Effect of Solvent on Stretching and Twisting of DNA

Rizki Khairani\(^1\), Donny Dwiputra\(^1\), Wahyu Hidayat\(^1,2\), Freddy P. Zen\(^1,2\)

\(^1\) Theoretical Physics Laboratory, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung Jl. Ganesha no. 10 Bandung, 40132, Indonesia
\(^2\) Indonesian Center of Theoretical and Mathematical Physics (ICTMP), Indonesia

E-mail: rizkikhairani028@gmail.com

Abstract. We extend the Peyrard-Bishop model into helicoidal DNA case in presence of solvent. Our model focuses on twisting and stretching DNA base pair motions. The dynamics of this system are drawn from the solitary wave solution by using multiple scale expansion to the case of the vectorial lattice. The breather soliton is the solution for the stretching motion and the kink soliton is the solution for the twisting motion. A positive value of the solvent strength factor stretches the hydrogen bond and the strand arc length which makes the DNA more winded.

Keywords: DNA-Solvent interaction, breather-soliton, kink-soliton.

1. Introduction
The conformational changes of DNA due to the environment is not fully understood yet. The complex dynamics of DNA have been investigated. Peyrard-Bishop (PB) presented simple model focusing on the denaturation which is initialized by the bubble formation, without considering the DNA helicity in a vacuum [1]. In this model, PB interpret the hydrogen bond as a Morse potential. Next, Dauxois extended the PB model by considering the DNA helicity, assumed the bases in a different strand do not only interact in transversal direction but also in longitudinal direction [2]. Still, in helicoidal DNA case, Barbi et al [3] extended the PB model by generalizing the curvature of the DNA double helix, assuming the helicity can be presented by observing the interaction of three body diks (nucleotida). Because of that, Barbi uses two degrees of freedom, radial and angular (polar coordinates) and modifies the Hamiltonian PB model by adding the helicoidal term [3].

In this paper, we include the effect of the environment, in our case is the solvent, to realistically model the dynamics of DNA. PB, Dauxois, and Barbi have modeled the DNA dynamics but still cannot explain the real DNA condition. In fact, DNA is immersed in solvent, one of the environment factors affecting the dynamics of DNA. According to Zoli, solvent will stabilize the open state or the base pairs opening [4]. Solvation causes weakening and lengthening of the hydrogen bond from base pairs [5] while experimentally it is the transition of DNA from B-DNA to A-DNA due to increased solvent concentration [6]. Some experimental facts of DNA-solvent interaction cannot be fully explained. Because of that, we present a model of DNA in solvent for helicoidal DNA by adopting the helicoidal model of Barbi by adding the solvent potential term in Lagrangian.
2. Lagrangian Model

The helicoidal model of DNA is described in polar coordinate system with two degrees of freedom: \( r_n \) dan \( \varphi_n \), radial dan angular variable, both of them will produce the stable helical (variable coupled) structure. The distance of a base from the hydrogen bond center in equilibrium condition (B-DNA) is denoted by \( R_0 \). The distance between the neighbouring base pairs denoted by \( h \). In our model, the adjacent site height \( h \) is assumed constant. The length of backbone between disks in one strand when in equilibrium state is denoted by \( L \). The helicoidal model of DNA is described in polar coordinate system with two degrees of freedom \( r \) dan \( \varphi \), position, yielding \( R \) value of parameter geometry for B-DNA are adopted from the helicoidal DNA model \[3\]. The values of \( h, R_0, \) and \( \Theta_0 \) adopted from B-DNA parameters \[3\]. When the opening base pair processing, the backbone interpreted by a spring will stretch, \( L_d \). The strands are hydrophobic, making the water molecules around base pairs eliminated from the core of DNA, by forming helicoidal backbone. The competition between the rigidity of backbone and the stacking interaction of the neighboring base pairs stabilizes the double helix structure of DNA. To get the effect of helicity, we should take three body curvature, and adding to the Lagrangian in third terms (1), where \( G_0 \) is a curvature constant.

The DNA-solvent interaction described by the solvent potential, in the last terms of the Lagrangian (1) with \( f_s \) is the solvent strength, and \( l_s \) is the width of solvent potential \[4\] that expresses the effective range. Thus the Lagrangian of our model is

\[
\mathcal{L} = \sum_n m \left( \dot{r}_n^2 + r_n^2 \dot{\varphi}_n^2 \right) - \sum_n D \left( e^{-\alpha(r_n-R_0)} - 1 \right)^2 - \sum_n G_0 \left( \varphi_{n+1} + \varphi_{n-1} - 2\varphi_n \right)^2 - \sum_n K \left( \sqrt{h^2 + r_n^2 - 2r_n \cos (\varphi_n - \varphi_{n-1})} - L^2 \right) + \sum_n D f_s \left( \tanh \left( \frac{r_n - R_0}{l_s} \right) - 1 \right) .
\]

The mass of a disk, \( m \), the depth of the Morse potential, \( D \), and the width of the Morse potential, \( \alpha \), are adopted from the PB model \[3\], \( m = 300 \) u.m.a., \( D = 0.04 \) eV, \( \alpha =4.45\)Å\(^{-1}\), while \( K = 1.0 \) eVÅ\(^{-2}\) and \( G_0 = KR_0^2/2 \) are adopted from the helicoidal DNA model \[3\]. The value of parameter geometry for B-DNA are \( R_0 \approx 10.0\) Å, \( \Theta_0 \approx 36^\circ \) and \( h = 3.4\) Å. In this model, the assumed constant \( h \) means that before and after solvation \( h \) is constant.

To simplify the Lagrangian, we present some dimensionless variables, \( r'_n = \alpha r_n \), \( t' = t(\sqrt{m/\alpha}) \), \( K' = \frac{K}{\alpha} \), \( G' = \frac{G_0}{\alpha} \), \( R'_0 = \alpha R_0 \), \( h' = \alpha h \), \( L' = \alpha L \), \( l'_s = \alpha l_s \). For simplicity hereafter, the new variables are written without prime. The shift of base, \( y_n \), or the arc length \( y_n = r_n - R_0 \) and the length of the disk position in angular, \( \phi_n = R_0(\varphi_n - n\Theta_0) \).

The stacking interaction (fourth term) is expanded to second order around the equilibrium position, yielding \( \frac{K_{yy}^2}{2} \left( y_{n+1} + y_{n-1} \right) (1 - \cos \Theta_0) + \left( \phi_{n+1} - \phi_{n-1} \right) \sin \Theta_0 \right)^2 \). The Morse potential and the solvent potential are expanded to forth order, so we get the equation of motion:

\[
\ddot{y}_n = \left( \frac{y_n}{R_0} + 1 \right) \frac{\partial^2 \mathcal{L}}{\partial \dot{y}_n^2} - \left( y_n - \frac{3}{2!} y_n^2 + \frac{7}{3!} y_n^3 \right) - K_{yy} \left( y_{n+1} + y_{n-1} \right) + \frac{f_s}{2l_s} \left( 1 - y_n^2 \right) ,
\]

and

\[
\ddot{\phi}_n = K_{\phi\phi} \left( \phi_{n+1} + \phi_{n-1} - 2\phi_n \right) + \frac{K_{y\phi}}{2} \left( y_{n+1} - y_{n-1} \right)
\]
The 6th International Conference on Mathematics and Natural Sciences

IOP Conf. Series: Journal of Physics: Conf. Series 1127(2019) 012013
doi:10.1088/1742-6596/1127/1/012013

\begin{align}
-2 & G (6 \phi_n - 4 \phi_{n+1} - 4 \phi_{n-1} + \phi_{n+2} + \phi_{n-2}) \\
-2 & R_0 \frac{\partial}{\partial t} y_n \phi_n - 2 R_0 \frac{\partial}{\partial t} y_n \phi_n - 2 R_0 \frac{\partial}{\partial t} y_n \phi_n - \frac{1}{R_0^2} y_n^2 \phi_n, \\
\end{align}

where $K_{\phi \phi} = \frac{K R_0^2}{L} \sin^2 \Theta_0; K_{y \phi} = 2 \frac{K R_0^2}{L} \sin \Theta_0 (1 - \cos \Theta_0) + K_{yy} = \frac{K R_0^2}{L} (1 - \cos \Theta_0)^2; G = \frac{G}{R_0}$. $K_{yy}$ and $K_{\phi \phi}$ are the effective elastic constant from the shifting or stretching of disk and twist rotation, $K_{y \phi}$ are the coupling constant between the stretching and twisting motion.

3. The envelope soliton solution for small amplitude

The soliton solution is obtained by using the analytical method developed by Cocco in [7]. We use the well-known technic to get the solution of envelope soliton for small amplitude use the Multiple Scale Expansion (MSE), based on the perturbation theory. The model of helicoidal DNA in solvent has two component wave vectors $E; y_n$ and $\phi_n$. First, we solve the linear order. Here the wave is packed with a weak dispersion. From (2) and (3) the equation of motion is written:

\begin{align}
\frac{\partial^2}{\partial t^2} \left( \begin{array}{c}
y_n \\
\phi_n
\end{array} \right) &= - \sum_{n', \alpha} J (n - n', \alpha, \alpha') \left( \begin{array}{c}
y_n \\
\phi_n
\end{array} \right)
\end{align}

with

\begin{align}
\hat{J} (q) &= \begin{pmatrix}
1 + 2 K_{yy} (1 + \cos q) & i K_{y \phi} \sin q \\
- i K_{y \phi} \sin q & 2 K_{\phi \phi} (1 - \cos q) + G (6 - 8 \cos q + 2 \cos 2q)
\end{pmatrix} \equiv \begin{pmatrix}
a & c \\
e^* & b\end{pmatrix}.
\end{align}

Now, we take care of the nonlinearity of the system to get the small amplitude envelope soliton. To increase the accuracy of the nonlinear term gradually, we present an expansion parameter $\epsilon$. We introduce the coordinates $x_1 = \epsilon x$, $t_1 = \epsilon t$, $t_2 = \epsilon^2 t$. The nonlinearity solution form is

\begin{align}
\tilde{E} (n) &= \epsilon e^{i(q_0 n_0 - \omega_1 t_0)} \left( \tilde{V}^+ - i \epsilon \tilde{V} (1) + \frac{\partial}{\partial x_1} \right) A (x_1, t_1, t_2) + c.c + \epsilon \tilde{\sigma} (x_1, t_1, t_2) + \epsilon^2 c \tilde{\gamma} (x_1, t_1, t_2) + c.c + \epsilon^2 \tilde{\mu} (x_1, t_1, t_2) + \epsilon^2 \tilde{\eta} (x_0, x_1, t_0, t_1, t_2).
\end{align}

with $A (x_1, t_1, t_2), \tilde{\sigma} (x_1, t_1, t_2), \tilde{\gamma} (x_1, t_1, t_2), \tilde{\mu} (x_1, t_1, t_2), \tilde{\eta} (x_0, x_1, t_0, t_1, t_2)$ are the slowly varying amplitudes that envelope the wave packet. The velocity of the envelope is smaller than the velocity of the carrier wave. The carrier wave is $e^{i(q_0 n_0 - \omega_1 t_0)}$ in first order and $e^{2i(q_0 n_0 - \omega_1 t_0)}$ in second order.

The wave packet is interpreted by first order $O(\epsilon)$, the second term arise because the linear system produce the nonzero solution on the force constant matrices $\hat{J} (0)$ at $q_0 = 0$, $\hat{J} (0) = \begin{pmatrix}
1 + 4 K_{yy} & 0 \\
0 & 0
\end{pmatrix}$ where $1 + 4 K_{yy}$ is the non zero solution of $\hat{J} (0)$. Hence there must be another new basis for non oscillating term in the solution, denoted by $\tilde{\sigma} (x_1, t_1, t_2)$. The second order $O(\epsilon^2)$ consists of the second harmonic term and the constant term, these terms are due to the quadratic term of the equations of motion. The last term appears because in the equation of motion present $O(\epsilon^0)$ term (2).

From the $O(\epsilon^0)$ and $O(\epsilon^2)$ term, we obtain the solutions for $\eta_1$ and $\eta_2$ are $\eta_1 = \frac{f_1}{2 \epsilon (1 + 4 K_{yy})}, \eta_2 = \text{const}$. We choose $\eta_2 = 0$. From the second order, $O (\epsilon^2)$, we get the solution
for $\gamma$, is $\tilde{\gamma} = \gamma_c A^2$. The solution of the constant term ($\epsilon^0$), $\tilde{\mu}$ and $\tilde{\sigma}$, are obtain from the second order and the third order. Checking the second order, we get $-2K_{\omega} \frac{\partial}{\partial x} \sigma_1 = 0$ and $\sigma_1 = \text{const.}$. From the non oscillating term of third order we get the solution for $\mu_2$ can be written

$$
\mu_2 = \frac{4\omega u_+^{(1)}}{K_{\omega} R_0} |V^+_2| \left| q_1 \right| \int |A|^2 dx_1 = \mu_2 c \int |A|^2 dx_1,
$$

we get the solution $\mu_1$ and $\sigma_2$, $\mu_1 = \mu_{1c} |A|^2$, $\sigma_2 = \sigma_{2c} |A|^2 dx_1$.

Now we are done with the envelopes, $\eta$, $\sigma$, $\tilde{\gamma}$, $\tilde{\mu}$, unless those are dependent on $A$. The nonlinear term $O(\epsilon^3)$ in $\exp(iq_0 u_0 - \omega_+ t_0)$ gives the equation for $A$ in form of Nonlinear Schrödinger (NLS). We found the NLS equation for envelope expression $A$ in moving frame with velocity $\omega_+^{(1)} (S_1 = x_1 - \omega_+^{(1)} t_1, t_2)$,

$$
E \frac{\partial^2 A}{\partial S_1^2} + F i \frac{\partial A}{\partial t_2} + H i \frac{\partial A}{\partial S_1} + IA + Q |A|^2 A = 0,
$$

This simplifies to $PA_{s_1 s_1} + iA_{t_2} + HiA_{s_1} + IA + Q |A|^2 A = 0$, with $P = \frac{E}{\gamma}, H = \frac{H}{\eta}, I = \frac{I}{\gamma}, Q = \frac{Q}{\eta}$. When $PQ > 0$, the bright soliton case, the envelope solution will be [8]

$$
A (s_1, t_2) = A \sec h \left[ \frac{1}{L_e} (s_1 - u_e t_2) \right] \exp \left[ i \left( \frac{u_e - H}{2P} \right) (s_1 - u_e t_2) \right],
$$

where $u_e$ is the velocity of the envelope and $u_e$ is the velocity of carrier, while the amplitude is

$$
A \equiv \sqrt{\frac{u_e^2 - 2u_e u_+ + 2u_+ \eta - \frac{2}{\eta} + 4 \frac{I}{Q}}{2PQ}}, \quad \text{and} \quad \frac{1}{L_e} = \sqrt{\frac{u_e^2 - 2u_e u_+ + 2u_+ \eta - \frac{2}{\eta} + \frac{I}{P}}{2PQ}}
$$

where $L_e$ is the inverse width of the envelope.

The complete solution is

$$
y = 2 \epsilon V_+^{(1)} A \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \cos (K x - \Omega t) + \epsilon \sigma_1 + \eta_1
$$

$$
+ \epsilon^2 \gamma_1 \epsilon A^2 \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \cos (2K x - 2\Omega t)
$$

$$
+ \epsilon^2 \mu_1 \epsilon A^2 \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] + \epsilon^2 V_+^{(1)} A \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right]
$$

$$
\left\{ - \frac{1}{L_e} \tanh \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \sin (K x - \Omega t) + \left( \frac{u_e - H}{P} \right) \cos (K x - \Omega t) \right\} + \mathcal{O} \left( \epsilon^3 \right),
$$

and

$$
\phi = \epsilon [\sigma_2 + \epsilon \mu_2] \frac{1}{L_e} A^2 \tanh \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] + \eta_2
$$

$$
+ 2i \epsilon \mathcal{V}^+_2 \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \sin (K x - \Omega t)
$$

$$
+ 2i \epsilon^2 \gamma_2 \epsilon A^2 \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \sin (2K x - 2\Omega t) + i \epsilon^2 V_+^{(1)} A \sec h \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right]
$$

$$
\left\{ 2 \frac{1}{L_e} \tanh \left[ \frac{1}{L_e} \epsilon (x - V_e t) \right] \cos (K x - \Omega t) + \left( \frac{u_e - H}{P} \right) \sin (K x - \Omega t) \right\} + \mathcal{O} \left( \epsilon^3 \right)
$$

(10)

where $V_e = \omega_+^{(1)} + \epsilon u_e$; $V_e = \omega_+^{(1)} + \epsilon u_e$; $K = q_0 + \epsilon \frac{u_e - H}{2P}$; $\Omega = \omega_+ + \epsilon \frac{u_e - H}{2P} V_e$. The variables are still dimensionless. We take the wave vector, $q_0 = 0.1$ and we choose $\epsilon = 0.1$, $u_e = -0.1$, $u_e = 0.1$. 

4. Effect of solvent on the stretching motion $y_n$

The analytical method produces the soliton solution like a breather. The value of the non-oscillating factor $\sigma_{1c}$ is taken small to prioritize the solvent effect, $\sigma_{1c} = 0.001$. To understand the base pairs opening, the analytical solution is plotted in $(x, t)$ plot under the varying solvent strength factor $f_s$. The stretch-inverse stretch plot $y_n$ vs. $-y_n$ is plotted to give a clear view of the base opening process.

Figure 1. $y$ and its inverse $(-y)$ vs. $x$ for $t = 0$ with the variation of $f_s$. We obtain the positive value of a solvent strength factor make the hydrogen bond stretch.

Figure 1 shows the base pair opening after solvation with positive value of $f_s$. The base pair opening ($Y_n$) for a solvent having strength $f_s$ at $t = 0$ is $Y_n = y_n - (-y_n) = 2y_n$. The more positive value of the $f_s$ the more the hydrogen bonds are stretched. The value of the base pair opening in Figure 1 are still dimensionless. The real value of the base pair with $f_s = 0.1$ is $Y_n = 2 \times (0.0006 \text{ Å} < y_n < 0.025 \text{ Å})$. If we compare to the diameter of B-DNA, 20 Å, the solvent effect in the stretching motion is small.

For each $x = -50$ and $x = 50$ without a solvent $y = -y = 0$ that means the base pairs opening is zero $Y = 0$. After solvation $Y(x = 50) > 0$ indicated that the minimum point of the bound state (the bound state is the DNA breathing condition) no more at $Y = 0$ (the base pair in the stable condition). The shifting of the stable condition is caused by the Constant Force of the solvent, $S_F = f_s/2\alpha l_s$, the value of the constant force depends on the type of the solvent ($f_s$). The constant force is obtained from $O(\epsilon^0)$ of the equation of motion. It forces the base outward the DNA axis.

Effect of solvent on the twisting motion

The solvation affects DNA gets more winded (Figure 2) and increasingly steep of the kink-soliton solution. The arc length of the base pair for $x = 0$ is $\phi_0 = 0$, then the arc length of the neighbours have the same value but have an opposite direction.
5. Conclusion
The solvent effect on the helicoidal DNA dynamics especially for stretching motion changes the stability of the open state, it is showed in the shifting of the minimum point to a positive value, $y_n > 0$. The shifting is caused by the constant force due to the existence of solvent. For twisting motion, the solvation caused the arc length of the disk to be more positive in value and more negative in value with the initial condition at $x = 0$. This means that DNA after solvation is more winded for the more positive value of the solvent barrier factor.

The hydrophobic effect of DNA strand forces the structure to be double helix. If the environment of DNA consist of water effect the backbone winding. Because of the behaviour of the base pair, it can be concluded that DNA will be more winded if there is some water around the DNA. Related to the analytical solution of this model, that the positive value of $f_s$ lets the DNA more winded, we get can suggest that a positive $f_s$ correspond to the situation around DNA with high water content.

References
[1] Peyrard M 2004 *Nonlinearity* **17** R1
[2] Dauxois T 1991 *Physics Letters A* **159** 390–395
[3] Barbi M, Cocco S and Peyrard M 1999 *Physics Letters A* **253** 358–369
[4] Zoli M 2011 *The Journal of chemical physics* **135** 115101
[5] Poater J, Swart M, Guerra C F and Bickelhaupt F M 2012 *Computational and Theoretical Chemistry* **998** 57–63
[6] Fang Y, Hoh J H and Spisz T S 1999 *Nucleic acids research* **27** 1943–1949
[7] Cocco S, Barbi M and Peyrard M 1999 *Physics Letters A* **253** 161–167
[8] Remoissenet M 2013 *Waves called solitons: concepts and experiments* (Springer Science & Business Media)