Unique magnetic signatures of mismatched base pairs in DNA

Vadim Apalkov,¹ Julia Berashevich,² and Tapash Chakraborty²,a)

¹Department of Physics and Astronomy, Georgia State University, Atlanta, Georgia 30303, USA
²Department of Physics and Astronomy, University of Manitoba, Winnipeg R3T 2N2, Canada

(Received 14 January 2010; accepted 1 February 2010; published online 23 February 2010)

Magnetic properties of DNA containing mispairs, such as different conformations of the G·A mispair, or a G·T mispair inserted into the DNA chain, have been theoretically investigated. The essential ingredients for these studies, the charge transfer integrals, were evaluated from the DNA sequences containing the mispair and optimized in the solvent. We find that the magnetic susceptibilities of the host DNA chain containing a large number of Watson–Crick base pairs are significantly altered in the presence of the mispairs, and the effects depend on the choice of mispairs. In particular, insertion of even a single G·A mispair changes the nature of magnetization (sign of the susceptibility) of the host DNA. We propose that measurement of the magnetic properties of DNA might provide a direct route to detection and identification of those mispairs.


I. INTRODUCTION

Mismatched base pairs in DNA, which are directly linked to mutations, often appear as a result of misincorporation of bases during DNA replication, or from damage due to mutagenic chemicals and ionizing radiations. Failure of the mismatch repair system can lead to genomic instability, while damage in genes responsible for maintaining DNA stability can lead to the risk of developing cancer and many other debilitating diseases. Clearly, mismatch repair is of fundamental importance to reduce the accumulation of mutations and maintain the genome integrity. Understanding the complexity of recognition and repair of erroneous base pairing has been a major area of research in biological and medical communities for many decades. Although major advances have been made in understanding the properties of various mispairs, many crucial issues, such as how the DNA repair proteins locate the damaged sites in the vast expanse of defect-free DNA, or what are the observable physical characteristics of individual mispairs in DNA, are still poorly understood. The existing models designed to answer the first point are largely based on conjectures where the mismatch repair proteins supposedly exploit the differences in thermodynamic properties of the damaged and canonical bases that are often structurally very similar. The research on the second point has been primarily geared to develop sensors to detect a single base mismatch pair. The mispairs G·T and G·A are among the most stable mispairs. Studies have also indicated that the G·A mispair exists in several different forms depending on the base stacking environment, salt concentration and pH. The stability of a DNA duplex containing this mispair depends on the particular conformation adopted by the mispair.

Here we address some of the intrinsic physical properties of DNA mispairs. Specifically, our aim is to investigate if there are unique magnetic signatures of certain mispairs that could aid in their recognition. In particular, we propose that the magnetic susceptibility, a measurable physical quantity, might perhaps be useful in locating a mispair in DNA. Magnetic properties of DNA are generally quite intriguing, especially at low temperatures. Recent experiments on DNA at low temperatures have reported that the genomic length DNA in the B-form (wet DNA) at low temperatures (T=20 K) is in fact, paramagnetic while with increasing temperature or decreasing humidity it turns diamagnetic. The DNA molecule in the A form (dry DNA) remains diamagnetic at all temperatures. Subsequent theoretical studies have clearly indicated that, in determining the type of magnetization in DNA, intrastrand hopping of electrons between the nearest base pairs and also the interstrand hopping, play an important role. No such studies have been reported, as yet for the mispairs.

II. A DNA CHAIN IN A MAGNETIC FIELD

As in our earlier studies of mispair-free DNA, we begin with a double-stranded tight-binding model. In this work, we focus our attention primarily on homogeneous DNA sequences—a poly(G)–poly(C) chain or a poly(A)–poly(T) chain. In poly-DNA, the two strands correspond to highest occupied molecular orbitals (HOMOs) located on the purines and lowest unoccupied molecular orbitals (LUMOs), located on the pyrimidines of the base pairs and our scheme takes into account both the inter- and intras strand electron hopping. The interstrand hopping introduces coupling between the HOMO and LUMO states. This coupling is the most important term in our present study since it is responsible for the existence of closed loops containing magnetic fluxes. The properties of electrons in those loops are then influenced by an external magnetic field. The intras strand hopping integrals correspond to the overlap between the HOMOs of the nearest base pairs (i1), and similarly for the LUMOs (iL), within DNA (Fig. 1). The value of the hopping integral depends on the type of the base pair chosen. In what

---

a)Electronic mail: tapash@physics.umanitoba.ca.
also that of unchanged. Therefore, we can calculate the charge transfer noncanonical pairs and this mispair leaves the DNA helix almost unchanged. Therefore, for our investigations we chose the G·A mispair and also the G·T mispair, whose planar size is expected, the original geometry of the DNA molecule is not distorted.

III. BUILDING THE DNA SEQUENCES

Incorporation of mispairs into a DNA sequence can cause structural changes to its backbone or even global structural instability of the entire helix. Among all the mispairs the G·A mispair is known to cause the smallest structural changes. As an example, the planar size of the G(anti)·A(syn) conformation is similar to those for the canonical pairs and this mispair leaves the DNA helix almost unchanged. Therefore, we can calculate the charge transfer parameters for the DNA sequences containing a G·A mispair without taking into account the sugar-phosphate backbone whose contribution is known to be minimal for DNA containing only canonical base pairs. We also found that the sheared conformation adopted by this mispair is inappropriate in this regard as it causes significant shift in the sugar complex toward the intrastrand charge transfer pathway. Therefore, for our investigations we chose the G(anti)·A(anti) and G(anti)·A(syn) conformations of the G·A mispair and also the G·T mispair, whose planar size is identical to that of the canonical pairs.

The geometries of the DNA sequences in the B-form containing three base pairs were generated using the CHARMM program applying the adopted basis Newton–Raphson minimization procedure. The nonbonded interactions were calculated including both the electrostatic and the Lennard-Jones potentials. The nonbonded list of atom pairs were constructed using a cut-off radius of 14 Å, while the contribution from the atoms within a cut-off distance of 10 Å were used for the energy calculations. To describe the solvent, a uniform dielectric constant was used. The DNA sequences comprising of canonical base pairs were built using the structural parameters of 1bna. The G·A and G·T mispairs were inserted in the middle of the poly(G)–poly(C) sequence. The geometries of the G(anti)·A(anti) and G(anti)·A(syn) conformations are presented in Fig. 2(a).

We have applied a multistep optimization process with the empirical potential method developed in CHARMM to evaluate the energetically optimal positions for the incorporated mispairs. First, the structure of the sugar-phosphate backbone was restored in the optimization process where the geometry and positions of all base pairs were fixed. The next step was to optimize the position of the middle base pair in relation to its neighbors, which was done by applying the constraints to the location of the terminated base pairs. Finally, for all DNA sequences, the last optimization step was performed in the water box. As a result, we achieved the energetically optimal location of the middle pair, that is, the canonical pair or a mispair, in relation to its neighbors in the solvent environment. For calculation of the charge transfer parameters that follows, both the water content and the sugar-phosphate backbone were removed. We have presented in Fig. 2(b) a comparison of the original (G–C)₃ sequence and its change due to the incorporation of a G·A mispair. As expected, the original geometry of the DNA molecule is not significantly modified due to the incorporation of a G·A mispair while the G·T mispair was found to cause even less distortion.

The charge transfer parameters were computed with the ADF program within the TZ2p approximation of the density functional theory. The HOMO and LUMO energies (ε₉ and ε₈) were calculated for the isolated nucleobases optimized in vacuum. The results are presented in Table I.

The intrastand charge transfer integrals (t₉₈) were calculated for each base pair within the obtained DNA sequences. The interstrand charge transfer integrals (t₉₈) were computed between two nearest-neighbor base pairs. In this case, we have assumed that all the HOMO states are fully occupied while all the LUMO states are empty. Interestingly, for a large HOMO-LUMO gap, a perturbation theory calculation indicates that the magnetic susceptibility is proportional to $\chi \propto t_{HL}H_{tHL}^{-1}$, where $t_{HL}$ is the hopping integral for the HOMO (LUMO) orbitals and $H_{tHL}$ is that between the two orbitals. Therefore, magnetization in DNA (paramagnetic or diamagnetic) depends on the relative sign of $t_{HL}$ and $H_{tHL}$ and also that of $H_{tHL}$ in the presence of a mispair.

<table>
<thead>
<tr>
<th>TABLE I. The HOMO and LUMO energies (ε₉ and ε₈) of the single nucleobases (in electron volt).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>HOMO</td>
</tr>
<tr>
<td>LUMO</td>
</tr>
</tbody>
</table>
TABLE II. The values of the intra- and interstrand charge transfer integrals (in eV) for different DNA sequences, where the middle pair (X₁-X₂) has been replaced by G-C, A-T, or G(anti)-A(anti), G(anti)-A(syn), G(anti)-T mispairs.

<table>
<thead>
<tr>
<th>(X₁,X₂)</th>
<th>tₛ(1)</th>
<th>tₛ(2)</th>
<th>tᵢ(1)</th>
<th>tᵢ(2)</th>
<th>tᵢL(1)</th>
<th>tᵢL(2)</th>
<th>tᵢH(1)</th>
<th>tᵢH(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-C</td>
<td>−0.133</td>
<td>−0.133</td>
<td>−0.041</td>
<td>−0.041</td>
<td>−0.050</td>
<td>−0.050</td>
<td>−0.050</td>
<td>−0.050</td>
</tr>
<tr>
<td>A-T</td>
<td>−0.218</td>
<td>−0.102</td>
<td>0.066</td>
<td>0.067</td>
<td>−0.001</td>
<td>−0.063</td>
<td>−0.001</td>
<td></td>
</tr>
<tr>
<td>G-A(anti)</td>
<td>−0.213</td>
<td>0.146</td>
<td>−0.072</td>
<td>−0.071</td>
<td>0.024</td>
<td>0.020</td>
<td>−0.025</td>
<td></td>
</tr>
<tr>
<td>G-A(syn)</td>
<td>0.254</td>
<td>−0.136</td>
<td>−0.153</td>
<td>−0.073</td>
<td>0.024</td>
<td>0.091</td>
<td>−0.025</td>
<td></td>
</tr>
<tr>
<td>G-T</td>
<td>−0.197</td>
<td>−0.087</td>
<td>−0.003</td>
<td>−0.005</td>
<td>−0.025</td>
<td>0.074</td>
<td>−0.026</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(X₁,X₂)</th>
<th>tₛ(1)</th>
<th>tₛ(2)</th>
<th>tᵢ(1)</th>
<th>tᵢ(2)</th>
<th>tᵢL(1)</th>
<th>tᵢL(2)</th>
<th>tᵢH(1)</th>
<th>tᵢH(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-T</td>
<td>0.011</td>
<td>0.011</td>
<td>−0.075</td>
<td>−0.075</td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
<td>0.054</td>
</tr>
<tr>
<td>A(anti) - G</td>
<td>−0.002</td>
<td>−0.055</td>
<td>−0.079</td>
<td>−0.117</td>
<td>0.054</td>
<td>0.131</td>
<td>0.054</td>
<td>0.054</td>
</tr>
<tr>
<td>A(syn) - G</td>
<td>0.027</td>
<td>0.207</td>
<td>−0.092</td>
<td>−0.011</td>
<td>0.054</td>
<td>−0.161</td>
<td>0.053</td>
<td>0.053</td>
</tr>
</tbody>
</table>

scheme, each base pair was presented as a molecular fragment. The advantage of the fragment based approach is the possibility to get the overlap matrix S, the eigenvector C and the eigenvalue matrix E, directly in terms of the linear combinations of the fragment orbitals. The charge transfer integrals between the fragment orbitals ϕᵢ and ϕⱼ were evaluated through the matrix elements of the Kohn-Sham Hamiltonian HKS as tᵢ,j = ⟨ϕⱼ|HKS|ϕᵢ⟩. The resulting magnitudes of the intra-(tᵢ₁,tᵢ₂) and the interstrand (tᵢH₁) hopping integrals are given in Table II. The notations of the integrals are defined in Fig. 1.

For the poly(G)–poly(C) DNA, the intra-(tᵢ₁ or tᵢ₂) and the interstrand (tᵢH₁) charge transfer integrals do not alter significantly along the sequence. For the DNA sequences containing A-T, G-A or a G-T pair, the charge transfer between two different neighboring base pairs, for example, tᵢH₁(1) and tᵢH₁(2) in Fig. 1, experience change in magnitude and sign. The conformational changes in the DNA molecule also influence the magnitude of the charge transfer integrals. In Ref. 11, the effect of the twist angle between neighboring base pairs on the charge transfer integral was investigated. It was found that for the angle range 30°–50° the sign of the charge transfer integral remains the same for most of the intrastrand interactions, while its magnitude is only slightly altered. In Ref. 15, the role of DNA conformation and the environment on the magnitude of the charge transfer integrals has been investigated using a hybrid quantum mechanics–molecular mechanics coupling scheme. It was noticed that the environment plays only a minor role in evaluating the charge transfer integrals, while changes in the DNA conformation leads to some fluctuations in the magnitude of the charge transfer integral, which we found also to be not so crucial for evaluation of the DNA magnetization. Moreover, influence of a mispair on DNA magnetization is however expected not only from a change in the charge transfer integrals but also from that of the HOMO and LUMO energies, which are modified due to replacement of the canonical pair by the mispair.

**IV. MAGNETIZATION OF A DNA CHAIN WITH MISPAIRS**

With the parameters at hand we now proceed to evaluate the magnetization of the DNA chains with or without a mispair. We characterize the magnetic properties of a finite DNA molecule by its magnetic susceptibility at zero field and zero temperature. We consider a finite size DNA molecule within the double-stranded tight-binding model. The magnetic field, B, in this model is introduced through Peierls substitution, which modifies the hopping integrals between nearest neighbor sites. We assume that the magnetic field is orthogonal to the DNA chain. In this case the Peierls substitution modifies only the interstrand hopping integrals as follows:

\[ t_{iH,L}(B) = t_{iH,L} \exp[2\pi i f \Theta_i], \]

where \( t_{iH,L} \) is the interstrand hopping integral for the \( i \)th base pair, \( f = dB/\Phi_0 \) is the magnetic flux in units of the flux quantum through the untwisted plaquette of size \( a = 0.34 \text{ nm} \) (base pair spacing) and \( d = 1 \text{ nm} \) (DNA radius). Here \( \Theta_i = \sum_{j=1}^{1} \cos(\theta_{ij}) + (1/2) \cos(\theta_{ij}) \) with equilibrium twist angle \( \theta_0 = 36^\circ \). Then for a given value of the magnetic field, we calculate numerically the single-particle energy spectra, \( \epsilon_i \), of the DNA molecule. The free energy of the DNA system, which at low temperature becomes just the energy of the ground state of the system, can be found from the expression, \( F(B) = \Sigma_i \epsilon_i \), where the sum is taken over all occupied single-particle states. With the known free energy as a function of the magnetic field, \( B \), we calculate the magnetization of DNA, \( M(B) = -\partial F(B)/\partial B \), and then the magnetic susceptibility, \( \chi = \partial M(B)/\partial B \).

We consider first the poly(G)–poly(C) DNA with a finite number of base pairs, \( N \). The mispair is placed in the middle of the DNA chain. Due to the finite size of the DNA molecule, in what follows, we have considered the hard-wall boundary conditions for the electron wave functions. In fact, we found that the magnetic susceptibility has only a weak dependence on the choice of boundary conditions.
The magnetic susceptibility per base pair is shown in Fig. 3 as a function of the length of the DNA molecule, i.e., the number of base pairs, \( N \). The oscillating behavior of the magnetic susceptibility as a function of the number of the base pairs is due to the helix structure of DNA. For homogeneous DNA molecules without a mispair, the magnetic susceptibility is quasiperiodic with a period that is about half of the helix period, i.e., 5 base pairs. This is because the net magnetic susceptibility of the molecule can be roughly calculated as a sum of contributions due to each elementary closed loop formed by the nearest neighbor base pairs \([\text{HOMO}(1)–\text{HOMO}(2)–\text{LUMO}(2)–\text{LUMO}(1)–\text{HOMO}(1)\) loops]. The contribution from each of such loops depends only on the absolute value of the component of the magnetic field perpendicular to the loop. As a result, the magnetic susceptibility becomes quasiperiodic with the period corresponding to the rotation of the elementary loop by 180°, which is half the period of the helix. Such quasiperiodicity is visible in Fig. 3. Introduction of a mispair in the middle of the DNA molecule effectively divides the molecule into two submolecules with the lengths equal to half the length of the original molecule. Each of such submolecule is homogeneous and its contribution to magnetic susceptibility of the system is quasiperiodic with the period equal to 5 base pair. As a result, the whole DNA becomes quasiperiodic with period of about 10 base pairs. The period is clearly visible in Fig. 3. Our explanation of the quasiperiodic behavior also suggest that if the distance from the mispair to one end of the molecule is kept constant, then the magnetic susceptibility of such a molecule becomes quasiperiodic with a period of about 5 base pairs.

From our results it is quite clear that the effect of G·A mispair on the magnetic properties of the DNA molecule strongly depends on the type of conformation of the mispair. For the G(anti)·A(syn) conformation the magnetic susceptibility is strongly modified (see Fig. 3, top panel), while for the G(anti)·A(anti) conformation, the magnetic susceptibility of the molecule is only weakly changed by the presence of the mispair (see Fig. 3, middle panel). The susceptibility for G(anti)·A(syn) has an oscillatory behavior with large amplitudes, as opposed to the mild oscillatory behavior of the mispair-free G-C chain. Most importantly, the susceptibility of DNA with a G(anti)·A(syn) mispair has a different sign. Therefore, the host poly(G)–poly(C) chain that was initially paramagnetic now becomes diamagnetic in the presence of the G·A mispair in the G(anti)·A(syn) conformation. Magnetic susceptibilities for the poly(G)–poly(C) molecule with a G·T mispair at the center are presented in Fig. 3 (middle panel). In this case the magnetic susceptibility is also modified. However, there is no change in sign (and hence the nature) of magnetization of the host DNA. With an increasing size of the molecule the effect of the mispair on the magnetic properties of the molecule is diminished, but it is still visible for the DNA molecules with up to 40 base pairs. Based on these results for the poly(G)–poly(C) chain as the host, we conclude that while for all the mispairs considered here, the magnetic properties are considerably different from those of the host DNA, only the G·A mispair in the G(anti)·A(syn) conformation stands out as having the strongest influence. To illustrate how the strong effects of the mispair on the magnetic susceptibility of the DNA molecule compare with that in the presence of a defect (a different type of base pair) within the homogeneous molecule, we present in Fig. 3 (top panel) the results for a single A–T base pair in the middle of the poly(G)– poly(C) molecule. The results indicate that the change in susceptibility of the host DNA is the same as that of the G(anti)·A(anti) and the G·T mispair with no change in the nature of magnetization.

Finally, we present the results corresponding to the poly(A)–poly(T) chain (Fig. 3, bottom panel). The magnetic susceptibilities of the mispair-free chain, and those for two different conformations of the G·A mispair are shown here. In contrast to the case of the defect-free G–C chain, the defect-free A–T chain is diamagnetic and with an almost constant value of the susceptibility (Fig. 3). The insertion of a mispair clearly alters that behavior significantly with the appearance of strong oscillations and a change in sign. Here the mispair converts the diamagnetic behavior to a paramagnetic behavior!

To simulate the effect of a mispair on the magnetic properties of a ‘generic’ DNA, we considered a completely random sequence of DNA containing a mispair. We calculated the difference between the magnetic susceptibilities of DNA with a mispair and the same DNA with the mispair replaced by the corresponding Watson–Crick pair. We found that the maximum change in magnetic susceptibility was found to occur for the G(anti)·A(syn) mispair and only when its nearest neighbor pairs are G–C. Our calculations indicate that even for a random sequence of DNA the changes in magnetic susceptibility of DNA are in the same range as for a homogeneous DNA. The deviation from the results for homogeneous DNA is less than \( 1 \times 10^{-4} \mu_B/T \). Therefore, the correction to the magnetic properties of DNA is determined mainly by the nearest neighbors of the mispair.

In our calculations thus far, we assumed that the mispair is placed exactly in the middle of a DNA molecule. For small molecules such as 10–30 base pairs the actual position of the mispair within the molecule should modify the results. In Fig. 4(a) we show how the position of the mispair affects the
magnetic susceptibility of DNA. The dependence of susceptibility on the position of the mispair is strong for short molecules, i.e., for about 10 base pairs, and becomes weaker for longer molecules. Although the DNA susceptibility depends on the position of the mispair, this dependence is within the range of oscillations shown in Fig. 3 and even if the mispair is not in the middle of the DNA we see a strong effect of the mispair on magnetic susceptibility of DNA.

If a DNA contains more than one mispair then there is a coupling between electronic states of different mispairs. Such a coupling is strong only if the distance between the mispairs is relatively small. The coupling between the mispairs modifies the corrections to the magnetic properties of DNA due to mispairs. To characterize the strength of these corrections we considered a poly(G)–poly(C) DNA chain with two G(anti)-A(syn) mispairs. The magnetic susceptibility of that system is shown in Fig. 4(b) for different separations \( d \) between the mispairs. The coupling between mispairs, which is stronger for small \( d \), results in nontrivial oscillatory corrections to the magnetic susceptibility of DNA with two mispairs. In general, the coupling between the mispairs enhances the changes in magnetic susceptibility of DNA. By increasing the distance \( d \) between the mispairs this enhancement first increases, reaches its maximum at \( d = 4 \), and then decreases. Finally, for large distances between the mispairs (\( d > 10 \)) the mispairs can be considered as decoupled.

In the above analysis of the magnetic properties of DNA containing mispairs, we considered only the low temperature case and disregarded the dynamic disorder, which is due to finite temperature vibrations of the DNA structure. Such dynamic disorder changes the hopping integrals and can modify the magnetic properties of the system. In Ref. 7, it was shown that finite temperature corrections turn the system diamagnetic. These corrections are due to electron-vibration interactions and can be roughly understood within the perturbation theory, for which the finite temperature corrections are proportional to \( \gamma_H \gamma_L T \). Here the coefficients \( \gamma_H \) and \( \gamma_L \) are the dependence of the intrastrand hopping integrals on the deviation of the DNA twist angle from its equilibrium value. These coefficients have opposite signs, which results in a diamagnetic contribution to the magnetic susceptibility of the DNA. Similarly, for DNA with mispairs we should also expect the diamagnetic temperature correction to the magnetic properties. Concerning the change in magnetic susceptibility due to the presence of mispairs, i.e., the difference between the susceptibilities of DNA with mispairs and DNA without mispairs, the finite temperature correction can be roughly estimated from the perturbation theory as proportional to

\[
[\langle \gamma_H \gamma_L \rangle_{\text{mismatch}} - \langle \gamma_H \gamma_L \rangle_{\text{canonical}}] T
\]

Depending on the relative values of the coefficients \( \gamma_H \) and \( \gamma_L \) for mispairs and the canonical base pairs, the temperature correction can either suppress or enhance the effects of mispairs.

V. CONCLUSION

We have found some unique magnetic properties of DNA chains containing G·A or G·T mispairs. Magnetic susceptibilities of a mispair-free DNA chain are strongly modified even when a single mispair is inserted in DNA. Although all the mispairs considered here generally alter the magnetic properties of the host DNA, the strongest modifications of the magnetic properties originate from the G(anti)-A(syn) conformation of the G·A mispair. In poly(G)–poly(C) DNA, this mispair induces large oscillations in magnetic susceptibility and a change in sign for a small number of base pairs. This means that the paramagnetic poly(G)-poly(C) DNA converts to diamagnetic DNA in the presence of this mispair. In poly(A)-poly(T) DNA which is diamagnetic, both conformations of the G·A mispair induce strong oscillations in the susceptibilities and change the sign of susceptibility by making the system paramagnetic. The G·A mispair, which is structurally close to the canonical base pair (and hence can normally escape detection), has the strongest signature because it changes the nature of magnetization of the host DNA, and therefore should be susceptible to immediate detection. We propose that measurement of the magnetic properties of DNA might provide a direct route to detect the presence of a mispair and identify the type of mispairs from their unique signatures in the magnetic susceptibility.

ACKNOWLEDGMENTS

The work was supported by the Canada Research Chairs Program and the NSERC Discovery Grant.

---


---