrogen bond 3			$\phi_k(j) \pm 1.5621$	$\phi_k(j) \pm 1.2737$
Backbone connection	$\phi_k(j) \pm 3.3781$	$\phi_k(j) \pm 3.4285$	$\phi_k(j) \pm 3.4816$	$\phi_k(j) \pm 3.5354$
Hydrogen-bond length r_i^{eq}				
H bond 1	2.7786		2.6699	
H bond 2	2.9419		2.9021	
H bond 3			3.0051	
Distance between stacked base-pair planes	3.38		3.38	

TABLE II. Model force parameters in Eqs. (2)–(6).

Bond stretch	$K_b = 1.0\ \mathrm{kcal\,mol^{-1}\,\AA^{-2}}$
Bond bending	$K_\theta = 10\ \mathrm{kcal\,mol^{-1}\,rad^{-2}}$
Torsional energy	$K_\tau = 40\ \mathrm{kcal\,mol^{-1}}$
van der Waals nonbonded interaction	$\epsilon = 0.1\ \mathrm{kcal\,mol^{-1}}$, $\sigma = 4.0$ Å
Modification to the H-bond interaction	$\beta = 2\ \mathrm{\AA^{-1}}, r^* = 4$ Å
H-bond force parameters	See Ref. [14] for V_0 and α

fluctuation.

In order to examine the validity of the DNA model and to investigate its melting behavior, we have simulated chains of various lengths and base-pair compositions. In the following, the simulation results for three types of B-DNA fragments of 100 base pairs (100 bp) over a time of about 10 sec will be reported in detail. Simulations of shorter and longer sequences will also be briefly discussed. The three sequences are (i) poly(A)·poly(T), (ii) poly (G)·poly(C), and (iii) mixed content. Here the notation poly( ) represents a homopolymer consisting of one type of base as indicated in the parentheses. The fragment of 100 bp with mixed content has the following sequence: *ACTTATGCTG-ATATCCATGT-CGC***TATAAT***G-CGCGCAACGC-CTATAGGCAA-TCCTAGCATC-GGGTAAAGCC-TTCGCGGGGG-TTAAATAAAT-GCGCGCGCGC*, for one strand and the other strand is determined by the complementary rule. Note that some AT rich regions are deliberately introduced in this sequence. In particular, the **TATAAT** sequence, known as the "Pribnow box," is a highly conserved sequence appearing in many promoter sites in many related organisms [16]. This sequence and the other AT rich regions are expected to be more amenable to base-pair opening, as will be shown in our simulations.

The first feature of note in the dynamics of this model, is that the amplitude of the position fluctuations of different sites in the molecule increases from the inside to the outside of the double helix and that the sites near the chain ends move most (see Fig. 2). Moreover, the first and third hydrogen-bonding sites in GC base pairs move more than the middle (second) site. These results, which were also observed in our previous constant energy simulations of a similar DNA model over short time scales [17], are qualitatively in agreement with those reported in all-atom MD simulations of DNA [3, 4, 6]. Thus we have a picture of the base-pair motion in which the base-ribose bodies move in and away from their partners while undergoing hindered rotation. The overall picture is not one in which the base and ribose swing about a hinge supplied by the (fixed) backbone; it is rather the reverse, where the base, ribose, and backbone are hinged flexibly about the *hydrogen bonds.* This important qualitative aspect of the motion of the DNA was not taken into account in some previous attempts [18] to derive simple nonlinear models to study the base pair planary rotations.

Another qualitative feature observed in all the simulations of this model is that the double helical structure is basically stable at low temperatures. This can be clearly seen in Fig. 3(a), which plots the "helix repeat number" against the temperature. Here the helix repeat number is the average number of base pairs contained in a 360° twist. For the model B-DNA at the equilibrium geometry, the helix repeat number is exactly 10. A value higher than 10 indicates that the helix is "unwound" by comparison with the standard structure. The curves in Fig. 3(a) stay close to 10 at low temperatures, but display a sharp increase at higher temperatures that correspond to melting of the DNA molecules.

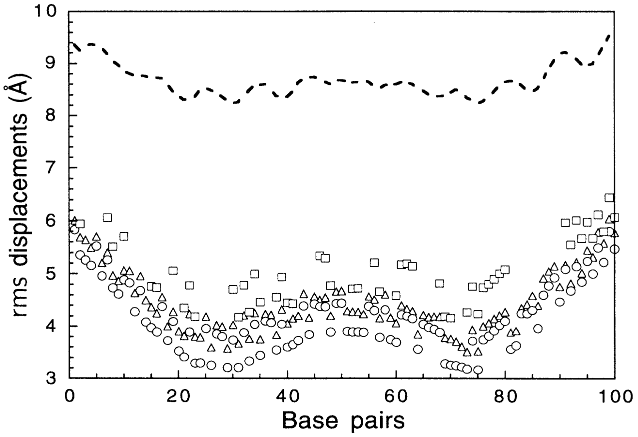

FIG. 2. Root-mean-square (rms) displacement of sites on the base pairs from their average positions for a DNA fragment of 100 bp with mixed content simulated over 10 nsec at temperature $T = 300$ K. The sites at the backbone connections (dashed line) move more than the H-bonding sites inside the helix (symbols). The first (triangles) and the third (squares) (for GC pairs) H-bonding sites move more than the second (circles). The chain ends move more than the middle.

A measure of the fluctuations in the unwinding of DNA is given by the root-mean-square average of the angle between successive interbase center of mass vectors, $\mathbf{R}_{\mathrm{c.m.}}(n)$ and $\mathbf{R}_{\mathrm{c.m.}}(n+1)$. This is shown in Fig. 3(b). In particular, the angle fluctuations in this model at T=300 K are about 8°, which is comparable to values observed in all-atom simulations of $d(CGCGAATTCGCG)_2$ by Rao and Kollman [5], and $d(CTGATCAG)_2$ by Briki and Genert [8], and close to values observed for $d(CGCAACGC)$-$d(GCGTTGCG)$ over 80 psec simulations by Van Gunsteren *et al.* [4]. There is no significant effect of base-pair composition apparent in these angular fluctuations at low temperatures.

Hydrogen-bond breaking is an important element in DNA thermal denaturation, replication, and transcription. Here we measure the mean displacements of the H-bond lengths as shown in Fig. 4(a), where the curves display behavior similar to that in Fig. 3(a): A very slow increase of the H-bond displacement indicates the stability of the model DNA at low temperatures and a rapid increase at temperatures above some transition point shows that the helix is actually melting.

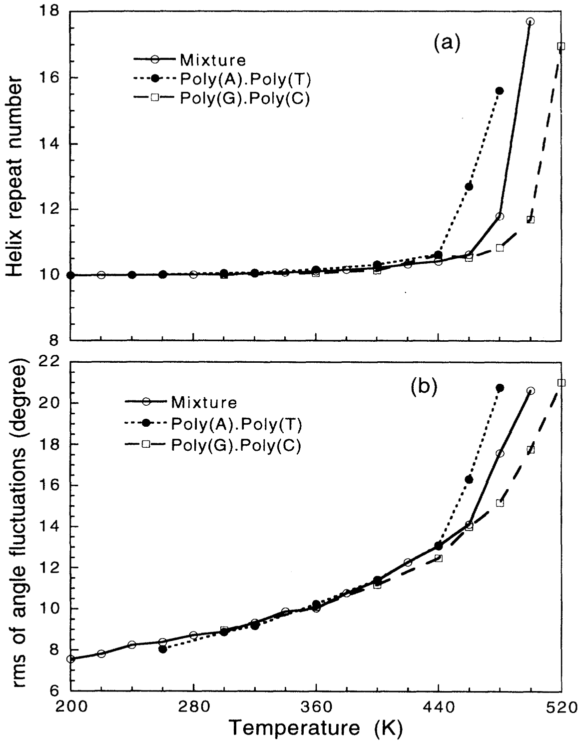

FIG. 3. (a) Helix repeat number for DNA segments of 100 base pairs and (b) rms of angle fluctuations (see the text).

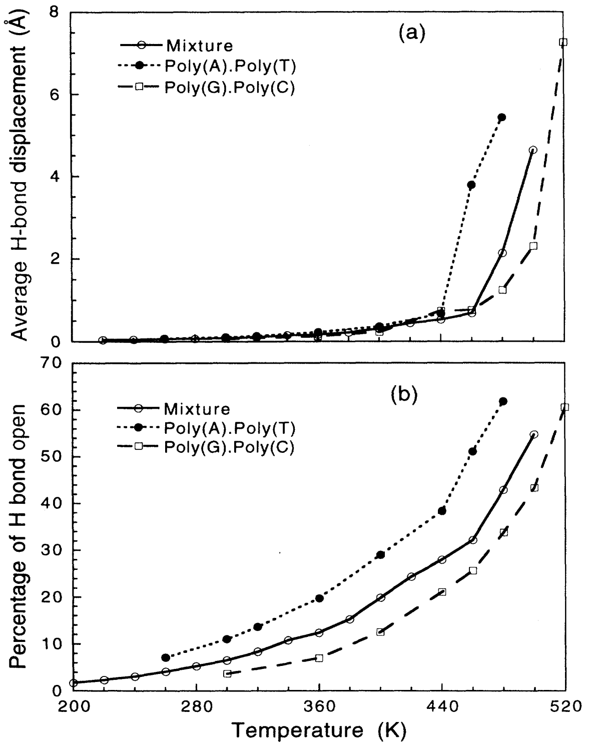

FIG. 4. (a) Average hydrogen-bond displacements and (b) percentage of H bond open, for simulations of DNA fragments of 100 bp over 10 nsec.

The percentage of H-bond "opening" has also been measured. Following Chen and Prohofsky [14] we say that a hydrogen bond is "broken" or "open" if its length exceeds its equilibrium value by a few tenths of an angstrom (see Table I of Ref. [14]). Figure 4(b) shows that, as expected, base pairs in poly(A)·poly(T) are the most amenable to opening, followed by the mixture and the poly(G)·poly(C). In fact, we have consistently found that even in the DNA fragment of mixed content, there is a much greater tendency for an AT pair to be broken than is the case for a GC pair.

Figures 3(a) and 4(a) clearly demonstrate that the present DNA model is able to show *sharp* melting transitions and that the DNA fragments of different compositions have *different melting temperatures.* The GC homopolymer has the highest melting temperature $T_{GC} \approx 500$ K, the AT homopolymer has the lowest $T_{AT} \approx 460$ K, while the sequence with mixed content has intermediate melting temperature $T_{\mathrm{mix}} \approx 480$ K. These characteristics are qualitatively in agreement with the results of thermal melting experiments for DNA [11], although the melting temperatures are about 100 K too high, indicating that the simplified DNA model is too "stiff" overall.

Most interestingly, we find that the melting transitions require long time scales to develop and that the larger the DNA sequence, the longer the time required for melting to occur. Note that we start each simulation from the ground state of the DNA with randomly distributed velocities (for X, Y, and ϕ coordinates) according to the Boltzmann law. At low temperatures, the DNA appears to reach equilibrium in less than 1 nsec: the energy of the molecule, the helix repeat number, the H-bond displacement (see Fig. 5), and the rms displacements of Fig. 2, for example, do not "drift" in value with time but merely fluctuate about a steady mean after less than 1 nsec. Thus the averages in Figs. 3 and 4, which are obtained from results of simulations over 10 nsec, may represent the thermal equilibrium values at low temperatures. However, at temperatures above the "melting

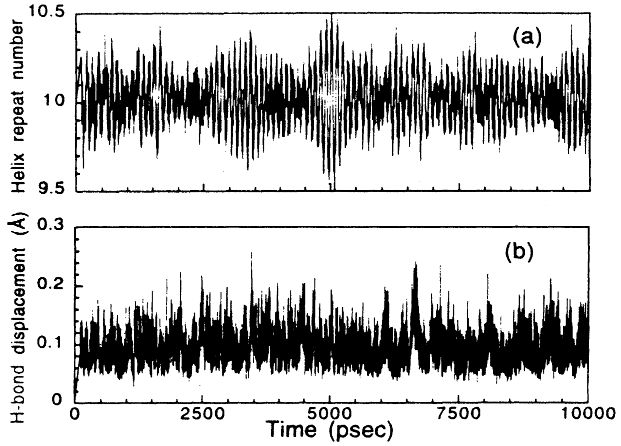

FIG. 5. (a) Helix repeat number and (b) average H-bond displacement, for 100 bp of mixed content at $T = 300$ K.

points," disorder in the model becomes apparent only after several nsec of simulation (Fig. 6). The large "steps" in Fig. 6 (also seen in the total energy and helix repeat number) might be interpreted as arising from large-amplitude distortions of the molecule or transitions of the molecule to high-energy configurations that were not easily accessed at low temperature. These transitions might involve crossing high-energy barriers and would therefore occur rarely even at high temperature. Thus, to reach thermal equilibrium in these simulations would require extremely long simulation times. For example, at temperatures above the melting point, 10-nsec simulations may only result in the melting of about 20 base pairs at both ends of the DNA helices (Fig. 7). This suggests that the unwinding kinetics of DNA double helix is rather slow, as also indicated in experiments [9]. Hence we do not suggest that our simulation results at high temperatures reflect the true equilibrium values. Indeed, the "melting temperatures" indicated here should be taken as upper bounds, as melting may possibly occur at lower temperatures over much longer time scales than were simulated.

We have also simulated B-DNA of shorter and longer sequences and obtained results similar to those reported above. In particular, we have simulated sequences of up to 400 base pairs with mixed AT and GC content and have found that the critical temperature for denaturation is about 480 K, *coinciding* with that for a 100-bp sequence. However, the time scale for denaturation seems to be longer: above the melting temperature, one must wait at least several nsec to see an apparent increase of disorder in the system (Fig. 6). A short sequence such as a dodecamer retains its stable double helix structure below 350 K, but denatures at only slightly higher temperatures, due to the chain-end effects.

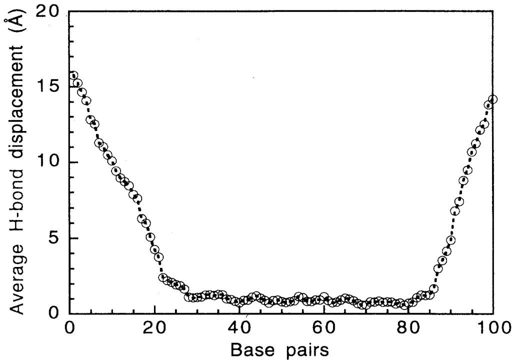

FIG. 7. Average H-bond displacement at $T = 480$ K, for simulation of a DNA fragment of 100 bp with mixed content over 10 nsec. The melting of the helix from both ends is seen.

Recently, it has been postulated that nonlinearity and inhomogeneity might play a role in energy localization process in DNA duplex [1]. In particular, simple phenomenological models were employed to investigate large-amplitude excitations in the hydrogen-bonds that connect the complementary base pairs. In the present simulation of DNA dynamics, the hydrogen bond displacements are monitored to observe possible large amplitude excitations and their dependence on base pair composition. Figure 8 shows that at temperatures below the melting point, the model DNA exhibits some large-amplitude, though short-lived, fluctuations in the H bonds and such excitations tend to repeat

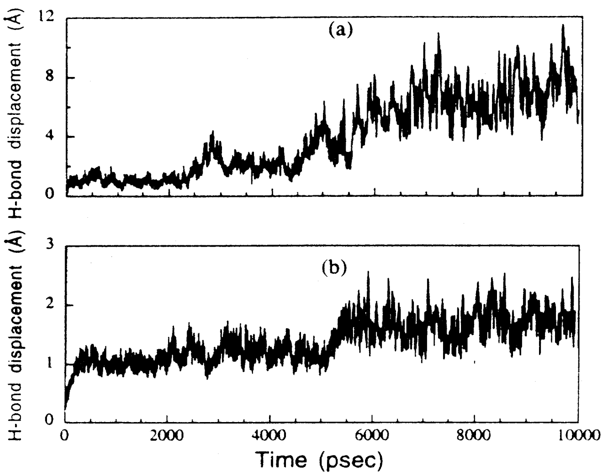

FIG. 6. Average H-bond displacement for DNA fragments with mixed content, showing the development of disorder at temperature $T = 480$ K: (a) 100 bp and (b) 400 bp. The larger the sequence, the longer the time scale required for melting to occur. Similar behavior is observed for the helix repeat number.

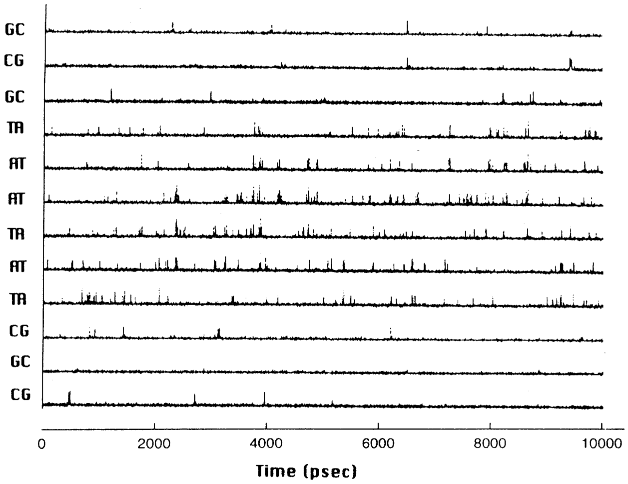

FIG. 8. Mean H-bond displacement from equilibrium for a sequence of base pairs near the "Pribnow box" **TATAAT** sequence, in the 100-bp DNA fragment with mixed content simulated at $T = 300$ K, showing that the AT pairs are more amenable to breaking. The largest amplitude in this figure is about 4 Å.

at nearby base pairs. Most interestingly, compared to *GC* pairs there are more frequent large-amplitude fluctuations in the H bonds of *AT* pairs and sequences of several contiguous *AT* and *TA* base pairs are found to be more than usually prone to hydrogen bond breaking. To analyze the nature of such large-amplitude fluctuations, we note that the length of each H bond can be expressed in terms of the dynamical variables $[X_1(n), Y_1(n), \phi_1(n); X_2(n), Y_2(n), \phi_2(n)]$, which are the coordinates of the two complementary bases. The large-amplitude excitations might be due to the interplay of these *six* variables, which are coupled strongly through the intramolecular potentials approximated by Eqs. (2)–(6). A straightforward derivation of an analytically tractable nonlinear model, taking into account all these variables, seems impossible. However, as the simulations indicate that the base pairs and the backbones are hinged flexibly about the H bonds, it might be reasonable to describe the base-pair motion with *three* characteristic degrees of freedom: the distance between the two complementary bases (rigid bodies) and the planary rotation of each base. It remains to be seen whether a simple phenomenological model based on these three degrees of freedom can be derived from the present DNA model.

IV. CONCLUDING REMARKS

The opening of base pairs in DNA is an important process since it is a prerequisite for replication, transcription, and DNA-drug binding. The kinetics of base-pair opening has been extensively studied by proton exchange experiments and the lifetime of base pairs has been determined to be of the order of milliseconds or longer [19]. However, little is known about the conformational structure of the transitory open states. Using molecular modelling, Ramstein and Lavery [21] have shown that the base-pair opening energetics are strongly coupled to DNA bending and there are many possible pathways for base-pair opening, although a preference for opening towards the major groove of the double helix was found. Figure 8 has been presented to show that the *AT* pairs are more vulnerable than *GC* pairs to H-bond stretching. The figure also indicates that the present model allows large-amplitude, though short-lived, extensions of the H bonds on the nsec time scale at temperatures around 300 K. *If* these extensions of the H bonds correspond to the "open states" that allow proton exchange with a water solvent to occur, then the model would appear to be too "loose" with respect to H-bond extension. However, the dynamics presented in Fig. 8 more closely represents a DNA molecule's internal motion in *vacuum* rather than in aqueous solution. While the modified Morse potential of Eq. (6) crudely mimics hydrogen bonding to the solvent, there is no account of viscous effects or random collisions with the solvent molecules [8, 20] in the current model. A viscous medium might well damp out the H-bond vibrations at an early stage of opening, so that the observed large-amplitude stretching motions may be exaggerated in the simple model. Of course, the lack of stacking forces and complex steric hindrance effects in our simple model may also contribute to an exaggerated H-bond extention.

However, Fig. 2 indicates that the rotational motion of the bases is qualitatively correct in some respects (see Sec. III). Moreover, Fig. 3 suggests that the magnitude of the angular fluctuations are quantitatively reasonable. Briki and co-workers [8, 20] have argued that angular fluctuations of similar magnitude are consistent with the measured rates of proton exchange. Hence the model appears to reasonably describe the gross angular motion of the bases (albeit only in the plane). Overall, however, the simple model appears to be too stiff since the onset of melting appears about 100 K above that observed. This may be primarily due to the restriction to two-dimensional motion in addition to the rather primitive description of the forces. Many very-low-frequency motions are excluded from the model, including relative motion of the strands along the helix axis, tilting, and relative twisting of the bases. Three-dimensional motion, which demands a more complete description of the forces between the stacked bases, ought to be essential in a quantitative description of melting. In this context, the effect of relatively long-range forces along the helix axis ought to be investigated.

In summary, we have investigated a relatively simple molecular model to study the dynamics of DNA molecules. The reduced model, which uses only relatively few degrees of freedom, has allowed simulations of much larger DNA sequences over much longer time scales than is possible with all-atom approaches. It has been shown that the present model can reproduce some but not all of the basic characteristics of the DNA dynamics. The model displays sharp thermal melting transitions for the DNA duplex, with melting temperature depending correctly on the base-pair composition. In addition, we have demonstrated that the complete melting transition may take place *only* over rather long time scales (say, many tens of sec) even for a DNA of just 100 base pairs. In addition, the present model provides an alternative framework in which the interplay of nonlinearity and inhomogeneity (disorder) might be studied and in which large-amplitude localized excitations might be identified in the DNA double helix. In this context, we have indeed observed localized fluctuating modes in H-bond displacements in the MD simulations, though this model appears to exaggerate their magnitude. The inhomogeneity due to the difference in the strength of H bonds between *AT* and *GC* pairs has been found to have a strong effect on energy localization processes.

ACKNOWLEDGMENT

We acknowledge the ANU Supercomputing Facility for an allocation of time on the Fijitsu VP2200 computer, on which part of the simulations was carried out.

[1] M. Peyrard and A.R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989); T. Dauxiois, M. Peyrard, and A.R. Bishop, Phys. Rev. E **47**, 684 (1993); V. Muto, P.S. Lomdahl, and P.L. Christiansen, Phys. Rev. A **42**, 7452 (1990); P.L. Christiansen and V. Muto, Physica D **68**, 93 (1993); M. Techera, L.L. Daemen, and E.W. Prohofsky, Phys. Rev. A **40**, 6636 (1989); **41** 4543 (1990); **42** 1008 (1990).

[2] J.A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).

[3] U.C. Singh, S. J. Weiner, and P. Kollman, Proc. Natl. Acad. Sci. U.S.A. **82**, 755 (1985).

[4] W.F. van Gunsteren, H.J.C. Berendsen, R.G. Geurtsen, and H.R.J. Zwinderman, Ann. N.Y. Acad. Sci. **482**, 287 (1986).

[5] S.N. Rao and P. Kollman, Biopolymers **29**, 517 (1990).

[6] S. Falsafi and N.O. Reich, Biopolymers **33**, 459 (1993).

[7] C. Prevost, S. Louise-May, G. Ravishanker, R. Lavery, and D.L. Beveridge, Biopolymers **33**, 335 (1993).

[8] F. Briki and D. Genest, J. Biomol. Struct. Dyn. **11**, 43 (1993).

[9] C.R. Cantor and P.R. Schimmel, *Biophysical Chemistry* (Freeman, San Francisco, 1980), p. 1224; R.M. Wartell and A.S. Benight, Phys. Rep. **126**, 67 (1985).

[10] S. Arnott and D.W.L. Hulins, Biochem. Biophys. Res. Commun. **47**, 1504 (1972); J. Mol. Biol. **81**, 93 (1973).

[11] W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).

[12] R.K.Z. Tan and S.C. Harvey, J. Mol. Biol. **205**, 573 (1989).

[13] M.D. Barkley and B.H. Zimm, J. Chem. Phys. **70**, 2991 (1979); P.J. Hagerman, Ann. Rev. Biophys. Biophys. Chem. **17**, 265 (1988).

[14] Y.Z. Chen and E.W. Prohofsky, Phys. Rev. E **47**, 2100 (1993); Biopolymers **33**, 797 (1993).

[15] S. Nosé, J. Chem. Phys. **81**, 511 (1984); Prog. Theor. Phys. **103**, 1 (1991); W.G. Hoover, Phys. Rev. A **31**, 1695 (1985); D.J. Evans and B.L. Holian, J. Chem. Phys. **83**, 4069 (1985).

[16] R. Kornberg, *DNA Replication* (Freeman, San Francisco, 1974), p. 242.

[17] M.A. Collins and F. Zhang, in *Nonlinear Excitations in Biomolecules*, edited by M. Peyrard (Springer-Verlag, Berlin, 1995), pp. 117–125.

[18] S.M. Englander *et al.*, Proc. Natl. Acad. Sci. U.S.A. **77**, 7222 (1980); S. Homma and S. Takeno, Prog. Theor. Phys. **72**, 679 (1984); C.T. Zhang, Phys. Rev. A **35**, 886 (1987); L.V. Yakushevich, Phys. Lett. A **136**, 413 (1989).

[19] M. Gueron, M. Kochoyan, and J.L. Leroy, Nature **328**, 89 (1987).

[20] F. Briki, J. Ramstein, R. Lavery, and D. Genest, J. Am. Chem. Soc. **113**, 2490 (1991).

[21] J. Ramstein and R. Lavery, Proc. Natl. Acad. Sci. U.S.A. **85**, 7231 (1988); J. Biomol. Struct. Dynam. **7**, 915 (1990).