J|A|C|S COMMUNICATIONS

Published on Web 11/24/2009

Identification of Single-Base Mismatches in Duplex DNA by EPR Spectroscopy

Pavol Cekan and Snorri Th. Sigurdsson*

University of Iceland, Science Institute, Dunhaga 3, 107 Reykjavik, Iceland

Received July 15, 2009; E-mail: snorrisi@hi.is

Single-base mutations, also called single-nucleotide polymorphisms (SNPs), are variations in the human genome that are a cause of many genetic diseases and disorders.[1] Most methods for detection of SNPs are based on monitoring hybridization of a nucleic acid probe to the sequence of interest, which requires a careful selection of probes and hybridization conditions.[2] A more direct approach is to detect the SNP-dependent change in the local microenvironment of the DNA duplex, for example by using fluorescence.[3,4] However, the nucleotides that flank the SNP site can have a large impact on the fluorescence detection and therefore limit the generality of this approach.[4]

Electron paramagnetic resonance (EPR) spectroscopy gives information about unpaired electrons and has been widely used to study the structure and dynamics of nucleic acids.[5] The study of dynamics by EPR spectroscopy can also reveal dramatic structure-dependent differences in the mobility of specific nucleotides. For example, the effects of metal ions, small organic molecules, and peptides on the dynamics of spin-labeled TAR RNA has given insights into the structures of these receptor−ligand complexes.[6] We were interested in using such EPR-detected structure-dependent dynamics to analyze single-base mismatches in duplex DNA. The recently described rigid spin-labeled nucleoside **Ç**, which forms a base pair with G (Figure 1), seemed like a promising probe since the nitroxide reports directly on the movement of the base to which it is fused.[7] However, very small differences between mismatches and the fully base-paired duplex were observed, indicating that **Ç** is not sensitive to changes in the motion resulting from mispairing (Figure S4 in the Supporting Information).[8] We hypothesized that connecting a nitroxide spin label with some degree of flexibility to an atom involved in hydrogen bonding would enable detection of different base pairings. Here we report that the nucleoside ^{T}**C**, containing 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) conjugated to the exocyclic amino group of C (Figure 1), can be used to readily detect mismatches in duplex DNA by EPR spectroscopy. Furthermore, the EPR spectra of all three mismatches are significantly different, and thus, the probe can identify its base-pairing partner in duplex DNA.

The phosphoramidite of 4-*N*-TEMPO-2′-deoxycitidine (^{T}**C**) was prepared by a slight modification of a previously reported synthesis[9] (Scheme 1) and subsequently used to synthesize the spin-labeled DNAs (see the Supporting Information). Incorporation of ^{T}**C** into the oligomer was confirmed by mass spectrometry and by HPLC analysis of an enzymatically digested spin-labeled DNA (see the Supporting Information). The melting temperature (T_M) of the ^{T}**C**-containing DNA duplex was only 3 °C lower than that of the unlabeled oligomer (Table S1). As expected, the T_M values were significantly lower when ^{T}**C** was paired with A, C, or T, another indication that ^{T}**C** forms a stable base pair with G. The CD spectrum of a spin-labeled DNA duplex containing a ^{T}**C**·**G** pair was consistent with B-form DNA (Figure S3). These studies indicate

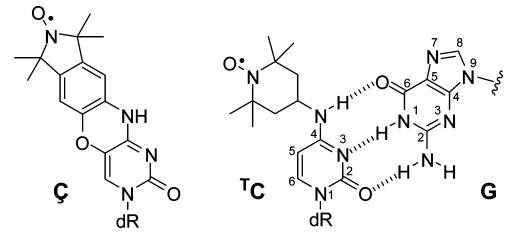
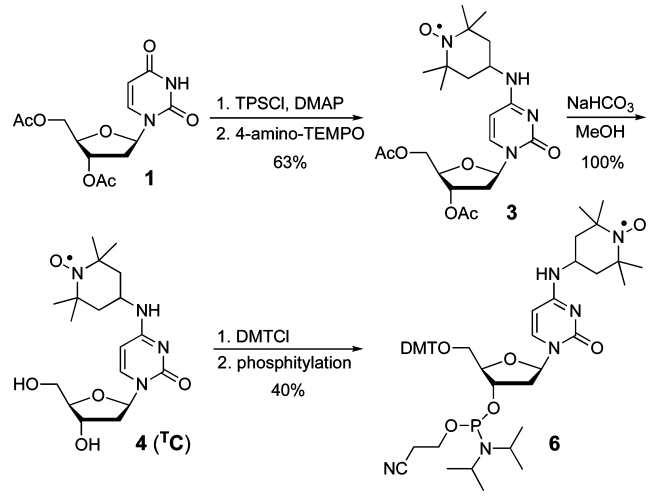


Figure 1. Nitroxide spin-labeled nucleosides **Ç** (left) and ^{T}**C** (right), which is shown base-paired to G.

Scheme 1. Synthesis of Spin-Labeled Nucleoside ^{T}**C** and Spin-Labeled Phosphoramidite **6**[a]



[a] Abbreviations: TPS, 2,4,6-triisopropylbenzenesulfonyl; DMT, dimethoxytrityl.

that the spin-labeled nucleoside ^{T}**C** does not significantly alter the stability or conformation of duplex DNA.

To determine whether ^{T}**C** could be used as a spin-labeled probe for the detection and discrimination of fully base-paired and mismatched duplexes, the EPR spectra of the four 14-mer duplexes ^{T}**C**·**G**, ^{T}**C**·**A**, ^{T}**C**·**C**, and ^{T}**C**·**T** were recorded (Figure 2A). The EPR spectra showed considerable variation, to the extent that they could all be visually distinguished from each other (Figure 2A). As a semiquantitative measure of mobility, the ratios of the low-field peak (h_l) and the center peak (h_c) of the EPR spectra were plotted for each duplex (Figure 2B); a higher h_l/h_c ratio indicates increased spin-label mobility. It should be noted that this representation of the data may not reflect an accurate description of mobility, especially when there are two or more components in the spectrum, as in the case of ^{T}**C**·**T** (see the Supporting Information for further analysis of the ^{T}**C**·**T** spectrum). However, this simplified analysis reflects differences in the mobilities of the four spin-labeled oligomers, yielding the following mobility order: ^{T}**C**·**A** > ^{T}**C**·**C** > ^{T}**C**·**T** > ^{T}**C**·**G**.

10.1021/ja905623k

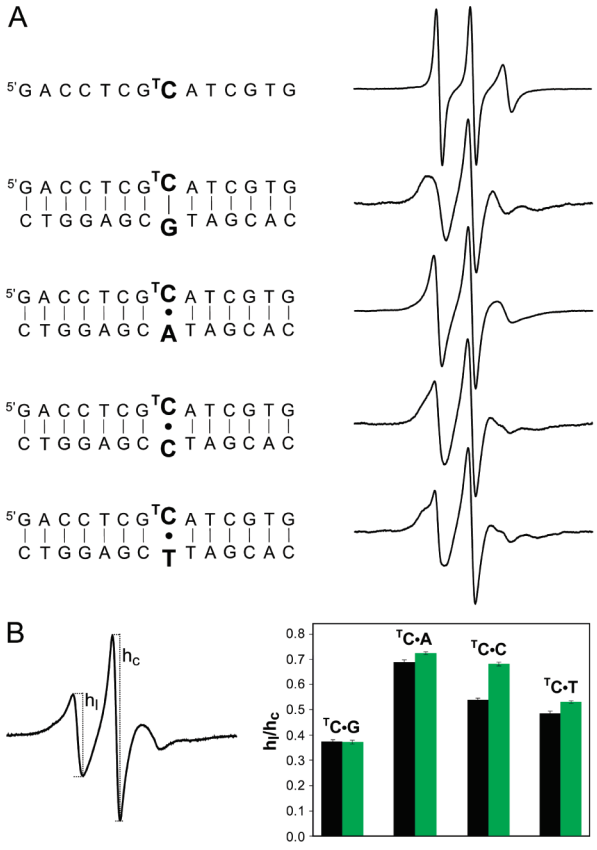


Figure 2. Mismatch detection by EPR spectroscopy. (A) Sequences and EPR spectra of single-stranded and duplex DNAs measured at 10 °C and pH 7.0. (B) h_l/h_c ratio of each spin-labeled duplex at pH 7.0 (black) and 5.0 (green), as determined from the EPR spectra. h_l and h_c are the heights of the low-field and center peaks, respectively.

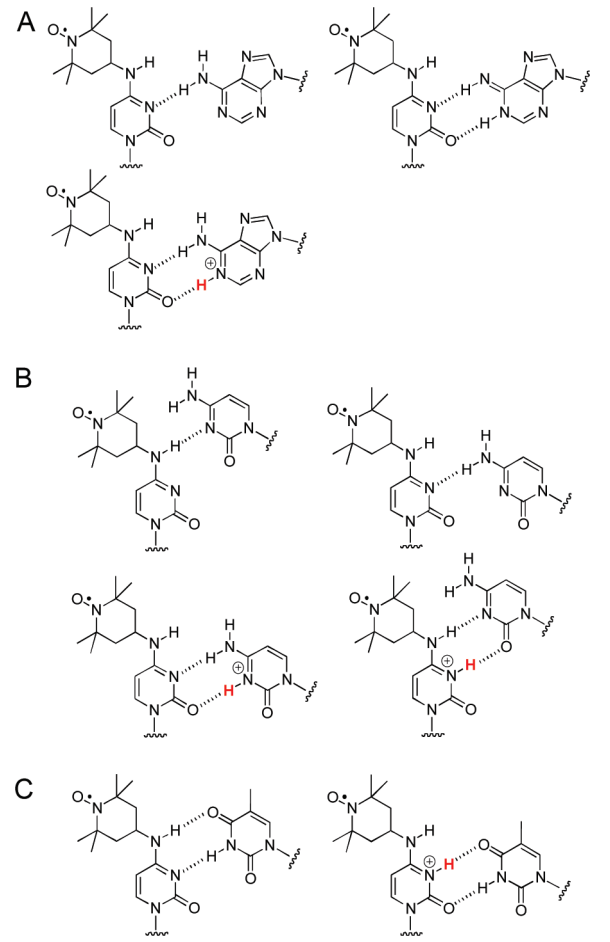


Figure 3. Possible base-pairing of ^{T}**C** with A (A), C (B), and T (C). Protons that are involved in protonation of the base pairs are shown in red.

In an attempt to explain this mobility order, we drew the base-pairing possibilities for ^{T}**C** with the natural bases (Figures 1 and 3) on the basis of NMR and thermodynamic studies of C mismatches.[10−13] ^{T}**C** exhibits the lowest mobility when base-paired to G (Figure 1), presumably because the hydrogen bond between N4 of ^{T}**C** and O6 of G slows the motion of the spin label by restricting the rotation around the ^{T}**C** C4−N4 bond. Such rotation is not restricted in ^{T}**C·A**, since N4 of ^{T}**C** does not participate in any hydrogen bonding[10,12] (Figure 3A), and this explains why ^{T}**C·A** shows the highest mobility of all the duplexes. In ^{T}**C·C** and ^{T}**C·T**, the mobility of the spin label lies somewhere between those of ^{T}**C·A** and ^{T}**C·G** (Figure 2B). Moreover, the spectra indicate that there are two types of motion present in the sample, a fast-motion (^{T}**C·A**-like) component and a slow-motion (^{T}**C·G**-like) component, which can be explained by the existence of two base-pairing possibilities for both ^{T}**C·C** and ^{T}**C·T** (Figure 3B,C). For both ^{T}**C·C** and ^{T}**C·T**, one base-pair structure has N4 of ^{T}**C** hydrogen-bonded, which should result in lower mobility (Figure 3B,C), while the other does not.

The apparent pK_as of the C-mismatched pairs have been determined to be 7.2,[10] 6.95,[11] and 5.65[11] for **C·A**, **C·C**, and **C·T**, respectively. With the assumption that the spin label has a negligible effect on the pK_a of the C to which it is attached, the spin-label mismatched pairs are protonated to some extent at pH 7.0. To investigate whether further protonation of the base pairs had an effect on their mobility, the EPR spectra of the duplexes were recorded at pH 5.0 (Figure S5A) and the h_l/h_c ratios plotted (Figure 2B). As expected, the mobility of fully base-paired duplex ^{T}**C·G** remained the same. Although the mobility order of the duplexes did not change from that at pH 7.0, the mobilities of ^{T}**C·A**, ^{T}**C·T**, and in particular ^{T}**C·C** were higher at lower pH (Figure 2B). The increase in the mobilities of both ^{T}**C·C** and ^{T}**C·T** is consistent with protonation of **C** and ^{T}**C**, respectively, to form a base pair where N4 does not participate in hydrogen bonding (Figure 3B,C). The data indicate that the favored structure for the protonated ^{T}**C·C** pair is one in which **C** is protonated rather than ^{T}**C** (Figure 3C), which pushes the spin label toward the major groove. The smaller increase in the mobility of ^{T}**C·T** relative to ^{T}**C·C** can be explained by a lower pK_a for **C·T**. Thus, at pH 5.0, ^{T}**C·T** still contains a considerable amount of neutral ^{T}**C·T**.

Modeling of ^{T}**C** in duplex DNA showed that the methyl group of T on the opposing strand in the 3′-direction [5′-d(^{T}**C**A)] and the one on the same strand in the 5′-direction [5′-d(T^{T}**C**)] were in close proximity to the spin label (Figure 4A). Therefore, in addition to the original sequence 5′-d(G^{T}**C**A), duplexes with the flanking sequences 5′-d(T^{T}**C**A) and 5′-d(G^{T}**C**C) were prepared for both the ^{T}**C·G** pair and all the mismatched sequences (Figure 4). Visual inspection of their EPR spectra (Figure S5B) showed that the individual mismatches can be detected for all three flanking sequences on the basis of the shape of the EPR spectra, which was also reflected in the h_l/h_c ratios (Figure 4B). The presence of one or two methyl groups clearly increases the discrimination between the different base pairs. In fact, very similar data were obtained for most mismatches in 5′-d(G^{T}**C**A) and 5′-d(T^{T}**C**A), which contain one and two methyl groups, respectively.

Two spin-labeled base pairs showed a substantial change in their EPR spectra upon a change in the flanking sequence. First, the

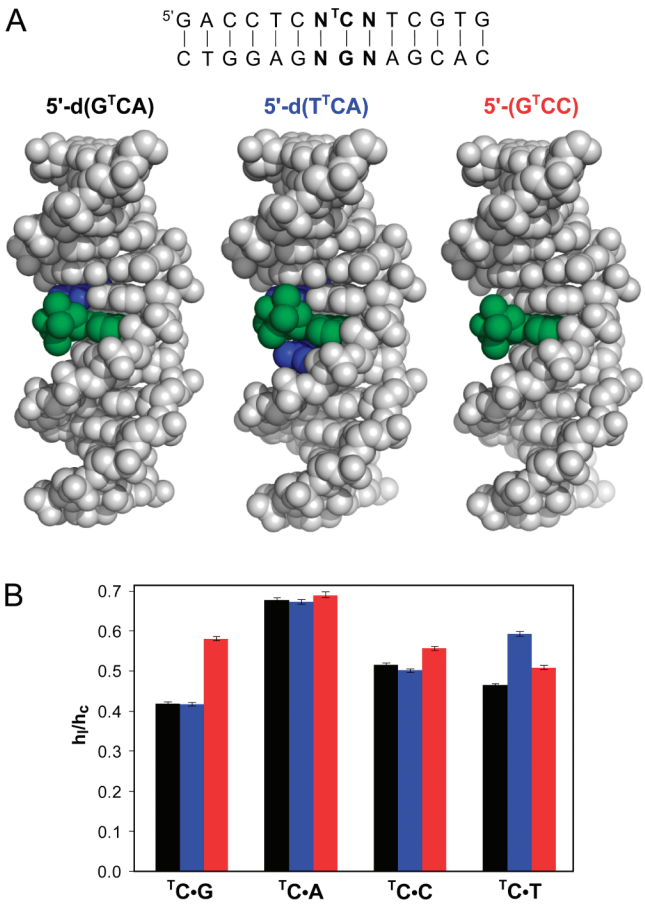


Figure 4. Mismatch detection in different flanking sequences. (A) Duplex sequences 5′-d(G^{T}CA) (black), 5′-d(T^{T}CA) (blue), and 5′-d(G^{T}CC) (red) and their molecular models for the ^{T}C·G duplexes. ^{T}C is colored green, and methyl groups in the base pairs flanking ^{T}C are shown in blue. (B) h_l/h_c ratios determined from the EPR spectra at 10 °C.

relatively large increase in the mobility of the ^{T}C·G pair in the sequence 5′-d(G^{T}CC) shows that factors other than H-bonding to N4 of ^{T}C can influence its mobility. Close inspection of the model containing this sequence (Figure 4A, right) reveals that the absence of flanking methyl groups allows free rotation around the bond connecting N4 to the spin label, a motion that now dominates the EPR spectrum. The reason why the increase does not occur to the same extent for the mismatches could be due to a minor movement of the spin-labeled base toward the center of the helix, which would restrict this rotation. The other somewhat surprising result in this flanking-sequence study was the increased mobility of the ^{T}C·T mismatch in the sequence 5′-d(T^{T}CA). An NMR structure of the ^{T}C·T mismatch in duplex DNA shows that the C is shifted toward the center of the helix.[11] Using this structure as a starting point for modeling showed that the spin label does not fit because of steric clashing with the methyl groups of the flanking base pairs. We hypothesize that this steric interaction results in displacement of the spin-labeled base toward the major groove, resulting in increased spin-label mobility.

In conclusion, single-base mismatches have been detected for the first time by EPR spectroscopy using a standard X-band continuous-wave EPR spectrometer. Furthermore, the spin label ^{T}C can report the identity of the base to which it is paired in duplex DNA and thereby give insights into hydrogen bonding between bases. This study illustrates that minor structural variations in nucleic acids can be detected with carefully chosen spin labels in conjunction with EPR spectroscopy.

Acknowledgment. We thank Lina Olofsson for technical assistance. This work was supported by The Icelandic Research Fund (080041022) and by a doctoral fellowship to P.C. from the Eimskip Fund of the University of Iceland.

Supporting Information Available: Supplementary figures and tables, experimental procedures, and characterization of compounds and oligomers, including spectral data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Nakatani, K. *ChemBioChem* **2004**, *5*, 1623–1633. (b) Olivier, M. *Physiol. Genomics* **2004**, *16*, 182–183. (c) Suh, Y.; Vijg, J. *Mutat. Res.* **2005**, *573*, 41–53.
(2) Twyman, R. M. *Curr. Top. Med. Chem.* **2004**, *4*, 1423–1431.
(3) (a) Hurley, D. J.; Seaman, S. E.; Mazura, J. C.; Tor, Y. *Org. Lett.* **2002**, *4*, 2305–2308. (b) Asseline, U. *Curr. Org. Chem.* **2006**, *10*, 491–518. (c) Bag, S. S.; Saito, Y.; Hanawa, K.; Kodate, S.; Suzuka, I.; Saito, I. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 6338–6341. (d) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265–4274.
(4) (a) Dohno, C.; Saito, I. *ChemBioChem* **2005**, *6*, 1075–1081. (b) Okamoto, A.; Saito, Y.; Saito, I. *J. Photochem. Photobiol., C* **2005**, *6*, 108–122. (c) Cekan, P.; Sigurdsson, S. T. *Chem. Commun.* **2008**, 3393–3395.
(5) (a) Klug, C. S.; Feix, J. B. *Methods Cell. Biol.* **2008**, *84*, 617–658. (b) Sowa, G. Z.; Qin, P. Z. *Prog. Nucleic Acid Res. Mol. Biol.* **2008**, *82*, 147–197.
(6) (a) Edwards, T. E.; Okonogi, T. M.; Sigurdsson, S. T. *Chem. Biol.* **2002**, *9*, 699–706. (b) Edwards, T. E.; Sigurdsson, S. T. *Biochemistry* **2002**, *41*, 14843–14847. (c) Edwards, T. E.; Sigurdsson, S. T. *Biochem. Biophys. Res. Commun.* **2003**, *303*, 721–725. (d) Edwards, T. E.; Robinson, B. H.; Sigurdsson, S. T. *Chem. Biol.* **2005**, *12*, 329–337.
(7) (a) Barhate, N.; Cekan, P.; Massey, A. P.; Sigurdsson, S. T. *Angew. Chem., Int. Ed.* **2007**, *46*, 2655–2658. (b) Cekan, P.; Smith, A. L.; Barhate, N.; Robinson, B. H.; Sigurdsson, S. T. *Nucleic Acids Res.* **2008**, *36*, 5946–5954. (c) Cekan, P.; Jonsson, E. O.; Sigurdsson, S. T. *Nucleic Acids Res.* **2009**, *37*, 3990–3995.
(8) Smith, A. L.; Cekan, P.; Brewood, G. P.; Okonogi, T. M.; Alemayehu, S.; Hustedt, E. J.; Benight, A. S.; Sigurdsson, S. T.; Robinson, B. H. *J. Phys. Chem. B* **2009**, *113*, 2664–2675.
(9) Giordano, C.; Fratini, F.; Attanasio, D.; Cellai, L. *Synthesis* **2001**, 565–572.
(10) Boulard, Y.; Cognet, J. A.; Gabarro-Arpa, J.; Le Bret, M.; Carbonnaux, C.; Fazakerley, G. V. *J. Mol. Biol.* **1995**, *246*, 194–208.
(11) Boulard, Y.; Cognet, J. A.; Fazakerley, G. V. *J. Mol. Biol.* **1997**, *268*, 331–347.
(12) Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1998**, *37*, 9435–9444.
(13) Allawi, H. T.; SantaLucia, J., Jr. *Nucleic Acids Res.* **1998**, *26*, 2694–2701. Peyret, N.; Seneviratne, P. A.; Allawi, H. T.; SantaLucia, J., Jr. *Biochemistry* **1999**, *38*, 3468–3477.

JA905623K