Base-pair opening and bubble transport in a DNA double helix induced by a protein molecule in a viscous medium

V. Vasumathi* and M. Daniel†

Centre for Nonlinear Dynamics, School of Physics, Bharathidasan University, Tiruchirappalli 620 024, India

(Received 29 October 2008; revised manuscript received 3 November 2009; published 9 December 2009)

The protein-DNA interaction dynamics is studied by modeling the DNA bases as classical spins in a coupled spin system, which are bosonized and coupled to thermal phonons and longitudinal motion of the protein molecule in the nonviscous limit. The nonlinear dynamics of this protein-DNA complex molecular system is governed by the completely integrable nonlinear Schrödinger (NLS) equation which admits *N*-soliton solutions. The soliton excitations of the DNA bases in the two strands make localized base-pair opening and travel along the DNA chain in the form of a bubble. This may characterize the bubble generated during the transcription process, when an RNA polymerase binds to a promoter site in the DNA double helical chain. When the protein-DNA molecular system interacts with the surrounding viscous solvating water medium, the dynamics is governed by a perturbed NLS equation. This equation is solved using a multiple scale perturbation analysis, by treating the viscous effect as a weak perturbation, and the results show that the viscosity of the solvent medium damps out the soliton as time progresses.

DOI: 10.1103/PhysRevE.80.061904 PACS number(s): 87.15.−v, 05.45.Yv, 02.30.Jr

I. INTRODUCTION

Cellular processes, such as gene expression, suppression, replication, transcription, recombination, repair, and several other processes start with the binding of an enzyme protein to the DNA. For instance, the transcription process starts with the binding of an RNA polymerase at a promoter site of the DNA, and this binding is known to change the conformation of DNA by opening the bases [1]. Experimentally, dynamic protein-DNA interaction has been directly visualized under physiological conditions with atomic force microscope [2]. The active role played by protein molecule in flipping the target base out of the DNA helix can be understood through NMR measurements [3]. Base flipping is a biphasic process with very fast flipping and slow binding of the flipped base into the active site of the protein (enzyme) [4]. The dynamic force microscopy combined with unzipping force analysis of protein association distinguishes protein-DNA complexes on a site-specific single-molecule basis [5]. A simple RNA-polymerase molecule can rotate the DNA for more than 100 revolutions over thousands of base pairs [6]. Interaction of protein with DNA facilitates and stabilizes the base flipping process, which occurs via the major groove of the DNA [7] and also through an intermediate state [8]. In the theoretical front, results obtained through free-energy calculations based on molecular-dynamics simulations suggest that sequence-specific recognition of protein is linked to the flipping event [9]. When a protein molecule binds to the DNA, it may constrain the DNA twist over a length and the local torsional stress in DNA can eject the DNA-bound protein [10]. DNA-bending proteins cause local DNA untwisting and generate an increase in the entropic elastic stiffness [11]. The recent results in the framework of equilibrium statistical mechanics of the protein-DNA system based on coarse-grained and wormlike chain models describe unzipping of the DNA chain through DNA-bending and protein binding fluctuations [11–14]. In spite of these developments, the nonlinear dynamics of protein-DNA interaction has not been understood well because of its high complex nature. Even the knowledge of nonlinear dynamics of DNA through solitonlike excitations of bases, which were based on rotational [15–22] and translational (longitudinal and transverse) [23–26] motions of bases, describing base flipping is very limited. Even though these solitonic excitations representing base-pair opening are capable of propagating for a very long distance and time, thermal fluctuation [27,28] and viscosity of the surrounding medium [29–34] damp the solitons and unzipping by limiting their propagation. The impact of regulatory proteins through hydrogen bonds on breather excitations in DNA was studied by Sataric and Tuszynski [35] by considering Davydov's model of amide-I vibration for protein dynamics [36,37] and Peyrard-Bishop's model of hydrogen bonds stretching for DNA dynamics [23], and they found that binding of protein to DNA generates breather soliton spontaneously. In the present paper, we examine the base flipping or base-pair opening in a DNA molecule interacting with a protein at the physiological temperature in a surrounding viscous medium by solving the underlying nonlinear dynamics of the protein-DNA complex. We use a different mechanism for protein-DNA interaction dynamics, by mapping the base pairs of the DNA to classical spins in a coupled spin system or spin ladder, and the amino acids of the protein molecule as a collection of mass points in a linear chain. Starting with the Heisenberg model of the Hamiltonian for the coupled spin system, we express it in terms of the rotational angles of DNA bases and the longitudinal motion of the protein molecule. The dynamics is then derived in the form of coupled nonlinear evolution equations after bosonizing the Hamiltonian, which finally reduces to the completely integrable nonlinear Schrödinger (NLS) equation that admits *N*-soliton solutions. The viscosity of the surrounding medium is treated as a perturbation to the above completely

*vasu@cnld.bdu.ac.in
†daniel@cnld.bdu.ac.in

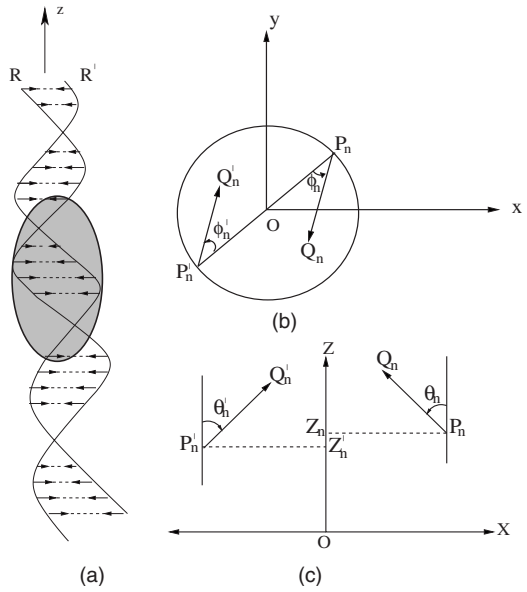

FIG. 1. (a) A schematic representation of DNA double helix with a specific segment (shaded region) showing binding of protein molecule. (b) A horizontal projection of the nth base pair of the DNA in the xy plane. (c) A horizontal projection of the nth base pair of the DNA in the xz plane.

integrable soliton equation. The paper is organized as follows. In Sec. II, the details of the model and the Hamiltonian for the DNA molecular chain are presented. The model Hamiltonian for protein-DNA interaction is given in Sec. III. In Sec. IV, we derive the equation of motion for the protein-DNA molecular system in the continuum limit and, in Sec. V, the dynamical equation is reduced to the NLS equation, the soliton solution of which represents opening of DNA-base pairs in a nonviscous medium. The effect of viscosity of the surrounding solvent medium, on the base-pair opening in DNA, is understood through a multiple scale perturbation analysis in Sec. VI. The results are concluded in Sec. VII.

II. MODEL HAMILTONIAN FOR DNA DYNAMICS

We consider the B form of a DNA double helical chain and a protein molecule (say RNA polymerase) at physiological temperature in a surrounding viscous medium and study the impact of the protein on the DNA molecule during the dynamical process. The model, we propose here, treats DNA as a set of two coupled linear molecular chains and treats protein as a single linear molecular chain interacting with the DNA through a linear coupling. A schematic representation of this protein-DNA molecular system is shown in Fig. 1(a). In the figure R and R' represent the two complementary strands of the DNA double helix. Each arrow represents the direction of the base attached to the strands, and the dots between arrows represent the net hydrogen bonding effect between the complementary bases. The shaded ellipse overlapping the DNA double helical structure represents the region where the protein molecule can bind to a specific segment of the DNA molecule. The conformation and stability of DNA double helix is mainly determined by the stacking of bases through intrastrand dipole-dipole interaction and through interstrand hydrogen bonds between the complementary bases. From a heuristic argument, it is assumed that the hydrogen bonding energy between the complementary bases depends on the distance between them. Generally, the distance between complementary bases can be expressed through longitudinal, transverse, and rotational motions of bases. Among them, the rotational motion of bases is found to contribute more toward the opening of bases pairs. Hence, it is appropriate to consider a plane-base rotator model for DNA [16,18], which the authors have extensively used in the study of pure DNA dynamics in the recent years [21,22]. In Figs. 1(b) and 1(c), the horizontal projections of the nth base pair in the xy and xz planes are presented, respectively. In the figures, Q_n and Q'_n denote the tip of the nth bases belonging to the complementary strands R and R' at P_n and P'_n, respectively, and $\theta_n(\theta'_n)$ and $\phi_n(\phi'_n)$ represent the angles of rotation of the bases in the xz and xy planes, respectively. By using the simple geometry in Figs. 1(b) and 1(c), the distance between the tips of bases is written as [19]

$$\begin{aligned}(Q_nQ'_n)^2 = {} & 2 + 4r^2 + (z_n - z'_n)^2 + 2(z_n - z'_n)(\cos\theta_n - \cos\theta'_n) \\ & - 4r[\sin\theta_n \cos\phi_n + \sin\theta'_n \cos\phi'_n] \\ & + 2[\sin\theta_n \sin\theta'_n(\cos\phi_n \cos\phi'_n + \sin\phi_n \sin\phi'_n) \\ & - \cos\theta_n \cos\theta'_n], \end{aligned} \tag{1}$$

where "r" is the radius of the circle depicted in Fig. 1(b).

The hydrogen bonding energy can be understood in a more clear and transparent way by introducing quasispin operators $\mathbf{S_n} \equiv (S_n^x, S_n^y, S_n^z) = (\sin\theta_n \cos\phi_n, \sin\theta_n \sin\phi_n, \cos\theta_n)$ and $\mathbf{S'_n} \equiv (S_n'^x, S_n'^y, S_n'^z) = (\sin\theta'_n \cos\phi'_n, \sin\theta'_n \sin\phi'_n, \cos\theta'_n)$ and using this Eq. (1) upon choosing $z_n = z'_n$ can be rewritten as

$$(Q_nQ'_n)^2 = 2 + 4r^2 + 2[S_n^x S_n'^x + S_n^y S_n'^y - S_n^z S_n'^z] - 4r[S_n^x + S_n'^x]. \tag{2}$$

It is interesting to note that the above form of $(Q_nQ'_n)^2$ matches with the Hamiltonian for a generalized Heisenberg spin model. This suggests that the intrastrand base-base interaction or stacking energy in DNA also can be written using the same consideration. It is reasonable to think that, if such a quasispin model can be used for this problem, the DNA molecule with two strands and runglike base pairs can be conceived as a coupled spin chain or a spin ladder system.

With the above consideration, we are at liberty to use the following Heisenberg model of the Hamiltonian for a two coupled spin chain model or spin ladder system for DNA:

$$H_D = -\sum_n [J(\mathbf{S_n} \cdot \mathbf{S_{n+1}} + \mathbf{S'_n} \cdot \mathbf{S'_{n+1}}) + \mu(\mathbf{S_n} \cdot \mathbf{S'_n})]. \tag{3}$$

In the case of spin ladder system $\mathbf{S_n}$ and $\mathbf{S'_n}$ represent the spins at the lattice site "n" in the two legs, with ferromagnetic coupling among the neighboring spins and ferromagnetic or antiferromagnetic rung coupling between spins in the two legs. Thus, the DNA double helical chain is mapped onto a two coupled spin chain model or a spin ladder system with ferromagnetic legs ($J > 0$) and ferromagnetic ($\mu > 0$) or

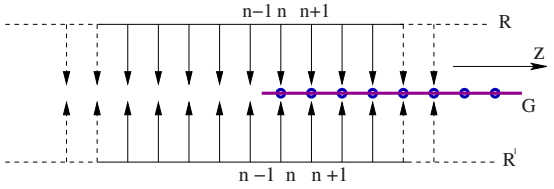

FIG. 2. (Color online) A sketch representing the protein-DNA molecular system.

antiferromagnetic ($\mu<0$) rungs. Therefore, in Hamiltonian (3), the terms proportional to J correspond to the stacking interaction between the nth base and its nearest neighbors in the two strands, and the last term which is proportional to μ corresponds to the interstrand interaction or hydrogen bond energy between the complementary bases. In equilibrium, the parameter μ is expected to be less than zero. As DNA functions at the biological temperature, the hydrogen atom attached to the bases are normally in a thermally excited state. Therefore, it is necessary to generalize the above model for a thermal DNA. Thermal phonons can be included in the system by adding the following forms of the Hamiltonian:

$$H_T = \sum_n \left[\frac{p_n^2}{2m_1} + k_1(X_n - X_{n+1})^2 \right], \tag{4a}$$

$$H_{D-T} = \alpha_1 \sum_n (X_{n+1} - X_{n-1})(\mathbf{S_n} \cdot \mathbf{S'_n}), \tag{4b}$$

where $p_n = m_1 \dot{X}_n$, with overdot representing the time derivative. The Hamiltonian (4a) corresponds to pure thermal phonons, where k_1 is the elastic constant and m_1 is the mass of the hydrogen atom attached to the base. X_n represents the displacement of the bases at the nth site along the direction of the hydrogen bond. The interaction Hamiltonian H_{D-T} given in Eq. (4b) corresponds to the coupling between the oscillation of the above hydrogen atom due to thermal fluctuation and the rotation of bases.

III. MODEL HAMILTONIAN FOR PROTEIN-DNA INTERACTION

When a protein molecule binds to DNA, it induces large mechanical stress on it, which causes conformational change in DNA. Even though protein is much larger in size than the DNA molecule, only a small portion of the protein molecule, namely, the active site is directly interacting with the DNA molecule. Hence, we treat the short active site region of the protein molecule that interacts with the DNA as a linear chain as has been considered earlier (see, e.g., Sataric *et al.* [38]). Based on this, we propose the model for the protein-DNA molecular system by treating DNA as a set of two coupled linear chains and the active site region of the protein molecule which slides on the DNA as a linear molecular chain interacting with the bases as shown schematically in Fig. 2. In the figure, the indices n and $n \pm 1$ represent the nth and $(n \pm 1)$th sites for the bases [similar to P_n, P'_n, $P_{n\pm1}$, and $P'_{n\pm1}$ in Figs. 1(b) and 1(c)]. G represents the active site region of the linear segment of the protein molecule that interacts with the bases of DNA. Specifically, the linear segment of the protein molecule is treated as a collection of mass points, with each mass point representing a peptide unit and connected by linear springs. It is assumed to exhibit longitudinal stretching parallel to the helical axis of the DNA, which couples linearly with the hydrogen bonds between bases. Hence, the model Hamiltonians for the longitudinal stretching motion of the protein molecule (H_P) and that for the coupling to the DNA chain (H_{D-P}) is written, respectively, as

$$H_P = \sum_n \left[\frac{q_n^2}{2m_2} + k_2(y_n - y_{n+1})^2 \right], \tag{5a}$$

$$H_{D-P} = \alpha_2 \sum_n (y_{n+1} - y_{n-1}) S_n^z S_n'^z, \tag{5b}$$

where $q_n = m_2 \dot{y}_n$ and m_2 is the mass of the peptide. y_n denotes the displacement of the nth peptide in the protein chain from its equilibrium position and k_2 represents the elastic constant associated with the small amplitude oscillation of the protein molecule. The interaction Hamiltonian H_{D-P} represents the change in hydrogen bonding energy, which is due to the impact of the linearly coupled oscillations of the peptide units of the protein molecule. In Eq. (5b), α_2 is the coupling coefficient. As the protein molecule is assumed to slide along the DNA chain, the interaction energy along the direction of the helical axis (z direction) is expected to be dominant over the energy in the normal xy plane [see Eq. (5b)]. Thus, the total Hamiltonian for our model can be written using Eqs. (3), (4a), (4b), (5a), and (5b) as

$$\begin{aligned} H &= H_D + H_p + H_{D-p} + H_T + H_{D-T} \\ &= \sum_n \Big[-\{J(\mathbf{S_n} \cdot \mathbf{S_{n+1}} + \mathbf{S'_n} \cdot \mathbf{S'_{n+1}}) + \mu(\mathbf{S_n} \cdot \mathbf{S'_n})\} \\ &\quad + \frac{p_n^2}{2m_1} + \frac{q_n^2}{2m_2} + k_1(X_n - X_{n+1})^2 + k_2(y_n - y_{n+1})^2 \\ &\quad + \alpha_1(X_{n+1} - X_{n-1})(\mathbf{S_n} \cdot \mathbf{S'_n}) + \alpha_2(y_{n+1} - y_{n-1}) S_n^z S_n'^z \Big]. \end{aligned} \tag{6}$$

The above Hamiltonian can be expressed in terms of the angles of rotation of bases of DNA as

$$\begin{aligned} H = \sum_n \Big[& -J\{\sin\theta_n \sin\theta_{n+1} \cos(\phi_{n+1} - \phi_n) \\ & + \cos\theta_n \cos\theta_{n+1} + \sin\theta'_n \sin\theta'_{n+1} \cos(\phi'_{n+1} - \phi'_n) \\ & + \cos\theta'_n \cos\theta'_{n+1}\} - [\mu - \alpha_1(X_{n+1} - X_{n-1})] \\ & \times \{\sin\theta_n \sin\theta'_n \cos(\phi_n - \phi'_n) + \cos\theta_n \cos\theta'_n\} \\ & + \frac{p_n^2}{2m_1} + \frac{q_n^2}{2m_2} + k_1(X_n - X_{n+1})^2 + k_2(y_n - y_{n+1})^2 \\ & + \alpha_2(y_{n+1} - y_{n-1}) \cos\theta_n \cos\theta'_n \Big]. \end{aligned} \tag{7}$$

In spin systems, when the spin value is large, the dynamics is investigated under classical or semiclassical approximation by bosonizing the Hamiltonian (see, e.g., Ref. [39]). Also, creation and annihilation operators were used to represent excitations in the dynamics of hole transfer in short fragments of DNA [40,41] and in the dynamics of alpha helical protein based on the model proposed by Davydov [36,37]. Therefore, to understand the underlying nonlinear dynamics of the protein-DNA molecular system, we bosonize Hamiltonian (6) using Holstein-Primakoff (H-P) representation [42] by writing $S_n^+ = \sqrt{2}[1-\epsilon^2 a_n^\dagger a_n]^{1/2}\epsilon a_n$, $S_n^- = \sqrt{2}\epsilon a_n^\dagger[1-\epsilon^2 a_n^\dagger a_n]^{1/2}$, and $S_n^z = [1-\epsilon^2 a_n^\dagger a_n]$, where $S_n^\pm = S_n^x \pm i S_n^y$. In the low-temperature limit, $a_n^\dagger a_n \ll 2S$, and hence the H-P transformation can be expanded in a power series in terms of the parameter $\epsilon = 1/\sqrt{S}$ as

$$S_n^+ = \sqrt{2}\epsilon\left[1 - \frac{\epsilon^2}{4}a_n^\dagger a_n - \frac{\epsilon^4}{32}a_n^\dagger a_n a_n^\dagger a_n - O(\epsilon^6)\right]a_n, \quad (8a)$$

$$S_n^- = \sqrt{2}\epsilon a_n^\dagger\left[1 - \frac{\epsilon^2}{4}a_n^\dagger a_n - \frac{\epsilon^4}{32}a_n^\dagger a_n a_n^\dagger a_n - O(\epsilon^6)\right], \quad (8b)$$

and similar expansions for $S_n'^+$, $S_n'^-$, and $S_n'^z$ in terms of $b_n(b_n^\dagger)$. Here, $a_n^\dagger(b_n^\dagger)$ and $a_n(b_n)$ are creation and annihilation operators satisfying the usual commutation relations $[a_m, a_n^\dagger] = [b_m, b_n^\dagger] = \delta_{mn}$ and $[a_m, a_n] = [b_m, b_n] = [a_m^\dagger, a_n^\dagger] = [b_m^\dagger, b_n^\dagger] = 0$. Substituting Eqs. (8a) and (8b) in Eq. (6), up to $O(\epsilon^2)$, we have

$$\begin{aligned} H = \sum_n \Bigg[& \frac{p_n^2}{2m_1} + \frac{q_n^2}{2m_2} + k_1(X_n - X_{n+1})^2 + k_2(y_n - y_{n+1})^2 \\ & + \epsilon^2\{-J(a_n a_{n+1}^\dagger + a_n^\dagger a_{n+1} - a_n^\dagger a_n - a_{n+1}^\dagger a_{n+1} + b_n b_{n+1}^\dagger \\ & + b_n^\dagger b_{n+1} - b_n^\dagger b_n - b_{n+1}^\dagger b_{n+1}) \\ & - [\mu - \alpha_1(X_{n+1} - X_{n-1})](a_n b_n^\dagger + a_n^\dagger b_n - a_n^\dagger a_n - b_n^\dagger b_n) \\ & - \alpha_2(y_{n+1} - y_{n-1})(a_n^\dagger a_n + b_n^\dagger b_n)\}\Bigg]. \end{aligned} \quad (9)$$

IV. DYNAMICAL EQUATIONS

Having written down the Hamiltonian, to understand the dynamics, we construct the following equations of motion for a_n and b_n as well as for X_n and y_n:

$$i\hbar\frac{\partial A_n}{\partial t} = [A_n, H] = F(A_n^\dagger, A_n, A_{n+1}^\dagger, A_{n+1}). \quad (10)$$

Here, A_n strands for a_n and b_n. The equations of motion for X_n and y_n are written using the Hamilton equations of motion, $\partial X_n/\partial t = \partial H/\partial p_n$, $\partial p_n/\partial t = -\partial H/\partial X_n$, $\partial y_n/\partial t = \partial H/\partial q_n$, and $\partial q_n/\partial t = -\partial H/\partial y_n$. The explicit form of the equations of motion is written down by substituting Hamiltonian (9) in the above equations of motion for a_n, b_n, X_n, and y_n. Thus, we get

$$\begin{aligned} i\frac{\partial a_n}{\partial t} = & -J(a_{n+1} - 2a_n + a_{n-1}) \\ & - [\mu + \alpha_1(X_{n+1} - X_{n-1})](b_n - a_n) - \alpha_2(y_{n+1} - y_{n-1})a_n, \end{aligned} \quad (11a)$$

$$\begin{aligned} i\frac{\partial b_n}{\partial t} = & -J(b_{n+1} - 2b_n + b_{n-1}) \\ & - [\mu + \alpha_1(X_{n+1} - X_{n-1})](a_n - b_n) - \alpha_2(y_{n+1} - y_{n-1})b_n, \end{aligned} \quad (11b)$$

$$\begin{aligned} m_1\frac{\partial^2 X_n}{\partial t^2} = & k_1(X_{n+1} - 2X_n + X_{n-1}) \\ & + \alpha_1[a_{n-1}^\dagger a_{n-1} - a_{n+1}^\dagger a_{n+1} + b_{n-1}^\dagger b_{n-1} \\ & - b_{n+1}^\dagger b_{n+1} a_{n+1} b_{n+1}^\dagger - a_{n-1} b_{n-1}^\dagger + a_{n+1}^\dagger b_{n+1} \\ & - a_{n-1}^\dagger b_{n-1}], \end{aligned} \quad (11c)$$

$$\begin{aligned} m_2\frac{\partial^2 y_n}{\partial t^2} = & k_2(y_{n+1} - 2y_n + y_{n-1}) \\ & + \alpha_2[a_{n-1}^\dagger a_{n-1} - a_{n+1}^\dagger a_{n+1} + b_{n-1}^\dagger b_{n-1} - b_{n+1}^\dagger b_{n+1}]. \end{aligned} \quad (11d)$$

While writing the above equations of motion (11a)–(11d), the time is rescaled and m_1, m_2, k_1, and k_2 are redefined. In order to represent the large amplitude collective modes by coherent states, we introduce Glauber's coherent state representation [43] for boson operators $a_n^\dagger|u\rangle = u_n^*|u\rangle, a_n|u\rangle = u_n|u\rangle, |u\rangle = \Pi_n|u_n\rangle$ and $b_n^\dagger|v\rangle = v_n^*|v\rangle, b_n|v\rangle = v_n|v\rangle, |v\rangle = \Pi_n|v_n\rangle$ with $\langle u|u\rangle = 1$ and $\langle v|v\rangle = 1$, where u_n and v_n are the coherent amplitudes of the operators a_n and b_n for the system in the states $|u\rangle$ and $|v\rangle$, respectively. As the length of the DNA and the protein chains are large compared to the lattice parameter, we make continuum approximation by introducing the following fields $u_n \to u(z,t)$, $v_n(t) \to v(z,t)$, $X_n(t) \to X(z,t)$, and $y_n(t) \to y(z,t)$, where $z = nl$ and the expansions $u_{n\pm1} = u(z,t) \pm l(\partial u/\partial z) + (l^2/2!)(\partial^2 u/\partial z^2) \pm O(l^3)$ and similar ones for $v_{n\pm1}$, $X_{n\pm1}$, and $y_{n\pm1}$. Under the above approximations, the equations of motion (11a)–(11d) up to $O(l^2)$, after suitable rescaling of z and redefining of the parameters α_1, α_2, k_1 and k_2, become

$$iu_t = -u_{zz} - (\mu - \alpha_1 X_z)(v - u) - \alpha_2 y_z u, \quad (12a)$$

$$iv_t = -v_{zz} - (\mu - \alpha_1 X_z)(u - v) - \alpha_2 y_z v, \quad (12b)$$

$$m_1 X_{tt} = k_1 X_{zz} - \alpha_1[|u|^2 + |v|^2 - uv^* - u^* v]_z, \quad (12c)$$

$$m_2 y_{tt} = k_2 y_{zz} - \alpha_2[|u|^2 + |v|^2]_z. \quad (12d)$$

In the above equations, the subscripts "t" and "z" represent partial derivatives with respect to time and the spatial variable, respectively. On adding and subtracting Eqs. (12a) and (12b) and by defining $v = -u$, Eqs. (12a)–(12d) can be rewritten as

$$iu_t + u_{zz} + (2\alpha_1 X_z + \alpha_2 y_z)u = 0, \tag{13a}$$

$$X_{tt} - \frac{k_1}{m_1}X_{zz} = -\frac{4\alpha_1}{m_1}[|u|^2]_z, \tag{13b}$$

$$y_{tt} - \frac{k_2}{m_2}y_{zz} = -\frac{2\alpha_2}{m_2}[|u|^2]_z. \tag{13c}$$

While writing Eq. (13a), the term proportional to μ was transformed away using the transformation $u(z,t) = \hat{u}(z,t)e^{-2i\mu t}$ and finally the hat was dropped. The set of coupled equations (13b) and (13c) describe the dynamics of our protein-DNA molecular system at the physiological temperature, when the protein molecule binds to the DNA double helical chain through linear harmonic coupling. The dynamics is governed by the rotation of DNA bases and thermal vibration of the bases along the hydrogen bonds, combined with the longitudinal motion of the peptide units of the binding protein. Here, we are specifically concerned with the nonlinear excitation of bases, induced by protein and thermal fluctuations, in which a cluster of DNA bases may undergo a large excursion as compared to the rest of the bases.

When $\alpha_1 = \alpha_2 = 0$, Eqs. (13b) and (13c) are decoupled and reduced to a set of linear equations. Thus, when the protein molecule detaches from the DNA chain, the dynamics of the system is governed by the following set of linear equations:

$$iu_t + u_{zz} = 0, \tag{14a}$$

$$X_{tt} - \frac{k_1}{m_1}X_{zz} = 0, \quad y_{tt} - \frac{k_2}{m_2}y_{zz} = 0. \tag{14b}$$

While Eq. (14a) is the time-dependent Schrödinger equation for a free particle, the equations in Eq. (14b) are homogeneous linear wave equations. Equation (14a) admits plane transverse wave solution of the form $u = u_0 e^{i(\kappa z - wt)}$ with the dispersion relation $w = \kappa^2$, where u_0 is the constant amplitude. On the other hand, Eq. (14b) admits linear nondispersive wave solutions given by $X = f_1(z - v_1 t) + g_1(z + v_1 t)$ and $y = f_2(z - v_2 t) + g_2(z + v_2 t)$, where f_1, g_1 and f_2, g_2 are arbitrary functions and $v_1 = \sqrt{k_1/m_1}$, $v_2 = \sqrt{k_2/m_2}$ represent the constant phase velocities of the waves. When the protein molecule binds to the DNA molecule at the physiological temperature, i.e., when $\alpha_1, \alpha_2 \neq 0$ the excitation energy of the protein-DNA molecular system due to thermal fluctuation increases, and nonlinearity started playing its role. At this point, the nonlinear dynamics of the system is governed by the complete set of coupled nonlinear equations (13a)–(13c).

V. SOLITON, BASE-PAIR OPENING, AND BUBBLE TRANSPORT

In order to solve the set of coupled nonlinear equations (13a)–(13c), we differentiate Eqs. (13b) and (13c) with respect to z once and define two fields $\hat{X}(z,t) = X_z$ and $Y(z,t) = y_z$, so that Eqs. (13a)–(13c) are written as

$$iu_t + u_{zz} + (2\alpha_1 \hat{X} + \alpha_2 Y)u = 0, \tag{15a}$$

$$\hat{X}_{tt} - v_1^2 \hat{X}_{zz} + \frac{4\alpha_1}{m_1}[|u|^2]_{zz} = 0, \tag{15b}$$

$$Y_{tt} - v_2^2 Y_{zz} + \frac{2\alpha_2}{m_2}[|u|^2]_{zz} = 0. \tag{15c}$$

Defining the wave variable $\zeta = z - v_3 t$, where v_3 is the velocity of the wave, and writing $\hat{X}(z,t) \to \hat{X}(\zeta)$, $Y(z,t) \to Y(\zeta)$, Eqs. (15b) and (15c) can be written as

$$\hat{X}_{\zeta\zeta} - 2\beta_1[|u|^2]_{\zeta\zeta} = 0, \tag{16a}$$

$$Y_{\zeta\zeta} - 2\beta_2[|u|^2]_{\zeta\zeta} = 0, \tag{16b}$$

where $\beta_1 = 2\alpha_1/[m_1(v_1^2 - v_3^2)]$ and $\beta_2 = \alpha_2/[m_2(v_2^2 - v_3^2)]$. On integrating Eqs. (16a) and (16b) with respect to ζ twice and assuming both the integration constants to be zero, we get $\hat{X} = 2\beta_1|u|^2$ and $Y = 2\beta_2|u|^2$, which upon using in Eq. (15a) gives

$$iU_t + U_{zz} + 2|U|^2 U = 0, \tag{17}$$

where $U(z,t) = (2\alpha_1\beta_1 + \alpha_2\beta_2)^{1/2}u(z,t)$. Equation (17) is the well-known completely integrable NLS equation, which has been solved for N-soliton solutions using inverse scattering transform method [44]. For instance, the explicit form of the one soliton solution is written as

$$U = \eta\,\mathrm{sech}[\eta(z - 2\xi t - \theta_0)]\exp\{i\xi(z - 2\xi t - \theta_0) + i[(\eta^2 + \xi^2)t - \sigma_0]\}, \tag{18}$$

where η, ξ, θ_0, and σ_0 are four real parameters which determine the propagating amplitude, velocity, initial position, and initial phase of the soliton. The solitons generated in the protein-DNA molecular system, at the physiological temperature, are formed as a result of the dynamical balance between the dispersion due to interaction of intrastrand dipole vibrations (stacking) and the nonlinearity provided by the hydrogen bonds, coupled to the local displacement of the peptides in the protein molecule and to the thermal phonons. The waves that arise in the protein molecule and in the bases of DNA provide a potential well that prevents dispersion of the rotational energy of the bases in DNA. Thus, the propagation of rotation of bases in DNA is coupled to the longitudinal waves in protein, and the coupled excitations propagate as a localized dynamically self-sufficient entity called solitons which travel along each strand of the DNA chain. The soliton solution describes an open state configuration in the individual strands of the DNA double helix, which collectively represents a bubble. Thus, the protein molecule induces opening of the bases in DNA during the process of transcription. From the expression for the one soliton solution found in Eq. (18), namely, $u = (2\alpha_1\beta_1 + \alpha_2\beta_2)^{-1/2}\eta\,\mathrm{sech}[\eta(z - 2\xi t - \theta_0)]\exp\{i\xi(z - 2\xi t - \theta_0) + i[(\eta^2 + \xi^2)t - \sigma_0]\}$, it is understood that the amplitude of the soliton depends on the strength of the coupling of DNA excitations to the thermal phonons (α_1) and to the molecular vibrations of the protein molecule (α_2), which also contrib-

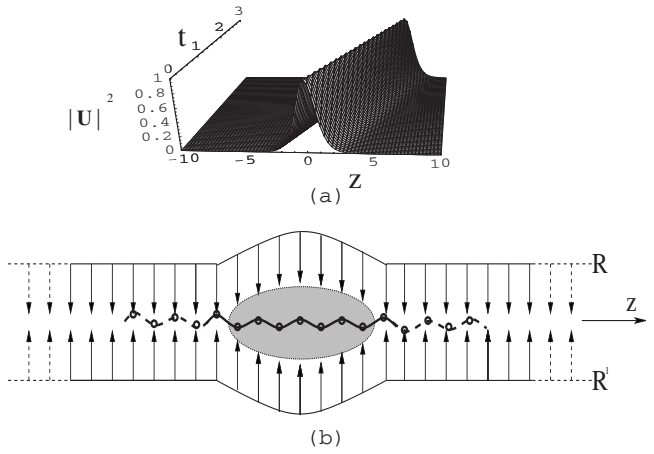

FIG. 3. (a) One soliton solution [Eq. (18)] of the NLS equation. (b) A schematic representation of the formation of bubble with the solitons and its propagation along DNA.

ute to nonlinearity in the formation of solitons. When the strength of the coupling increases, the amplitude of the soliton decreases. This result is in good agreement with the result of simulation studies by Campa [27] who discovered that, in the case of large thermal coupling, the bubble travels only for a short distance with decreasing amplitude. In Fig. 3(a), the square of the absolute value of the one soliton solution U, i.e., $|U|^2$, as given in Eq. (18) is plotted. A schematic representation of the coherent base excitations in DNA—in terms of rotation of bases—induced by the protein molecule in the form of solitons propagating along the two strands, which collectively form a traveling bubble created by energy delocalization due to nonlinear effects, is shown in Fig. 3(b). In the figure, the shaded ellipse represents the region of interaction of the protein molecule with the DNA, inducing the formation of bubble. With the distance between adjacent bases in DNA as equal to 0.34 nm and the approximate length of the bubble from Fig. 3(a) as 5 units, it is estimated that nearly 14 bases in each strand will contribute to the formation of the bubble and hence will participate in the base-pair opening. In the past, gel electrophoresis [45] and uv absorption experiments [46] as well as radiogram [47] measurements were performed to study the unwinding of DNA double helix, due to the disruption of a short helical segment, when RNA polymerase binds to it. The data had shown that the number of base pairs that are disrupted and participate in the unwinding process ranges from 7 to 15, which has a strong overlap with our result of roughly 14 base pairs, obtained through analytical calculations. Thus, our results show a relevant length scale for real protein-DNA interaction.

VI. EFFECT OF VISCOSITY

In a more realistic description of the dynamics of protein-DNA system, it is important to include the influence of the surrounding medium on the excitations. In the case of a protein-DNA system, the solvating water acts as a viscous medium that makes the nucleotide oscillations to damp out [48]. Therefore, essentially the effect of the surrounding medium on the excitations of protein-DNA molecular system reduces to viscous damping effect. This effect can be taken into account in the dynamics by adding a term of the form $-i\gamma U$, where γ is a coupling constant which determines the strength of viscous damping, to the right-hand side of the dynamical equation of the molecular system (18) which now becomes the following perturbed nonlinear Schrödinger equation:

$$iU_t + U_{zz} + 2|U|^2U = -i\gamma U. \tag{19}$$

As the viscosity of the solvating water is temperature dependent, from a simple fluid mechanics argument, one can estimate the magnitude of the damping coefficient, which is very small at the physiological temperature [31]. Hence, we treat the term proportional to γ in Eq. (19) as a weak perturbation and the small coupling constant γ as a perturbation parameter. When $\gamma=0$, Eq. (19) reduces to the completely integrable NLS equation, as found in Eq. (17), and the associated one soliton solution as given in Eq. (18) is rewritten, for convenience, in the form

$$U = \eta \,\mathrm{sech}\, \eta(\theta-\theta_0)\exp[i\xi(\theta-\theta_0)+i(\sigma-\sigma_0)], \tag{20}$$

where $\frac{\partial\theta}{\partial t}=-2\xi$, $\frac{\partial\theta}{\partial z}=1$, $\frac{\partial\sigma}{\partial t}=\eta^2+\xi^2$, and $\frac{\partial\sigma}{\partial z}=0$.

We carry out a perturbation analysis [49] to understand the impact of the viscous force from the surrounding medium, by introducing a slow time variable $T=\gamma t$, and treat the quantities η, ξ, θ_0, and σ_0 as functions of this time scale. Hence, the envelope one soliton solution (20) is written as

$$U = \hat{U}(\theta,T;\gamma)\exp[i\xi(\theta-\theta_0)+i(\sigma-\sigma_0)]. \tag{21}$$

Under the assumption of quasistationarity, Eq. (19) reads

$$\eta^2\hat{U} + \hat{U}_{\theta\theta} + 2|\hat{U}|^2\hat{U} = \gamma F(\hat{U}), \tag{22}$$

where

$$F(\hat{U}) = [(\theta-\theta_0)\xi_T - \xi\theta_{0T} - \sigma_{0T}]\hat{U} - i[\hat{U}_T + \hat{U}]. \tag{23}$$

We assume a Poincaré-type asymptotic expansion for $\hat{U}$ by writing $\hat{U}(\theta,T;\gamma)=\Sigma_{n=0}^{\infty}\gamma^n\hat{U}_n(\theta,T)$ and further restrict ourselves to the calculation of order (γ), such that $\hat{U}(\theta,T;\gamma)=\hat{U}_0(\theta,T)+\gamma\hat{U}_1(\theta,T)$, where $\hat{U}_0=\eta\,\mathrm{sech}[\eta(\theta-\theta_0)]$. Assuming $\hat{U}_1=\phi_1+i\psi_1$, where ϕ_1 and ψ_1 are real and on substituting the above in Eqs. (22) and (23), we obtain

$$L_1\phi_1 \equiv -\eta^2\phi_1 + \phi_{1\theta\theta} + 6|\hat{U}_0|^2\phi_1 = \mathrm{Re}\, F(\hat{U}_0), \tag{24a}$$

$$L_2\psi_1 \equiv -\eta^2\psi_1 + \psi_{1\theta\theta} + 2|\hat{U}_0|^2\psi_1 = \mathrm{Im}\, F(\hat{U}_0), \tag{24b}$$

where

$$\mathrm{Re}\, F(\hat{U}_0) = [(\theta-\theta_0)\xi_T - \xi\theta_{0T} - \sigma_{0T}]\hat{U}_0, \tag{25a}$$

$$\mathrm{Im}\, F(\hat{U}_0) = -[\hat{U}_{0T} + \hat{U}_0]. \tag{25b}$$

In Eqs. (24a) and (24b), L_1 and L_2 are self-adjoint operators. It may be verified that the solutions of the homogeneous

parts of Eqs. (24a) and (24b) are $\hat{U}_{0\theta}$ and $\hat{U}_0$, respectively, and hence we have the following secularity conditions:

$$\int_{-\infty}^{\infty} \hat{U}_{0\theta} \,\mathrm{Re}\, F(\hat{U}_0) d\theta = 0, \tag{26a}$$

$$\int_{-\infty}^{\infty} \hat{U}_0 \,\mathrm{Im}\, F(\hat{U}_0) d\theta = 0. \tag{26b}$$

On evaluating the above integrals after substituting the values of $\hat{U}_{0\theta}$, $\hat{U}_0$, Re $F(\hat{U}_0)$, and Im $F(\hat{U}_0)$, we obtain $\xi_T=0$ and $\eta_T=-2\eta$, which can be written after integrating once as

$$\xi = \xi_0, \quad \eta = \eta_0 e^{-2\gamma t}, \tag{27}$$

where ξ_0 and η_0 are the initial velocity and amplitude of the soliton. The first of Eq. (27) says that when the surrounding viscous solvating water medium interacts with the protein-DNA molecular system, the velocity of the soliton is unaffected by it. However, from the η equation of (27), one understands that the viscous effect of the surrounding medium diminishes the amplitude of the soliton excitations. In other words, the surrounding solvent medium damps out the soliton excitations exponentially and hence it will travel only for a limited distance. The η equation of Eq. (27) tells that the amplitude of the soliton reduces to $1/e$ or 0.369 times the initial amplitude η_0 after a duration of $T=1/2\gamma$, from when the perturbation due to the viscous effect is switched on.

Now, we construct the perturbed soliton solutions by solving Eqs. (24a) and (24b). For that, first we solve the homogeneous part of Eq. (24a), which admits the following two particular solutions:

$$\phi_{11} = \mathrm{sech}\ \eta(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0), \tag{28a}$$

$$\phi_{12} = \frac{1}{\eta}\left[\frac{3}{2}\eta(\theta-\theta_0)\mathrm{sech}\ \eta(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0) + \frac{1}{2}\tanh\ \eta(\theta-\theta_0)\sinh\ \eta(\theta-\theta_0) - \mathrm{sech}\ \eta(\theta-\theta_0)\right]. \tag{28b}$$

Knowing two particular solutions, the general solution can be found out by using the following expression:

$$\phi_1 = C_1\phi_{11} + C_2\phi_{12} - \phi_{11}\int_{-\infty}^{\theta} \phi_{12}\,\mathrm{Re}\, F(\hat{U}_0)d\theta + \phi_{12}\int_{-\infty}^{\theta} \phi_{11}\,\mathrm{Re}\, F(\hat{U}_0)d\theta, \tag{29}$$

where C_1 and C_2 are arbitrary constants. The solution ϕ_1 is constructed by substituting the expressions for ϕ_{11}, ϕ_{12}, and Re $F(\hat{U}_0)$ given in Eqs. (28a), (28b), and (25a) in Eq. (29) and by evaluating the integrals. The result reads

$$\phi_1 = -\frac{1}{\eta}\left[C_2 + \frac{1}{2}(\xi\theta_{0T} + \sigma_{0T})\right]\mathrm{sech}\ \eta(\theta-\theta_0) + \left[C_1 + \frac{3C_2}{2}(\theta-\theta_0) + \frac{1}{2}(\theta-\theta_0)(\xi\theta_{0T}+\sigma_{0T})\right] \times \mathrm{sech}\ \eta(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0) + \frac{C_2}{2\eta}\sinh\ \eta(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0). \tag{30}$$

The last term in Eq. (30) is a secular term which makes the solution unbounded and, hence, it is removed by assuming the arbitrary constant $C_2=0$. By using the boundary conditions $\phi_1|_{\theta=\theta_0}=\mathrm{const}=c$ and $\phi_{1\theta}|_{\theta=\theta_0}=0$, we obtain $\frac{1}{\eta}(\xi\theta_{0T}+\sigma_{0T})=-c$ and $C_1=0$. Using the above results in Eq. (30), the general solution ϕ_1 is written as

$$\phi_1 = c[1-(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0)]\mathrm{sech}\ \eta(\theta-\theta_0). \tag{31}$$

Next, we solve Eq. (24b), the homogeneous part of which admits the following particular solutions:

$$\psi_{11} = \mathrm{sech}\ \eta(\theta-\theta_0), \tag{32a}$$

$$\psi_{12} = \frac{1}{2\eta}[\eta(\theta-\theta_0)\mathrm{sech}\ \eta(\theta-\theta_0) + \sinh\ \eta(\theta-\theta_0)]. \tag{32b}$$

The general solution can be found from

$$\psi_1 = C_3\psi_{11} + C_4\psi_{12} - \psi_{11}\int_{-\infty}^{\theta} \psi_{12}\,\mathrm{Im}\, F(\hat{U}_0)d\theta + \psi_{12}\int_{-\infty}^{\theta} \psi_{11}\,\mathrm{Im}\, F(\hat{U}_0)d\theta, \tag{33}$$

where C_3 and C_4 are arbitrary constants. The explicit form of ψ_1 is constructed by substituting the values of ψ_{11}, ψ_{12}, and Im $F(\hat{U}_0)$ given in Eqs. (32a), (32b), and (25b) and by evaluating the integrals,

$$\psi_1 = \left(C_3 + \frac{C_4}{2}(\theta-\theta_0) - \frac{\eta}{2}\left\{(\theta-\theta_0)\left[\frac{\eta_T}{2\eta}(\theta-\theta_0) - \theta_{0T}\right] + \theta_{0T}\tanh\ \eta(\theta-\theta_0) + \theta_{0T}(\theta-\theta_0)\mathrm{sech}^2\ \eta(\theta-\theta_0)\right\}\right)\mathrm{sech}(\theta-\theta_0) + \frac{C_4}{2\eta}\sinh\ \eta(\theta-\theta_0). \tag{34}$$

In the above solution for ψ_1, the term proportional to $\sinh\ \eta(\theta-\theta_0)$ is secular, which can be removed by choosing $C_4=0$. We also get $C_3=0$ and $\theta_{0T}=0$ upon using the boundary conditions $\psi_1|_{\theta=\theta_0}=0$ and $\psi_{1\theta}|_{\theta=\theta_0}=0$. On using the above results in Eq. (34), the final form of ψ_1 is obtained as

$$\psi_1 = \frac{\eta}{2}(\theta-\theta_0)^2\mathrm{sech}\ \eta(\theta-\theta_0). \tag{35}$$

Using the results given in Eqs. (31) and (35), we write down the final form of the first-order perturbed soliton

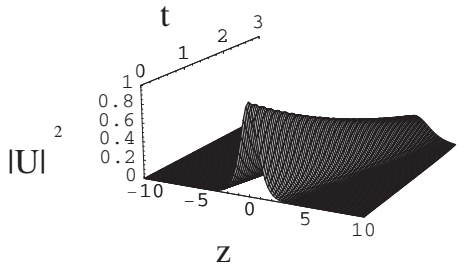

FIG. 4. Square of absolute value of the perturbed soliton solution [Eq. (36)], under viscous damping.

$U=[\hat{U}_0+\gamma(\phi_1+i\psi_1)]\exp[i\xi(\theta-\theta_0)+i(\sigma-\sigma_0)]$ (by choosing $\gamma=1$) as

$$U=\left[\eta\ \mathrm{sech}\ \eta(\theta-\theta_0)+c[1-(\theta-\theta_0)\tanh\ \eta(\theta-\theta_0)]+i\frac{\eta}{2}(\theta-\theta_0)^2\ \mathrm{sech}\ \eta(\theta-\theta_0)\right]\exp[i\xi(\theta-\theta_0)+i(\sigma-\sigma_0)]. \quad (36)$$

In Fig. 4, the square of the absolute value of the perturbed soliton, i.e., $|U|^2$ from Eq. (36) is plotted. From the figure, we observe that the amplitude of the soliton decreases as time progresses, in agreement with the variation of soliton amplitude as per Eq. (27), which is due to viscosity of the surrounding medium. Therefore, when the viscosity of the surrounding medium is high the soliton is expected to travel only for a short duration and will stop thereafter. Similar results were also obtained by Yakushevich *et al.* [30] through numerical analysis. It is found that the soliton passes more than 3000 chain links in DNA like a heavy Brownian particle when the viscosity is low, and it stops after a few chain links in the high viscosity limit.

VII. CONCLUSIONS

In this paper, we investigated the nonlinear dynamics of a protein-DNA molecular system, under thermal fluctuation in a viscous surrounding medium, by considering DNA as a set of two coupled linear chains and protein as a single linear molecular chain interacting through linear coupling. In the nonviscous limit, the dynamical equation for the system is derived from a Hamiltonian, modeled by treating the DNA bases in the double helical chain as classical spins of a coupled chain model, which are then expressed in terms of the angles of rotation of bases, and bosonized. Thermal phonon energy, longitudinal motion of the protein molecule, and their coupling to the motion of DNA bases are also included in the Hamiltonian suitably. The dynamical equations obtained from the Hamiltonian for the protein-DNA molecular system form a set of coupled equations, in which the DNA dynamics is governed by a nonlinear equation for rotation of bases, coupled to two inhomogeneous linear wave equations, representing thermal vibration of the bases and the longitudinal motion of peptide groups in the protein molecule. When the protein molecule detaches from DNA, the above equations decouple and reduce to time-dependent Schrödinger equation for a free particle and one-dimensional homogeneous linear wave equations. While the former one admits dispersive plane transverse wave solution, the latter ones admit nondispersive wave solutions. When the protein molecule and the thermal phonons started interacting with the DNA chain, the couplings introduce nonlinearity into the dynamics of bases in DNA, and the set of coupled equations reduces to the completely integrable NLS equation. During interaction, the energy of the excited DNA molecule increases and the nonlinearity localizes the energy, thus forming localized solitons. These solitons represent opening of base pairs in both the strands, which collectively form a bubble, traveling along the DNA double helical chain at the physiological temperature. Thus, the protein molecule acts as a zip runner that opens the base pairs, which close when the protein molecule progresses along the DNA chain. For strong coupling, the amplitude of the soliton decreases and the bubble travels only for a short distance. The size of the soliton indicates that nearly 14 base pairs participate in the opening of bases and in the formation of the bubble, when the protein molecule interacts with the DNA. The above length scale is in comparison with the length of the active site region for protein-DNA interaction. In a more realistic situation, when the protein-DNA molecular system interacts with the surrounding viscous solvating water medium, the dynamics of the system is governed by a perturbed nonlinear Schrödinger equation. The viscous effect of the medium makes the nucleotide oscillations in DNA to damp out. The soliton damps out quickly if the viscosity of the medium is high. The soliton moves over a definite length along the DNA chain if the viscosity of the surrounding medium is low. However, the velocity of the soliton is unaffected by the viscous effect. The events that happen in the present study may represent binding of an RNA polymerase to a promoter site in the DNA during the transcription process. Our results have very strong coincidence with the experimental data [45–47] that the binding of RNA polymerase to the promoter site in DNA is accompanied by a local distortion of the DNA bases in the form of solitons, which can propagate along the DNA double helix.

Finally, we present the possibility of testing our results experimentally. In the proposed experiment, a small DNA molecule is stretched and bound to fixed surfaces on both the ends. A single molecule of RNA polymerase is then allowed to move linearly along the DNA molecule. During this process, the polymerase molecule will make diffusional search for a promoter site. Upon identifying the promoter site, the polymerase binds to the DNA and a short segment of DNA will be disrupted, resulting in the opening of base pairs and formation of a bubble. In nature, protein binds to DNA in a very specific site such as promoter, coding, or terminator, which has a specific sequence of bases, and this makes the strands site dependent or inhomogeneous. Hence, it is important to understand the nonlinear dynamics of inhomogeneous DNA, with the protein molecule binding to specific site of the DNA, and the study is under progress.

ACKNOWLEDGMENT

The work of M.D. and V.V. forms part of a major DST project.

[1] L. Stryer, *Biochemistry*, 4th ed. (W. H. Freeman and Company, New York, 1995).
[2] S. J. van Noort, K. O. van der Werf, A. P. M. Eker, C. Wyman, B. G. de Grooth, N. F. van Hulst, and J. Greve, Biophys. J. **74**, 2840 (1998).
[3] S. Klimasauskas, T. Szyperski, S. Serva, and K. Wuthrich, EMBO J. **17**, 317 (1998).
[4] K. Liebert, A. Hermann, M. Schlickenrieder, and A. Jeltsch, J. Mol. Biol. **341**, 443 (2004).
[5] S. J. Koch and M. D. Wang, Phys. Rev. Lett. **91**, 028103 (2003).
[6] Y. Harada, O. Ohara, A. Takatsuki, H. Itoh, N. Shimamoto, and K. Kinoshita, Jr., Nature (London) **409**, 113 (2001).
[7] N. Huang, N. K. Banavali, and A. D. MacKerell, Jr., Proc. Natl. Acad. Sci. U.S.A. **100**, 68 (2003).
[8] J. R. Horton, G. Ratner, N. K. Banavali, N. Huang, Y. Choi, M. A. Maier, V. E. Marquez, A. D. MacKerell, Jr., and X. Cheng, Nucleic Acids Res. **32**, 3877 (2004).
[9] N. Huang and A. D. MacKerell, Jr., J. Mol. Biol. **345**, 265 (2005).
[10] A. Sarkar and J. F. Marko, Phys. Rev. E **64**, 061909 (2001).
[11] J. Yan and J. F. Marko, Phys. Rev. E **68**, 011905 (2003).
[12] M. D. Betterton and F. Julicher, Phys. Rev. Lett. **91**, 258103 (2003).
[13] S. M. Bhattacharjee, EPL **65**, 574 (2004).
[14] F. Habib and R. Bundschuh, Phys. Rev. E **72**, 031906 (2005).
[15] S. W. Englander, N. R. Kallenbanch, A. J. Heeger, J. A. Krumhansl, and S. Litwin, Proc. Natl. Acad. Sci. U.S.A. **77**, 7222 (1980).
[16] S. Yomosa, Phys. Rev. A **27**, 2120 (1983).
[17] S. Yomosa, Phys. Rev. A **30**, 474 (1984).
[18] S. Takeno and S. Homma, Prog. Theor. Phys. **70**, 308 (1983).
[19] S. Homma and S. Takeno, Prog. Theor. Phys. **72**, 679 (1984).
[20] L. V. Yakushevich, Nanobiology **1**, 343 (1992).
[21] M. Daniel and V. Vasumathi, Physica D **231**, 10 (2007).
[22] M. Daniel and V. Vasumathi, Phys. Lett. A **372**, 5144 (2008).
[23] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989).
[24] J. A. González and M. M. Landrove, Phys. Lett. A **292**, 256 (2002).
[25] M. Peyrard, Nonlinearity **17**, R1 (2004).
[26] P. L. Christiansen, P. C. Lomdahl, and V. Muto, Nonlinearity **4**, 477 (1991).
[27] A. Campa, Phys. Rev. E **63**, 021901 (2001).
[28] S. Cocco, R. Monasson, and J. F. Marko, Phys. Rev. E **65**, 041907 (2002).
[29] S. Cocco, R. Monasson, and J. F. Marko, Phys. Rev. E **66**, 051914 (2002).
[30] L. V. Yakushevich, A. V. Savin, and L. I. Manevitch, Phys. Rev. E **66**, 016614 (2002).
[31] S. Zhravković and M. V. Satarić, Phys. Scr. **64**, 612 (2001).
[32] S. Zdravković and M. V. Satarić, Int. J. Mod. Phys. B **17**, 5911 (2003).
[33] C. B. Tabi, A. Mohamadov, and T. C. Kofane, Math. Biosci. Eng. **5**, 205 (2008).
[34] M. Daniel and V. Vasumathi, Phys. Rev. E **79**, 012901 (2009).
[35] M. V. Satarić and J. A. Tuszynski, Phys. Rev. E **65**, 051901 (2002).
[36] A. S. Davydov, Ukr. Fiz. Zh. (Russ. Ed.) **20**, 179 (1975).
[37] A. S. Davydov, *Solitons in Molecular Systems* (Reidel, Dordrecht, 1985).
[38] M. V. Satarić, L. Matsson, and J. A. Tuszyński, Phys. Rev. E **74**, 051902 (2006).
[39] M. Daniel and L. Kavitha, Phys. Lett. A **295**, 121 (2002), and references therein.
[40] N. S. Fialko and V. D. Lakhno, Phys. Lett. A **278**, 108 (2000).
[41] A. M. Guo and H. Xu, Phys. Lett. A **364**, 48 (2007).
[42] T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1940).
[43] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
[44] V. E. Zakharov and A. B. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1972).
[45] J. C. Wang, J. H. Jacobsen, and J. M. Saucier, Nucleic Acids Res. **4**, 1225 (1977).
[46] T.-s. Hsieh and J. C. Wang, Nucleic Acids Res. **5**, 3337 (1978).
[47] U. Siebenlist, Nature (London) **279**, 651 (1979).
[48] M. Peyrard, J. Biol. Phys. **27**, 217 (2001).
[49] M. Daniel and J. Beula, Phys. Rev. B **77**, 144416 (2008).