

# Equilibrium and Nonequilibrium Statistical Mechanics of a Nonlinear Model of DNA

Mario Techera

*Max Planck Inst. For Biophysical Chemistry,  
Dept. of Molecular Biology,  
W-3400 Göttingen, Germany*

L.L.Daemen

*Theoretical Division, Los Alamos National Laboratory,  
Los Alamos, New Mexico 87545*

E.W.Prohofsky

*Department of Physics, Purdue University,  
West Lafayette, Indiana 47907*

Experimental and theoretical studies indicate that the hydrogen bond stretch mode dominates DNA dynamics close to denaturation temperatures. We analyze a simplified model for DNA which retains only this (nonlinear) degree of freedom. The dynamics and thermodynamics of the system are discussed. In particular, the analytical and numerical results do not exhibit a melting transition but instead a state of pseudo-equilibrium distinct from the state expected from equilibrium thermodynamics. Finally numerical results show that energy transport is unlikely at biological temperatures.

## I. THE MODEL

Deoxyribonucleic acid, or DNA, is conceivably the most important biomolecule. Its double stranded helical structure is of particular interest since the four bases (Adenine, Thymine, Guanine and Cytosine or A,T,G and C), whose sequence determines the genetic code, are projected inward toward the helix axis. On the outside of the double helix is found the backbone formed by two strands consisting of alternating phosphate groups and deoxyribose sugars. An excellent overview of DNA structure and function can be found in Saenger.<sup>1</sup> As such, the geometry of the double helix requires that the two complementary strands come apart in order for the base sequence to be read by other molecules. This melting, or denaturation of DNA has been the study of intensive experimental and theoretical investigation because of its biological importance.<sup>2</sup> In the present paper we present the motivation for a very simple model of DNA along with an analysis of its thermodynamics.

The infrared transmission spectrum of DNA has shown the existence of soft mode around  $85 \text{ cm}^{-1}$ .<sup>3</sup> This mode is seen to drop in frequency as the melting temperature is approached. Using normal mode analysis Awati was able to characterize this mode as a collective motion of the bases that stretch the interbase hydrogen bonds (HBs).<sup>4</sup> With the use of the Modified Self-Consistent Phonon Approximation, MSPA, he was able to predict the temperature dependence of this mode along with the fact that it also gained further HB stretch character as the melting temperature is approached.<sup>4</sup>

The following simplified geometry is considered for DNA; the molecule is first untwisted and each strand is then represented by a set of point masses (the nucleotides) connected by linear springs. The intrastrand interactions (*i.e.* the HBs between base pairs) are modeled by a Morse potential. Schematically this can be represented as in Fig.1(a). The displacement from equilibrium of the  $n^{\text{th}}$  mass point is denoted by  $u_n(v_n)$  in the top (bottom) chain respectively. Only transverse motions are considered. The equations of motion for  $u_n$  and  $v_n$ :

$$m\ddot{u}_n = k(u_{n+1} + u_{n-1} - 2u_n) - \frac{\partial\phi}{\partial(u_n - v_n)}, \quad (1)$$

$$m\ddot{v}_n = k(v_{n+1} + v_{n-1} - 2v_n) + \frac{\partial\phi}{\partial(u_n - v_n)}, \quad (2)$$

where  $\phi$  is the non-linear potential describing the HB interaction. At this point, it should be emphasized again that the source of the nonlinearity in the model lies in the coupling between the strands not between adjacent particles on the same strand.

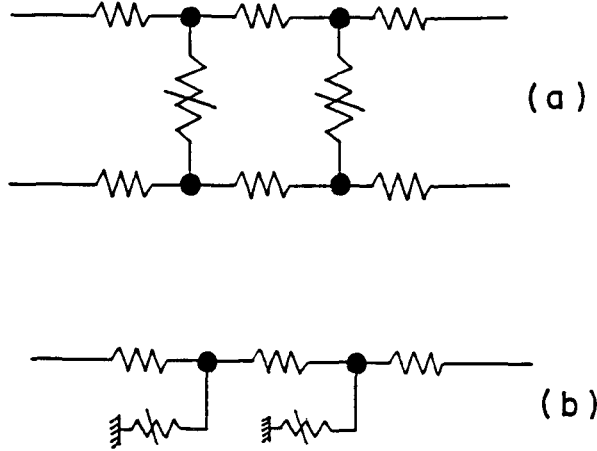


Figure 1

Since the individual masses of the four different nucleotides differ by at most 13%, the masses of the particles on each strand have been made equal in the above equations.<sup>1</sup> Furthermore, for the sake of simplicity, it has been assumed that the force constants  $k$  are the same throughout the chain. These assumptions permit a transformation to center of mass coordinates:

$$x_n = \frac{1}{\sqrt{2}}(u_n + v_n), \quad (3)$$

$$y_n = \frac{1}{\sqrt{2}}(u_n - v_n), \quad (4)$$

Eqns.(1) and (2) become:

$$m\ddot{x}_n = k(x_{n+1} + x_{n-1} - 2x_n), \quad (5)$$

$$m\ddot{y}_n = k(y_{n+1} + y_{n-1} - 2y_n) - \frac{\partial\phi}{\partial y_n}. \quad (6)$$

$x_n$  is the motion of the center of mass and  $y_n$  describes the motion about the center of mass (a positive  $y_n$  represents a stretch). The potential energy  $\phi$  is chosen to model the nonlinear HB interaction between base pairs. This is typically done by using the Morse potential:

$$\phi_M(y_n) = V_\infty(1 - e^{-\sqrt{2}ay_n})^2, \quad (7)$$

which, with suitably chosen parameters, can provide a good description of HB potentials in DNA. These equations may be obtained from the following Hamiltonian:

$$\begin{aligned} H &= \sum_n \frac{p_n^2}{2m} + \frac{k}{2}(x_{n+1} - x_n)^2 + \frac{q_n^2}{2m} + \frac{k}{2}(y_{n+1} - y_n)^2 + V_\infty[1 - \exp(-\sqrt{2}ay_n)]^2 \\ &\equiv H_x + H_y. \end{aligned} \quad (8)$$

We note here for future reference, that  $H_y$  can be viewed as an ensemble of Morse oscillators with nearest-neighbour harmonic coupling in the displacements and that in the limit  $k \rightarrow 0$  we have an ensemble of independent Morse oscillators. The equations of motion are now uncoupled. Eqn.(5) represents a pure harmonic lattice with plane wave solutions. In what follows, all the attention will be focused on the motion about the center of mass Eqn.(6). This equation can also be viewed as describing longitudinal displacements in the one-dimensional chain shown in Fig.1(b). An analysis of the dynamics of this model has been presented elsewhere.<sup>5</sup>

## II. EQUILIBRIUM THERMODYNAMICS

An initial attempt at calculating the thermodynamics for this model was presented by Peyrard and Bishop in 1989.<sup>6</sup> Their idea was to apply the transfer integral method to evaluate the partition function in the canonical ensemble.<sup>7</sup> Here we will show that strictly speaking this approach is incorrect due to the fact that the Morse potential is bounded for large stretches. To make this statement as clear as possible we consider first an ensemble of independent Morse oscillators i.e.  $H_y$  of Eqn.(8), in the limit  $k \rightarrow 0$ .  $H_y$  then becomes:

$$H_y = \sum_n \frac{q_n^2}{2m} + V_\infty[1 - \exp(-\sqrt{2}ay_n)]^2. \quad (9)$$

In order to calculate the thermodynamic properties in the canonical ensemble, the usual procedure is to calculate the partition function  $Z_y$ :<sup>8</sup>

$$Z_y = \frac{1}{h^N} \int_{-\infty}^{+\infty} \prod_{n=1}^N dy_n dq_n \exp(-\beta H_y), \quad (10)$$

$$= \frac{1}{h^N} \left( \frac{2\pi m}{\beta} \right)^{\frac{N}{2}} I^N(-\infty, +\infty), \quad (11)$$

where

$$I(-\infty, +\infty) \equiv \int_{-\infty}^{+\infty} dy_n \exp \left\{ -\beta V_\infty [1 - \exp(-\sqrt{2}ay_n)]^2 \right\}, \quad (12)$$

and where  $N$  is the number of particles,  $\beta = 1/k_B T$  and the integration is performed over all the available phase space *i.e.*  $y_n \in (-\infty, +\infty)$ . Everything appears to be in order until one notices that for  $y_n \rightarrow +\infty$  the integrand in Eqn.(12) is *bounded*. In other words, *the integral  $I(-\infty, +\infty)$  and consequently the partition function  $Z_y$  are divergent!* Furthermore, if one considers

$$\lim_{y_u \rightarrow +\infty} I(-\infty, y_u), \quad (13)$$

it can be shown that the divergence is linear in  $y_u$ , meaning that the integral diverges linearly with volume. This result *per se* may not be distressing since the partition function has no direct physical meaning, however on further examination the consequences become clear. Consider the average displacement from the Morse well,  $\langle y_n \rangle$  (which for nonzero values of  $k$  in Eqn.(8) represents the average HB stretch of a base pair in the molecule), for a given particle in the ensemble:

$$\begin{aligned} \langle y_n \rangle &= \frac{1}{Z_y h^N} \int_{-\infty}^{+\infty} \prod_{n=1}^N dy_n dq_n y_n \exp(-\beta H_y) \\ &= \left[ \frac{\int_{-\infty}^{+\infty} dy_n y_n \exp \left\{ -\beta V_\infty [1 - \exp(-\sqrt{2} a y_n)]^2 \right\}}{I(-\infty, +\infty)} \right]^N. \end{aligned} \quad (14)$$

The numerator in the above expression is also infinite but diverges as the square of the volume and the denominator has the linear volume divergence mentioned previously. The average position of a particle in the ensemble is then  $\langle y_n \rangle = +\infty$ . *Thus the particles are at equilibrium only when they have escaped the well.* This interpretation can be made more rigorous by considering finite upper boundaries on the integrals and taking the limit of the boundary to  $+\infty$  as in Eqn.(13). As long as the boundary remains finite, so does the quantity  $\langle y_n \rangle$ , but this quantity diverges in the limit of a boundary at infinity. We thus conclude that a meaningful physical treatment of such Hamiltonians in the canonical ensemble requires the explicit introduction of a cutoff in the integral which of course one hopes can be interpreted physically.

The Morse oscillator illustrates a much more general feature of systems of particles evolving under the influence of long-range, unscreened forces which asymptotically tend to a finite value. Considerable care must be exerted in applying the methods of statistical mechanics to these systems.<sup>9</sup> In particular, concepts such as thermodynamical equilibrium, statistical ensemble, thermodynamic limit, and the meaning of averages of various thermodynamic quantities must be thoroughly examined. In particular, several conceptual difficulties arise when the methods of statistical mechanics are applied in the canonical ensemble. Other physical systems of interest also involve long-range forces. Recently, an extensive study of the statistical mechanics of gravitating systems was performed by Padmanabhan.<sup>10</sup> Many of the conclusions he drew for gravitating systems are similar to our results for the Morse potential.<sup>11</sup>

One can now go back to the full Hamiltonian  $H_y$  as defined in Eqn.(8) and show that it exhibits the same problem. The partition function is given by:

$$Z_y = \frac{1}{h^N} \left( \frac{2\pi m}{\beta} \right)^{N/2} \int_{-\infty}^{+\infty} \prod_{n=1}^N dy_n \exp \left\{ -\beta \left[ \frac{k}{2} (y_{n+1} - y_n)^2 + V_\infty [1 - \exp(-\sqrt{2} a y_n)]^2 \right] \right\} \quad (15)$$

which again is not finite, since the integrand does not vanish at infinity. In particular, one can easily verify this from the above equation by considering the line  $y_1 = y_2 = \dots = y_n$  as  $y_n \rightarrow +\infty$ . A dramatic consequence of this divergence, from the point of view of the model is that the the average stretch of a base pair is infinite at all nonzero temperatures at equilibrium. The implication is that the thermal equilibrium state of the the chain, is one with all the bonds stretched to infinity. This is a direct consequence of the boundedness of the Morse potential as stretches become large (even infinite). This straightforward analysis does not, however, give us an idea of the time it takes for the stretches to become infinite, a question which will also be analyzed further in the next section.

For the sake of completion, we present in Fig.2 the results for the average HB stretch,  $\langle y_n \rangle$ , as a function of temperature, in the thermodynamic limit for various values of the cutoff,  $y_u$ , i.e. the Morse potential is valid for  $y_n < y_u$  and there is an infinite barrier for stretches greater than  $y_u$ . These results were calculated by solving numerically a transfer integral equation.<sup>11</sup> We stress that the transfer integral method is valid for this calculation only because we have added an explicit cutoff. The depth of the well,  $V_\infty$  is set to 0.2eV,  $k = 0.3\text{eV}/\text{\AA}^2$  and  $a = 2.77\text{\AA}^{-1}$  (this puts the inflexion point of the Morse at about  $0.2\text{\AA}$ ) which is a typical parametrization of the Morse potential.<sup>12</sup>

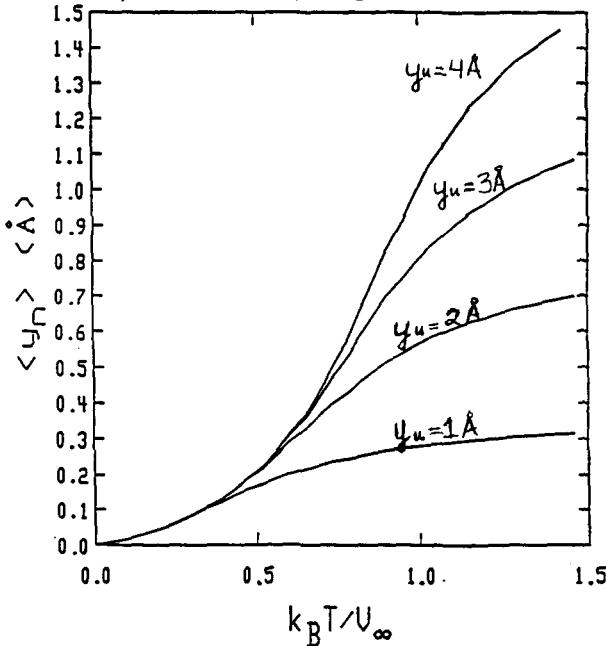


Figure 2

### III. NONEQUILIBRIUM THERMODYNAMICS

Based on the aforementioned results in the canonical ensemble and as a consequence of treating the whole Morse potential with no cutoff, one can then ask the following question: *If we don't restrict the Morse potential with a cutoff how much time is*

required for this divergence in the average HB stretch to occur at some temperature  $T$ , given some generic initial condition of the chain? Of course the obvious answer is an infinite amount of time since the particle must travel to  $+\infty$ . One can, nevertheless, define some finite stretch as representing dissociation, in which case the required time is finite. In the present model, this stretch can be identified with the stretch required for imino proton exchange in DNA to occur. By the argument presented in the last section, the above question is one of nonequilibrium thermodynamics.

In order to answer this question, one approach consists in using the Langevin equations.<sup>13</sup> The equations of motion are modified to include a heat bath with which the system is in contact and which is at a constant temperature (this can be associated to the solution in which the molecule is typically found). We add two additional terms to simulate stochastic collisions and dynamical friction of the molecule with the environment.<sup>14</sup> Eqn.(6) now becomes a system of stochastic differential equations given by:

$$m\ddot{y}_n = k(y_{n+1} + y_{n-1} - 2y_n) - \frac{\partial\phi}{\partial y_n} - m\gamma\dot{y}_n + \eta_n(t). \quad (16)$$

Characteristic of such equations, the influence of the surrounding medium is split into two parts: 1) a systematic term  $-m\gamma\dot{y}_n$  which causes the friction and 2) a fluctuating part  $\eta_n(t)$  which represent the 'random' collisions of the constituents of the medium with the system in question. This fluctuating term will be assumed to possess certain properties, namely:

$$\langle \eta_n(t) \rangle = 0 \quad (17)$$

$$\langle \eta_n(t_1)\eta_m(t_2) \rangle = q\delta(t_1 - t_2)\delta_{nm}, \quad (18)$$

where  $q = 2mk_B T\gamma$ .<sup>15</sup> The details of how this system is solved numerically along with further results are presented elsewhere. <sup>11</sup> The lattice consists of  $N = 125$  particles with circular boundary conditions *i.e.*  $y_1 = y_{126}$ . The dissipative constants are assumed to be equal at all the lattice sites with a value of  $\gamma^{-1} = 20\tau$ , where  $\tau$  is the small amplitude period of the Morse oscillators. Thus, the system will be underdamped and not dominated by friction, otherwise, the parameters will be taken to be similar to those used in the previous section.

The procedure to be followed will be to start the lattice at its dynamical equilibrium position *i.e.* with all particles at the bottom of their respective Morse wells, with no kinetic energy, and then to evolve the system according to Eqns.(16). In all runs, the average kinetic energy of a base pair on the lattice,  $\langle K_n \rangle$ , always thermalized to  $\frac{1}{2}k_B T$  as expected (after about 50ps). For the above parametrization of the Morse potential, a temperature of  $T_c = 4642K$  would correspond to the average kinetic energy being equal to the depth of the Morse well, nevertheless, at temperatures well below  $T_c$  it was possible to observe the system denature in a brief period of time as seen in Figure 3(a) which plots the average HB stretch vs time for a  $T = 3000K$ . The vertical scale is important since the inflexion point of the Morse is at about  $y_n = 0.2\text{\AA}$  as mentioned above. Our calculations have shown that the lower the temperature the longer the required time for the divergence discussed in the previous section to become

evident. In fact, at  $T = 300K$ , room temperature, we performed several runs of  $10ns$  without ever observing any such divergence. Figure 3(b) shows a portion of such a run. Once again notice the vertical scale. We can therefore see that even though the system is not in equilibrium, that at low temperatures the system behaves as if it were in a *pseudo-equilibrium* because its average kinetic energy is at the right value but other quantities, such as the average HB stretch  $\langle y_n \rangle$ , are not.

Finally it should be mentioned that the Langevin approach was used to study the effects of thermalization on the dynamics of this model that had been observed previously, particularly on the quasi-solitonic modes that exist.<sup>5</sup> The essential effect was that for realistic Morse parameters, nonlinear wave propagation was strongly hindered at temperatures above  $10K$ .<sup>11</sup>

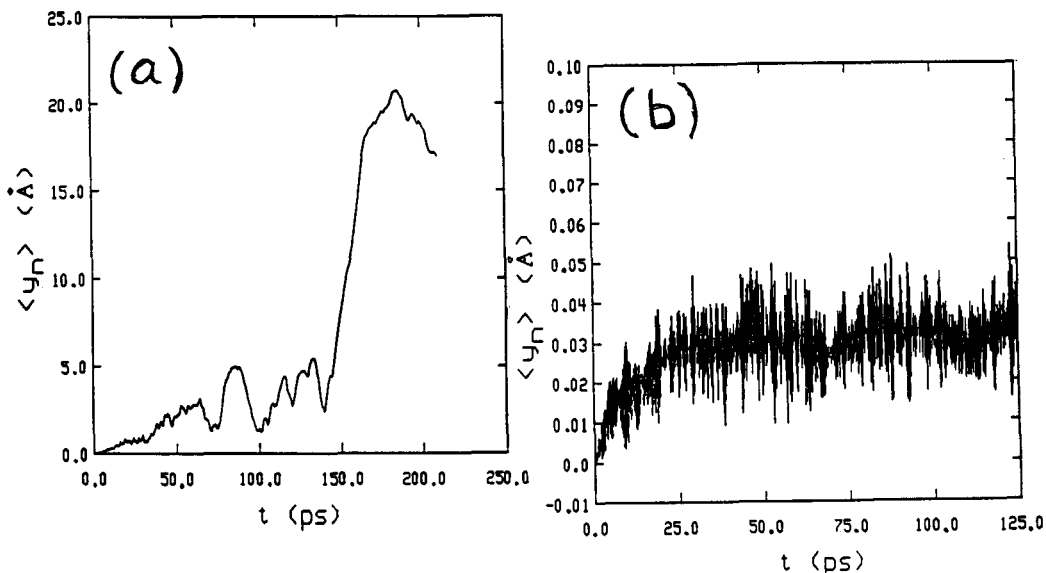


Figure 3

#### IV. CONCLUSION.

We have shown that the thermodynamics of this simple DNA model can be treated meaningfully in the canonical ensemble, even though the partition function diverges - one just has to interpret carefully the nature of the thermodynamic equilibrium state. Furthermore the denaturation predicted for all temperatures has been shown to be directly related to the boundedness of the Morse potential which represents the HB interaction. This problem has been encountered recently in the study of Coulomb gases<sup>9</sup> and the statistical mechanics of gravitating systems.<sup>10</sup> We have presented two ways of circumventing this problem: 1) avoid the divergences by putting a cutoff to the Morse potential i.e. allow only for a certain maximum stretch 2) accept the divergence as a final equilibrium state in the average stretch for all  $T$  and study how the system in a nonequilibrium state evolves to this final state. In the latter case, we presented

briefly results on the dynamical approach to equilibrium and showed that the time required for the bases to separate can be very long for low temperatures.

DNA is known to undergo a structural phase transition during thermal denaturation<sup>1</sup>, but this model showed no such clear cut behaviour at any particular temperature. We can therefore conclude that more physics needs to be incorporated such as next-to-nearest neighbour coupling and the helical twist. The HBs alone seem to account for energy localization due to nonlinearity and that might be an important precursor to thermal denaturation, but the transition itself requires that more phenomenology be added to such a model.

- <sup>1</sup>W.Saenger, 'Principles of Nucleic Acid Structure'(S pringer-Verlag,New York,1983).
- <sup>2</sup>V.Muto,J.Halding,P.L.Christiansen and A.C.Scott, J.Biom.Struct.Dyn. 5,873 (1988);V.Muto, A.C.Scott and P.L. Christiansen, Phys.Let.A 136,33 (1989); S.Yomosa,Phys.Rev.A 27,2120 (1983);S.Takeno,Prog.Theor.Phys.71,395 (1984); C.Zhang,Phys.Rev.A 35, 886 (1987).
- <sup>3</sup>J.W.Powell,G.S.Edwards,L.Genzel,F.Kremer,A.Wittlin, W.Kubasek and W.Peticolas, Phys.Rev.A 35, 9 (1987).
- <sup>4</sup>K.Awati, Ph.D. Thesis, Purdue University, West Lafayette, Indiana, USA (1989).
- <sup>5</sup>M.Techera, L.L.Daemen and E.W.Prohofsky, Phys.Rev.A 40, 6636 (1989);M.Techera, L.L.Daemen and E.W.Prohofsky, Phys.Rev.A 41, 4543 (1990);M.Techera, L.L.Daemen and E.W.Prohofsky, Phys.Rev.A 42, 1008 (1990).
- <sup>6</sup>M.Peyrard and A.R.Bishop, Phys.Rev.Let. 62, 2755 (1989).
- <sup>7</sup>D.J.Scalapino,M.Sears,and R.A.Ferrel, Phys.Rev.B 11,3535(1975).
- <sup>8</sup>K.Huang, 'Statistical Mechanics', 2<sup>nd</sup> Edition (Wiley,Chichester,1987).
- <sup>9</sup>M.K.H.Kiessling, J.Stat.Phys.59, 1157, (1990).
- <sup>10</sup>T.Padmanabhan, Phys.Reports, 188, 285 (1990).
- <sup>11</sup>M.Techera, Ph.D. Thesis, Purdue University, West Lafayette, Indiana, USA (1991).
- <sup>12</sup>R.M.Wartell and A.S.Benight, Phys.Rep.126, 67 (1985).
- <sup>13</sup>J.McLennan,'Introduction to Non-Equilibrium Statistical Mechanics' (Prentice Hall,New Jersey,1989).
- <sup>14</sup>S.Chandrasekhar, Rev.Mod.Phys. 15,1 (1943).
- <sup>15</sup>T.Schneider and E.Stoll, Phys.Rev.B 17, 1302 (1978).