# Soliton excitations in deoxyribonucleic acid (DNA) double helices 

S. Yomosa<br>Department of Physics, Nagoya University, Nagoya 464, Japan

(Received 15 December 1982)


#### Abstract

We present in this paper a soliton theory for the open states in deoxyribonucleic acid (DNA) and synthetic polynucleotide double helices. Kink and antikink solutions for the equation of motion of the sine-Gordon form correspond to the open states with positive and negative helicities. The energy of open form and the length of the open configuration which are theoretically estimated are in the same order with the values inferred from kinetic experimental data.


## I. INTRODUCTION

It has been known that hydrogens in deoxyribonucleic acid (DNA) and synthetic polynucleotide double helices exchange with solvent hydrogens under conditions in which these molecules are in ordered states. ${ }^{1-4}$ Recently, the existence of open states in DNA and synthetic polynucleotide double helices has been demonstrated by hydrogen-deuterium exchange measurements. ${ }^{5-7}$ However, there is no direct information about the conformation of the open form. Assuming a mobile open unit ${ }^{8}$ diffusing along the double helix, Englander et al. ${ }^{9}$ concluded from the data of kinetic and thermodynamic experiments ${ }^{5}$ that these open configurations can consist of mobile segments on the order of 10 base pairs in length. As a model for open configuration of the double-helical polynucleotide chains, they proposed two parallel rods on which pendula are attached and suggested that the open state is consistently described as a solitary excitation. In this paper we present a soliton theory which describes the open states in DNA and synthetic polynucleotide double helices.

## II. HAMILTONIAN AND EQUATION OF MOTION

The $B$ form of DNA and polynucleotide double helices in the ordered arrangement of the WatsonCrick model is schematically represented in Fig. 1(a), where each arrow shows the direction of the base attached to the strand, and the complementary base pairs are indicated by the conjugated arrows arranged in horizontal parallel planes in a distance of $a=3.4 \AA$. The $z$ axis in the figure is a tenfold screw axis. In Fig. 1(b), each base is projected as a small ellipse with an arrow in the $x y$ plane, where $B_{n}$ denotes the $n$th base belonging to one of the two
strands, and $B_{n}^{\prime}$ denotes the complementary base belonging to the other strand, and these form H bonded base pairs in an ordered state. The directions of $B_{n}$ and $B_{n}^{\prime}$ in the horizontal plane are specified by the rotational angles $\theta_{n}$ and $\theta_{n}^{\prime}$ around the axes $P_{n}$ and $P_{n}^{\prime}$ which pass through the points where the bases $B_{n}$ and $B_{n}^{\prime}$ attach to the strands and are parallel to the $z$ axis, respectively.

The conformation and the stability of DNA and the polynucleotide double helices are mainly determined by the energy of H bonds between interstrand complementary base pairs and the stacking energy between intrastrand adjacent bases. These interaction energies consist of mainly the electrostatic, the exchange, and the induction and dispersion interaction energies. The molecular-orbital studies on the


FIG. 1. (a) Schematical representation of the WatsonCrick model. (b) Horizontal projections of the complementary base pairs.
hydrogen-bonding and stacking energies have been carried out by many authors, ${ }^{10-14}$ and it has been known from the results of the calculations that the magnitudes of the average hydrogen-bonding and stacking energies per base pair of double-stranded DNA-like polymers are in good parallelism compared with the melting temperatures of the polymers, and the intrastrand stacking bases have the potential minimum at the distances near $3.4 \AA$, and the minimum of stacking energy of the doublestranded polymer for rotation of the base pair around the helix axis exists near $36^{\circ} .^{14}$ Although in the above calculations they considered the relative rotation of normally H -bonded base pairs around the helix axis neglecting the constraint arising from the strands, we consider here independent rotation of the base $B$ around the axis $P$ which passes through the point where the base attaches to the strand and is parallel to the $z$ axis [see Fig. 1(b)] neglecting the torsional constraint of the strands, since we can assume the free rotations around the single bonds of sugars and phosphates constracting the strands of DNA.

Assuming that the H -bonding and the stacking energies, which consist of the electrostatic, the ex-
change, the charge-transfer, and the induction and dispersion interactions, are roughly proportional to the overlaps of molecular orbitals, the energy of $\mathbf{H}$ bonds between interstrand complementary base pairs may be represented as

$$
\begin{equation*}
\sum_{n} B\left[1-\cos \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)\right] \tag{2.1}
\end{equation*}
$$

and the stacking energy between intrastrand adjacent bases may be written as

$$
\begin{align*}
\sum_{n}\{ & S\left[1-\cos \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)\right] \\
& \left.+S\left[1-\cos \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)\right]\right\} \tag{2.2}
\end{align*}
$$

where $\alpha_{0}=36^{\circ}$. Here the mean energy of distorted double and triple H bonds in $A-T$ (adenine-thymine) and G-C (guanine-cytosine) base pairs is approximately represented in a similar form with the Pople's formula ${ }^{15}$ for the energy of a distorted single $H$ bond. The zero level of the energies in Eqs. (2.1) and (2.2) are taken for the $B$ form of DNA and polynucleotide duplexes. Then the Hamiltonian of DNA and synthetic polynucleotide double helices is written as

$$
\begin{align*}
H=\sum_{n}\{ & \frac{1}{2} I\left[\dot{\theta}_{n}^{2}+\dot{\theta}_{n}^{2}\right]+B\left[1-\cos \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)\right] \\
& \left.+S\left[1-\cos \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)\right]+S\left[1-\cos \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)\right]\right\} \tag{2.3}
\end{align*}
$$

where the first term represents the rotational kinetic energy, in which $I$ is the mean value of the moments of inertia of the bases for the rotations around the axes $P$.

Denoting by $\Pi_{n}$ and $\Pi_{n}^{\prime}$ the conjugated momenta for $\theta_{n}$ and $\theta_{n}^{\prime}$, respectively,

$$
\begin{equation*}
\Pi_{n}=\partial L / \partial \dot{\theta}_{n}=I \dot{\theta}_{n}, \quad \Pi_{n}^{\prime}=\partial L / \partial \dot{\theta}_{n}^{\prime}=I \dot{\theta}_{n}^{\prime} \tag{2.4}
\end{equation*}
$$

our Hamiltonian and the Hamilton equation of motion are written as follows:

$$
\begin{align*}
& H\left(\theta_{n}, \theta_{n}^{\prime} ; \Pi_{n}, \Pi_{n}^{\prime}\right)=\sum_{n}\left[\frac{1}{2 I}\left(\Pi_{n}^{2}+\Pi_{n}^{2}\right)+B\left[1-\cos \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)\right]\right. \\
&  \tag{2.5}\\
& \left.\quad+S\left[1-\cos \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)\right]+S\left[1-\cos \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)\right]\right], \\
& \dot{\theta}_{n}=\partial H / \partial \Pi_{n}=\Pi_{n} / I, \\
& \dot{\Pi}_{n}=-\partial H / \partial \theta_{n}=-B \sin \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)-S\left[\sin \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)-\sin \left(\theta_{n+1}-\theta_{n}-\alpha_{0}\right)\right], \\
& \dot{\theta}_{n}^{\prime}=\partial H / \partial \Pi_{n}^{\prime}=\Pi_{n}^{\prime} / I,  \tag{2.6}\\
& \dot{\Pi}_{n}^{\prime}=-\partial H / \partial \theta_{n}^{\prime}=B \sin \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)-S\left[\sin \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)-\sin \left(\theta_{n+1}^{\prime}-\theta_{n}^{\prime}-\alpha_{0}\right)\right] .
\end{align*}
$$

From Eq. (2.6), we obtain a coupled set of equations of motion

$$
\begin{align*}
& \ddot{\theta}_{n}+B \sin \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)+S\left[\sin \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)-\sin \left(\theta_{n+1}-\theta_{n}-\alpha_{0}\right)\right]=0  \tag{2.7}\\
& I \ddot{\theta}_{n}^{\prime}-B \sin \left(\theta_{n}-\theta_{n}^{\prime}-\pi\right)+S\left[\sin \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)-\sin \left(\theta_{n+1}^{\prime}-\theta_{n}^{\prime}-\alpha_{0}\right)\right]=0
\end{align*}
$$

Assuming the smallness of the angles $\theta_{n}-\theta_{n-1}-\alpha_{0}$ and $\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}$, we may approximate

$$
\begin{aligned}
& \sin \left(\theta_{n}-\theta_{n-1}-\alpha_{0}\right)-\sin \left(\theta_{n+1}-\right.\left.\theta_{n}-\alpha_{0}\right) \\
& \simeq 2 \theta_{n}-\theta_{n+1}-\theta_{n-1}, \\
& \sin \left(\theta_{n}^{\prime}-\theta_{n-1}^{\prime}-\alpha_{0}\right)-\sin \left(\theta_{n+1}^{\prime}-\theta_{n}^{\prime}-\alpha_{0}\right) \\
& \simeq 2 \theta_{n}^{\prime}-\theta_{n+1}^{\prime}-\theta_{n-1}^{\prime} .
\end{aligned}
$$

If $\theta_{n}$ and $\theta_{n}^{\prime}$ do not vary too rapidly with $n$, we can introduce the fields of rotational angles with continuum approximation,

$$
\begin{equation*}
\theta_{n}(t) \rightarrow \theta(z, t), \quad \theta_{n}^{\prime}(t) \rightarrow \theta^{\prime}(z, t), \tag{2.9}
\end{equation*}
$$

and we may write

$$
\begin{align*}
& 2 \theta_{n}-\theta_{n+1}-\theta_{n-1}=-a^{2} \frac{\partial^{2} \theta}{\partial z^{2}} \\
& 2 \theta_{n}^{\prime}-\theta_{n+1}^{\prime}-\theta_{n-1}^{\prime}=-a^{2} \frac{\partial^{2} \theta^{\prime}}{\partial z^{2}} \tag{2.10}
\end{align*}
$$

Then Eqs. (2.7) reduce to

$$
\begin{align*}
& I \frac{\partial^{2} \theta}{\partial t^{2}}-S a^{2} \frac{\partial^{2} \theta}{\partial z^{2}}+B \sin \left(\theta-\theta^{\prime}-\pi\right)=0 \\
& I \frac{\partial^{2} \theta^{\prime}}{\partial t^{2}}-S a^{2} \frac{\partial^{2} \theta^{\prime}}{\partial z^{2}}-B \sin \left(\theta-\theta^{\prime}-\pi\right)=0 \tag{2.11}
\end{align*}
$$

Taking the difference of the two equations (2.11), we obtain the sine-Gordan equation

$$
\begin{equation*}
S a^{2} \frac{\partial^{2} \phi}{\partial z^{2}}-I \frac{\partial^{2} \phi}{\partial t^{2}}=2 B \sin \phi \tag{2.12}
\end{equation*}
$$

or

$$
\begin{equation*}
\frac{\partial^{2} \phi}{\partial z^{2}}-\frac{1}{v_{0}^{2}} \frac{\partial^{2} \phi}{\partial t^{2}}=\frac{1}{l^{2}} \sin \phi \tag{2.13}
\end{equation*}
$$

where

$$
\begin{align*}
& \phi=\theta-\theta^{\prime}-\pi \\
& v_{0}=\left(\frac{S}{I}\right)^{1 / 2} a, \quad l=\left(\frac{S}{2 B}\right)^{1 / 2} a . \tag{2.14}
\end{align*}
$$

It has the ground-state solutions

$$
\begin{equation*}
\phi(z)=0, \pm 2 \pi, \pm 4 \pi, \ldots, \tag{2.15}
\end{equation*}
$$

which correspond to the $B$ form of DNA.

## III. SOLITON SOLUTIONS

If we look for the solution of Eq. (2.13) which depends upon only a single variable

$$
\begin{equation*}
\xi=z-v t \tag{3.1}
\end{equation*}
$$

then Eq. (2.13) reduces to

$$
\begin{equation*}
\frac{d^{2} \phi}{d \zeta^{2}}=\frac{\gamma^{2}}{l^{2}} \sin \phi \tag{3.2}
\end{equation*}
$$

where

$$
\begin{equation*}
\gamma=\left(1-v^{2} / v_{0}^{2}\right)^{-1 / 2}, \tag{3.3}
\end{equation*}
$$

where we assume $v<v_{0}$. We impose the boundary condition

$$
\begin{equation*}
\cos \phi=1 \quad(\phi=2 \pi n, \quad n=0, \pm 1, \ldots) \tag{3.4}
\end{equation*}
$$

at

$$
\zeta= \pm \infty \quad(\text { at } z= \pm \infty \text { for all } t) .
$$

With this boundary condition, Eq. (3.2) is integrated at once,

$$
\begin{equation*}
\left(\frac{d \phi}{d \xi}\right)^{2}=4 \frac{\gamma^{2}}{l^{2}} \sin ^{2} \frac{\phi}{2} \tag{3.5}
\end{equation*}
$$

which gives

$$
\begin{equation*}
\int \frac{d \phi}{2 \sin \frac{\phi}{2}}= \pm \frac{\gamma}{l}\left(\zeta-\zeta_{0}\right) \tag{3.6}
\end{equation*}
$$

Then we have the solitary wave solutions of the form

$$
\begin{equation*}
\phi=4 \tan ^{-1} e^{ \pm(\gamma / l)\left(\zeta-\zeta_{0}\right)} . \tag{3.7}
\end{equation*}
$$

The numerical results of the solutions are shown in Fig. 2. The Lorentz-type contraction of a moving soliton with a velocity $v\left(v<v_{0}\right)$ is represented through the parameter $\gamma$ which is related to the velocity $v$. The upper limit of soliton velocity will be estimated later to be $v_{0} \simeq 1.1 \times 10^{4} \mathrm{~cm} / \mathrm{sec}$. Kink $(+)$ and antikink $(-)$ solutions represent localized disturbances in the $\phi$ field. The plus and minus signs correspond to the two possible helicities of the solutions. These kink and antikink solitons predict the existence of the positive and the negative open states in DNA and synthetic polynucleotide double helices.

In order to obtain direct information about the conformation of the kink and antikink open forms, we consider static solitons by setting $\gamma=1$ in Eq. (3.7). In the static case, Eq. (2.11) gives


FIG. 2. Kink and antikink soliton solutions.

$$
\begin{equation*}
\frac{\partial^{2} \theta}{\partial z^{2}}=\frac{B}{S a^{2}} \sin \phi \tag{3.8}
\end{equation*}
$$

Inserting Eq. (3.7) into Eq. (3.8) and using Eq. (2.14) we obtain

$$
\begin{equation*}
\frac{\partial^{2} \theta}{\partial z^{2}}=\frac{1}{2 l^{2}} \sin \left(4 \tan ^{-1} e^{ \pm\left(z-z_{0}\right) / l}\right) \tag{3.9}
\end{equation*}
$$

## Setting

$$
\begin{equation*}
p=\tan ^{-1} e^{ \pm\left(z-z_{0}\right) / l} \tag{3.10}
\end{equation*}
$$

and considering the boundary condition, Eq. (3.19) is integrated

$$
\begin{equation*}
\frac{\partial \theta}{\partial z}=\frac{\alpha_{0}}{a} \pm \frac{1}{l} \sin \left(2 \tan ^{-1} e^{ \pm\left(z-z_{0}\right) / l}\right) \tag{3.11}
\end{equation*}
$$

which gives

$$
\theta=\frac{\alpha_{0}}{a} \int_{-\infty}^{z} d z+\left\{\begin{array}{l}
2 \tan ^{-1} e^{\left(z-z_{0}\right) / l} \text { (for kink) }  \tag{3.12}\\
2 \tan ^{-1}\left(e^{-\left(z-z_{0}\right) / l}\right)-\pi \quad(\text { for antikink) }
\end{array}\right.
$$

On the other hand, Eq. (2.11) gives the similar equation for $\theta^{\prime}$ with that for $\theta$ in the static case

$$
\begin{equation*}
\frac{\partial^{2} \theta^{\prime}}{\partial z^{2}}=-\frac{1}{2 l^{2}} \sin \left(4 \tan ^{-1} e^{ \pm\left(z-z_{0}\right) / l}\right) \tag{3.13}
\end{equation*}
$$

which is integrated to give

$$
\begin{align*}
& \frac{\partial \theta^{\prime}}{\partial z}=\frac{\alpha_{0}}{a} \mp \frac{1}{l} \sin \left(2 \tan ^{-1} e^{ \pm\left(z-z_{0}\right) / l}\right),  \tag{3.14}\\
& \theta^{\prime}=-\pi+\frac{\alpha_{0}}{a} \int_{-\infty}^{z} d z-\left\{\begin{array}{l}
2 \tan ^{-1} e^{\left(z-z_{0}\right) / l} \quad \text { (for kink) } \\
2 \tan ^{-1}\left(e^{-\left(z-z_{0}\right) / l}\right)-\pi \quad \text { (for antikink) }
\end{array}\right. \tag{3.15}
\end{align*}
$$

Then, from Eqs. (3.12) and (3.15), we have

$$
\phi=\theta-\theta^{\prime}-\pi=\left\{\begin{array}{l}
4 \tan ^{-1} e^{\left(z-z_{0}\right) / l} \quad(\text { for kink })  \tag{3.16}\\
4 \tan ^{-1}\left(e^{-\left(z-z_{0}\right) / l}\right)-2 \pi \quad(\text { for antikink }) .
\end{array}\right.
$$

$\theta, \theta^{\prime}$, and $\phi$ fields are shown in Fig. 3, where the positive and the negative open states which have different helicities are denoted by $(+)$ and $(-)$ signs.

In order to visualize the conformation of the open state, we denote the bases belonging to one strand by the numbers $0,1,2, \ldots$, and the complementary bases belonging to another strand by the numbers $0^{\prime}, 1^{\prime}, 2^{\prime}, \ldots$ If we suppose the positive static open state which is centering at the base pair $5-5^{\prime}$, the $z$ coordinate of the base pair 5-5' corresponds to $z_{0}$ in the horizontal coordinate $\left(z-z_{0}\right) / l$ of Fig. 3. Assuming $l=0.5 a$, the horizontal coordinate of the $n$th base is written as $\left(z_{n}-z_{5}\right) / l=2\left(z_{n}-z_{5}\right) / a$. Then, the direction of the arrow attached to each base in the open region which deviates from that in the ordered state can be obtained for each strand graphically by using Fig. 3. The directions of the bases in the positive open state thus obtained are described for each strand separately in Figs. 4(a) and 4(b) in horizontal projections. The dotted lines in Figs. 4(a) and 4(b) indicate the directions of bases in an or-
dered state. Using these directional arrows in Figs. 4(a) and 4(b) and considering low-energy arrangements of complementary base pairs, we can describe the directional configuration of $2 \pi$ soliton of field $\phi$ or $\pi$ soliton of fields $\theta$ and $\theta^{\prime}$ in Fig. 4(c). This figure [4(c)] is the theoretical representation of the model of the open state of DNA proposed by Teitel-


FIG. 3. $\pi$ soliton in fields $\theta$ and $\theta^{\prime}$, or $2 \pi$ soliton in field $\phi$.

(a)

(b)

(c)

(d)

FIG. 4. Directionals only of the bases in the static positive open state are described for each strand separately in (a) and (b). In (c), these directional arrows of both strands are described together in the same horizontal projection, considering low-energy arrangements of complementary base pairs. Model of open state of DNA proposed by Teitelbaum and Englander is shown in (d).
baum and Englander, ${ }^{8}$ which is schematically shown in Fig. 4(d).
Returning to Eq. (2.11), if we take the sum of two equations of (2.11), we obtain a wave equation

$$
\begin{equation*}
\frac{\partial^{2} \varphi}{\partial z^{2}}-\frac{1}{v_{0}^{2}} \frac{\partial^{2} \varphi}{\partial t^{2}}=0 \tag{3.17}
\end{equation*}
$$

where

$$
\begin{equation*}
\varphi=\theta+\theta^{\prime}+\pi, \quad v_{0}=\left(\frac{S}{I}\right)^{1 / 2} a \tag{3.18}
\end{equation*}
$$

Then we can see that the upper limit of the velocity of the $\phi\left(\equiv \theta-\theta^{\prime}-\pi\right)$ soliton is given by the velocity $v_{0}$ of the wave equation for $\varphi\left(\equiv \theta+\theta^{\prime}+\pi\right)$.

We add here the consideration about small oscillations around $\phi=0$ which correspond to the phonon mode. Assuming $\phi$ is small, we obtain for Eq. (3.2)

$$
\begin{equation*}
\frac{d^{2} \phi}{d \zeta^{2}}=-\frac{\left(v^{2} / v_{0}^{2}-1\right)}{l^{2}} \phi \tag{3.19}
\end{equation*}
$$

In the case of $v>v_{0}$, the solutions are small oscillations in the form

$$
\begin{align*}
\phi & =A \sin [q(z-v t)+\theta] \\
& =A \sin (q z-\omega t+\theta), \tag{3.20}
\end{align*}
$$

where $A$ and $\theta$ are constants, and $q$ is given by

$$
\begin{equation*}
q^{2}=\frac{v^{2} / v_{0}^{2}-1}{l^{2}}, v>v_{0} \tag{3.21}
\end{equation*}
$$

The dispersion relation of the phonon mode is obtained as

$$
\begin{equation*}
\omega=\left(1+l^{2} q^{2}\right)^{1 / 2} v_{0} q \tag{3.22}
\end{equation*}
$$

## IV. ESTIMATION OF THE PARAMETER VALUES AND DISCUSSIONS

Our theory contains two parameters $B$ and $S$. These parameter values can be estimated by using the results of the molecular-orbital (MO) studies on the hydrogen-bonding and the stacking interaction energies. In the calculations of Devoe and Tinoco ${ }^{10}$ and of Claverie, Pullman, and Caillet, ${ }^{11}$ ionization potentials or polarizabilities were used in the calculation of the induction and dispersion energies. Rein and Pollak ${ }^{12}$ calculated these energies by the Pariser-Parr-Pople (PPP) method, and Bertran ${ }^{13}$ by the complete neglect of differential overlap (CNDO) method. Fujita et al. ${ }^{14}$ performed more detailed calculations by CNDO and PPP methods and reported that the hydrogen-bonding energies in the $B$ form were -0.48 eV (CNDO) and -0.39 eV (PPP) for the $G-C$ base pair, and -0.076 eV (CNDO) and -0.070 eV (PPP) for the $A-T$ base pair. Then mean value is -0.28 eV (CNDO) and -0.23 eV (PPP). The mean stacking energy between base pairs in the $B$ form is calculated from their results to be -0.15 eV (CNDO) and -0.05 eV (PPP). Therefore two parameters in the Hamiltonian (2.3) can be estimated by using these values (in eV )

$$
\begin{align*}
2 B & = \begin{cases}0.28 & (\mathrm{CNDO}) \\
0.23 & (\mathrm{PPP})\end{cases} \\
4 S & = \begin{cases}0.15 & (\mathrm{CNDO}) \\
0.05 & (\mathrm{PPP}) .\end{cases} \tag{4.1}
\end{align*}
$$

Then we can estimate the values of $l$ and $v_{0}$ by Eq. (2.14),

$$
\begin{align*}
l & =\left(\frac{S}{2 B}\right)^{1 / 2} a= \begin{cases}0.37 a & (\mathrm{CNDO}) \\
0.24 a & (\mathrm{PPP})\end{cases} \\
v_{0} & =\left(\frac{S}{I}\right)^{1 / 2} a=\left(\frac{S}{M}\right)^{1 / 2} \frac{a}{R}=0.8 \\
& \sim 1.3 \times 10^{4} \mathrm{~cm} / \mathrm{sec}, \tag{4.2}
\end{align*}
$$

where we used $M=130 m_{p}$ and $R=4.0 \AA$ as the mean mass of the four kinds of bases and the mean radius of gyration of the bases for the rotation around the axes $P$, respectively.

The length of the open configuration in the case of a slow moving or static soliton is estimated from Fig. 2, where static kink size is given by $\Delta \zeta / l$
$=\Delta z / l \simeq 14$; then the length of the open state is estimated as follows by Eq. (4.2):

$$
\Delta z \simeq 14 l= \begin{cases}5.2 a & (\mathrm{CNDO})  \tag{4.3}\\ 3.4 a & (\mathrm{PPP}) .\end{cases}
$$

For the purpose of approximate estimation of the energy of the open form, we consider slow moving or static solitons, where $\left(\xi-\zeta_{0}\right) / l \simeq\left(z-z_{0}\right) / l$, and we can neglect the kinetic part of solitons. The energy of the static soliton is calculated as follows by using Eqs. (3.7) and (3.11),

$$
\begin{align*}
E_{p} & =\int \frac{d z}{a}\left[B(1-\cos \phi)+4 S a^{2}\left[\frac{d \theta}{d z}-\frac{\alpha_{0}}{a}\right)^{2}\right] \\
& =4 B \frac{l}{a}+2 S \frac{a}{l} \tag{4.4}
\end{align*}
$$

and is estimated (in eV ) by using Eqs. (4.1) and (4.2),

$$
E_{p} \simeq \begin{cases}0.41 & (\mathrm{CNDO})  \tag{4.5}\\ 0.21 & (\mathrm{PPP})\end{cases}
$$

If we consider the fact that the energy calculated by the CNDO method is usually overestimated about two times and the fact that the absolute values of $2 B$ or $4 S$ correspond to the difference of the interaction energy between complementary bases or intrastrand bases and the interaction energy between bases and solvent molecules, including distorted hydrogen-bonding interactions with solvent molecules, then the absolute values (in eV ) of $2 B$ and $4 S$ are inferred to be reduced to about

$$
\begin{equation*}
2 B \simeq 0.10, \quad 4 S \simeq 0.10 \tag{4.6}
\end{equation*}
$$

Then the $l$ value, $v_{0}$, and the length $\Delta z$ of the open configuration may be about

$$
l=\left(\frac{S}{2 B}\right)^{2} a \simeq 0.5 a
$$

$$
\begin{align*}
& v_{0}=\left(\frac{S}{I}\right)^{1 / 2} a \simeq 1.1 \times 10^{4} \mathrm{~cm} / \mathrm{sec}  \tag{4.7}\\
& \Delta z \simeq 14 l \simeq 7 a
\end{align*}
$$

then, the energy (in eV ) of the static open state $E_{p}$ reduces to

$$
\begin{equation*}
E_{p} \simeq 0.20 \tag{4.8}
\end{equation*}
$$

On the other hand, Englander et al. ${ }^{9}$ and Nakanishi and Tsuboi ${ }^{7}$ estimated the energy of the open state from the kinetic experiments. They wrote the scheme for conformationally limited chemical exchange, when a deuteron is exchanged for a proton, as follows:

$$
{ }^{2} \mathrm{H}_{\text {closed }} \underset{k_{\mathrm{cl}}}{\stackrel{k_{\mathrm{op}}}{\stackrel{ }{\longrightarrow}}{ }^{2} \mathrm{H}_{\mathrm{open}} \xrightarrow{k_{\mathrm{chem}}}{ }^{1} \mathrm{H}_{\text {exchanged }}, \quad K_{\mathrm{eq}}=k_{\mathrm{op}} / k_{\mathrm{cl}}, \quad, \quad \text {, }}
$$

in which the conformational rate constants $k_{\text {op }}$ and $k_{\mathrm{cl}}$ represent opening and closing rates, and $k_{\mathrm{chem}}$ is the chemical rate constant for transfer of an NH proton from a nucleotide in the open state to the solvent. From the studies of the temperature dependence of $K_{\text {eq }}$, they estimated the formation enthalpy $\Delta H_{\mathrm{op}}$ for opening to be $6 \mathrm{kcal} / \mathrm{mol}(0.26 \mathrm{eV})$ and 8.1 $\mathrm{kcal} / \mathrm{mol}(0.35 \mathrm{eV}),{ }^{7}$ respectively. Thus the length of the open configuration and the energy of open form which are theoretically estimated are in agreement in order with the values estimated from the kinetic experimental data. In this paper, we neglected the constraints arising from sugar-phosphate strands of DNA and neglected the coupling with the longitudinal motion (accordian motion) of the bases, and have considered only the horizontal motion of the bases in order to extract a solitary wave solution in DNA by a highly simplified model. More detailed treatment will be developed in the future.
${ }^{1}$ M. Printz and P. H. von Hippel, Proc. Natl. Acad. Sci. U.S.A. 53, 361 (1965).
${ }^{2}$ S. W. Englander and J. J. Englander, Proc. Natl. Acad. Sci. U.S.A. 53, 370 (1965).
${ }^{3}$ J. J. Englander and P. H. von Hippel, J. Mol. Biol. 63, 171 (1972).
${ }^{4}$ S. W. Englander and J. J. Englander, Methods Enzymol. 49, 24 (1978).
${ }^{5}$ C. Mandal, N. R. Kallenbach, and S. W. Englander, J. Mol. Biol. 135, 391 (1979).
${ }^{6}$ M. Nakanishi and M. Tsuboi, Y. Saijo, and T. Nagamura, FEBS Lett. 81, 61 (1977).
${ }^{7}$ M. Nakanishi, and M. Tsuboi, J. Mol. Biol. 124, 61 (1978).
${ }^{8}$ H. Teitelbaum and S. W. Englander, J. Mol. Biol. 92, 55 (1975); 92, 79 (1975).
${ }^{9}$ S. W. Englander, N. R. Kallenbach, A. J. Heeger, J. A. Krumhansl, and S. Litwin, Proc. Natl. Sci. U.S.A. 77, 7222 (1980).
${ }^{10} \mathrm{H}$. Devoe and I. Tinoco, J. Mol. Biol. 4, 500 (1962).
${ }^{11}$ P. Claverie, B. Pullman, and J. Caillet, J. Theor. Biol. 12, 419 (1966).
${ }^{12}$ R. Rein and M. Pollak, J. Chem. Phys. 47, 2039 (1967).
${ }^{13}$ J. Bertran, J. Theor. Biol. 34, 353 (1972).
${ }^{14}$ H. Fujita, A. Imamura, and C. Nagata, J. Theor. Biol. 45, 411 (1974).
${ }^{15}$ J. A. Pople, Proc. R. Soc. London Ser. A 205, 163 (1951).

