Hole Transfer Energetics in Structurally Distorted DNA: The Nucleosome Core Particle

Alexander A. Voityuk^{*,†} and William B. Davis^{*,‡}

Institució Catalana de Recerca i Estudis Avançats and Institute of Computational Chemistry, Universitat de Girona, 17071 Girona, Spain, and School of Molecular Biosciences, Washington State University, Pullman, Washington 99164-4660

Received: October 2, 2006; In Final Form: December 28, 2006

The dynamics of long-range hole transport (HT) through DNA are critically dependent on the relative energies of guanine radical cation states. Electrostatic contacts with protein fragments and changes in the secondary structure of the DNA helix are expected to directly influence the stability of a guanine radical cation. This expectation is especially relevant when considering DNA HT in the eukaryotic nucleus, where DNA is condensed into nucleosome core particles (NCPs), the fundamental building blocks of chromatin. Using quantum-chemical calculations, we consider how the electrostatic interactions between the DNA nucleobases and the surrounding protein and water atoms and the structural changes in DNA arising from compaction into a NCP affect the energetics of hole transfer between guanine sites. We find that structural distortions of DNA can have dramatic consequences for the stability of a guanine radical cation, and therefore, these effects must be taken into account during the modeling of in vivo DNA HT and in the interpretation of experimental findings. When the electrostatic potential arising from the water and basic histone proteins is included we find that DNA–histone contacts, particularly between arginine residues and the DNA minor groove, destabilize the hole state on specific guanine residues. Therefore, contacts between the DNA nucleobases and basic amino acids have the potential to perturb the sites of preferred hole stability in DNA.

1. Introduction

In vivo, DNA oxidative damage arises from many sources. One common example is the removal of a single electron from DNA by an exogenous oxidant resulting in the formation of an electron deficient site, or hole. A hole generated in DNA is expected to quickly localize at the nearest guanine residue to form a guanine radical cation ($G^{\bullet+}$) since guanine is the site of lowest oxidation potential in DNA.¹ The dynamics of the initiation step in DNA oxidation have been extensively studied through single-step hole transfer reactions in isolated DNA duplexes.^{2–7} Once formed, a $G^{\bullet+}$ has several fates awaiting it. The first is that the hole may become irreversibly trapped by chemical reaction with various oxygen and free radical species to form mutagenic oxidative lesions such as 8-hydroxyguanine, oxazolone, and imidazolone.⁸ However, since the rate of trapping is relatively slow⁹ the hole can become mobile in DNA via hole transport (HT). In HT electron transfer from a distant G residue to a $G^{\bullet+}$ leads to charge migration, and this process will continue until the competing irreversible trapping reactions occur. In vitro, hole hopping can lead to charge migration in DNA over distances > 200 Å.^{10,11} The nature of hole migration in DNA is still a subject of considerable discussion, with competing pictures of polaron-like mechanisms^{12–15} and superexchange-mediated hole transfer reactions^{3,16–21} as the reigning paradigms.

For a deeper understanding of the experimental results, theoretical assessments of the electronic properties of DNA are necessary. These efforts go back to the early calculations of Pullman and Pullman²² who found that guanine is the nucleobase

of lowest ionization energy (IE) in isolation. Modern theoretical treatments of DNA charge transport (CT)^{17,23,24} have taken into account the effects of base pairing,²⁵ base stacking,^{26,27} conformational dynamics,^{28,29} polar environment,^{30–32} reorganization energies,^{33–35} and DNA structural parameters on electronic coupling matrix elements.^{36,37} Many different levels of theoretical calculations have been used to understand the relative IEs of guanine residues in duplex DNA.^{24,26,27,38,39} Saito and co-workers reported IEs for XGY triplets calculated at the HF/6-31G* level,^{26,38} and they concluded that the oxidation potential of G is strongly influenced by the adjacent base pairs. NDDO-G semiempirical calculations of all possible nucleobase triplets 5'-XBY-3' demonstrated that the oxidation potential of B in 5'-XBY-3' is considerably affected by the identity of Y and becomes smaller in the order $C \approx T > A > G$ while the effect of the nature (or variation) of the preceding base pair X is rather small.²⁷ The results suggest that the 5'-G in GG and both G's on the 5'-side in GGG have the lowest oxidation potentials and, thus, are the best hole traps in B-form DNA. Similar estimates were derived from density functional theory (DFT) calculations of triplets 5'-XBY-3'.^{24,39} However, almost all theoretical studies have solely focused on DNA HT in B-form DNA duplexes. While this is adequate for the majority of DNA hole transfer and hole transport studies reported for DNA systems free in solution, it may not be enough to understand HT in genomic DNA for reasons outlined below.

In general, DNA in the eukaryotic nucleus is not found in an extended conformation; rather, it is compacted by histone proteins into chromatin. The basic repeat unit of chromatin is the nucleosome core particle (NCP), which contains a central histone octamer built from two copies each of histones H2A, H2B, H3, and H4 and ~ 147 bp of DNA wrapped around the periphery of the protein core (Figure 1).⁴⁰ The DNA in a NCP

* Author to whom correspondence should be addressed. E-mail: alexander.voityuk@icrea.es; wbdavis@wsu.edu.

[†] Universitat de Girona.

[‡] Washington State University.

Figure 1. View of the nucleosome core particle rendered using the coordinates from PDB code 1KX5. The histone octamer is in gray, and only one-half of the palindromic DNA structure is shown. The DNA backbone is rendered in red, and the nucleobases are in gray. All 30 guanines are colored, with those in single G steps in dark blue and those in GG tracts in cyan. The locations of representative guanine residues are indicated. Also indicated on the figure are the locations of NCP structural features referred to in the text such as the NCP entrance, the dyad axis, and turns 1.0 and 2.0. This figure and Figure 3 were generated using the program VMD.⁷⁰

is wrapped 1.5 times around the histone “spool”, and the NCP has a 2-fold pseudo-symmetry axis, the dyad axis, passing through the central DNA base pair (Figure 1). To relieve the strain placed on the DNA by the negative supercoiling required for wrapping, the DNA in the nucleosome is on average overwound (10.0 bp/turn) in the entrance and exit regions and underwound (10.7 bp/turn) near the dyad axis. In addition, there are specific regions where the DNA is kinked, particularly at sites where the DNA minor groove makes contact with the core histones.⁴¹ Needless to say, the structural parameters (Figure 2) that characterize many of the individual DNA base pairs and base-pair steps are expected to vary considerably from the values associated with B-form DNA.

A second component to DNA packaging into a NCP is the perturbed electrostatic environment of the DNA. It has been demonstrated that the dynamics of the polar environment surrounding the DNA, especially the counterions and water molecules, can strongly influence hole transfer.^{28,30,32} The environmental effects in the NCP may be very important because of the intimate DNA–protein contacts made throughout this complex. In particular, there are several regions in the NCP crystal structure where amino acids, and in particular arginine (Arg) residues, make contact with the DNA base pairs through the minor groove. We note that two previous experimental studies of DNA CT have raised the possibility that interactions between guanines in DNA and Arg residues may perturb,⁴² or even shut down,⁴³ long-range CT.

In this paper, we investigate how the changes in DNA structure and electrostatic environment inherent to a NCP perturb the energetics of the nucleobases, and in particular the energetics of the guanines that could potentially be involved in long-range DNA HT. Our starting point is the 1.9 Å resolution crystal structure of the NCP reported by Davey et al.⁴¹ The DNA in this structure is palindromic, 147 bp in length, and its sequence was derived from a human α -satellite sequence. Using the DNA coordinates, we have calculated the energetics of all of the

Figure 2. Illustration of (A) the base-pair parameters and (B) the base-pair step parameters discussed throughout this paper.

guanine residues in a nucleosome core particle using the semiempirical NDDO-G method.⁴⁴ We find that there are significant perturbations in the IEs of guanine residues and that these changes can be correlated with specific parameters that describe the local DNA structure. In particular, we find that structural distortions of a G–C base pair are much more important than inter-base-pair structural perturbations such as rise and roll (Figure 2). When the electrostatic environment arising from the histone proteins, counterions, and water was included in the NDDO-G calculations roughly half of the guanine IE values were either stabilized or destabilized. Those guanines that were stabilized are largely solvent exposed, whereas specific DNA–histone contacts are responsible for guanine IE increases in the NCP. The implications of these results on DNA HT reactions in the eukaryotic nucleus are discussed in light of recent experimental^{42,43,45–48} and theoretical⁴⁹ investigations of DNA hole transport in DNA–protein complexes.

2. Computational Details

2.1 Quantum-Mechanical Calculations. Because electron holes are trapped at sites of minimum oxidation potential, calculations of the ionization energies (IEs) of base-pair sequences are useful for predicting the reactivity of different sites in DNA toward one-electron oxidation. In particular, the energy for hole transfer between two bases B and B' can be

estimated as the IE difference, or stabilization energy (SE), between these bases embedded in the duplex. To calculate the SE values for guanines in DNA wrapped around a nucleosome core particle, the semiempirical method NDDO-G was utilized for all of the calculations in this paper. NDDO-G was initially elaborated to simulate absorption spectra and ionization energies (within Koopmans' theorem) of large organic and biological molecules,⁴⁴ and NDDO-G HT energies have been shown in the past to be in good agreement with both DFT data^{24,39} and experimental estimates.⁵⁰

As an aside, it should be noted that charge-induced structural reorganization within (1) oxidized DNA base pairs and (2) the surrounding environment will affect the calculated ionization energies. The internal reorganization energy arising from structural relaxation of a GC base pair is expected to be very similar for all GCs (~ 0.3 eV); therefore, it will tend to cancel out and not substantially change the hole transfer free energies between guanine bases. The relaxation effects of the polar environment depend on the dynamics of the local environment surrounding each GC site and can be accurately estimated on the basis of combined molecular dynamics and quantum-chemical simulations. Such modeling of the NCP requires a huge amount of computer time and is near the upper limits of current computational power. Taking into account that the experimental value of the reorganization energy for HT in DNA is ~ 1 eV (arising from nearly equal ~ 0.5 eV contributions from the donor and acceptor reorganization) we expect that the difference of these terms should be smaller than 0.1 eV and in general will not dramatically change our estimates.

2.2. DNA Structure. The structure of DNA in a NCP was taken from the 1.9 Å crystal structure published by Davey et al. (PDB code 1KX5).⁴¹ Since there is pseudo-2-fold symmetry in the NCP, we have used only half of the DNA duplex in the structural analysis (chain I residues $-73-0$ and chain J residues $0-73$ in 1KX5). We have employed the program X3DNA⁵¹ to extract the structural parameters of the DNA in the NCP. The isolated base-pair parameters calculated were the translational factors shear, stretch, and stagger and the rotational parameters buckle, propeller twist, and opening. The base-pair step parameters included the translational variables shift, slide, and rise and the rotational parameters tilt, roll, and twist. Illustrations of these 12 parameters are provided in Figure 2, and a full discussion of them is presented in a paper by Dickerson.⁵² The structures of the four bases A, C, G, and T were generated with the help of averaged experimental atomic coordinates taken from high-resolution X-ray and neutron crystal structures.⁵³ With both the NCP base-pair structural parameters and the atomic coordinates of the isolated nucleobases in hand, X3DNA was used to rebuild the DNA fragments needed for the quantum-chemical calculations.

When discussing ideal B-form DNA structure, we assume that rise = 3.38 Å, twist = 36°, and all other structural parameters have a value of zero.

2.3. Electrostatic Calculations. NDDO-G calculations were carried out on base-pair triads built using the molecular coordinates from PDB code 1KX5. The triads were constructed with the GC pair of interest in the middle and the two flanking base pairs on either side. The rest of the NCP was treated as a system of point charges that creates an electrostatic potential on the atoms of the triad. In solution the histone tails make little to no contact with the packaged DNA since they are dispensable for NCP formation and stability,⁵⁴⁻⁵⁶ whereas there are nonsymmetric interactions between the histone tails and the core DNA in the NCP crystal structure, presumably arising from

TABLE 1: Values for the Local Base-Pair and Base-Pair Step Parameters in the NCP^{a,b}

parameter	NCP average	standard deviation	maximum	minimum
shear	0.04 Å	0.37 Å	1.35 Å	-1.25 Å
stretch	-0.07 Å	0.19 Å	0.38 Å	-0.58 Å
stagger	-0.01 Å	0.37 Å	1.14 Å	-0.97 Å
buckle	0.1°	7.8°	19.0°	-20.7°
propeller twist	-11.9°	7.0°	4.5°	-37.2°
opening	0.8°	3.4°	11.4°	-6.3°
shift	0.00 Å	0.69 Å	1.56 Å	-1.67 Å
slide	0.19 Å	0.78 Å	2.58 Å	-0.86 Å
rise	3.31 Å	0.24 Å	4.01 Å	2.75 Å
tilt	-0.1°	4.0°	8.3°	-13.4°
roll	2.5°	7.3°	23.8°	-18.4°
twist	34.7°	5.3°	50.0°	24.2°

^a Values calculated by the program X3DNA using the NCP crystal structure PDB code 1KX5. ^b Illustrations of these structural parameters are presented in Figure 2.

crystal packing forces. To remove this apparent crystal artifact the histones in PDB code 1KX5 were truncated by removing amino acids from the N-terminal and C-terminal tails leaving behind a "trypsinized",⁵⁷⁻⁵⁹ tailless histone octamer. Specifically, the following residues of each histone were included in the calculations: H2A 12-118, H2B 21-122, H3 27-129, and H4 20-102. The X-ray coordinates of these protein heavy atoms were extracted from PDB code 1KX5, and hydrogen atoms were added using appropriate bond lengths and angles. The resulting tailless NCP model has a total of 12 772 protein atoms and 3130 water molecules. We used the force field developed by Cornell et al.⁶⁰ to assign point charges to all surrounding protein and water atoms. Because a previous study⁶¹ has indicated that the dielectric constant (ϵ) for DNA varies between 3.4 and 20.6 (similar to the situation in proteins, where ϵ usually is taken to be between 4 and 20 for the protein environment and approximately 80 for the bulk water^{33,62}), we performed quantum-chemical calculations for each triad using a series of different values of the effective permittivity constant.

3. Results and Discussion

3.1 Structural Analysis of DNA Incorporated into a NCP.

Because of the palindromic sequence of the α -satellite DNA used in the NCP crystal structure, we chose to focus on only half of the DNA sequence. The resulting DNA construct has 74 base pairs of DNA since the nucleobases at the dyad axis were included. Within this sequence, there are 16 isolated G-C base pairs and 7 GG sites yielding a total of 30 G-C base pairs. The single guanines are numbered sequentially from the entrance of the nucleosome to the dyad axis as G1 through G16, while the GG steps are sequentially numbered from the NCP entrance to the dyad axis as GG1 to GG7. Figure 1 shows the location of all 30 guanine residues in the NCP structure.

Using the program X3DNA, we extracted the isolated base-pair and base-pair step parameters (Figure 2) from PDB code 1KX5 for all 74 DNA base pairs. The exact values of the structural parameters of all 74 base pairs can be found in the Supporting Information. Due to the overwinding of the DNA at the entrance of the NCP and the underwinding of the DNA near the dyad axis, we expected the DNA in the crystal structure to have some structural parameters that vary considerably from their average B-form values. Table 1 presents an overview of the values calculated by X3DNA for the local base-pair and base-pair step parameters of the base pairs in the NCP. Some of these structural parameters do indeed fall outside of the

Figure 3. Illustrative structures of distorted G–C base pairs in the NCP. Part A is the local structure around the 5′-AGT-3′ triplet containing site G6, and part B shows the local structure around the 5′-GGC-3′ triplet containing site GG6. In part B, GG6–5′-G is on the top right-hand side of the figure, and GG6–3′-G is the middle right-hand base.

normal limits for B-form DNA, while others are very similar. An example of a parameter well outside of normal B-form limits is roll, which normally falls between -11° and 7° in B-form DNA,⁶² while in the NCP it ranges from -18.4° to 23.8° . This difference is not unexpected since roll is one of the primary parameters involved in DNA bending. Another example is buckle, which is normally only a couple of degrees, but in the NCP it ranges over $\sim 40^\circ$. In contrast, the rise and twist values in the NCP fall within the B-form ranges of 2.5–4.4 Å and 24–51°, respectively.⁶³

To illustrate two types of structural perturbation observed in the NCP structure, Figure 3A shows the local structure around G6 (sequence 5′-AGT-3′) and Figure 3B shows the local structure around doublet GG6 (sequence 5′-GGC-3′). The structure of the G6 base pair (Figure 3A) is highly distorted, as characterized by its large values of buckle (-20.7°) and stagger (-0.97 Å). The two base pairs on either side of G6 have normal base-pair structural parameters. The base-pair step parameters around G6 are fairly normal as well, indicating that most of the structural distortion at this site is confined to the G–C base pair itself. Around site GG6 there are significant distortions of all three base pairs shown in the triplet structure (Figure 3B); e.g., the CG step has buckle = -9.6° , and the 3′-side GC pair has an opening of -6.30 Å. The base-pair step parameters in this region are also significantly different from ideal B-form DNA. For instance, roll = -12.4° for the 5′-AG-3′ step, while shift = -1.32 Å in the 5′-GG-3′ step.

3.2. Effects of G–C Base-Pair Structure on Guanine IEs.

Using the X3DNA-derived structures, we performed NDDO-G semiempirical calculations to obtain the IE values for the isolated guanine nucleobases (IE(G)), the isolated G–C base pairs (IE(GC)), and the guanine in the base-pair triplet setting 5′-X-G-Y-3′ + 5′-Y′-C-X′-3′ (where Y′ and X′ are the Watson–Crick complements to Y and X, respectively) (IE(triplet)).

TABLE 2: Stabilization Energies of the Single and GG Step Guanine Residues in the NCP^a

guanine ^b	BP no. ^c	5′-XGY-3′ ^d	SE _{total} ^e	SE _{bp} ^f	SE _{stack} ^g
G1	71	TGA	−0.655	−0.399	−0.256
G2	59	TGC	−0.688	−0.454	−0.234
G3	58	TGC	−0.607	−0.402	−0.205
G4	56	AGA	−0.601	−0.339	−0.262
G5	52	AGT	−0.515	−0.321	−0.194
G6	43	AGT	−0.528	−0.315	−0.213
G7	41	TGT	−0.588	−0.424	−0.164
G8	30	AGT	−0.531	−0.379	−0.152
G9	28	TGC	−0.705	−0.372	−0.333
G10	27	AGC	−0.772	−0.383	−0.389
G11	21	TGA	−0.734	−0.542	−0.192
G12	14	TGC	−0.712	−0.395	−0.317
G13	11	TGT	−0.672	−0.535	−0.137
G14	8	TGA	−0.714	−0.442	−0.272
G15	6	AGC	−0.712	−0.417	−0.295
G16	5	AGC	−0.829	−0.393	−0.436
GG1–5′-G	64	TGG	−0.733	−0.349	−0.384
GG1–3′-G	65	GGA	−0.842	−0.506	−0.336
GG2–5′-G	61	AGG	−0.997	−0.563	−0.434
GG2–3′-G	62	GGT	−0.736	−0.480	−0.256
GG3–5′-G	48	TGG	−0.826	−0.397	−0.429
GG3–3′-G	49	GGT	−0.571	−0.340	−0.231
GG4–5′-G	35	TGG	−0.930	−0.444	−0.486
GG4–3′-G	34	GGA	−0.927	−0.460	−0.467
GG5–5′-G	24	TGG	−0.713	−0.396	−0.317
GG5–3′-G	25	GGA	−0.949	−0.454	−0.495
GG6–5′-G	16	AGG	−0.611	−0.490	−0.121
GG6–3′-G	15	GGC	−0.511	−0.301	−0.210
GG7–5′-G	3	TGG	−0.844	−0.503	−0.341
GG7–3′-G	2	GGA	−0.687	−0.393	−0.294

^a Calculated using NDDO-G semiempirical theory. ^b See Figure 1 for the location of these guanine residues in the NCP. ^c Base pairs numbered from the dyad axis (0) to the NCP entrance (73). ^d Local DNA sequence around each G site. ^e SE_{total} is the total stabilization energy calculated for the base-pair triplet centered on the guanine of interest. ^f SE_{bp} is the stabilization energy calculated for the isolated GC base pair. ^g SE_{stack} is the stabilization energy *c* arising from interactions with the 5′ and 3′ neighboring base pairs.

Tables S3 and S4 in the Supporting Information have all of the calculated IEs for the 16 single guanines and the 14 guanines in GG steps. The total stabilization energy of a guanine radical cation in the nucleosome, SE_{total}, is defined as

$$SE_{\text{total}} = IE(\text{triplet}) - IE(G) \quad (1)$$

Obviously SE_{total} will depend on two factors: (1) structural distortion of the corresponding G–C pair and (2) the effects of the flanking base pairs. The contribution from structural distortion of the isolated G–C base pair is calculated as

$$SE_{\text{bp}} = IE(GC) - IE(G) \quad (2)$$

while the contribution arising from the effects of the flanking base pairs is calculated as

$$SE_{\text{stack}} = IE(\text{triplet}) - IE(GC) \quad (3)$$

The NDDO-G-calculated values for SE_{total}, SE_{bp}, and SE_{stack} for all of the guanines in the nucleosome are provided in Table 2.

The IE(G) values for all 30 guanines fall into a very narrow window centered around 8.083 eV. However, the SE_{bp} values show a much wider range of values spanning 0.279 eV, indicating that the hole stabilization on any guanine due to base pairing with cytosine is highly dependent upon the structure of the G–C base pair. Obviously, the structure of some G–C base

pairs is highly stabilizing to the hole state (e.g., G11, G13, and GG2–5'-G), whereas the structure of others does not lead to large stabilization of the hole state (e.g., G4, G5, G6, and GG6–3'-G). The average SE_{bp} value for all 30 guanines is -0.420 eV with a standard deviation of 0.069 eV. As a point of comparison, according to our calculations $IE(GC) = 7.695$ eV and $SE_{bp} \approx -0.4$ eV for a guanine residue in an average B-form G–C base pair. Thus, the average SE_{bp} values for B-form DNA and DNA in the NCP are very close, with the large distribution in the NCP most likely due to local structural changes leading to perturbation of the hydrogen bonding and electrostatic interactions between several guanines and cytosines. It has been found experimentally that the lowering of the oxidation potential of a guanine derivative due to base pairing is 0.34 eV.⁶⁴ According to B3LYP/6-311++G** calculations, the vertical IE of a G–C base pair in the gas phase is 0.63 eV lower than that of G.⁶⁴ Thus, the NDDO-G method provides very good estimates of the effects of base pairing on the IE of guanine.

3.3. Base-Pair Stacking Effects on Guanine IEs. In previous studies it was found that the IE of a guanine residue in B-form DNA is largely independent of the identity of base X in the general triplet 5'-XGY-3' but nucleobase Y has a large effect on hole stability.²⁷ For the overall stabilization of the IEs of the single guanines in the NCP, the trends generally follow those observed previously for B-form DNA structures. For instance, in the triplets 5'-AGC-3' (G10, G15, G16), 5'-AGA-3' (G4), and 5'-AGT-3' (G5, G6, G8) the SE_{total} values fall in the order $Y = C < A < T$, as expected. Also, the identity of X is not as important as Y in the NCP as demonstrated by the smaller differences between the SE_{total} values of the 5'-TGA-3' (G1, G11, G14) and 5'-AGA-3' (G4) sites. However, there are two groups of guanine triplets, 5'-TGC-3' (G2, G3, G9, and G12) and 5'-AGC-3' (G10, G15, G16), where the SE_{total} values vary by more than 0.1 eV for sites with identical local sequence. In the 5'-TGC-3' sites, G2, G9, and G12 have almost identical SE_{total} values, while SE_{total} for G3 is ~ 0.1 eV lower. In breaking the numbers down further one notices SE_{stack} (eq 3) for G3 is much lower than average, and hence the destabilization of this site is most likely due to effects attributable to base-pair stacking interactions. A similar observation can be made for the 5'-AGC-3' sequences, where all of the SE_{bp} values fall within a limited range, but the SE_{stack} value for G15 is much lower than the others.

To further understand the effects of base stacking on the IE of a guanine site, we assume that the contributions from the flanking base pairs in the triplet 5'-XGY-3' are additive, i.e.,

$$SE_{stack} = SE_5(X) + SE_3(Y) \quad (5)$$

where

$$SE_5(X) = IE(5'-XG-3') - IE(GC) \quad (6)$$

$$SE_3(Y) = IE(5'-GY-3') - IE(GC) \quad (7)$$

We have predicated this assumption on the observation that electrostatic interactions are mainly responsible for hole stabilization,⁶⁵ and therefore one may expect that the stabilization energy will be approximately additive. Figure 4 compares the stabilization energies calculated using eqs 3 and 5–7 for all G–C base pairs, single guanines, as well as GG tracts in the nucleosome. Indeed, the effects of the base pairs on the 5'- and 3'-sides are additive for all cases with one exception for GG4–5'-G (triplet 5'-TGG-3'). In this sequence a hole is

Figure 4. Comparison of the stabilization energy of a guanine site calculated using 5'-XGY-3' triplets (x-axis) and the sum of contributions found for the corresponding doublets 5'-XG-3' and 5'-GY-3'. Points 1 and 2 (shown as larger circles) represent the values for site GG4–5'-G before (1) and after (2) correction of the IEs values due to delocalization of the hole over the two guanines in this GG step (see text).

delocalized over two guanines ($\sim 64\%$ and 36% on GG4–5'-G and GG4–3'-G, respectively). In the triplet, there is a low lying excited state (the energy gap is 0.19 eV) with the reverse charge distribution ($\sim 36\%$ and 64% on GG4–5'-G and GG4–3'-G). A similar situation is also found for the normal doublet 5'-GG-3' in B-form geometry. Note that the effects of polar surroundings essentially suppress the charge delocalization over two guanines.^{31,66,67} Point 1 in Figure 4 is obtained if one estimates the IEs using the orbital energies of the highest occupied molecular orbitals in which the contribution of GG5–5'-G is greater than 50% . If corrected IE values, obtained by averaging the energy of the two states, are used, then one obtains point 2 in Figure 4, which falls on the trend line.

An analysis of the values of $SE_3(Y)$ for the guanines in the NCP indicates that the value of this parameter is highly dependent on the identity of Y. The values of $SE_3(A)$ and $SE_3(T)$ do not vary dramatically throughout the nucleosome— $SE_3(A)$ ranges from -0.203 (GG5–3'-G) to -0.124 eV (G4), and $SE_3(T)$ ranges from -0.044 (G6) to $+0.017$ (G8). These values are close to those that we calculate for doublets of ideal B-form structure, $SE_3(A) = -0.135$ eV and $SE_3(T) = -0.063$ eV. On the contrary, $SE_3(G)$ and $SE_3(C)$ exhibit much larger variations. In most cases $SE_3(G)$ has a negative value of less than -0.25 eV, similar to the reference value of -0.318 eV in a GG dimer with B-form structure. However, GG6–5'-G is one notable outlier to this trend since $SE_3(G) = +0.01$ eV for this site. Cytosine displays the largest distribution of $SE_3(Y)$ values. For instance, in some cases the values of $SE_3(C)$ are quite small, e.g., -0.065 and -0.039 for guanines G2 and G3, respectively, and are similar to the value calculated for a reference GC dimer, $SE_3(C) = 0.039$ eV. However, other sites such as G10, GG6–3'-G, and G16 possess quite large values for $SE_3(C)$, -0.257 , -0.303 , and -0.312 eV, respectively.

Interestingly, two of the anomalous values for $SE_3(C)$ and $SE_3(G)$ come from the same guanine doublet, GG6. Figure 3B shows the local structure around the 5'-GGC-3' triplet at this site, and we have already commented on the amount of local distortion in this region of the NCP. The question now becomes, given the perturbations in DNA structure present at GG6, which are more important for the stability of the guanine residues—

TABLE 3: Delineating the Effects of DNA Structure on the Hole Stabilization Energies of the Two G Residues at NCP Site GG6^{a,b}

system	base-pair parameters ^c	base-pair step parameters ^c	SE ₃ (G) (eV) ^d	SE ₃ (C) (eV) ^e
1	ref ^f	ref ^g	−0.318	0.037
2	ref ^f	str ^h	−0.232	−0.042
3	str ⁱ	ref ^g	−0.099	−0.234
4	str ⁱ	str ^h	−0.055	−0.303

^a NDDO-G semiempirical calculations were used. ^b Two base-pair doublets were used in these calculations, 5'-GG-3' and 5'-GC-3', and in each the 5'-G-C base pair is given ideal B-form structure. ^c Base-pair parameters and base-pair step parameters are defined in Figure 2. ^d Stabilization energy calculated for the 5'-G in the 5'-GG-3' base-pair doublet of site GG6. ^e Stabilization energy calculated for the 5'-G in the 5'-GC-3' base-pair doublet of site GG6. ^f Ref is the configuration where the 3'-base pair has ideal B-form structure. ^g Ref denotes ideal B-form base-pair step parameters. ^h Str is the structure of the 3'-base pair in the NCP. ⁱ Str is the structure where the base-pair step parameters in the NCP are used.

the base-pair structural parameters or the base-pair step parameters? To answer this question, let us first turn our attention to the anomalously small SE₃(G) observed for GG6-5'-G. Our calculation scheme allows us to separate the effects of structural distortions arising from changes in (i) intra-base-pair structural parameters and (ii) inter-base-pair structural parameters on the stabilization energy of a guanine nucleobase. To this end, we have generated four structures of a 5'-GG-3' doublet, and labeled them systems 1–4. The 5'-G-C base pair is given an ideal B-form DNA structure in all four constructs—hence only structural parameters of the 3'-G-C base pair vary between the different systems. We have independently changed the intra-base-pair parameters and the base-pair step parameters for the 3'-G-C base pair to be either ideal B-form “ref” or the values obtained using X3DNA on the NCP X-ray structure “str”. After generating systems 1–4, we next ran NDDO-G calculations on them, and the obtained SE₃(G) values are reported in Table 3. In the ideal doublet (system 1), the stabilization of the radical cation is the strongest. Changes in base-pair step parameters only (system 2) lead to a smaller stabilization energy, but the changes are not dramatic. On the contrary, changing only the intra-base-pair parameters of the 3'-G-C base pair (system 3) results in a substantial decrease in SE₃(G) as compared to system 1. In system 4, the value of SE₃(G) is comparable to that obtained from system 3. Finally, if we take system 4 and set the intra-base-pair structural parameters in the 5'-G-C base pair to those found in the X-ray structure, we calculate SE₃(G) = 0.01 eV. The corresponding data for SE₃(C) calculated for the 5'-GC-3' step containing GG6-3'-G are presented in Table 3. We note that the results for SE₃(C) and SE₃(G) are very similar at site GG6, as are the results for the other anomalous SE₃ values (data not shown). To summarize these results, (1) the IE of a guanine residue is very sensitive to intra-base-pair distortions of the 3'-flanking base pair; (2) the effects of base-pair step parameters are also important, but they do not lead to large changes in hole stability; (3) inherent distortions of the 5'-G-C base pair also affect the value of SE₃(G). One caveat to point 2 is that the two different sets of parameters are interrelated to some extent, so a true separation of the inter- and intra-base-pair effects may not be realistic.

What about the stabilization of G^{•+} due to 5'-X neighbors? In the case of the ideal dimers 5'-XG-3' we find the following stabilization energies: SE₅(X) is −0.183, −0.158, −0.151, and −0.115 eV for X = G, A, T, and C, respectively. These data show that the stabilization energy for B-form reference structures is largely independent of the identity of the 5'-neighbor, as

expected.²⁷ In general, this is also true for DNA in the NCP. SE₅(G) ranges from −0.166 to −0.332 eV, with the largest values of −0.332 and −0.306 eV found for GG4-3'-G and GG5-3'-G. For all SE₅(A) values (with the only exception for GG2-5'-G where SE₅(A) = −0.072 eV) we found that −0.188 eV ≤ SE₅(A) ≤ −0.133 eV. There are 14 5'-TG-3' steps in the NCP, and for 11 of them −0.197 ≤ SE₅(T) ≤ −0.118 eV. For G11, G12, and GG7-5'-G the stabilization energies are remarkably smaller, −0.060, −0.013, and −0.060 eV, respectively. In the DNA sequence of interest there are no 5'-CG-3' sequences present.

Let us briefly consider the differential effects of the base-pair and base-pair stacking parameters for the anomalous value of SE₅(T) calculated for G12. The following analysis follows a similar logic to that discussed earlier for the anomalous SE₃(G) and SE₃(C) values in GG6. Starting with an ideal base-pair structure and base-pair stack for the 5'-TG-3' sequence containing G12 results in a −0.151 eV reference value for SE₅(T). Changing only the TA base-pair structure from ref to str decreases SE₅(T) to −0.113 eV. Changing only the base-pair stacking parameters from ref values to str results in SE₅(T) = −0.104 eV. Finally, changing both sets of parameters for the TA base pair to those found in the NCP results in a value of SE₅(T) = +0.002 eV. Thus, the changes in SE₅(T) for site G12 are due to a convolution of effects arising from the base-pair parameters and base-pair stacking parameters describing the structure of the 5'-flanking TA base pair.

To better understand why changes in the base-pair step geometry play a minor role in the stabilization of G^{•+}, we carried out NDDO-G calculations on base-pair doublets in which we started out with an ideal B-form structure and varied the six structural parameters shift, slide, rise, tilt, roll, and twist individually. Table 4 has the data for base-pair doublets of the form 5'-GY-3', and Table 5 shows the results for 5'-XG-3' doublets. Note that in general the isolated effects of changes in the base-pair step parameters on the SE₅(X) and SE₃(Y) values of any nucleobase site are predicted to be rather small. This is consistent with the results of our structural interrogation of site GG6 in Table 3. The overall trends observed in this data are as follows. When Y is a purine, the two parameters that can most affect the guanine IE are rise and tilt. When Y is a pyrimidine, the two most important parameters are slide and tilt, with shift and rise also contributing in the case of Y = T. The IE of a guanine nucleobase is strongly dependent on the parameters shift and rise for all four nucleobases at position X. For a X = A, the IE is also dependent upon roll and twist. In the case of a pyrimidine at position X, the guanine's IE is dependent on all base-pair step parameters except tilt.

The data in Table 2 indicate that for most of the guanines in the NCP the stability of the hole state on any site is dictated more by SE_{bp} and local base-pair structure than base-pair stacking interactions and SE_{stack}. Even in the exceptions, e.g., G10 and G16, the base-pair stacking interactions contribute less than 53% to SE_{total}. We note that the majority of guanines that possess SE_{stack} ≥ SE_{bp} occur in GG tracts.

3.4. Effects of DNA Structure on GG Sites in the Nucleosome. When present in a B-form geometry, the 5'-G's of the seven GG steps are expected to be the sites of lowest SE_{total} in the studied human α-satellite sequence. Figure 5 shows the SE values for all 30 guanine residues in this DNA sequence, normalized by calculating IE(triplet) − ⟨IE(triplet)⟩. The guanines are numbered according to their distance from the dyad axis in the NCP, with base pair 0 located at the dyad axis, and the correspondence between base pair number and the guanine

TABLE 4: Changes in the Ionization Energy of a Guanine Residue Arising from in the Base-Pair Step Parameters of Nucleobases in the 3'-Direction^{a,b}

structure	5'-GC-GC-3'	5'-GC-AT-3'	5'-GC-TA-3'	5'-GC-CG-3'
reference ^c	7.377	7.560	7.632	7.732
shift = -1.0Å	-0.017	-0.002	-0.026	-0.021
shift = +1.0Å	-0.006	-0.002	0.030	-0.027
slide = -1.0Å	-0.018	0.010	0.056	0.085
slide = +1.0Å	0.030	-0.011	-0.054	-0.090
rise = 2.88Å	-0.082	-0.034	-0.038	-0.009
rise = 3.88Å	0.044	0.014	0.007	-0.013
tilt = -10°	-0.095	-0.037	-0.046	-0.047
tilt = +10°	0.055	0.008	-0.013	-0.047
roll = -10°	0.004	0.000	0.004	-0.003
roll = +10°	-0.026	-0.009	-0.014	-0.003
twist = 26°	0.010	0.004	0.018	0.015
twist = 46°	-0.016	-0.005	-0.021	-0.019

^a NDDO-G semiempirical theory was used for all calculations. ^b All energies are given in units of eV. ^c The reference values for IE were calculated for base-pair doublets where all base-pair step parameters have ideal B-form values (rise = 3.38 Å, twist = 36°, all other parameters zero); the other values in the table are the calculated differences in IE.

TABLE 5: Changes in the Ionization Energy of a Guanine Residue Arising from Changes in the Base-Pair Step Parameters of Nucleobases in the 5'-Direction^{a,b}

structure	5'-GC-GC-3'	5'-AT-GC-3'	5'-TA-GC-3'	5'-CG-GC-3'
reference ^c	7.509	7.534	7.541	7.578
shift = -1.0Å	0.021	-0.036	-0.047	-0.041
shift = +1.0Å	0.033	0.029	0.065	0.071
slide = -1.0Å	0.015	-0.010	-0.024	-0.022
slide = +1.0Å	0.020	0.027	0.050	0.052
rise = 2.88Å	-0.029	-0.046	-0.037	-0.052
rise = 3.88Å	0.022	0.022	0.025	0.021
tilt = -10°	-0.020	-0.015	-0.001	-0.001
tilt = +10°	0.019	0.003	0.004	-0.002
roll = -10°	-0.007	-0.038	-0.041	-0.054
roll = +10°	0.016	0.034	0.038	0.033
twist = 26°	0.021	0.031	0.048	0.055
twist = 46°	0.010	-0.020	-0.027	-0.023

^a NDDO-G semiempirical theory was used for all calculations. ^b All energies are given in units of eV. ^c The reference values for IE were calculated for base-pair doublets where all base-pair step parameters have ideal B-form values (rise = 3.38 Å, twist = 36°, all other parameters zero); the other values in the table are the calculated differences in IEs.

Figure 5. Comparison of normalized IP(triplet) values for α -satellite DNA in B-form geometry (■) vs in the NCP (●).

nomenclature of this manuscript is indicated in Table 2. As expected, the 5'-G residues of GG1–GG7 are indeed predicted to be the sites of lowest IE in naked DNA. Figure 5 also shows the values of $IE(\text{triplet}) - \langle IE(\text{triplet}) \rangle$ calculated in the NCP structure. Just as in the B-form structure, the 5'-G's located in GG steps are predicted to be the overall sites of lowest IE in the NCP. However, there are several differences between GG energetics in the B-form and NCP DNA structures. First, from the SE_{total} values in Table 2 and Figure 5 one can immediately

see that the commonly observed rule that the 5'-G is the preferred site of oxidation in a GG tract does not always hold true in the NCP. For instance, the 3'-G is the site of lowest IE in GG1 and GG5, while the two guanines are almost isoenergetic in GG4. Second, in one region of this NCP a GG step is NOT predicted to be the preferred local oxidation site (i.e., site of lowest IE). In particular, the $IE(\text{triplet})$ values for the 5'- and 3'-G's of GG6 are larger than the $IE(\text{triplet})$ of any of the single guanines (G11, G12, and G13) within 5 base pairs (Figure 5). GG6 lies in a region of the nucleosome ~ 1.5 turns from the dyad axis, a region that is known to be distorted because of kinks at turns 1.0 and 2.0 (Figure 1).⁴⁰ The changes in the local DNA base-pair step structure accompanying these kinks are manifested in the relatively small values of SE_{stack} calculated for both of these guanine residues. Interestingly, the two single guanine residues (G6 and G7) near the other major site of kinking in the NCP at turn 4.0 also have relatively small values of SE_{stack} . These results imply that GG sites in these two regions of the nucleosome, and perhaps other regions as well, may not act as local hole traps due to changes in the local DNA structure arising from DNA compaction. Furthermore, if a hole distributes itself over a DNA sequence in accordance with Boltzmann statistics, then the observed damage distribution in these distorted DNA sequences, for instance, after piperidine treatment, could also vary from that observed in B-form DNA.

3.5. Effects of Electrostatics on Guanine Ionization Potentials in a NCP. In addition to the structural changes in the DNA arising from DNA–histone contacts in the NCP, there

also is the perturbed electrostatic environment of the DNA base pairs. These electrostatic effects are expected to arise from (at least) two specific sources: (1) ordered water molecules in regions with extensive DNA–histone contacts and (2) unique electrostatic contacts between the DNA nucleobases and basic residues of the histone proteins. These latter interactions occur at 6 out of 7 of the sites where the DNA minor groove is rotated toward the histone core (Figure 1), and they are predominantly interactions between Arg guanidinium groups and the minor groove face of the nucleobases. To investigate these environmental effects we therefore decided to calculate the NDDO-G IP values of all of the guanine triads discussed in the previous sections in the electrostatic context of a NCP. All DNA atoms associated with the G-centered triad of interest were included in the calculations, and all protein and water atoms were included in the calculations as a set of AMBER-derived atomic point charges. The coordinates of the water oxygens and the heavy atoms of the histones were taken from the crystal structure of Davey et al.⁴¹ Hydrogens were added to the heavy atoms before calculations were performed. Since (1) Mn^{2+} ions are present in the X-ray structure as opposed to sodium ions and (2) the dynamics of the counterions and phosphate groups affects the electrostatics, we decided not to include the charges of counterions and the sugar–phosphate backbone in this model to avoid artificial effects (and uncertainties).

Two aspects of our electrostatic model demand further discussion. The first is the utilization of histone proteins whose N- and C-terminal tails have been truncated via trypsinization. Experimentally, several studies have established that the highly basic histone tails play little if any role in the formation of the NCP particle or in its salt-induced dissociation, presumably because of limited DNA–protein contacts.^{54–56} In contrast, crystal packing forces lead to random association of the histone tails with the core DNA in PDB code 1KX5. To remove any spurious data arising from artificial histone tail–DNA interactions, we therefore removed the histone tails at known sites of trypsin cutting in NCPs in solution.^{57–59} This approach is further justified since a prior report of DNA HT dynamics in a NCP showed that there is little to no effect of the histone tails on DNA CT in a NCP.⁴² The second aspect to discuss is the utilization of a range of ϵ values in our electrostatic calculations. To begin with, the use of macroscopic ϵ values for microscopic calculations is a rather crude model but one that has been utilized previously in studying electrostatic effects in DNA.^{33,68} A wide range of “reasonable” values of ϵ are expected to exist in DNA–protein complexes, depending on the nature of the complex. Therefore, to avoid drawing any ϵ -specific conclusions we carried out calculations using a series of ϵ values to extract trends in the electrostatic calculations. Finally, we analyzed $\text{IE}(\text{triplet})$ values normalized via the calculation of $\text{IE}(\text{triplet}) - \langle \text{IE}(\text{triplet}) \rangle$ since (i) a decrease in ϵ led to a global upward shift in the calculated G $\text{IE}(\text{triplet})$ values (data not shown) and (ii) we are only interested in relative changes in the $\text{IE}(\text{triplet})$ values of the 30 guanine sites.

Figure 6 shows the results of calculations carried out with ϵ values of 20, 10, and 5. Calculations with $\epsilon = 80$ yield very similar results to those obtained for $\epsilon = 20$ (data not shown) and are omitted from Figure 6 to save transparency. As in Figure 5, the data are presented as a function of the base-pair location within the NCP (Table 2), with site 0 corresponding to the base pair at the dyad axis and base pair 73 located at the entrance to the NCP. Remarkably, the IE plots as a function of base-pair site look qualitatively similar for NCP–DNA in vacuum and for the three ϵ values studied, indicating that the

Figure 6. Effects of electrostatics on the $\text{IE}(\text{triplet})$ values for all 30 guanine residues in the NCP. Shown are the data for the isolated, NCP structure DNA (■) and the full electrostatic calculations using $\epsilon = 5$ (●), $\epsilon = 10$ (▲), and $\epsilon = 20$ (▼).

structural and electrostatic effects are not independent, and accounting for the charged and polar surroundings is not expected to lead to new qualitative changes in the energetic landscape of DNA CT in a NCP. If one looks in more detail, then the data illustrate that the effects of electrostatics on the $\text{IE}(\text{triplet})$ of any guanine residue depend on its location in the NCP structure. In general, there are three types of guanine sites in the NCP. (1) Out of the 30 guanines 11 are located at sites where electrostatics essentially do not essentially change the relative values of IE (the changes are smaller than ~ 0.1 eV at $\epsilon = 10$). At these locations in the NCP, the energetics of hole transfer between guanine residues are arise almost exclusively from differences in the local DNA structure. (2) Nine guanine residues are found to have a stabilization (i.e., decrease) of the normalized $\text{IE}(\text{triplet})$ values > 0.1 eV at $\epsilon = 10$. All of these sites are located either near turn 1.0 (GG6–5′-G, GG6–3′-G, and G12; Figure 1) or within 1.5 turns of the NCP entrance (GG1–5′-G, GG1–3′-G, GG2–5′-G, GG2–3′-G, G2, and G3) in regions where DNA makes little to no contact with the histone proteins. The stabilization of these sites most likely arises from the extensive interaction of DNA with bulk solvent in these regions. (3) Nine guanine residues are destabilized by electrostatic interactions in the NCP as ascertained by an increase in their normalized $\text{IE}(\text{triplet})$ values > 0.1 eV at $\epsilon = 10.0$. Five of these guanines (GG3–5′-G, GG3–3′-G, GG4–5′-G, GG4–3′-G, and G15) are located near regions where the DNA minor groove is rotated in toward the histone surface, while the other four (G6, G8, G9, and G14) are in regions where the DNA minor groove makes contact with the histone tails as they enter the NCP. In particular, G14 is located at the position where the H3 N-terminal tail enters the NCP, G8 and G9 are at the site where the H2B N-terminal tail enters the NCP, and G6 is at the site where the H2A N-terminal tail enters the NCP. A common element found for 7 of these guanine sites is the presence of an arginine side chain inserted into the DNA minor groove. In particular, the following interactions are noted: H4 Arg45 with G15; H3 Arg40 with G14; either H2B Arg29 or H2B Arg30 with G9 and G8; H2A Arg43 with GG4–5′-G, H2B Arg33 with GG3–5′-G and GG3–3′-G; and H2B Arg30 with GG3–3′-G in one of the two symmetry-related sites in the NCP. Other electrostatic interactions include one between H2B Arg 86 and GG4–3′-G via the DNA major groove and a minor groove interaction between H2B Lys27 and GG3–3′-G, and site G6 is surrounded by three basic amino acids from H2A Lys14 near

the minor groove and Arg18 and Arg21 in the major groove. Surprisingly, GG5 does not fall into this category of guanine sites in the NCP. Both guanines in GG5 are in contact with H3 Arg83, yet there is a slight stabilization of the normalized IE(triplet) values for these nucleobases. From this detailed structural analysis, it is clear that placing basic amino acids near guanine residues in DNA–protein structures is predicted to lead to a destabilization of the guanine radical cation the vast majority of the time. Note that the calculated electrostatic effects are due to different values of the local potential at specific guanine locations in the NCP. Some additional corrections (which are not considered here) can arise because of the inhomogeneity in the dielectric constant of the environment, as demonstrated by a previous model³⁴ that incorporates several dielectric zones in DNA.

3.6. Implications of this Study for Previously Reported Studies on HT in Protein–DNA Complexes. The effects of DNA–protein interactions on DNA CT have been investigated in several reports to date.^{42,43,45–48} Studies involving the binding of proteins like the Antennapedia Homeobox,⁴⁶ restriction endonuclease *PvuII*,⁴⁶ the glycosylase MutY,⁴⁸ and the NCP histone octamer⁴⁵ reported little to no changes in the yields of DNA CT after protein binding. In contrast, there are several reports of DNA–protein complex formation decreasing the yields of long-range DNA CT. In the cases of methyltransferase *HhaI*⁴⁷ and TATA box binding protein⁴⁶ the observed decrease in DNA CT was assigned to interruption of the DNA π -stack in the DNA–protein complex. In another study, Nakitani et al.⁴³ found that binding of restriction endonuclease *BamHI* to DNA completely inhibited long-range DNA CT through the restriction site. The authors attributed this shutdown in DNA CT to a specific interaction between a charged Arg residue and guanine residues in the restriction site. The electrostatic calculations on the NCP support this hypothesis since most Arg–G interactions increase the IE of the involved guanine. A recent study⁴² of DNA CT in a NCP found that the relative yield of hole transport to one GG site (GG2-B) was decreased in the NCP versus naked DNA, while the relative yields at all other guanine sites were relatively constant. The location of GG2-B in the NCP was mapped out using a combination of DNA footprinting reactions, and the local environments of these guanines are expected to be most similar to residue G4 in PDB code 1KX5. From the results of our calculations, we predict that if a change in the IE of GG2-B is responsible for the decrease in hole transport to this site (and not a competing process such as trapping), then local DNA structural changes, and not electrostatics, are responsible. In Figure 5, there is a relative increase in the IE of G4 arising from structural perturbation of the GC base pair due to interactions between H2A Arg78 and the DNA backbone. However, in Figure 6 it is apparent that even though H2A Arg78 is nearby electrostatic perturbation of the IE of G4 (site 56) is minimal.

4. Conclusions

In this paper we have investigated how changes in DNA structure and electrostatic environment arising from DNA–protein interactions can influence the relative stability of guanine radical cation states. Our model system is the nucleosome core particle, the fundamental DNA–protein complex found in the eukaryotic nucleus. We have found that the IE of a guanine base can be strongly dependent on the local helix structure, to the extent that the stabilization of a hole state on any guanine may differ dramatically from the results predicted for ideal B-DNA.

In variance with polaron models,^{12,13} which assume that the excess positive charge is extended over several nucleobases in the stack, our current calculations predict that the radical cation states are essentially localized on individual guanines. In other words, the energetics of the hole states are well described by the ionization energies of the guanine sites within the NCP. Our result that the radical cation states are quite localized is in agreement with two recent findings: (1) High-level quantum-mechanical calculations (at the CASPT2 level) show that even in isolated π -stacks the excess charge is quite localized,⁶⁹ (2) the polar environment essentially suppresses charge delocalization in DNA^{23,31} so that even in homogeneous sequences consisting only of GC pairs hole states are localized on individual guanines.

While the importance of intra-GC base-pair structural changes might have been expected, the strong influence of intra-base-pair structural distortions in the flanking base pairs is novel. In general, changes in the structure of 3'-base pairs have a larger effect than 5'-base pairs, and fluctuations of 3'-G–C base pairs have more influence than 3'-A–T base pairs. Surprisingly, changes in the base-pair step parameters such as roll and rise have a smaller influence on the IE of guanine than the inter-base-pair parameters of a 3'-neighbor. Because HT energetics are very sensitive to the local helical structure of DNA, we suggest that thermal fluctuations in DNA possess the ability to change the free energy of HT. In the extreme, DNA structural fluctuations may even be capable of reversing the expected direction of HT due to a change in the sign of the driving force. This last observation suggests that controlled structural perturbations, such as those arising from DNA–protein interactions, might serve as switching elements in DNA-based molecular electronics. Additionally, one could expect that DNA regions possessing high structural flexibility will exhibit anomalous behavior because of free energy variation (among other factors). Obviously, description of the HT mechanism becomes more complicated when accounting for these effects.

Electrostatic interactions in the NCP can also determine the overall landscape of guanine IEs, however not to as large of an extent as the local structural changes. In particular, there are specific sites where the IE of a guanine residue increases due to DNA–protein contacts in the NCP. Most of these sites are characterized by contact between the guanidinium group of an arginine residue and the minor groove surface of a GC base pair and are largely localized at locations where (i) the DNA minor groove is rotated in toward the histone octamer and (ii) the histone tails enter or exit the NCP. At a higher level of complexity, we note that the electrostatic effects due to the dynamics of the counterions³⁰ and water²⁸ molecules will be superimposed on the static picture above.

From a biological standpoint, the results of this investigation warrant a re-examination of the question, “What constitutes a DNA oxidation hot spot in vivo?” Prior to this work, it has been assumed that hot spots for DNA oxidative damage would occur primarily in guanine-rich regions, particularly in regions with poly-G tracts (e.g., GG and GGG). The large effects of local DNA structure on guanine IEs, coupled with the fact that a vast majority of genomic DNA in the eukaryotic nucleus is packaged into NCPs, and thus distorted away from normal B-form, requires that we look anew at this assumption. The idea that genomic oxidation hot spots could be dependent on local DNA structure (in particular on structural fluctuations of the neighboring base pairs) to an extent equal to, or in some cases greater than, DNA sequence is a novel concept and has important implications. One consequence is that the de novo

hunt for oxidative hot spots in vivo must rely not only on the human genomic sequence, but also a mapping of local chromatin structure and its remodeling during times of cell cycle progression and transcriptional activation.

Acknowledgment. We thank Chad Bjorklund for comments on this manuscript. Funding for this work was provided by the Spanish Ministerio de Educación y Ciencia (Project No. CTQ2005-04563 (A.A.V.)). This material is based upon work supported by the National Science Foundation (CAREER Award No. 0347370) and startup funds provided by Washington State University (W.B.D.).

Supporting Information Available: X3DNA-calculated values of the structural parameters for all 74 DNA base pairs and NDDO-G-calculated values of IE(G), IE(GC), and IE(triplet) for all guanines. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. *J. Phys. Chem.* **1996**, *100*, 5541.
- (2) Takada, T.; Kawai, K.; Tojo, S.; Majima, T. *J. Phys. Chem. B* **2003**, *107*, 14052.
- (3) Giese, B. *Acc. Chem. Res.* **2000**, *33*, 631.
- (4) Treadway, C. R.; Hill, M. G.; Barton, J. K. *Chem. Phys.* **2002**, *281*, 409.
- (5) Lewis, F. D.; Letsinger, R. L.; Wasielewski, M. R. *Acc. Chem. Res.* **2001**, *34*, 159.
- (6) Hess, S.; Gotz, M.; Davis, W. B.; Michel-Beyerle, M. E. *J. Am. Chem. Soc.* **2001**, *123*, 10046.
- (7) Beljonne, D.; Pourtois, G.; Ratner, M. A.; Bredas, J. L. *J. Am. Chem. Soc.* **2003**, *125*, 14510.
- (8) Neeley, W. L.; Essigmann, J. M. *Chem. Res. Toxicol.* **2006**, *19*, 491.
- (9) Giese, B.; Spichty, M. *ChemPhysChem* **2000**, *1*, 195.
- (10) Nunez, M. E.; Hall, D. B.; Barton, J. K. *Chem. Biol.* **1999**, *6*, 85.
- (11) Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8353.
- (12) Conwell, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8795.
- (13) Schuster, G. B. *Acc. Chem. Res.* **2000**, *33*, 253.
- (14) O'Neill, M. A.; Barton, J. K. *J. Am. Chem. Soc.* **2004**, *126*, 11471.
- (15) Valis, L.; Wang, Q.; Raytchev, M.; Buchvarov, I.; Wagenknecht, H.-A.; Fiebig, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10192.
- (16) Jortner, J.; Bixon, M.; Langenbacher, T.; Michel-Beyerle, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12759.
- (17) Berlin, Y. A.; Kurnikov, I. V.; Beratan, D.; Ratner, M. A.; Burin, A. L. DNA electron transfer processes: Some theoretical notions. In *Long-Range Charge Transfer in DNA II*; Schuster, G. B., Ed.; Topics in Current Chemistry 237; Springer: Berlin, 2004; pp 1.
- (18) Bixon, M.; Jortner, J. *J. Am. Chem. Soc.* **2001**, *123*, 12556.
- (19) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *J. Am. Chem. Soc.* **2001**, *123*, 260.
- (20) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *J. Phys. Chem. A* **2000**, *104*, 443.
- (21) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *Chem. Phys.* **2002**, *275*, 61.
- (22) Pullman, B.; Pullman, A. *Rev. Mod. Phys.* **1960**, *32*, 428.
- (23) Tong, G. S. M.; Kurnikov, I. V.; Beratan, D. N. *J. Phys. Chem. B* **2002**, *106*, 2381.
- (24) Senthikumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Lewis, F. D.; Berlin, Y. A.; Ratner, M. A.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14894.
- (25) Voityuk, A.; Jortner, J.; Bixon, M.; Rösch, N. *J. Chem. Phys.* **2001**, *114*, 5614.
- (26) Yoshioka, Y.; Kitagawa, Y.; Takano, Y.; Yamaguchi, K.; Nakamura, T.; Saito, I. *J. Am. Chem. Soc.* **1999**, *121*, 8712.
- (27) Voityuk, A.; Jortner, J.; Bixon, M.; Rösch, N. *Chem. Phys. Lett.* **2000**, *324*, 430.
- (28) Voityuk, A.; Siri Wong, K.; Rösch, N. *Angew. Chem., Int. Ed.* **2004**, *43*, 624.
- (29) Grozema, F. C.; Siebbeles, L. D. A.; Berlin, Y. A.; Ratner, M. A. *ChemPhysChem* **2002**, *3*, 536.
- (30) Barnett, R. N.; Cleveland, C. L.; Joy, A.; Landman, U.; Schuster, G. B. *Science* **2001**, *294*, 567.
- (31) Voityuk, A. *J. Chem. Phys.* **2005**, *122*, 204904.
- (32) Liu, C.-S.; Hernandez, R.; Schuster, G. B. *J. Am. Chem. Soc.* **2004**, *126*, 2877.
- (33) Tavernier, H. L.; Fayer, M. D. *J. Phys. Chem. B* **2000**, *104*, 11541.
- (34) Siri Wong, K.; Voityuk, A. A.; Newton, M. D.; Rösch, N. *J. Phys. Chem. B* **2003**, *107*, 2595.
- (35) LeBard, D. N.; Lilichenko, M.; Matyushov, D. V.; Berlin, Y. A.; Ratner, M. A. *J. Phys. Chem. B* **2003**, *107*, 14509.
- (36) Troisi, A.; Orlandi, G. *J. Phys. Chem. B* **2002**, *106*, 2093.
- (37) Voityuk, A. A.; Siri Wong, K.; Rösch, N. *Phys. Chem. Chem. Phys.* **2001**, *3*, 5421.
- (38) Saito, I.; Nakamura, T.; Nakatani, K.; Yoshioka, Y.; Yamaguchi, K.; Sugiyama, H. *J. Am. Chem. Soc.* **1998**, *120*, 12686.
- (39) Senthikumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2003**, *125*, 13658.
- (40) van Holde, K. E. *Chromatin*; Springer-Verlag: New York, 1989.
- (41) Davey, C. A.; Sargent, D. F.; Luger, K.; Maeder, A. W.; Richmond, T. J. *J. Mol. Biol.* **2002**, *319*, 1097.
- (42) Bjorklund, C. C.; Davis, W. B. *Nucleic Acids Res.* **2006**, *34*, 1836.
- (43) Nakatani, K.; Dohno, C.; Ogawa, A.; Saito, I. *Chem. Biol.* **2002**, *9*, 361.
- (44) Voityuk, A. A.; Zerner, M. C.; Rösch, N. *J. Phys. Chem. A* **1999**, *103*, 4553.
- (45) Nunez, M. E.; Noyes, K. T.; Barton, J. K. *Chem. Biol.* **2002**, *9*, 403.
- (46) Rajsiki, S. R.; Barton, J. K. *Biochemistry* **2001**, *40*, 5556.
- (47) Rajsiki, S. R.; Kumar, S.; Roberts, R. J.; Barton, J. K. *J. Am. Chem. Soc.* **1999**, *121*, 5615.
- (48) Boon, E. M.; Pope, M. A.; Williams, S. D.; David, S. S.; Barton, J. K. *Biochemistry* **2002**, *41*, 8464.
- (49) Cramer, T.; Krapf, S.; Koslowski, T. *Phys. Chem. Chem. Phys.* **2004**, *6*, 3160.
- (50) Lewis, F. D.; Liu, J. Q.; Zuo, X. B.; Hayes, R. T.; Wasielewski, M. R. *J. Am. Chem. Soc.* **2003**, *125*, 4850.
- (51) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5109.
- (52) Dickerson, R. E. *EMBO J.* **1989**, *8*, 1.
- (53) Clowney, L.; Jain, S. C.; Srinivasan, A. R.; Westbrook, J.; Olson, W. K.; Berman, H. W. *J. Am. Chem. Soc.* **1996**, *118*, 509.
- (54) Ausio, J.; Dong, F.; van Holde, K. E. *J. Mol. Biol.* **1989**, *206*, 451.
- (55) Hayes, J. J.; Clark, D. J.; Wolffe, A. P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6829.
- (56) Polach, K. J.; Lowary, P. T.; Widom, J. *J. Mol. Biol.* **2000**, *298*, 211.
- (57) Bohm, L.; Briand, G.; Sautiere, P.; Crane-Robinson, C. *Eur. J. Biochem.* **1981**, *119*, 67.
- (58) Bohm, L.; Briand, G.; Sautiere, P.; Crane-Robinson, C. *Eur. J. Biochem.* **1982**, *123*, 299.
- (59) Bohm, L.; Crane-Robinson, C.; Sautiere, P. *Eur. J. Biochem.* **1980**, *106*.
- (60) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (61) Yang, L.; Weerasinghe, S.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **1995**, *69*, 1519.
- (62) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books: Sausalito, CA, 2000.
- (63) Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67.
- (64) Caruso, T.; Carotenuto, M.; Vasca, E.; Peluso, A. *J. Am. Chem. Soc.* **2005**, *127*, 15040.
- (65) Prat, F.; Houk, K. N.; Foote, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 845.
- (66) Kurnikov, I. V.; Tong, G. S. M.; Madrid, M.; Beratan, D. N. *J. Phys. Chem. B* **2002**, *106*, 7.
- (67) Voityuk, A. *J. Phys. Chem. B* **2005**, *109*, 10793.
- (68) Antosiewicz, J.; Briggs, J. M.; Elcock, A. H.; Gilson, M. K.; McCammon, J. A. *J. Comput. Chem.* **1996**, *17*, 1633.
- (69) Blancafort, L.; Voityuk, A. A. *J. Phys. Chem. A* **2006**, *110*, 6426.
- (70) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.