DNA-DNA interaction beyond the ground state

D. J. Lee, A. Wynveen,* and A. A. Kornyshev
Department of Chemistry, Faculty of Physical Sciences, Imperial College London, SW7 2AZ London, United Kingdom
(Received 14 April 2004; published 23 November 2004)

The electrostatic interaction potential between DNA duplexes in solution is a basis for the statistical mechanics of columnar DNA assemblies. It may also play an important role in recombination of homologous genes. We develop a theory of this interaction that includes thermal torsional fluctuations of DNA using field-theoretical methods and Monte Carlo simulations. The theory extends and rationalizes the earlier suggested variational approach which was developed in the context of a ground state theory of interaction of nonhomologous duplexes. It shows that the heuristic variational theory is equivalent to the Hartree self-consistent field approximation. By comparison of the Hartree approximation with an exact solution based on the QM analogy of path integrals, as well as Monte Carlo simulations, we show that this easily analytically-tractable approximation works very well in most cases. Thermal fluctuations do not remove the ability of DNA molecules to attract each other at favorable azimuthal conformations, neither do they wash out the possibility of electrostatic "snap-shot" recognition of homologous sequences, considered earlier on the basis of ground state calculations. At short distances DNA molecules undergo a "torsional alignment transition," which is first order for nonhomologous DNA and weaker order for homologous sequences.

DOI: 10.1103/PhysRevE.70.051913 PACS number(s): 87.15.Kg, 87.15.Ya

I. INTRODUCTION

DNA is a polyelectrolyte molecule; in solution it dissociates leaving negative charges on phosphates and positive charges around the molecule in the electrolyte or readsorbed onto the DNA surface, that compensate partially the charge of the phosphates. Obviously, electrostatics should be important in the interaction between DNA molecules. This should affect the structure of DNA mesophases in solution, and, more so, the properties of DNA crystals or aggregates [1]. With appreciation of the importance of electrostatics, the so-called "polyelectrolyte model" treats a DNA molecule as a cylinder with a net charge homogenously distributed along its surface [2]. Such a model appeared to be insufficient in describing a number of properties of DNA in assemblies. Notably, Ref. [3] has shown that the electrostatic interaction between DNA duplexes depends dramatically on surface charge patterns, thereby invalidating the "polyelectrolyte model" and providing a plausible alternative. Taking into account that the negative charges follow the double helical symmetry of the molecule, and assuming that the readsorbed cations are distributed either homogeneously along the molecular surface or in a certain proportion between major or minor grooves, one obtains helical motifs with distinct separation between the positive and negative charges. Adsorption into the grooves causes the strongest charge separation. Two such distributions on juxtaposing molecules will attract each other under a favorable azimuthal alignment of the molecules [3]. This gives rise to a concept of the "electrostatic zipper motif for DNA aggregation" [4]. A number of consequences have been explored in the context of DNA assembling in columnar aggregates [5,6], including the measured decay length of short range repulsion, the laws of attraction, the transition from 10.5 base pairs per turn in solution to 10 bp in assemblies, and from *B* to *A* form in dry aggregates, etc.

Furthermore, the analytical solution of the problem of interaction of long DNA with nonparallel main axes was obtained [7,8]. One of the results was the discovered chiral torque which tends to twist the angles between the main axes of the molecules. The extension of this theory to assemblies of molecules of finite length [9] helped to rationalize a number of observed properties of DNA chiral liquid crystals.

These works triggered papers on the statistical theory of many particle DNA assemblies [10–13].

In Ref. [14] it was argued that the helical electrostatic zipper may be responsible for a *snapshot* recognition of homologous DNA sequences from a distance, the possibility of which was conjectured in molecular genetics [15,16]. Indeed, DNA is not an ideal staircase. Step angles are distorted for each step, and the pattern of distortions is related to the text of the sequence [17,18]. This affects the interaction between nonhomologous sequences. The first exploration of this hypothesis was based on the theory of electrostatic interaction of two *torsionally rigid* DNA fragments in parallel juxtaposition [3,4]. In Ref. [14] it was found that the interaction energy between two DNA fragments of uncorrelated sequences are considerably higher than two homologous sequences, and the difference grows with the juxtaposition length.

A simple interpretation was given to this effect in the case of rigid duplexes [14]. Two *homologous* duplexes have almost identical patterns of twist angle distortions, and they can be azimuthally aligned in such a way that the motifs of positive and negative charges will stay in register along the whole juxtaposition length. This causes attraction. On the contrary, two *nonhomologous* sequences have texts which are random relative to each other. Their step-angle distortions are, therefore, unrelated. The quasihelical charge distributions on the molecules can be positioned in register over a

*Electronic address: a.wynveen@imperial.ac.uk

section of a certain length, but they will lose register between the rest of the sequences, if they are long enough. Upon accumulation of the relative distortion, the two sides of the zipper will not match, making the attraction much weaker or even resulting in repulsion. The difference in the interaction energy between homologous and nonhomologous fragments allows the intact homologous sequences to recognize each other, and this difference will be more pronounced the longer the sequence.

The length over which two juxtaposing torsionally rigid duplexes completely lose register was called the *helical coherence* length. It is equal to $\lambda_c = h/\Delta\Omega^2$ [14], where h is the step rise and $\Delta\Omega$ is the standard deviation of the twist angle in one molecule. For *B*-DNA $h \approx 3.4$ Å, $\Delta\Omega \approx 0.07-0.1$ rad [19–21] so that $\lambda_c \approx 300-700$ Å. On length scales larger than λ_c the torsional mismatch accumulates as a random walk. This positive difference of the interaction between nonhomologous and between homologous duplexes of the same length was called the *recognition energy*. It was found to be $\gtrsim k_B T$ for sequences longer than 50 bp at DNA-DNA interaxial separations of $R = 30$ Å [14]. At closer separations the interaction strength increases nearly exponentially with diminishing R and so does the recognition energy.

This conclusion did not take into account the torsional elasticity of DNA, which allows for a relaxation of the mismatch, which may reduce the electrostatic energy, but costing an energy of torsional deformation. The Euler equation including both the distortion of twist angle patterns and torsional elasticity has been derived in Ref. [14] but was not explored there. It has the form of a sine-Gordon equation with frustration (caused by asymmetry between the major and minor grooves in a double stranded DNA) and with a nonlocal external random field, the consequence of a sequence-determined pattern of twist angle distortions. (Simpler forms of this equation appear in different physical contexts, viz. in the theory of Josephson junctions, nonlinear pendulum dynamics, and commensurate-incommensurate transitions, Refs. [22,23].) Its solution has been investigated in Refs. [24,25]. It was shown that the torsional softness is important, and that in many situations the mismatch does not accumulate but relaxes continuously or in an abrupt manner. This reduces but still does not eliminate the recognition energy.

At considerable interaxial separations, the recognition energy drops down to $k_B T$ per pair of duplexes. The same refers to the difference in the energies of columnar mesophases of ideal helical DNA [13,24]. This forces us to explore the role of thermal fluctuations, i.e., to move the theory from the ground state level (Euler equation) towards a calculation of path integrals determining the free energy. Our goal is to understand the role of fluctuations and to reveal the conditions where the conclusions of the ground state theory remain valid. Such an exploration will be performed in this paper for a pair of interacting DNAs in solution. We limit the analysis to parallel juxtaposition, which is sufficient for further applications of this pair potential to a description of columnar aggregates.

We first consider the case of two homologous (identical) DNA, and solve the problem first analytically in the Hartree approximation and then exactly, formulating a quantum mechanical (QM) analog of the problem and computing the spectrum of the eigenstates.

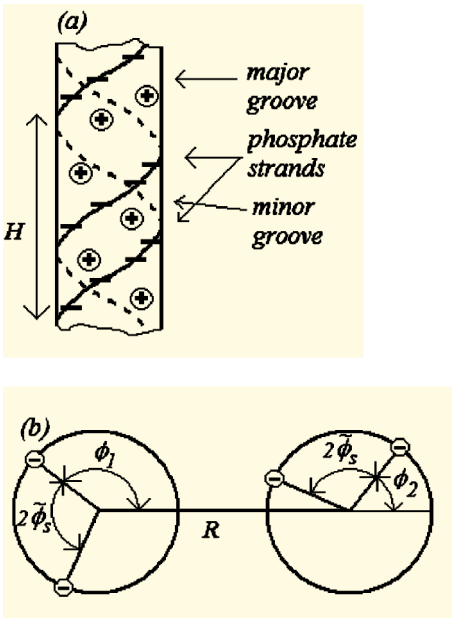

FIG. 1. The charge distribution of a double helix (a) and the horizontal cross section of two identical double helices in parallel juxtaposition (b). The double helix consists of two spiraling negative phosphate strands with cations readsorbed in its minor and major grooves. The pitch H of the helix for *B*-DNA is approximately 34 Å. The angle of the minor groove between the phosphate strands $2\tilde{\phi}_s$ is about 0.8π for *B*-DNA. The parallel DNA are separated by an interaxial spacing R and the relative azimuthal angle between them is $\phi = \phi_1 - \phi_2$.

Next, we extend the Hartree result to the case of nonhomologous duplexes, thus obtaining a combined description of both the effects of torsional thermal fluctuations and nonhomology. We did not find any possibility for a similar extension of the exact QM analog. However, our extended Hartree approximation, which reproduces the variational solution of Ref. [24] as a particular case, is good for a large class of systems. We discuss in detail its advantages and limitations.

Finally, we compare the Hartree results with finite temperature Monte Carlo simulations of torsionally flexible homologous and nonhomologous DNAs using a specially developed discrete analog of the model.

Most of the "field-theoretical" algebra is appended [26], but its key points are left in the main text. We do our best, however, to outline the derivation so that the reader not interested in mathematical details will be able to follow the ideas of the work presented in the main text.

II. PATH INTEGRAL FORMULATION FOR IDENTICAL DNA

A. The partition function of two DNA in parallel juxtaposition

For each of the two molecules, we shall assume a negative charge density of phosphates spiraling along the two strands of the double helix and a positive charge density of

readsorbed cations spiraling along the major and minor groves (Fig. 1). We consider the patterns of adsorbed counterions as fixed, irreversible and unaffected by the separation between DNA or by torsional fluctuations. This is the case of strong chemisorption of DNA condensing counterions onto the DNA surface, seemingly typical for spermine, spermidine, cobalt-hexamine and some other DNA condensers.

Mainly, this work will focus on the situation when DNA molecules have ideal helical structure or they are nonideal, but identical (same sequence of base pairs). However, in the end we will extend this to different sequences on the opposing DNA.

The ground state DNA-DNA electrostatic interaction energy per unit length of juxtaposition given as a function of interaxial separation, R and azimuthal angle, ϕ, has the form [4,14]

$$w_{\rm int}(R,\phi) \approx a_0(R) - a_1(R)\cos(\phi) + a_2(R)\cos(2\phi). \tag{2.1}$$

The expressions for a_0, a_1, and a_2 are provided in the Appendix; for a detailed discussion we refer the reader to previous work [4,14]. What is important is that these coefficients fall off exponentially at large R. Simply, by minimizing this energy with respect to ϕ, we get $\phi=0$ as the preferred relative azimuthal angle at $a_1/4a_2>1$, and when $a_1/4a_2<1$ this angle is $\phi \equiv \pm\phi_0 = \pm\cos^{-1}(a_1/4a_2)$. The ratio of $a_1/4a_2$ depends on R. It is large at large R, but small at small R. Below a critical value R_* this ratio is small enough for ϕ to be nonzero.

For identical sequences or ideally helical DNA, in the ground state, ϕ does not depend on the position along the helices, z. However, our aim is to include thermal fluctuations due to the twisting of one DNA molecule with respect to another; then ϕ will depend on z. To incorporate this effect, we must include a torsional energy term [14,24,25]

$$w_{tor} = \frac{C\phi'^2}{4}, \tag{2.2}$$

where C is the torsional elasticity modulus of the helices and $\phi' = d\phi/dz$.

We are now able to write down a partition function as a path integral (defined as the continuum limit of a product of integrals over ϕ_z, one for each point z along the helices),

$$Z = \int \mathcal{D}\phi \, \exp\left(-\frac{E[\phi]}{k_BT}\right),$$

$$E[\phi] = \int_{-L/2}^{L/2} dz (w_{\rm int}[R,\phi(z)] + w_{tor}[\phi'(z)]). \tag{2.3}$$

We shall, for the moment, assume the length of the helices, L, is very large and ignore all finite size corrections. In addition to the partition function, we may calculate the thermal average of any "ϕ-dependent" observable quantity

$$\langle \mathcal{A} \rangle = \frac{1}{Z}\int \mathcal{D}\phi \, \mathcal{A}[\phi]\exp\left(-\frac{E[\phi]}{k_BT}\right). \tag{2.4}$$

B. Perturbation theory: Gaussian approximation

The first step beyond the ground state is the Gaussian approximation which is good when the fluctuations are small. Below we will write a criterion when this approximation holds.

In order to build the perturbation theory around the mean field solution, $\phi(z)=\phi_*$ we utilize the following expansion:

$$w_{\rm int}[R,\phi(z)=\phi_*+\varphi(z)] = \sum_{n=0}^{\infty} \frac{(-1)^n[a_1 \sin(\phi_*)\varphi^{2n+1} - a_2 \sin(2\phi_*)(2\varphi)^{2n+1}]}{(2n+1)!} - \sum_{n=0}^{\infty} \frac{(-1)^n[a_1 \cos(\phi_*)\varphi^{2n} - a_2 \cos(2\phi_*)(2\varphi)^{2n}]}{(2n)!}, \tag{2.5}$$

and consider all terms of higher order than $O(\varphi^2)$ as perturbations.

Let us start by calculating the free energy in the Gaussian approximation, in which we discard in Eq. (2.5) all terms of higher order than quadratic and set $\phi_*=\phi_0$, the value of ϕ "at $T=0$;" more precisely, in the *ground state*, because the electrostatic interaction coefficients a_0, a_1, and a_2 are also functions of temperature. Here, we may write

$$E[\varphi(z)] = Lw_{\rm int}(R,\phi_0) + \int_{-L/2}^{L/2} dz \left[\left(\frac{a_1}{2}\cos(\phi_0) - 2a_2\cos(2\phi_0) \right)\varphi(z)^2 + w_{tor}[\varphi'(z)] \right]. \tag{2.6}$$

The term linear in φ vanishes. The path integral is, now, of Gaussian form and so φ may be integrated over (cf. Appendix B of Ref. [26]), giving us the free energy

$$F = Lw_{\rm int}(R,\phi_0) + \frac{LC}{4\lambda\lambda_p} - k_BT \ln \Theta, \tag{2.7}$$

where $\lambda = \sqrt{\dfrac{C}{2[a_1\cos(\phi_0) - 4a_2\cos(2\phi_0)]}}$.

The *thermal persistence length* is defined as $\lambda_p = C/(2k_BT)$. Θ is an R -independent constant which is poorly defined in the path integral formulation [27].

The validity criterion for the Gaussian approximation is $\lambda/\lambda_p \ll 1$. As we move towards R_*, it will break down, as $\lambda \to \infty$ when $R \to R_*$. Therefore, generally we need a better approximation.

C. Hartree approximation

The simplest self-consistent approximation is the Hartree approximation. This has proved successful in various many-body problems; a notable example is BCS theory [28]. We shall present a schematic discussion of how one obtains the results in this approximation (details given in Appendix B of Ref. [26]). In the context of our model, this approximation is by no means intuitive, and in order to derive it we will have to use diagrammatric expansion. This clarifies which contributions due to thermal excitations it takes into account. This also gives a basis for any further refinements, in case one wants to improve the accuracy or extend the model.

All diagrammatic expansions will involve the "bare" correlation function, defined as $G_0(z-z')=\langle\varphi(z)\varphi(z')\rangle_0$ where $\langle\cdots\rangle_0$ denotes thermal averaging over Gaussian fluctuations [where Eq. (2.6) for $E[\phi]$ is used in Eq. (2.4)]. The Feynman rules for such an expansion are shown in Fig. 2.

Our task will be first to calculate the full correlation function $G(z-z')=\langle\varphi(z)\varphi(z')\rangle$, for which the diagrammatic expression is shown in Fig. 3.

We are not able to sum up all contributions in this expansion, and so have to simplify it. Our choice will be the Hartree approximation for $G(z-z')$. The first step is to sum up the infinite series of graphs that contain only $G_0(0)$ in their loops (correlation functions coming back on themselves) to give us the result which will be called $G_1(z-z')$ [an approximation of $G(z-z')$ that contains only these graphs]; for illustration, we show three of such graphs on the first line of Fig. 3. Next, we replace $G_0(0)$ in the loops with $G_1(0)$ and repeat the process to obtain $G_2(0)$. We may see what kinds of graphs are contained in $G_2(0)$ by representing $G_1(0)$ and $G_2(0)$ by the first line of Fig. 3. However, each line in each loop in $G_2(0)$ is replaced with the diagrammatic representation for $G_1(0)$. By doing this we generate new graphs. In the Hartree approximation we do this kind of iteration an infinite number of times, and so calculate $G_\infty(0)$. Now, $G(0)=\langle\varphi(z)^2\rangle$ is determined self-consistently through an equation which may be obtained by setting $G(0)=G_\infty(0)=G_{\infty-1}(0)$. Effectively, we are summing over a class of diagrams which are usually referred to as "tadpole" graphs [29].

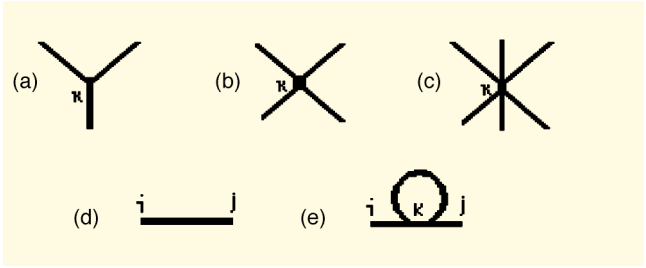

FIG. 2. The Feynman rules. Expanding $E[\phi]$ [cf. Eq. (2.5)], each φ^n term, for $n>2$, may be represented by a diagram (or vertex). For instance the vertices in (a), (b), and (c) represent the coefficients that sit in front of the φ^3, φ^4, and φ^6 terms, respectively, in $E[\phi]$. The straight line in (d) represents the "bare" correlation function $G_0(z_i-z_j)$ (see the text). Each Feynman graph is then constructed from vertices and straight lines. For each vertex positioned at a point k, there must be an integration over z_k. There is also a symmetry factor which accounts for how many ways a term (graph) in the expansion may be generated. As an example we show one Feynman graph in (e). It contains one integration over z_k, three Gaussian correlation functions $G_0(z_i-z_k)$, $G_0(z_k-z_k)$, and $G_0(z_k-z_j)$ (see the text), the coefficient that sits in front of the φ^3 term, and a symmetry factor. An analytical expression for (e) can be found in Appendix B of Ref. [26].

Introducing a new variable, $\langle\varphi(z)^2\rangle=\lambda_h/2\lambda_p$, this derivation yields an equation on λ_h,

$$\lambda_h=\sqrt{\frac{C}{2\left[a_1\cos(\phi_*)\exp\left(-\frac{\lambda_h}{4\lambda_p}\right)-4a_2\cos(2\phi_*)\exp\left(-\frac{\lambda_h}{\lambda_p}\right)\right]}}. \tag{2.8}$$

Calculating the free energy, we sum up all the graphs in Fig. 4, as well as taking care in replacing λ by λ_h. The last step corresponds to replacing $G_0(0)$ by $G(0)$. This leads to

$$F=L\left[a_0-a_1\cos(\phi_*)\exp\left(-\frac{\lambda_h}{4\lambda_p}\right)+a_2\cos(2\phi_*)\times\exp\left(-\frac{\lambda_h}{\lambda_p}\right)+\frac{C}{8\lambda_h\lambda_p}\right]-k_BT\ln\Theta. \tag{2.9}$$

Since we have built our expansion around an arbitrary value of ϕ_* we must minimize this result with respect to ϕ_* [30]. This should be performed separately above and below a new critical separation R_*',

$$\phi_*=0,\quad R>R_*',$$

$$\phi_*=\cos^{-1}\left[\frac{a_1}{4a_2}\exp\left(-\frac{3\lambda_h}{4\lambda_p}\right)\right],\quad R<R_*'. \tag{2.10}$$

There is a striking resemblance of these results to those of Ref. [24] for the average interaction energy between two nonhomologous DNA molecules, which uses a variational

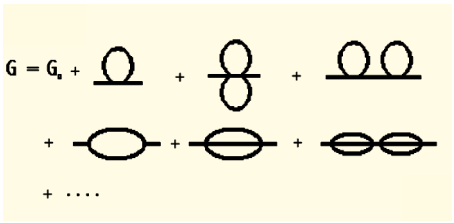

FIG. 3. Diagrammatic expansion for the full correlation function $G(z-z')$ in Hartree approximation.

approach, if λ_p is simply replaced by the helical coherence length λ_c [14]. This might be due just to the mere fact that both the patterns of twist angle mismatch between nonhomologous DNA and thermal fluctuations are random. In Sec. IV, where we consider the case of nonhomologous DNA, we shall exploit this similarity to arrive at an approximation which combines both effects.

Similar to the results of Ref. [24], the Hartree result gives a first order transition at R'_*: a discontinuity in ϕ_*, the "order parameter" in this context. It appears that $R'_* < R_*$, i.e., the "phase transition" in the Hartree approximation occurs at smaller separation than in the ground state.

In this section we shall defer from plotting the Hartree results for the free energy and ϕ_*. We leave this to the next section where we present the exact solution of the problem with which we compare the Hartree results.

III. "QM" FORMULATION FOR IDENTICAL DNA: EXACT RESULTS

A. The quantum mechanical analogy

Instead of calculating the path integral Eq. (2.3), we reformulate the problem into a quantum mechanical (QM) one [31] that describes a particle with a momentum p_ϕ in a periodic potential. This not only provides an exact numerical solution for free energy but also gives a framework for a physical interpretation of results.

In dimensionless units this potential has the form

$$V(\phi) = -\frac{\lambda_p^2}{\lambda_0^2}\cos(\phi) + \frac{a_2\lambda_p^2}{a_1\lambda_0^2}\cos(2\phi), \tag{3.1}$$

where $\lambda_0^2 = C/(2a_1)$.

Our first step is to rewrite the integral as

$$Z \propto \int\int \mathcal{D}p_\phi \mathcal{D}\phi \, \exp\, i\int_{-\tilde{t}}^{\tilde{t}} dt(p_\phi(t)\phi'(t) - H[p_\phi(t),\phi'(t)]), \tag{3.2}$$

where $H[p_\phi\phi(t)] = (p_\phi(t))^2/2 - V[\phi(t)]$ and $L/(2\lambda_p) = i\tilde{t}$. Here, we have performed a *Wick rotation* ($z \to it$) into a timelike variable t [29]. Following Feynman's path integral formulation of quantum mechanics [32] one deduces that the free energy, for the fixed end boundary conditions $\phi(-i\tilde{t}) = \phi_-$ and $\phi(i\tilde{t}) = \phi_+$, is (cf. Appendix C of Ref. [26])

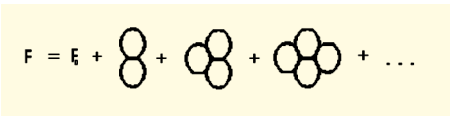

FIG. 4. Diagrammatic expansion for the free energy. Here, F_0 is the free energy calculated in the Gaussian approximation.

$$F = a_0 L - k_B T \ln\sum_E \langle\phi_+|\mathcal{E}\rangle\exp\left(-\frac{\mathcal{E}L}{\lambda_p}\right)\langle\mathcal{E}|\phi_-\rangle - k_B T \ln\Theta. \tag{3.3}$$

The energy eigenvalues are then determined by the Schrödinger equation

$$-\frac{1}{2}\frac{d^2\psi(\phi)}{d\phi^2} + V(\phi)\psi(\phi) = \mathcal{E}\psi(\phi). \tag{3.4}$$

If we take length of the helices to be very long, we may further simplify the free energy

$$F = a_0 L - \frac{\mathcal{E}_0 CL}{2\lambda_p^2} - k_B T \ln\Theta, \tag{3.5}$$

where $\mathcal{E}_0$ is the ground state energy. One may also show that for a system of infinite length the thermal average of an observable depending on $\phi(z)$ is the same as its QM average. In Appendix D of Ref. [26] we show that this principle for thermal averaging [as well as Eq. (3.5)] remains valid for molecules with a finite length of juxtaposition if $L/\lambda \gg 1$ for the Gaussian approximation, and $L/\lambda_h \gg 1$ for the Hartree approximation. So, in this limit, we should interpret the probability density $|\psi(\phi)|^2$ of the ground state as the "physical" probability distribution of ϕ at a given temperature.

Before solving Eq. (3.4) let us discuss what we expect to get. It is useful, first, to present a schematic drawing of both the potential and the probability density (Fig. 5). We can see from Eq. (3.1) that λ_p controls the height of the potential barrier; as we reduce λ_p the barrier diminishes. This increases the amount of tunneling between minima, thereby causing the probability density to become flatter. In the extreme limit of high temperatures the probability density is the same everywhere. In other words, at these temperatures, fluctuations in ϕ are so large that the angular dependence of $w_{\rm int}$ is effectively washed out.

As we increase a_2, when we enter the regime $R < R_*$, the potential splits into two minima, centered at $-\phi_0$ and ϕ_0. If we increase a_2 further, we find that $|\psi(\phi)|^2$ will develop two peaks as seen in Fig. 5. This occurs below a separation which we define as R''_*. Importantly, we expect the mean "quantum average" $\langle\phi\rangle$ (thermal average along the whole molecule) to be *zero* for all values of a_1 and a_2.

Now for $R > R''_*$, even when $\lambda/\lambda_p \ll 1$, there is an equal probability, as we go along the molecule, of ϕ being any integer of 2π. This might be due to a very dilute gas of thermally excited solitons along the juxtaposing duplexes [25]. Whereas, for $T=0$ we expect $\phi(z)=0$ at large R. Each soliton corresponds to a kink in the angle ϕ from $2\pi n$ to $2\pi(n+1)$ or in the opposite direction, from $2\pi n$ to $2\pi(n$

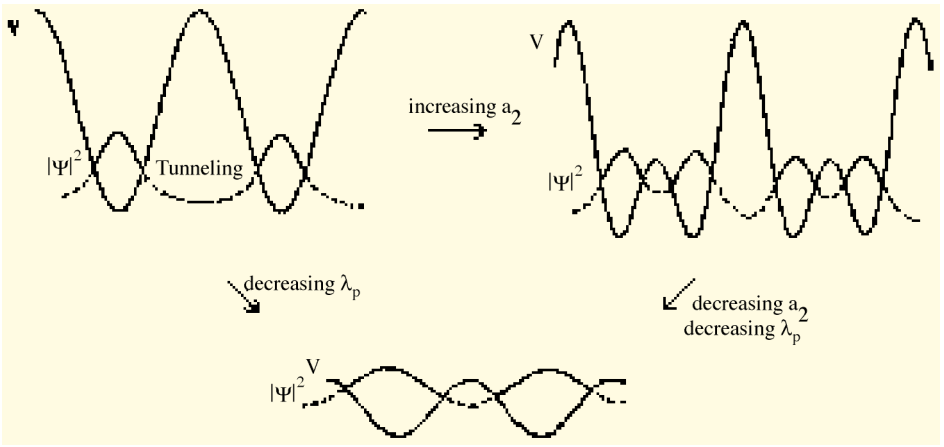

FIG. 5. Schematic drawing showing how $V(\phi)$ and $|\psi(\phi)|^2$ vary with respect to λ_p and a_2.

-1). At $T=0$, each soliton is a nontrivial solution of the Euler equation which is obtained by minimizing $E[\phi_0]$ with respect to ϕ_0; each having $\Delta E \sim 1/\lambda$ more energy than the trivial solution. The probability of finding at least one soliton of "length" λ can be estimated as

$$P_s \sim \exp\left(-\frac{\Delta E}{k_B T}\right). \tag{3.6}$$

At least for infinitely long molecules there are always some solitons present. However, there are too few solitons at $\lambda/\lambda_p \ll 1$ to give any substantial contribution to the free energy.

When $R < R_*''$, there is another type of soliton corresponding to a twist in the angle from $-\phi_0$ to ϕ_0 or in the opposite direction. For $\lambda/\lambda_p \ll 1$, we expect a small thermally excited density of this type of solitons, and this causes $\langle\phi\rangle$ to be *zero* for all separations. A rough estimate of the density of solitons should be the amount of QM tunnelling through potential barriers, Eq. (3.1).

B. Limiting cases

For our confidence, let us check the limiting cases. We first consider $\lambda/\lambda_p \ll 1$, where we must recover the Gaussian result. In the "QM" picture, this regime corresponds to a large magnitude of the potential. Here, we expect $\psi(\phi)$ to be localized; for $R > R_*$, around $\phi_0 = 0$, and for $R < R_*$ around $\pm\phi_0$. Therefore, we should Taylor-expand in $\delta\phi = \phi \pm \phi_0$ around these respective values. At small values of $\delta\phi$, one readily obtains the Schrödinger equation for a simple harmonic oscillator ($\hbar = 1$)

$$-\frac{1}{2}\frac{d^2\tilde{\psi}(\delta\phi)}{d(\delta\phi)^2} - \frac{\omega^2\delta\phi^2}{2}\tilde{\psi}(\delta\phi) = \tilde{\mathcal{E}}\tilde{\psi}(\delta\phi). \tag{3.7}$$

Here $\omega^2 = (\lambda_p^2/\lambda_0^2)\cos\phi_0 - (4a_1\lambda_p^2/a_2\lambda_0^2)\cos 2\phi_0$, and $\tilde{\mathcal{E}} = \mathcal{E} + (\lambda_p^2/\lambda_0^2)\cos(\phi_0) - (a_1\lambda_p^2/a_2\lambda_0^2)\cos(2\phi_0)$. The ground state energy is of course $\mathcal{E}_0 = \omega/2$ and the wave function takes the form

$$\tilde{\psi}(\delta\phi) = c_0 \exp\left(-\frac{\omega(\phi_0 + \delta\phi)^2}{2}\right), \tag{3.8}$$

where c_0 is the normalization factor. Then using the formula given in Eq. (3.5) it is easy to recover Eq. (2.7).

What is the place of the Hartree approximation in the "QM" reformulation? It is equivalent to a variational principle for the free energy where now Eq. (3.8) is its trial function and ω a variational parameter. Now $\delta\phi = \phi \pm \phi_*$, and ϕ_* is treated as a variational parameter as well. By computing $\langle\tilde{\mathcal{E}}|\hat{H}|\tilde{\mathcal{E}}\rangle$ and setting $\omega = \lambda_p/\lambda$ it is not hard to recover the Hartree result (2.9).

In one detail, however, the Hartree and Gaussian approximations are incorrect (even when $\lambda/\lambda_p \ll 1$); namely, in the value of $\langle\phi\rangle$. In the Gaussian approximation we find that $\langle\phi\rangle = \pm\phi_0$, and in the Hartree approximation $\langle\phi\rangle = \pm\phi_*$, whereas in fact it should be zero. This is due to both *Gaussian and Hartree approximations not containing solitons* [33]. To remedy this, we should treat the ground state wave function as a superposition of Gaussian wave functions, each centered about minima in the potential $V(\phi)$ when $\lambda/\lambda_p \ll 1$:

$$\psi(\phi) \simeq \lim_{N\to\infty} \frac{c_0}{\sqrt{N_p}} \sum_{n=-N_p/2}^{N_p/2} \exp(-\omega(\phi - 2\pi n)^2), \quad R > R_*, \tag{3.9}$$

$$\psi(\phi) \simeq \lim_{N\to\infty} \frac{c_0}{\sqrt{2N_p}} \sum_{n=-N_p/2}^{N_p/2} [\exp(-\omega(\phi - 2\pi n - \phi_0)^2) + \exp(-\omega(\phi - 2\pi n + \phi_0)^2)], \quad R < R_*.$$

If the degree of overlap between the Gaussian wave functions is small, i.e., the number of solitons is negligible which is true when $\lambda/\lambda_p \ll 1$ for $R > R_*$ and $\lambda/(\lambda_p\phi_0) \ll 1$ for $R < R_*$, then by using Eq. (3.5) we are able to show again that the Gaussian result (2.7), is essentially correct. By similar reasoning, we may show that the Hartree result (2.9), is recovered by applying the variational principle to QM formulation.

Now, for the wave functions given in Eq. (3.9) we find $\langle\phi\rangle=0$ for all R, in accordance with our previous discussion. Thereby, the analysis of the limiting cases is mathematically consistent with the QM picture.

C. Exact solution: Comparison with approximate results

To solve Eq. (3.4) exactly for the ground state, utilizing that the ground state wave function is periodic, we write

$$\psi(\phi)=\sum_n A_n \exp(-in\phi). \tag{3.10}$$

Equation (3.4) then reduces to an infinite system of algebraic equations given by

$$\left[\frac{n^2}{2}-\mathcal{E}_0\right]A_n-\frac{\lambda_p^2}{2\lambda_0^2}[A_{n-1}+A_{n+1}]+\frac{a_1\lambda_p^2}{2a_2\lambda_0^2}[A_{n-2}+A_{n+2}]=0. \tag{3.11}$$

We then solve Eq. (3.11) numerically for $\mathcal{E}_0$ and A_n subject to normalization of the wave function. The relevant quantities resulting from this calculation are shown in Figs. 6–8.

In Fig. 6 the free energy—the exact calculation and the Hartree approximation—are compared. The general trend is that it increases with temperature. We see that agreement between the two is quite good, above and below R_*'', for the temperatures plotted. There is, however, a small discrepancy near R_*'', where the Hartree approximation is less accurate.

In Fig. 7 we compare ϕ_* R_*' calculated in the Hartree approximation with $\phi_{\max}$, the angle of the maximum probability density in one period, and R_*'' the critical interaxial separation from the exact solution. Again, we find that agreement is good except near the "phase transition." Here, the exact calculation gives no discontinuity at R_*'', unlike the Hartree approximation.

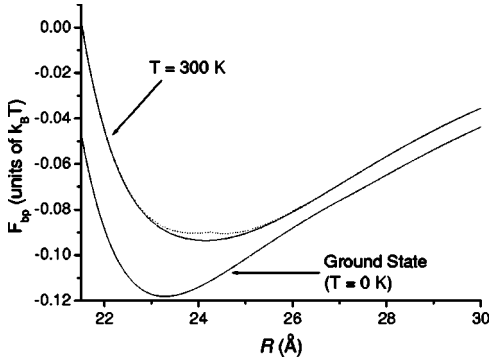

FIG. 6. Free energy per base pair for the Hartree (dotted line) and QM (solid line) formulations at $T=300$ K as a function of separation, R. The electrostatic coefficients, which vary with R, used for this plot are for helices with 90% counterion neutralization and a 30%/70% distribution of these ions between the minor and major grooves [14]. The torsional modulus is given the experimental value of $C=3.0\times10^{-19}$ erg cm. (These values will be used throughout the paper.) As we see, the Hartree value for the free energy deviates from the one calculated within the QM formulation only near the transition (around 24–25 Å for both temperatures).

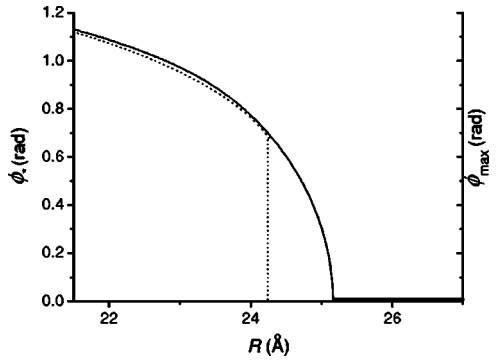

FIG. 7. Approximate and exact solutions for the azimuthal angle as a function of separation. ϕ_* (dotted line) and $\phi_{\max}$ (solid line) are shown for the Hartree and QM formulations, respectively. Whereas the Hartree formulation predicts a first order transition in the optimum phase angle, the most probable phase angle from the QM formulation continuously decreases to zero as R increases.

In Fig. 8 we compare $\langle\phi^2\rangle$, the measure of fluctuations as a function of separation in the Hartree approximation with the exact calculation. Here, again, there is good agreement away from R_*'. Near R_*', as expected, the Hartree approximation breaks down. However, the exact calculation has a cusp in $\langle\phi^2\rangle$ at R_*'. So there might be *some kind of phase transition* going on, albeit not a conventional one.

IV. EXTENSION OF THE HARTREE APPROXIMATION FOR NONHOMOLOGOUS PAIRS

Now let us consider two DNA molecules with different base pair sequences. Following Ref. [14] we write the torsional energy in a continuum approximation as

$$w_{tor}(\phi',\delta\Omega)=\frac{C}{4}\left(\frac{d\phi}{dz}-\frac{\delta\Omega(z)}{h}\right)^2. \tag{4.1}$$

$\delta\Omega(z)=\Omega_1(z)-\Omega_2(z)$ is a random field that corresponds to the preferred value of the derivative of ϕ when there is no interaction. Here, $\Omega_1(z)$ and $\Omega_2(z)$ correspond to the preferred values of the derivatives of ϕ_1 and ϕ_2, respectively.

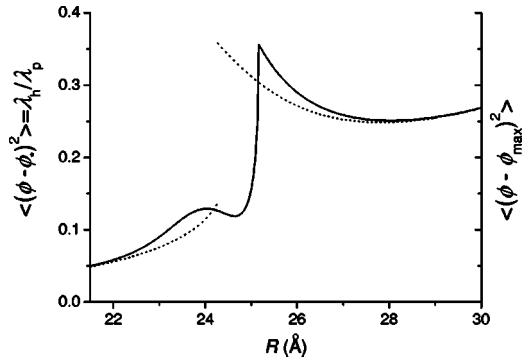

FIG. 8. Mean square amplitude of phase angle fluctuations for the Hartree case (dotted line) and for the QM formulation (solid line). The Hartree case shows a first order transition whereas there is no discontinuity for the QM formulation.

Both $\Omega_1(z)$ and $\Omega_2(z)$ are sequence dependent. Therefore, a particular nonzero value of $\delta\Omega(z)$ corresponds to different sequences along each double helix. For a detailed discussion and a derivation of Eq. (4.1), the reader is referred to Ref. [24]. As in Ref. [24], we assume that the statistical distribution of $\delta\Omega(z)$ is Gaussian with $\overline{\delta\Omega(z)\delta\Omega(z')}=2\Delta\Omega^2 h\delta(z-z')$ (throughout, we use a bar to denote ensemble averaging so as not to confuse with thermal averaging).

Our aim will be to calculate the ensemble-averaged free energy $\overline{F}$. The main problem is how to include a random field in Eq. (2.3). Indeed, ϕ_* will now depend on both z and $\delta\Omega(z)$. In the Gaussian approximation we may take $\phi_* = \phi_0[\Omega]$ where $\phi_0[\Omega]$ satisfies the saddle point equation [14]

$$\frac{d^2\phi_0}{dz^2} - \frac{1}{\lambda_0^2}\sin(\phi_0)\left(1 - \frac{4a_2}{a_1}\cos(\phi_0)\right) = \frac{1}{h}\frac{d\delta\Omega}{dz}, \quad (4.2)$$

the solutions of which are discussed in Refs. [24,25]. To calculate the free energy in both the Gaussian approximation and perturbation theory without further approximation is a difficult task. Essentially, one has to solve this equation and then calculate the correlation function for Gaussian approximations around this mean field, before even being able to perform perturbation theory.

At this point, it is worth mentioning why we have been unable to find a QM formulation for nonhomologous pairs. When one tries to reformulate the problem as a path integral over $\phi_0[\Omega]$ (as well as over φ), the integrand is not of closed form and is inherently nonlocal. This makes it impossible to write this path integral in the form of Eq. (3.2); therefore no "QM" approach can be formulated when $\Delta\Omega \neq 0$.

A way of calculating $\overline{F}$ is to introduce field replicas, usually employed in the field theoretical literature when handling a random field [34,35]. But, it is easier to employ the following approximation instead. The key step will be to replace w_{tor}, $\sin\phi_*$, $\sin 2\phi_*$, $\cos\phi_*$, and $\cos 2\phi_*$ by their ensemble averages, so that

$$\overline{F} \approx -k_B T \ln \int \mathcal{D}\phi \exp\left(-\frac{E_{approx}[\varphi]}{k_B T}\right), \quad (4.3)$$

where

$$E_{approx}[\varphi] = \int_{-L/2}^{L/2} dz(w_{approx}(R,\varphi) + \overline{w}_{tor}(\varphi' + \phi_*'[\delta\Omega], \delta\Omega))$$

and

$$w_{approx}(R,\varphi(z)) = \sum_{n=0}^{\infty} \frac{(-1)^n[a_1\,\overline{\sin(\phi_*)}\varphi^{2n+1} - a_2\,\overline{\sin(2\phi_*)}(2\varphi)^{2n+1}]}{(2n+1)!} - \sum_{n=0}^{\infty} \frac{(-1)^n[a_1\,\overline{\cos(\phi_*)}\varphi^{2n} - a_2\,\overline{\cos(2\phi_*)}(2\varphi)^{2n}]}{(2n)!}.$$

One can see that $w_{approx}(R,\varphi(z))$ is simply the form given for $w_{\rm int}$ in Eq. (2.5), but with all the quantities that depend on ϕ_* replaced by their ensemble averages; varying them so as to minimize $\overline{F}$. Now, things are a lot easier if the helices are assumed to be very long, as these ensemble averages will not depend on z. We can now use the Hartree approximation to obtain

$$F = L\left(a_0 - a_1\,\overline{\cos(\phi_*)}\exp\left(-\frac{\lambda_h}{4\lambda_p}\right) + a_2\,\overline{\cos(2\phi_*)}\exp\left(-\frac{\lambda_h}{\lambda_p}\right) + \frac{C}{4}\left[\frac{1}{2\lambda_h\lambda_p} + \overline{\left(\frac{d\phi_*}{dz} - \frac{\delta\Omega}{h}\right)^2}\right]\right) - k_B T \ln\Theta. \quad (4.4)$$

All that is left is to compute the averages, using the same type of trial function as given in Ref. [24]

$$\phi_*(z) = \overline{\phi}_* - \frac{\lambda_h}{2h}\int_{-\infty}^{\infty} dz' \frac{d\delta\Omega}{dz'}\exp\left(-\frac{|x-x'|}{\lambda_h}\right). \quad (4.5)$$

An important difference is that the trial function is for ϕ_*, not for ϕ_0, and λ_h will be chosen to minimize $\overline{F}$ at finite temperature, instead of $\overline{E}$ at zero temperature. Then using the results of Ref. [24] we are able to recast $\overline{F}$ in the form

$$\overline{F} = L\left[a_0 - a_1\cos(\overline{\phi}_*)\exp\left(-\frac{\lambda_h}{4\lambda_s}\right) + a_2\cos(2\overline{\phi}_*) \times \exp\left(-\frac{\lambda_h}{\lambda_s}\right) + \frac{C}{8\lambda_h\lambda_s}\right] - k_B T \ln\Theta, \quad (4.6)$$

where we have introduced an effective coherence length λ_s

$$\frac{1}{\lambda_s} = \frac{1}{\lambda_c} + \frac{1}{\lambda_p}, \quad (4.7)$$

with $\lambda_c = h/\Delta\Omega^2$, the helical coherence length in the ground state [14] discussed in the Introduction. This should be valid when $\lambda_h/\lambda_s < 1$. Then minimizing Eq. (4.6) with respect to λ_h and $\overline{\phi}_*$ we recover Eqs. (2.8) and (2.10), but with λ_p replaced by λ_s.

V. SPIN-SPRING MODEL: MONTE CARLO SIMULATIONS

In order to test the results of the Hartree approximation for both homologous and nohomologous pairs we have developed a discrete computational model. Performing Monte Carlo simulations for this model we have been able to determine over what range of parameters we may use Eqs. (4.6) and (4.7). These simulations also provide an additional test of the QM results.

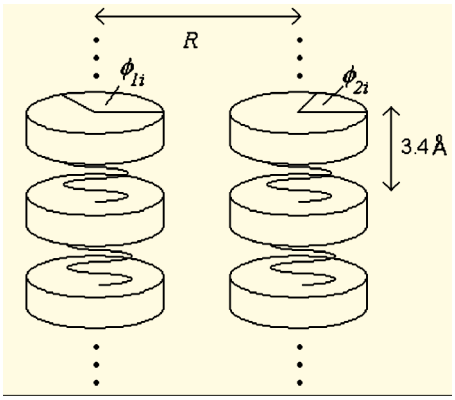

FIG. 9. "Discrete spin model" of two interacting DNA duplexes. Each spin represents an individual base pair which is coupled to its neighboring base pairs along the same duplex by a spring with a torsional elasticity modulus C. Each spin also interacts electrostatically with the spin adjacent to it in the other duplex.

A. Discrete model

In previous work [25], numerical simulations were carried out to find the ground state energies of interacting DNA duplexes and to discover how nonlinear effects observed in their torsional readjustment influence the interaction. To incorporate temperature, we have developed a discrete spin model of interacting base pairs. This discrete model yields results for the ground states identical to those obtained in the continuum model used in Ref. [25].

In the discrete model each duplex is treated as a series of disks where the spin variable corresponds to the azimuthal orientation of a disk at site i. Each disk or spin represents an individual base pair in the duplex and is coupled to its neighboring disks in the same DNA duplex by torsional springs with the experimentally determined force constant C. For perfectly rigid helices, the phase angle difference between neighboring disks, $\phi_{1i}-\phi_{1(i-1)}$, is determined by the preferential twist angle, Ω_{1i}, between the base pairs. For B-DNA this preferential twist angle has an average value of 34° with a sequence-dependent variation $\Delta\Omega=4°-6°$ [19–21]. Any deviation of the phase angle difference from the corresponding preferential twist angle difference yields an increase in the torsional energy of the duplex. Each disk also interacts electrostatically with the disk adjacent to it in the second DNA duplex. This electrostatic interaction between them depends on their relative azimuthal angle, $\phi_i=\phi_{1i}-\phi_{2i}$, and the interaxial separation, R. This configuration is sketched in Fig. 9.

The total energy is then simply a discretized form of Eqs. (2.1) and (4.1)

$$E = h\left[\sum_{i=1}^{N}\left[a_0 - a_1\cos(\phi_i) + a_2\cos(2\phi_i)\right] + \frac{C}{4h^2}\sum_{i=2}^{N}\left[(\phi_i-\phi_{i-1}) - \delta\Omega_i\right]^2\right], \tag{5.1}$$

where h is the helical rise per base pair ($\approx$3.4 Å), N is the total number of base pairs (disks), and $\delta\Omega_i$ is the difference in the preferential twist angles at site i between duplexes 1 and 2.

B. Finite temperature Monte Carlo simulations

The ground state energy can then be determined by minimizing Eq. (5.1). To do this, we employed a simple finite temperature Monte Carlo routine in which a spin is randomly chosen and perturbed by a small angle. The perturbation is dynamically altered during the routine so that half of the Monte Carlo moves are accepted.

The inclusion of temperature into the model is thus a simple task. However, we were required to modify our original Metropolis [36] algorithm in order to build up a canonical distribution of the spin configurations. For a given Monte Carlo move, the change in the electrostatic energy tends to be dominated by the change in the torsional energy, thus biasing the sampling. In our modified algorithm, we would first address only the change in the electrostatic energy for a given move to determine if it was accepted. Afterwards, the combined change in the torsional and electrostatic energy, δE, for the move was compared against the Boltzmann factor $\exp(-\delta E/k_BT)$ in the standard Metropolis fashion. This sampling yielded the correct canonical distribution of the configurations at a given temperature so that the thermal average of the energy is simply the arithmetic average of the energy of each sampled configuration. In practice, we allow the system to thermally equilibrate (at least 10^4 MC sweeps) and then begin averaging over subsequent Monte Carlo steps.

We used this model to find the thermal average of the energy of interacting homologous DNA, where $\delta\Omega=0$, and of interacting nonhomologous pairs, in which we averaged over a number of randomly generated sets of $\delta\Omega$. Many simulations using different numbers of base pairs were carried out to take into consideration end effects, i.e., we assumed the ends of the DNA to be free to twist rather than being fixed. We found that as long as fifty base pairs or more were used, the energy per base pair was independent of the length of the duplexes. For both homologous and nonhomologous pairs, we also averaged over many different initial spin configurations as to remove any bias from our sampling.

To compare the calculated energy from the previous sections with that found from the numerical simulations, we subtract the energy determined at each interaxial separation by the energy found at a large separation to eliminate the ill defined Θ term. As shown in Fig. 10, the results of these simulations are in excellent agreement with the free energy found from the Hartree calculations. Along with the criteria outlined in Appendix D of Ref. [26], this implies that the Hartree and QM formulations are good descriptions of the interaction.

VI. CONCLUSION

In all the problems of biological importance, DNA operates at physiological temperatures and hence it is important to understand how thermal fluctuations will affect the DNA-DNA interaction. This article provides this understanding.

We have extended the theory of electrostatic interactions between DNA from the ground state to account for thermal

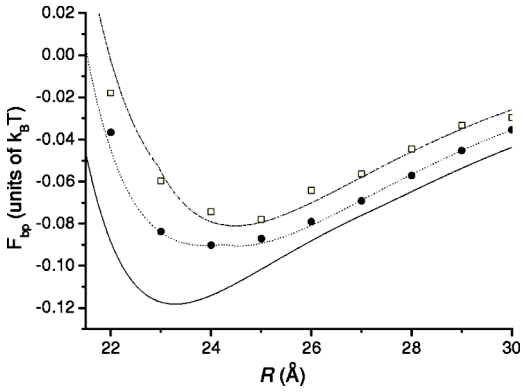

FIG. 10. The free energy per base pair for the ground state of homologous sequences (solid line) and for the Hartree calculation at $T=300$ K for homologous (dotted line) and nonhomologous sequences (dot-dashed line). Numerical results of the discrete spin model for homologous sequences at $T=300$ K are shown as filled circles. For nonhomologous sequences (open squares), the data points were found by averaging over many different realizations of the random field $\delta\Omega$. The error bars (not shown for clarity's sake) of the numerical simulations are found from averaging over many initial configurations for the homologous case and averaging over different realizations of the random field for nonhomologous cases.

fluctuations. We have developed three levels of description: path integral formulations on the level of Gaussian and Hartree approximations, a quasi "QM" formulation giving us an exact solution for the interaction of homologous DNA [37], and finite temperature Monte Carlo simulations of a spin-spring model which permitted verification of the exact and approximate solutions.

We have shown that a heuristic, ground state variational solution of Ref. 24 is reproduced by the Hartree self-consistent field approximation, where temperature plays the role of nonhomology. Subsequently, we have extended our approximations to embrace both the nonhomology and thermal effects in the interaction.

The Hartree approximation predicts a first order transition close to the ground state frustration point instead of a weaker transition given by the exact solution. This discrepancy is less noticeable for nonhomologous DNA. The Hartree approximation thus seems to be a reasonable tool for the description of torsionally flexible DNA.

There are also two "practical" lessons from this study. First, as shown in Fig. 6, thermally induced torsional fluctuations increase the energy of interaction of DNA duplexes relative to the ground state by roughly $0.03k_BT$ per base pair ($\sim 4k_BT$ per persistence length), but they do not remove the attraction minimum. Second, recognition between homologous and nonhomologous duplexes via the difference in electrostatic attraction [14] remains possible. Figure 10 gives typical energy differences from $1.6k_BT$ to $6k_BT$ per persistence length, which could cause recognition.

It should be recalled that in this work the counter ions have been assumed to be irreversibly adsorbed and immobile on the DNA surface. There were a number of theoretical works exploring the opposite limiting cases when all the counterions are floating around DNA or rearranging on the surfaces subject to DNA-DNA interactions [38,39]. Thus our work is limited to strong chemisorption of counterions, not untypical for many DNA condensers.

The parameters of the effective Hamiltonian (2.3) are temperature dependent. Indeed, the electrostatic coefficients describing the pair interaction, a_0, a_1, and a_2 depend on temperature through the dielectric constant of the solvent and the Debye screening length, whereas the torsional elasticity modulus decreases dramatically approaching the DNA melting point. The degrees of freedom that control temperature dependence of these parameters are much faster than DNA torsional fluctuations, so that we are not obliged to consider both on the same footing. But in the calculation of thermodynamic properties of DNA assemblies this additional and equally important source of temperature dependence must be parametrically taken into account. And with these findings concerning the pair interaction it will be possible to start investigating the statistical thermodynamics of columnar assemblies. This will be a rich object for theoretical studies due to the unique properties of the pair potential. What is also interesting is that by the above discussed spin-analogies, the problem could be mapped onto unexplored models of exotic magnetic systems, an area of possible theoretical interest.

ACKNOWLEDGMENTS

Numerous discussions with Dr. S. Leikin are gratefully acknowledged. This work was made possible via the Royal Society Wolfson Research Merit Grant to A.A.K. and the EPSRC Grant No. GR/S31068/01.

APPENDIX: THE ELECTROSTATIC COEFFICIENTS

The coefficients derived in Ref. [3] have the form

$$a_0 = \frac{8\pi^2\sigma^2}{\varepsilon}\left\{\frac{(1-\theta^2)K_0(\kappa R)}{\kappa^2[K_1(\kappa r)]^2} - \sum_{n,j=-\infty}^{\infty}\frac{[f(n,\theta)]^2}{\kappa_n^2}\left[\frac{[K_{n-j}(\kappa_n R)]^2 I_j'(\kappa_n r)}{[K_n'(\kappa_n r)]^2 K_j'(\kappa_n r)}\right]\right\}, \quad \text{(A1)}$$

$$a_{n=1,2} = \frac{16\pi^2\sigma^2}{\varepsilon}\frac{[f(n,\theta)]^2}{\kappa_n^2}\frac{K_0(\kappa_n R)}{[K_n'(\kappa_n r)]^2}, \quad \text{(A2)}$$

where

$$f(n,\theta) = e^{-\pi^2 n^2 w^2/2H^2}[f_1\theta + f_2(-1)^n\theta - (1-f_3\theta)\cos(n\tilde{\phi}_s)] \quad \text{(A3)}$$

and

$$\kappa_n = \sqrt{\kappa^2 + n^2\left[\frac{2\pi}{H}\right]^2}. \quad \text{(A4)}$$

$I_n(x)$, $K_n(x)$, $I_n'(x)$, and $K_n'(x)$ are the modified Bessel functions and their derivatives, respectively. R, again, is the interaxial spacing between two parallel helices and r is the radius of the cylindrical surface formed by the centers of the phosphates. σ ($\approx$16.8 μC/cm^2) is the surface charge density of the phosphates. θ is the fraction of phosphate charge neutralized by bound counterions. f_i are the fractions of counterions bound in the minor groove (f_1), the major groove (f_2), and on the phosphate strands (f_3) so that $f_1+f_2+f_3=1$. $\tilde{\phi}_s$ ($\approx 0.4\pi$ for B-DNA) is the azimuthal half-width of the minor groove. w ($\approx$5 Å) is the width of the charge density distribution across each phosphate strand or groove. ε ($\approx$80) is the dielectric constant of water. And, finally, κ^{-1} ($\approx$7 Å in physiological solution) is the Debye screening length.

[1] W. M. Gelbart, R. F. Bruinsma, P. A. Pincus, and V. A. Parsegian, Phys. Today **53**, 38 (2000).

[2] D. M. Frank-Kamentskii, V. V. Anshelevich, and A. V. Lukashin, Sov. Phys. Usp. **30**, 317 (1987).

[3] A. A. Kornyshev and S. Leikin, J. Chem. Phys. **107**, 3656 (1997); **108**, 7035(E) (1998).

[4] A. A. Kornyshev and S. Leikin, Phys. Rev. Lett. **82**, 4138 (1999).

[5] A. A. Kornyshev and S. Leikin, Biophys. J. **75**, 2513 (1998).

[6] A. A. Kornyshev and S. Leikin, Proc. Natl. Acad. Sci. U.S.A. **95**, 13579 (1998).

[7] A. A. Kornyshev and S. Leikin, Phys. Rev. E **62**, 2576 (2000).

[8] A. A. Kornyshev and S. Leikin, Phys. Rev. Lett. **84**, 2537 (2000).

[9] A. A. Kornyshev, S. Leikin, and S. Malinin, Eur. Phys. J. E **7**, 83 (2002).

[10] H. H. Strey, J. Wang, R. Podgornik, A. Rupprecht, L. Yu, V. A. Parsegian, and E. B. Sirota, Phys. Rev. Lett. **84**, 3105 (2000).

[11] V. Lorman, R. Podgornik, and B. Žekš, Phys. Rev. Lett. **87**, 218101 (2001).

[12] H. M. Harreis, A. A. Kornyshev, C. N. Likos, H. Löwen, and G. Sutmann, Phys. Rev. Lett. **89**, 018303 (2002).

[13] H. M. Harreis, C. N. Likos, and H. Löwen, Biophys. J. **84**, 3607 (2003).

[14] A. A. Kornyshev and S. Leikin, Phys. Rev. Lett. **86**, 3666 (2001); DNA need not unzip, Phys. Rev. Focus (http://focus.aps.org/v7/st19.html).

[15] B. M. Weiner and N. Kleckner, Cell **77**, 977 (1994).

[16] S. M. Burges, N. Kleckner, and M. Weiner, Genes Dev. **13**, 1627 (1999).

[17] R. R. Sinden, *DNA Structure and Function* (Academic, San Diego, 1994).

[18] S. Neidle, *DNA Structure and Recognition* (IRL Press at Oxford University Press, Oxford, 1994).

[19] A. A. Gorin, V. B. Zhurkin, and W. K. Olson, J. Mol. Biol. **247**, 34 (1995).

[20] W. Kabsch, C. Sander, and E. N. Trifonov, Nucleic Acids Res. **10**, 1097 (1982).

[21] W. K. Olson, A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin, Proc. Natl. Acad. Sci. U.S.A. **95**, 11163 (1998).

[22] H. Strogaz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering* (Addison-Wesley, Reading, MA, 1994).

[23] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 1995).

[24] A. Cherstvy, A. A. Kornyshev, and S. Leikin, J. Phys. Chem. B **108**, 6508 (2004).

[25] A. A. Kornyshev and A. Wynveen, Phys. Rev. E **69**, 041905 (2004).

[26] See EPAPS Document No. E-PLEEE8-70-093410 for appendices which contain more detailed derivations of the formulas given in the text. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.

[27] A possible physical interpretation is that we may consider ln Θ as the entropy of the two double helices at $R=\infty$. One should point out that the thermodynamic quantities that we might be interested in calculating (for instance the average interaction energy or a contribution to the specific heat) will not depend on Θ in the continuum model. Such a term may be, indeed, eliminated by subtracting off the free energy at $R=\infty$, where a_0, a_1, and a_2 are all zero.

[28] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[29] For example, P. Ramond, *Field Theory: A Modern Primer* (Addison–Wesley, Reading, MA, 1990); D. J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena* (World Scientific, Singapore, 1984).

[30] By minimizing the free energy with respect to λ_h we simply get back Eq. (2.8).

[31] This approach has been used in the past to solve other problems involving path integrals for 1D systems D. J. Scalapino, M. Sears, and R. A. Ferrell, Phys. Rev. B **6**, 3409 (1972); G. F. Mazenko and P. S. Sahni, *ibid.* **18**, 6139 (1978); I. M. Gel'fand and A. M. Yaglom, J. Math. Phys. **1**, 48 (1960).

[32] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1981).

[33] Or to put it another way, the ground state wave function should be periodic in ϕ, since our starting potential was periodic.

[34] I. V. Yurkevich and I. V. Lerner, Phys. Rev. B **63**, 064522 (2001), and references contained therein.

[35] I. F. Herbut, Phys. Rev. B **61**, 14723 (2000), and references contained therein.

[36] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[37] In addition, the "QM" analysis has clarified the generic fea-

tures of excitation of soliton modes and their contribution to the interaction.

[38] A theory of Shklovskii *et al.* [B. I. Shklovskii, Phys. Rev. E **60**, 5802 (1999); V. I. Perel and B. I. Shklovskii, Physica A **274**, 446 (1999); T. T. Nguyen, I. Rouzina, and B. I. Shklovskii, J. Chem. Phys. **112**, 2562 (2000)] has claimed that triply charged counterions of the buffer electrolyte are capable of building nonlinear patterns: "Wigner crystals." Such patterns can provide charge separation motifs around negatively charged macromolecules, even if the latter were homogeneously charged rods. When formed, these nonlinear structures can cause attraction between the parallel rods, as positively charged sections around one rod can set in front of bare, negatively charged sections of another rod. These findings had a resonance in relation with a more general phenomenon, the so-called "anomalous" attraction between similarly charged colloid particles [K. S. Schmitz, *Macroions in Solution and Colloidal Suspension* (VCH, New York, 1993); *Ordering and Phase Transitions in Charged Colloids*, edited by A. K. Arora and B. V. R. Tata (VCH, New York, 1996)]. The stratified counterion clouds surrounding the charged rods *are not pinned* and are free to slide along the rod axis. The assemblies built of such molecules will be therefore lacking *any azimuthal correlations*. However, it is well known that such correlations exist, otherwise, it would not have been possible to observe the helical structure of DNA in columnar phases, as it has been done in classical experiments [D. Rhodes and A. Klug, Nature (London) **286**, 573 (1980); S. B. Zimmermann and B. H. Pfeiffer, Proc. Natl. Acad. Sci. U.S.A. **76**, 2703 (1979)]. In reality, the helical structure of the charge distribution of phosphates of the DNAs may cause pinning, and the two effects—Wigner crystal formation and DNA helical structure—may produce DNA-structure-sensitive nonlinear patterns. The study of such phenomena is a challenging task for theory and simulations. However, we believe that (i) the helical structure of DNA is always there, and this alone will cause charge separation motif for the chemisorbed counterions; (ii) the attraction exists not only for triply charged ions but for singly and doubly charged ions (Mn^{2+}, Cd^{2+} are good DNA condensers), here Wigner crystals are unlikely; (iii) the polyanions such as spermine, spermidine, and protamine cannot be considered as point charges but rather a system of connected single charges envisaged as adsorbing in a wormlike fashion into the DNA grooves. Thus experimental data can be explained by a simpler model, not involving Wigner crystal formation.

[39] The counterion distributions might not necessarily be rigid. A model of the temperature-assisted rearrangements of counterions between the minor and major grooves has been suggested [A. Cherstvy, A. A. Kornyshev, and S. Leikin, J. Phys. Chem. B **106**, 13362 (2002)] offering an explanation of the temperature induced DNA condensation observed in solution of manganese salts [D. C. Rau and V. A. Parsegian, Biophys. J. **61**, 260 (1992); S. Leikin, D. C. Rau, and V. A. Parsegian, Phys. Rev. A **44**, 5272 (1991)].