DNA Duplex Stability: The Role of Preorganized Electrostatics

Urban Bren,[†] Jurij Lah,*[,‡] Matevž Bren,[§,⊥] Václav Martínek,[#,∥] and Jan Florián*[,∥]

National Institute of Chemistry, Hajdrihova 19, 1001 Ljubljana, Slovenia, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia, Institute of Mathematics, Physics and Mechanics, Jadranska 19, 1000 Ljubljana, Slovenia, and Department of Chemistry, Loyola University Chicago, Chicago, Illinois 60626, FOV and FVV, University of Maribor, Slomškov trg 15, 1000 Maribor, Slovenia

Received: July 7, 2009; Revised Manuscript Received: November 12, 2009

The insertion of a DNA base moiety at the end of a DNA duplex to form a Watson−Crick or wobble pair during DNA annealing or replication is a step of fundamental biological importance. Therefore, we investigated the energetics of a formation of the terminal G·C, G·T, and G·A base pairs in DNA containing a 5′-dangling G adjacent to the base insertion point using differential scanning calorimetry and computer simulations. The energies calculated along classical molecular dynamics trajectories in aqueous solution were analyzed in the framework of linear-response approximation (LRA) to obtain relative free energies for the base insertion and their electrostatic, van der Waals, and preorganization components. Using the generic set of LRA parameters, the calculated free energies disfavored the mispair formation by 2.5 (G·C → G·T) and 1.7 (G·C → G·A) kcal/mol, in reasonable agreement with the experimental free energy differences of 1.8 and 1.4 kcal/mol, respectively. The calculated preorganization components of these free energies of 0.6 (G·C → G·T) and −0.1 (G·C → G·A) kcal/mol show that electrostatic preorganization, which is an important source of DNA replication fidelity, plays a lesser role in the mispair destabilization in the absence of DNA polymerase.

1. Introduction

The detailed knowledge of factors determining the structure and stability of DNA in aqueous solution is important for understanding and control of the transcription, replication, recombination, and preservation of genetic information.[1] DNA stability can be quantitatively characterized by the free-energy change for the formation (folding) of DNA duplex from two single-stranded DNAs or, equivalently, by the free-energy change for the reverse process, called DNA melting or denaturation.

The stability of a DNA duplex with N base pairs can be decomposed into the free energy for the formation of the nucleus of the first M base pairs, $\Delta G_{\mathrm{nuc}} = \Delta G_{\mathrm{nuc},1} + \Delta G_{\mathrm{nuc},2} + ... + \Delta G_{\mathrm{nuc},M}$, plus the sum of $N - M$ contributions of inserting the base at the end of the growing duplex, $\Delta G_{\mathrm{ins},i}$, $i = 1, ..., N - M$. The magnitude of $\Delta G_{\mathrm{ins},i}$ which depends on the DNA sequence and ionic strength,[2] is determined by a delicate balance of the hydrogen bonding, stacking, torsional, and solvation free energies.[1,3–5]

The formation of the nucleus of the first $M = 3-5$ base pairs represents the rate-limiting step in DNA folding.[1] The nucleus promotes formation of adjacent base pairs so that $\Delta G_{\mathrm{ins},i} < 0$ and $\Delta G_{\mathrm{ins},i} \ll \Delta G_{\mathrm{nuc},j}$, which constitutes the cooperativity of the folding process.[6] Since the stacking of adjacent base pairs, being a short-range interaction,[7,8] is already present for $M = 2$, the cooperativity of DNA annealing cannot be directly attributed to DNA stacking interactions, but rather, to the coupling among base stacking, base pairing, and DNA solvation in the framework of the dynamics of a solvated short-length DNA. One possible contribution lies in the preorganized electrostatics, whereby already formed base pairs, deoxyribose rings and phosphodiester bridges have dipoles oriented in directions that promote formation of adjacent base pairs. Although this contribution is explicitly determined by evaluating charge−charge interactions in the studied system (see eq 5d in section 2.2.), the structural underpinning of electrostatic preorganization may include van der Waals interactions and the hydrophobic effect. The electrostatic preorganization concept,[9,10] which was initially advanced by likening an enzyme active site to a "supersolvent",[11] is the main factor in the catalytic rate-enhancement of many enzymes[9,12–14] and represents an important contribution to DNA replication fidelity achieved by DNA polymerases.[15,16]

To evaluate the contribution of electrostatic preorganization to DNA stability, we performed a series of molecular dynamics (MD) simulations to determine ΔG_{ins} using a linear-response approximation (LRA)/α method,[15,17] which combines the electrostatic and preorganization terms of Warshel's LRA[18,19] with the van der Waals energy component of the linear-interaction energy (LIE) method of Åqvist and co-workers.[20] We also assessed the reliability of our LRA/α calculations by determining the corresponding ΔG_{ins} values using differential scanning calorimetry (DSC). Although DSC experiments do not provide ΔG_{ins} directly, the experimental difference in melting free energies of duplex DNAs with and without the terminal base in one of its strands can be converted to ΔG_{ins} using a thermodynamic cycle (Figure 1). Abasic sites induce significant sequence-dependent reduction in thermal and thermodynamic stability of DNA[21–23] while retaining the overall B-DNA conformation.[24–26] To work with a consistent data set, experimental duplex stability parameters ΔG, ΔH, and ΔS need to be

* Corresponding author. E-mail: jfloria@luc.edu or jurij.lah@fkkt.uni-lj.si.
[†] National Institute of Chemistry.
[‡] University of Ljubljana.
[§] Institute of Mathematics, Physics and Mechanics.
[∥] Loyola University Chicago.
[⊥] University of Maribor.
[#] Permanent address: Department of Biochemistry, Faculty of Science, Charles University, Albertov 2030, 12840 Prague, Czech Republic.

10.1021/jp9064246
Published on Web 02/04/2010

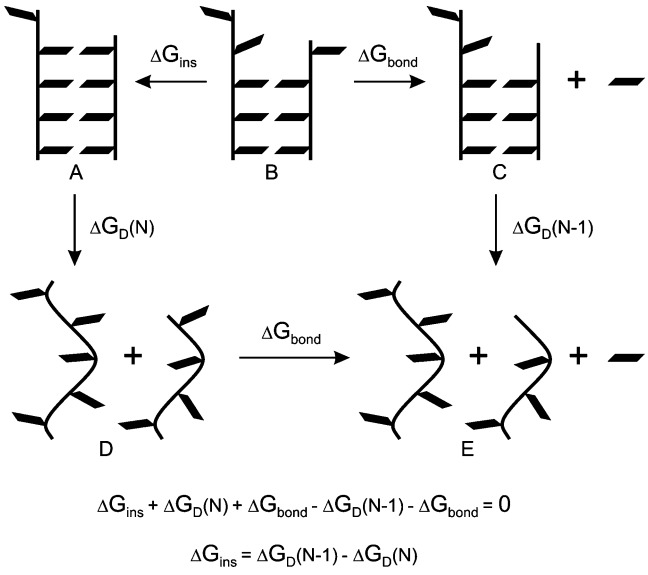

Figure 1. Thermodynamic cycle to determine a base insertion free energy (ΔG_{ins}) from the observed melting free energies for the full length DNA ($\Delta G_D(N)$) and DNA containing terminal abasic site ($\Delta G_D(N-1)$).

extrapolated to a common reference temperature by taking into account their temperature dependences reflected in changes of the heat capacity ΔC_P.[27–32] This extrapolation can be rigorously achieved only for thermodynamic data obtained by calorimetric measurements.[33]

In this study, we used DSC to monitor the denaturation of DNA duplexes with $^{5'G\mathbf{G}CGCATCAGC}_{3'\text{-}\mathbf{B}GCGTAGTCG}$ sequences. These sequences contained the terminal Watson−Crick (B = C) or wobble (B = T) base pairs, or an abasic site (B = 0), and included a 5′ dangling G nucleotide. The dangling nucleotide was included to model the product of the DNA polymerization reaction in the absence of the protein.[34] This model provides an important reference state for the proofreading selectivity of DNA polymerases.[35] We determined complete thermodynamic stability profiles and described the nature of the helix-to-coil transitions of these model oligonucleotides.

In the theoretical part of this study, the MD and LRA/α[15,19] calculations were applied to determine the relative stability of $^{5'G\mathbf{G}CGCATCAGC}_{3'\text{-}\mathbf{B}GCGTAGTCG}$ (where B = 0, C, A, or T) DNA duplexes. The LRA/α method has been employed because it allows for assessment of the influence of preorganized electrostatics on the DNA duplex stability. Additionally, the LRA/α can provide the electrostatic and van der Waals free-energy components for this stabilization.[36]

On the other hand, the LRA/α is an empirical method, the accuracy of which relies on the parameters α and β. These parameters may differ for structurally dissimilar ligands, for different host molecules, or for both.[37] In this light, it is important that using the generic LRA/α parameters optimized for calculations of solvation free energies of small polar molecules[15] revealed good agreement with ΔG_{ins} determined experimentally. Nevertheless, to further improve the accuracy of the LRA/α method and to extend its applicability to simulations of nucleic acids, we adjusted its empirical α and β parameters by minimizing the root-mean-square difference between the calculated and the experimental results.

2. Methods

2.1. Experimental Materials and Methods.

Materials. Oligonucleotides used in this work were synthesized and purified by Invitrogen (Germany). The concentrations of single-stranded oligonucleotides were determined by measuring absorbance at 260 nm and 25 °C (Cary Bio 100 UV spectrophotometer; Varian, Australia) using the extinction coefficients that had been calculated according to the nearest neighbor procedure.[38] We used the following extinction coefficients (M^{-1} cm^{-1}): 102 600 for 5′-GGCGCATCAGC-3′ and 89900 (B = C), 92 300 (B = T), and 83 800 (*B* = 0) for the complementary 5′-GCTGATGCG**B**-3′ strands. B = 0 represents the tetrahydrofuran analog of the abasic site. Biochemical studies show that DNA containing this tetrahydrofuran derivative retains its activity for DNA polymerases[39] and for apurinic/apyrimidinic endonucleases,[40] thus confirming its value as a model for recognition patterns associated with the natural abasic site. Buffer solutions (10 mM phosphate, [NaCl] = 100 mM, [EDTA] = 1 mM, adjusted to pH 7.0) containing equimolar amounts of the complementary single-stranded oligomers were mixed to obtain the corresponding $^{5'G\mathbf{G}CGCATCAGC}_{3'\text{-}\mathbf{B}GCGTAGTCG}$ duplex solutions. The 1:1 stoichiometry of formation was for all studied duplexes verified by Job plots.[41] The duplex solutions were extensively dialyzed against the appropriate buffer (see above) and degassed at room temperature (~25 °C) for 20 min before conducting any calorimetric experiments.

Differential Scanning Calorimetry. DSC measurements were performed with a Nano-II DSC scanning calorimeter (Calorimetry Sciences Corporation, Lindon, UT). For each duplex (~0.24 mM), the first DSC scan was performed at a heating rate of 0.25 °C min^{-1}. Since essentially the same thermograms were obtained by repetitive experiments at a heating rate of 1 °C min^{-1}, the observed transitions may be considered as equilibrium processes. The DSC thermograms are presented as $\Delta C_P = \bar{C}_{P,2} - \bar{C}_{P,AB}$ versus *T* curves ($\bar{C}_{P,2}$ = partial molar heat capacity of DNA obtained from the raw signal corrected for the buffer contribution and normalized per mole of duplex contained in the measuring cell; $\bar{C}_{P,AB}$ = partial molar heat capacity of the native (AB duplex) state extrapolated over the measured temperature range). To determine complete thermodynamic stability profiles of the model oligonucleotides, the DSC thermograms were analyzed in a model-independent way. In addition, model-dependent analysis of DSC data was performed to get a deeper insight into the nature of helix-to-coil transitions for the studied oligonucleotides.

Model-Independent Analysis. Heat capacity of denaturation, $\Delta C_{P,D}$, was determined as a difference in the pre- and posttransition baselines extrapolated to a reference temperature, which was for comparative purposes chosen as the melting temperature, T_m, obtained by model-dependent analysis of DSC data (see below). The enthalpy of denaturation at T_m, $\Delta H_{D,T_m}$, was calculated by integration of the area enclosed by the ΔC_P-versus-*T* curve, the pre- and posttransition baselines, and T_m. Assumptions that $\Delta C_{P,D}$ is a temperature-independent quantity and that $\Delta H_{D,T_m}$ and $\Delta C_{P,D}$ are not functions of DNA concentration and thus equal to their values in the standard state ($\Delta H_{D,T_m} = \Delta H_{D,T_m}^\circ$ and $\Delta C_{P,D} = \Delta C_{P,D}^\circ$) enabled the calculation of standard Gibbs free energy of denaturation, ΔG_D°, as a function of *T* from the integrated Gibbs−Helmholtz equation.

$$\Delta G_D^\circ = \Delta G_{D,T_m}^\circ(T/T_m) + \Delta H_{D,T_m}^\circ[1 - T/T_m] + \Delta C_{P,D}^\circ[T - T_m - T\ln(T/T_m)] \quad (1)$$

Since the observed helix-to-coil transition is an equilibrium process, one can express the standard Gibbs free energy of denaturation at T_m, $\Delta G_{D,T_m}^\circ$, that corresponds to the AB ↔ A + B reaction performed in the 1 M standard state as $\Delta G_{D,T_m}^\circ =$

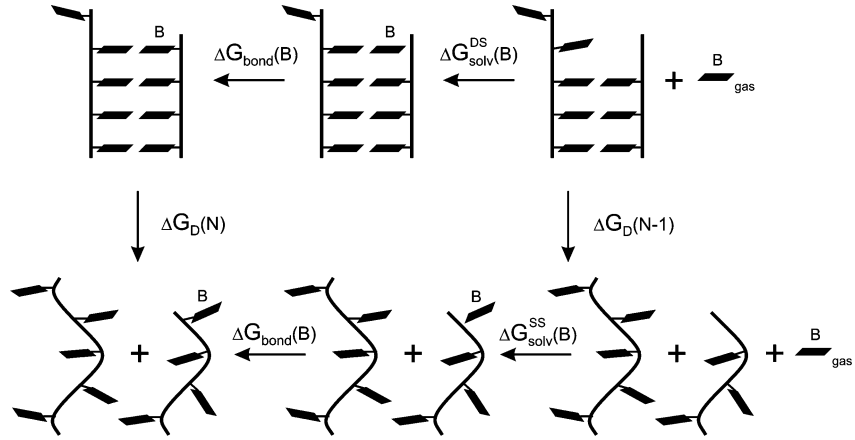

Figure 2. Thermodynamic cycle for the calculation of the free energy for the DNA duplex destabilization due to the creation of a terminal abasic site, $\Delta G_D(N) - \Delta G_D(N-1)$.

$-RT_m \ln(C_T/2)$, where C_T represents the total molar concentration of DNA expressed in moles of duplex per liter. The corresponding standard thermodynamic quantities (ΔG_D°, ΔH_D° and ΔS_D°) as functions of temperature were calculated from eq 1, from the Kirchhoff's law $\Delta H_D^\circ = \Delta H_{D,T_m}^\circ + \Delta C_{P,D}^\circ(T - T_m)$ and from the general relation $\Delta S_D^\circ = (\Delta H_D^\circ - \Delta G_D^\circ)/T$. The obtained thermodynamic parameters for denaturation of the model duplexes (B = C, T, 0) were used for calculation of the corresponding differential parameters $\Delta\Delta G_{B\to X}$, $\Delta\Delta H_{B\to X}$, $\Delta\Delta S_{B\to X}$, and $\Delta\Delta C_{P,B\to X}$ (X = C, T, 0). Since the first term of eq 1 cancels out when $\Delta\Delta$ values are calculated, these differential parameters depend on the choice of neither the standard state nor the type of chemical process and can be directly compared to the complementary theoretical data. Moreover, because these differential parameters depend on only ΔH_D and $\Delta C_{P,D}$ determined directly from experimental DSC data, they can be considered as model-independent quantities.

Model-Dependent Analysis. The monitored DNA denaturation can be described in the simplest way in terms of a reversible, two-state process in which a DNA molecule can exist either in the native duplex AB macro state or in the denatured, single-strand A, B macro state. Such a process can be presented as AB $\leftrightarrow$ A + B. According to this model, one can express partial molar enthalpy of DNA, $\bar{H}_2$, in terms of the corresponding contributions $\bar{H}_{AB}$, $\bar{H}_A$, and $\bar{H}_B$ that characterize species AB, A, and B as $\bar{H}_2 = \bar{H}_{AB} + \alpha_D(\bar{H}_A + \bar{H}_B - \bar{H}_{AB})$, where α_D represents the fraction of DNA molecules in the denatured state. The relative partial molar enthalpy, ΔH, can be expressed as[42]

$$\Delta H = \bar{H}_2 - \bar{H}_{AB} = \alpha_D(\bar{H}_A + \bar{H}_B - \bar{H}_{AB}) = \alpha_D \Delta H_D = \alpha_D \Delta H_D^\circ \quad (2)$$

where ΔH_D is defined as the enthalpy change accompanying the transition of 1 mol of the native AB duplex into the denatured single strands A and B. We considered ΔH_D to be independent of concentration and, thus, equal to its value in the standard state ($\Delta H_D = \Delta H_D^\circ$). The model function that describes the measured relative partial molar heat capacity of the DNA, ΔC_P, can be derived from the first partial derivative of eq 2 with respect to T at constant pressure.[27,28,33,43–46]

$$\Delta C_P = \alpha_D \Delta C_{P,D}^\circ + (\partial \alpha_D/\partial T)_P \Delta H_D^\circ \quad (3)$$

According to the AB $\leftrightarrow$ A + B model, α_D and $(\partial\alpha_D/\partial T)_P$ can be at any temperature expressed analytically in terms of parameters $\Delta H_{D,T_m}^\circ$, $\Delta C_{P,D}^\circ$ and T_m (temperature at which $\alpha_D = 0.5$) by combination of eq 1, equation $\Delta H_D^\circ = \Delta H_{D,T_m}^\circ + \Delta C_{P,D}^\circ(T - T_m)$, and the relation $\Delta G_D^\circ = -RT \ln[\alpha_D^2 C_T/(1 - \alpha_D)]$.[42] For this reason, the model-dependent values of $\Delta H_{D,T_m}^\circ$, $\Delta C_{P,D}^\circ$, and T_m can be obtained from the model function eq 3 fitting to the experimental DSC thermograms using the Levenberg−Marquardt nonlinear χ^2 regression procedure.[47]

2.2. LRA/α Calculations. The insertion free-energy, ΔG_{ins}, could be calculated by running long MD simulations on the DNA duplex (state A in Figure 1) and counting the number of occurrences of opening of the terminal base pair to form state B (Figure 1). However, such a direct approach would require sampling of extremely long MD trajectories. Perhaps more importantly, the free energy obtained in this way would lack information about its components, including the electrostatic preorganization. Thus, we determined ΔG_{ins} for the terminal G·B base pair by combining thermodynamic cycles in Figures 1 and 2 as

$$\Delta G_{ins}(B) = \Delta G_D(N-1) - \Delta G_D(N) = \Delta G_{solv}^{DS}(B) - \Delta G_{solv}^{SS}(B) \quad (4)$$

where $\Delta G_{solv}^{DS}(B)$ and $\Delta G_{solv}^{SS}(B)$ represent the solvation free energy of the base B in a hydrated double- or single-stranded environment. Magnitudes of $\Delta G_{solv}^{DS}(B)$ and $\Delta G_{solv}^{SS}(B)$ were calculated using the LRA/α method; the nucleobase B is in its terminology referred to as a probe region.[15]

The LRA/α solvation free energy of the nucleobase B in the solvent S (either hydrated DNA duplex, DS, or hydrated single stranded DNA, SS) is calculated as the sum of electrostatic ($\Delta G_{ES}^S(B)$), van der Waals ($\Delta G_{vdW}^S(B)$), and preorganization ($\Delta G_{PRE}^S(B)$) contributions,[15,19]

$$\Delta G_{solv}^S(B) = \Delta G_{ES}^S(B) + \Delta G_{vdW}^S(B) + \Delta G_{PRE}^S(B) \quad (5a)$$

$$\Delta G_{ES}^S(B) = \beta \langle U_{ES}^S(B) \rangle_Q \quad (5b)$$

$$\Delta G_{vdW}^S(B) = \alpha \langle U_{vdW}^S(B) \rangle_Q \quad (5c)$$

$$\Delta G_{PRE}^S(B) = \beta \langle U_{ES}^S(B) \rangle_0 \quad (5d)$$

where $\langle U_{ES}^S(B) \rangle_Q$ and $\langle U_{vdW}^S(B) \rangle_Q$ represent the average electrostatic and van der Waals interaction energy between the probe region and its surroundings calculated on a MD trajectory simulated using the fully charged probe. α and β denote the LRA/α empirical parameters, and $\langle U_{ES}^S(B) \rangle_0$ presents the average electrostatic interaction energy between the probe region and

its surroundings calculated on a MD trajectory (ensemble) generated using the uncharged probe (nucleobase B).

The linear response part of the LRA/α expression (i.e., the sum of eqs 5b and 5d for $\beta = 0.5$) can be derived by approximating the potential energy for a charging process in solution by the intersection of two parabolas of equal curvature[15,18−20] or by expanding the free-energy perturbation formula for two infinitesimally close states by the Taylor series.[20,49] The van der Waals term (eq 5c) represents an approximation to the solvation free energy of an uncharged probe[15,19] that was in the LRA/α method conveniently chosen to be the same as the corresponding term in the LIE method.[20] Thus, it is only the presence of the preorganization term (eq 5d) that distinguishes the LRA/α method from the linear interaction energy method of Åqvist and co-workers.[20] It has been argued that the preorganization term is not necessary in calculations of ligand−protein dissociation constants, since its effect can be simply represented by an increased value of the coefficient β,[50] but this conclusion may not be valid for all combinations of ligands and enzyme active sites. Indeed, the preorganization term did show an important contribution to binding fidelity of DNA polymerases β[15] and T7.[16]

The LRA/α method does not include an explicit van der Waals preorganization term because it is impossible to carry out MD simulations with the whole van der Waals term turned off. Simulations carried out with only the London term absent would, apart from their unphysical nature, likely bring no qualitatively new contributions that could not be obtained by simple readjustment of the coefficient α. Similarly, a more reasonable option of adding the van der Waals term sampled on the state with zero charges would likely not improve the performance of the method.[19] Nevertheless, steric and dispersion forces are implicitly included in the electrostatic preorganization via the generation of the statistical ensemble used for averaging electrostatic energies in eq 5d.

It should be noted that the LRA approximation, despite its acronym's implying a broad reach, represents a specialized version of the time-independent linear response theory:[51] specific to ligand, substrate, or transition state binding and electron transfer processes.[48,52] The linear response theory of hydration of ions[37,53–55] established that upon the process of solute (B) charging in water or other polar solvents, approximately one-half of the potential energy gained from the charge dipole interaction ($\langle U^{S}_{ES}(B)\rangle_{Q}$) is spent on reorienting the water dipoles.[9,56] For example, the water hydrogen bonding pattern has to be broken to form good hydrogen bonds with the charged solute. This finding is reflected in the value of 0.5 chosen for the LRA/α β parameter. The generic value of 0.56 (optimized for hydration free energies of imidazole, difluorotoluene, and protonated cytosine) was used for the empirical α parameter.[15]

Since both $\Delta G^{DS}_{solv}(B)$ and $\Delta G^{SS}_{solv}(B)$ terms can be decomposed into three contributions according to eq 5, also their difference representing ΔG_{ins} can be subjected to such decomposition; that is,

$$\Delta G_{ins}(B) = \Delta G^{ES}_{ins}(B) + \Delta G^{vdW}_{ins}(B) + \Delta G^{PRE}_{ins}(B) \quad (6a)$$

where

$$\Delta G^{ES}_{ins}(B) = \Delta G^{DS}_{ES}(B) - \Delta G^{SS}_{ES}(B) \quad (6b)$$

$$\Delta G^{vdW}_{ins}(B) = \Delta G^{DS}_{vdW}(B) - \Delta G^{SS}_{vdW}(B) \quad (6c)$$

$$\Delta G^{PRE}_{ins}(B) = \Delta G^{DS}_{PRE}(B) - \Delta G^{SS}_{PRE}(B) \quad (6d)$$

Hydrated single-stranded DNA is very flexible; thus, proper sampling of its configuration space would be computationally extremely demanding. To avoid this problem, a thermodynamic cycle was devised that approximates the nucleobase B embedded in the hydrated single-stranded DNA with the same base included in a hydrated nucleoside.[57] This reference system is denoted here dN. To retain the stacking interaction to the neighboring nucleobase, we used a hydrated dinucleoside monophosphate 3′-BG as an alternative reference system and denoted it as dNpN.

Free energy is a state function. Its difference is, therefore, dependent on only the choice of the initial and final states. Thus, the difference in free-energy of DNA duplex annealing contributed by the formation of terminal Watson−Crick (G•C) and wobble (G•X) base pairs can be calculated as[34]

$$\Delta\Delta G_{X\rightarrow C} = \Delta G_{ins}(C) - \Delta G_{ins}(X) \quad (7a)$$

and analogous expressions can be formulated for its electrostatic, van der Waals, and preorganization components.

$$\Delta\Delta G^{ES}_{X\rightarrow C} = \Delta G^{ES}_{ins}(C) - \Delta G^{ES}_{ins}(X) \quad (7b)$$

$$\Delta\Delta G^{vdW}_{X\rightarrow C} = \Delta G^{vdW}_{ins}(C) - \Delta G^{vdW}_{ins}(X) \quad (7c)$$

$$\Delta\Delta G^{PRE}_{X\rightarrow C} = \Delta G^{PRE}_{ins}(C) - \Delta G^{PRE}_{ins}(X) \quad (7d)$$

2.3. Simulation Conditions. The configurational ensembles for the evaluation of free energies were generated from molecular dynamics (MD) trajectories using the AMBER[58] force field implemented in the program Q.[59] The simulated solute molecules were immersed in a sphere (24 Å radius) of TIP3P water molecules subjected to the surface-constraint, all-atom solvent (SCAAS)-type boundary conditions,[60,61] as implemented in the program Q.[59] These constraints were designed to mimic infinite aqueous solution. Thus, the spherical models (water droplets) do not require large dimensions of the simulated systems (in comparison to simulation boxes utilizing periodic boundary conditions) to produce stable and realistic free energies.

The simulations of the DNA duplexes were initiated using DNA geometry from the X-ray structure of a ternary complex of DNA polymerase β (pdb code 1BPY; the structure is predominantly in the B-form, which is retained during our simulations),[62] which was modified by removing all the protein, dCTP substrate, water, and metal atoms. We also removed all DNA atoms upstream from guanine nucleotide that forms a Watson−Crick base pair with dCTP and, for simulations of the mispairs, manually mutated the 3′ terminal base of the primer strand to T or A. These modifications resulted in the $^{5'\mathbf{G}\mathbf{G}CGCATCAGC}_{3'\text{-}\mathbf{B}GCGTAGTCG}$ DNA sequence, where B stands for the C, T, or A nucleotides (Figure 3). DNA atoms protruding beyond the sphere boundaries were restrained to their coordinates in the 1BPY crystal structure using harmonic restraints. Nonbonding interactions of these atoms were turned off. Nonbonding interactions between other atoms inside the simulation sphere

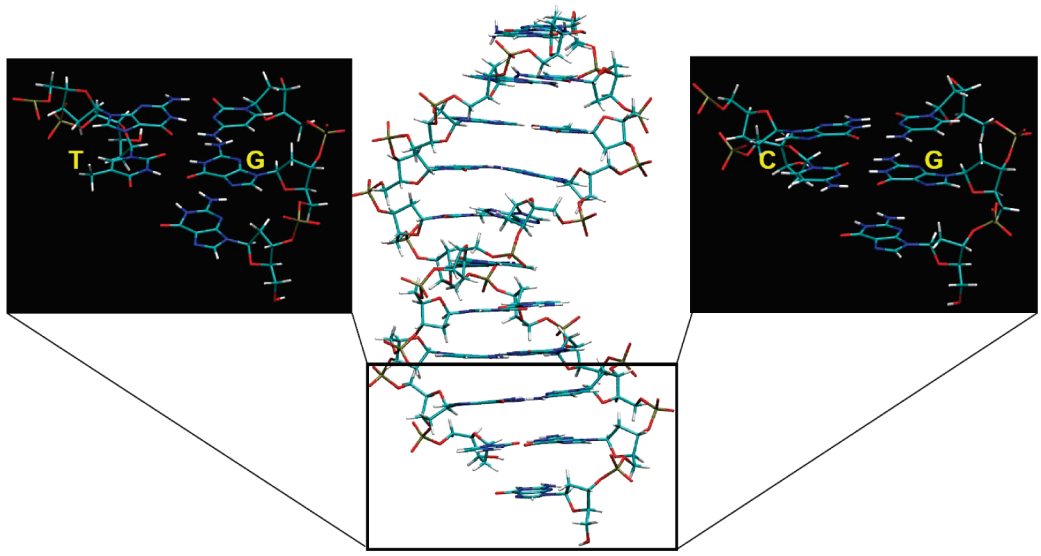

Figure 3. Representative structures of the model systems used to calculate DNA melting free energy for T·G (mispair) vs C·G. Explicit water and Na^+ counterions are not shown. During the production MD trajectories, the 5′-dangling guanosine was observed to sample both the stacked and unstacked configurations, while the DNA remained in the B-DNA form.

were subjected to a 10 Å cutoff. The local-reaction field method was used to treat long-range electrostatic interactions for distances beyond this cutoff.[48,61] Nonbonding interactions of atoms forming the probe region, which was defined as a nucleobase of B, were explicitly evaluated for all distances.

The structure of all simulated systems was relaxed in a series of four 10 000-and one 20 000-step MD simulations at 5 K with increasing stepsize from 0.01 to 1 fs. Subsequent gradual heating of the system from 5 to 250 K in a series of five 10 000-step MD simulations with 1 fs stepsize was followed by a 40 ps simulation at 298 K using a 2 fs stepsize. All production trajectories, the lengths of which were 4 ns, generated constant-temperature ensembles. The SHAKE algorithm[63] was used for bonds involving hydrogen atoms. The structure and trajectory analyses were carried out using the program VMD 1.8.3.[64] Energies were sampled every 10 steps. The reported free energies and standard deviations reflect an average over four independent simulations that differed in their starting geometries; that is, the total simulation length of 16 ns.

Sodium cations were added to the simulated system to achieve its electroneutrality. Positions of all Na^+ atoms were restrained by a flat-bottom harmonic potential (force constant of 50 kcal mol^{-1} $Å^{-2}$) that was zero for the distances less than 20 Å from the center of the simulation sphere. These potentials were applied to prevent the diffusion of the sodium ions toward the edge of the simulation sphere. No additional restraints were applied. The center of the simulation sphere was placed in the center of the studied base pair. This selection enclosed the studied base pair plus another five base pairs downstream inside the simulation sphere. Phosphate groups that were closer to the sphere boundary than 3 Å were made electroneutral by decreasing the negative charge on the oxygen atoms, as described previously.[57]

dNpN reference systems for simulations in aqueous solution consisted of 3′-CG, 3′-TG, and 3′-AG dideoxyribonucleoside monophosphate monoanions. Each of these dinucleoside monophosphates contained 3′OH and 5′OH terminal groups and one negatively charged phosphate group neutralized by Na^+. Simulations in water were also carried out for cytidine, thymidine, and adenosine deoxyribonucleosides (dN). In all simulations of dNpN or dN, the C1′ atom of B was constrained in the center of the simulation sphere using a harmonic potential defined by a force constant of 50 kcal mol^{-1} $Å^{-2}$ to prevent its diffusion toward the edge of the simulation sphere.

3. Results and Discussion

Helix-to-Coil Transitions of the Oligonucleotides. The position of the peaks of the DSC thermograms accompanying denaturation of the DNA duplexes (Figure 4) shows that the duplex with a terminal Watson−Crick base pair (B = C) exhibits the highest thermal stability (highest T_m). The T_m's for a mismatched duplex (B = T) and a duplex containing the abasic lesion (B = 0) are similar, but significantly lower than the T_m of the B = C duplex. The two state model (eqs 2 and 3) describes very well the transitions of B = C and B = T duplexes (Figure 4). In the case of B = 0, we observed a slight discrepancy between the model and experimental data.

Alternatively, the nature of a duplex denaturation process can be examined by comparing a model-dependent $\Delta H_{D,T_m}^{\circ}$ (van't Hoff enthalpy obtained from the model function fitting) with the corresponding model-independent enthalpy. Their ratio is close to unity if the observed transition is an equilibrium, two-state process (AB ↔ A + B). On the other hand, the value of the ratio <1 signifies for the presence of intermediates occurring between pure duplex and pure single-stranded DNA populations, whereas the ratio >1 is indicative of a process involving double-stranded aggregates. In the case of the model oligonucleotides, the $\Delta H_{D,T_m}^{\circ}$ (model-dependent)/$\Delta H_{D,T_m}^{\circ}$ (model-independent) ratio is equal to 1.0 (B = C, B = T), and 1.1 (B = 0). These ratios and a good agreement between the two-state model and experiment (Figure 4) indicate that B = C duplex denaturation can be described as a cooperative process, the nature of which is not changed by the C → T mutation, and is not significantly altered by the introduction of the abasic C → 0 lesion. These findings are supported by good agreement of the two-state model with the B = C, B = T, and B = 0 UV-melting curves (not shown), which results in about the same $\Delta H_{D,T_m}^{\circ}$ values as those obtained by model-dependent analysis of DSC data. On the basis of the analysis of thermodynamic and structural data for DNA containing abasic sites by Breslauer and co-workers,[22] the observed (two-state) nature of the B = C, B = T, and B = 0 denaturations suggests that the phosphodiester backbone rather

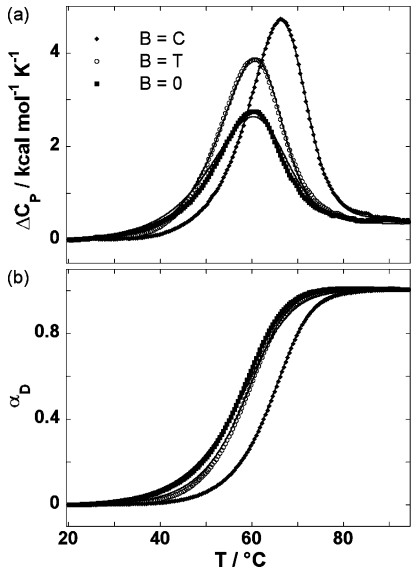

Figure 4. Experimentally monitored helix-to-coil transitions of the oligonucleotides $^{5'GGCGCATCAGC}_{3'\text{-}BGCGTAGTCG}$ (B = C, T, 0). Panel a: DSC thermograms (symbols, every fourth experimental point) and the best fits of the model function (lines, see eq 3). The thermodynamic parameters characterizing the transitions monitored by DSC are presented for B = C, T, and 0, respectively: T_m (°C) = 64.4, 58.6, 57.6; $\Delta H_{D,T_m}$ (kcal mol^{-1}) = 76.1 (+0.1), 67.2 (+1.2), 47.7 (+6.4); $\Delta C_{P,D}$ (kcal mol^{-1} K^{-1}) = 0.43 (+0.02), 0.41 (+0.00), 0.44 (−0.10). The values in the round brackets are the differences between the model-dependent eq 3 and the presented model-independent thermodynamic parameters. Panel b: Fraction of the oligonuceotides in the denatured state, α_D, calculated either directly from DSC thermograms (panel a) as a model-independent $\Delta H/\Delta H_D$ ratio (symbols: ΔH was obtained by integration of experimental ΔC_P from the reference temperature at which all DNA molecules exist in the initial (duplex) macrostate ($\alpha_D = 0$) to a given T, at which a certain fraction of DNA molecules exist in the final denatured (single-stranded) macrostate; ΔH_D was calculated as $\Delta H_D = \Delta H_{D,T_m} + \Delta C_{P,D}(T - T_m)$, see Experimental Materials and Methods) or as the model-dependent quantity obtained from the best fit model parameters (lines).

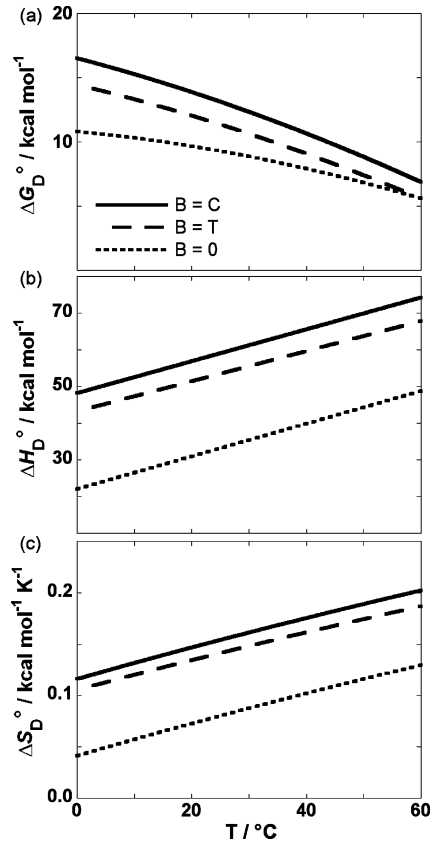

Figure 5. Thermodynamic profiles of helix-to-coil transitions of the oligonucleotides $^{5'GGCGCATCAGC}_{3'\text{-}BGCGTAGTCG}$ (B = C, T, or 0) obtained by the model-independent analysis of DSC thermograms.

than the base−sugar network serves as the primary propagation path for the communication between the cooperative DNA melting units.

Thermodynamic Stability of the Duplexes. The values of $\Delta G_D°$, $\Delta H_D°$, and $\Delta S_D°$ are strongly temperature-dependent (Figure 5). In a temperature range of 0−60 ^{0}C, B = C is the most stable duplex (the highest $\Delta G_D°$ value, Figure 5a), followed by the mismatched B = T duplex and the B = 0 duplex containing the abasic lesion. The observation that folding of all studied duplexes in the standard state at physiological temperatures is an enthalpy-driven process accompanied by an unfavorable entropy contribution and negative heat capacity change is a general feature of polymeric and oligomeric nucleic acids. In this light, $\Delta H_D° > 0$ observed at physiological temperatures can be qualitatively explained by more energetically favorable (stacking and H-bonding) interactions formed in the duplex state, in comparison to the single-stranded thermally denatured state. The corresponding $\Delta S_D° > 0$ can be explained by an increase in conformational, translational, and rotational freedom of DNA upon denaturation that is accompanied by a release of a significant amount of counterions. On the other hand, $\Delta C_{P,D}° > 0$ is probably mainly a consequence of exposure of aromatic surfaces to water upon denaturation,[65] although this contribution to $\Delta C_{P,D}°$ is opposed by the simultaneous exposure of polar surface area (e.g., Watson−Crick face of bases). Although hydration clearly plays a major role in determining $\Delta C_{P,D}°$ values of DNA denaturation, its exact origins still remain a matter of debate.[65] In this light, statistical mechanical treatments of explicitly simulated hydration may represent a promising way to the molecular understanding of $\Delta C_{P,D}°$ values accompanying denaturation of nucleic acids and proteins.[66] The average value of $\Delta C_{P,D}°$ for denaturation of B = C, B = T, and B = 0 is 0.43 (±0.03) kcal mol^{-1} K^{-1} (or 0.043 (±0.003) kcal (mol of base pairs)$^{-1}$ K^{-1}), which is consistent with the average $\Delta C_{P,D}°$ value accompanying the denaturation of oligomeric duplexes obtained by DSC[29] and somewhat lower than the average $\Delta C_{P,D}°$ observed for helix-to-coil transitions of polymeric nucleic acids.[32]

The contributions to the thermodynamic stability of the model duplexes can be discussed also in terms of the corresponding differential parameters $\Delta\Delta G_{B\to X}$, $\Delta\Delta H_{B\to X}$, $\Delta\Delta S_{B\to X}$, and $\Delta\Delta C_{P,B\to X}$ (X = C, T, 0) extrapolated to 25 °C using the equations described in the Experimental Materials and Methods section (DSC model-independent analysis; see also Figures 4 and 5). The $\Delta\Delta G_{0\to C}$, $\Delta\Delta H_{0\to C}$, and $\Delta\Delta S_{0\to C}$ values (Table 1) for contributions of G·C base pair formation to DNA duplex annealing are in accordance with the duplex stabilizing effects presented by Vesnaver et al.[22] ($\Delta\Delta G_{0\to T} = -6.5$ kcal mol^{-1}, $\Delta\Delta H_{0\to T} = -27.2$ kcal mol^{-1}, and $\Delta\Delta S_{0\to T} = -69.2$ cal mol^{-1} K^{-1}). Since, in their study, the A·T base pair was formed, the gain of two hydrogen bonds should result in a less negative $\Delta\Delta G$ value than obtained in our study (gain of three hydrogen bonds). This apparent discrepancy can be explained by considering differences in stacking interactions. That is, since their A·T base pair is located in the middle of the duplex, its two stacking interactions lower $\Delta\Delta G$ and $\Delta\Delta H$ values more than one stacking interaction in the terminal G·C base pair. Vesnaver et al.[22] also suggested that base pair formations exert stabilizing influences that cannot be explained by simple nearest-neighbor

TABLE 1: Differential Thermodynamic Parameters for $^{5'GGCGCATCAGC}_{3'-BGCGTAGTCG}$ and $^{5'GGCGCATCAGC}_{3'-XGCGTAGTCG}$, X = C or T, Helix-to-Coil Transitions Obtained from Model-Independent Analysis of DSC Data at 25 °C

	B = 0, X = C	B = 0, X = T	B = T, X = C
$\Delta\Delta G_{B\rightarrow X}$/kcal mol^{-1}	−3.9 ± 0.8	−2.1 ± 0.5	−1.8 ± 0.5
$\Delta\Delta H_{B\rightarrow X}$/kcal mol^{-1}	−25.9 ± 3.2	−20.3 ± 2.5	−5.6 ± 2.0
$\Delta\Delta S_{B\rightarrow X}$/cal mol^{-1} K^{-1}	−73.9 ± 11.1	−61.1 ± 8.6	−12.8 ± 6.9
$\Delta\Delta C_{P,B\rightarrow X}$/kcal mol^{-1} K^{-1}	0.01 ± 0.08	0.03 ± 0.06	−0.02 ± 0.05

TABLE 2: Average[a] Electrostatic, van der Waals, and Preorganization Interaction Energies between the Probe Region and Its Surrounding 'Solvent', and the Resulting Solvation Free Energies[b]

		interaction energy (kcal/mol)			
solvent[c]	probe (B)	$\langle U^S_{ES}(B)\rangle_Q$	$\langle U^S_{vdW}(B)\rangle_Q$	$\langle U^S_{ES}(B)\rangle_0$	ΔG_{solv} (kcal/mol)
DS	C	−49.6 ± 0.2	−15.7 ± 0.2	0.5 ± 0.2	−33.3 ± 0.3
	T	−36.9 ± 0.2	−18.6 ± 0.1	−2.3 ± 0.5	−30.0 ± 0.4
	A	−36.2 ± 0.2	−18.6 ± 0.1	−1.0 ± 0.5	−29.0 ± 0.4
dN	C	−45.0 ± 0.1	−7.1 ± 0.1	1.5 ± 0.2	−25.7 ± 0.2
	T	−36.0 ± 0.1	−9.4 ± 0.1	−1.4 ± 0.2	−24.0 ± 0.2
	A	−33.6 ± 0.1	−11.1 ± 0.1	−0.1 ± 0.2	−23.1 ± 0.2
dNpN	C	−45.0 ± 0.2	−11.1 ± 0.4	1.5 ± 0.2	−28.0 ± 0.4
	T	−35.1 ± 0.1	−14.8 ± 0.1	−2.4 ± 0.1	−27.0 ± 0.2
	A	−32.7 ± 0.1	−16.0 ± 0.1	0.2 ± 0.5	−25.2 ± 0.4

[a] Average and standard deviation from four independent simulations. [b] Equation 5, $\alpha = 0.56$, $\beta = 0.5$. [c] DNA duplex (DS), deoxynucleoside (dN), or dinucleoside monophosphate (dNpN) in aqueous solution.

TABLE 3: The Calculated Insertion Free Energies and Their Components

		free energy (kcal/mol)[c]				
SS model[a]	B[b]	ΔG^{ES}_{ins} (B)	ΔG^{vdW}_{ins} (B)	ΔG^{PRE}_{ins} (B)	ΔG_{ins}(B)	experiment[d]
dN	C	−2.3 ± 0.1	−4.8 ± 0.1	−0.5 ± 0.2	−7.6 ± 0.4	−3.9 ± 0.8
	T	−0.5 ± 0.1	−5.2 ± 0.1	−0.4 ± 0.3	−6.1 ± 0.5	−2.1 ± 0.5
	A	−1.3 ± 0.1	−4.2 ± 0.1	−0.5 ± 0.3	−6.0 ± 0.5	
dNpN	C	−2.3 ± 0.2	−2.6 ± 0.3	−0.5 ± 0.2	−5.4 ± 0.7	−3.9 ± 0.8
	T	−0.9 ± 0.1	−2.1 ± 0.1	0.1 ± 0.3	−2.9 ± 0.5	−2.1 ± 0.5
	A	−1.7 ± 0.1	−1.4 ± 0.1	−0.6 ± 0.5	−3.7 ± 0.7	

[a] Model used to calculate ΔG^{SS}_{solv}(B) (see Methods). [b] Inserted nucleobase (Figure 2). [c] Equation 6, $\alpha = 0.56$, $\beta = 0.5$. [d] DSC (Table 1).

effects and may be caused by changes in solvation. In this light, the value of $\Delta\Delta C_{P,0\rightarrow C} \sim 0$ estimated in our work (Table 1) may indicate that the formation of a terminal G·C base pair does not significantly alter the duplex solvation.

Energetic Contributions to DNA Duplex Stability. The calculated average electrostatic, vdW, and preorganization interaction energies of the C, T, or A base moieties (denoted as "probes") in solvated nucleoside, dinucleoside monophosphate, or DNA environments are presented in Table 2. Regardless of the composition of the environment, the interaction energy of each probe is dominated by its electrostatic component, followed by the van der Waals and preorganization energies. The electrostatic interaction energy of cytosine is always more favorable, in agreement with its larger polarity.[4] On the other hand, van der Waals interactions, which are dominated by the dispersion attraction, favor thymine and adenine due to their larger polarizability.[67]

The average preorganization energies, which reflect the interaction of the probe with its preorganized environment (i.e., environment with the geometry and polarization established in the presence of zero charges on all probe atoms), are presented in the last column of Table 2. If an uncharged solute B is immersed in water, the solvent dipoles are with respect to it randomly oriented. If the water and the solute geometries would remain unchanged during the charging process of the solute (i.e., if this process was instantaneous), the overall charge dipole interaction energy would be zero ($\langle U^S_{ES}(B)\rangle_0 = 0$). Thus, we can say that polar solvents are not electrostatically preorganized to accommodate solute molecules. On the other hand, in proteins, the binding site dipoles associated with polar groups and ionized residues may be already partially oriented toward the bound uncharged ligand B. The analogous instantaneous charging of the bound ligand may therefore yield favorable electrostatic interactions between the protein and its ligand ($\langle U^S_{ES}(B)\rangle_0 < 0$), in which case one can claim that proteins are electrostatically preorganized to accommodate their ligands. Here, although the preorganization concept has been most frequently invoked for transition states of enzymatic reactions, the term "ligand" may represent a substrate, transition state, or an inhibitor. The electrostatic preorganization represents the main source of the catalytic power of many enzymes,[52] although other factors may sometimes play an important role.[68] Its origin can be traced to the process of protein folding, whereby the protein attains its 3D structure and its dipoles get oriented.[12,13]

The calculated preorganization contributions are negligible compared to the sum of electrostatic and van der Waals energies (Table 2). A positive preorganization energy obtained for C can be attributed to the base−deoxyribose intramolecular interactions because it has the same magnitude (1.5 kcal/mol) in both the solvated nucleoside and dinucleoside monophosphate.

Using the nucleoside model of single-stranded DNA, we obtained insertion free energies (Table 3) that were significantly more negative than experimental ΔG_{ins}. This discrepancy is likely due to the absence of the base-stacking and base−phosphate interactions in the nucleoside model. Indeed, the agreement improved for the dinucleoside monophosphate model, although C and T bases remained too stable in the duplex environment. Even though this overstabilization may still reflect the simplified

nature of the reference system, it is more likely due to our use of generic LRA/α parameters, which were not optimized for the DNA environment.

Although it is, strictly speaking, not possible to decompose free energies into additive contributions from individual groups of atoms or force-field terms, LRA/α approximation allows one to carry out such a decomposition with high rigor. The van der Waals term is a major contributor to the calculated ΔG_{ins}, pointing to dispersion attraction between stacked DNA bases as important source of the duplex stability. This conclusion is consistent with the results of previous computational studies, including the decomposition of the free energy change for the mutation of a terminal base of a DNA duplex to "nothing"[34] and the loss of the helical structure of DNA duplex in simulations missing the London term.[69] The electrostatic term also stabilizes the DNA duplex, but more so when a correct Watson−Crick base pair is formed. The preorganization term is a negligible contributor of the stability of the duplex containing the terminal G·T base pair, but contributes −0.5 kcal/mol; that is, ∼10% of the free energy of inserting C or A opposite G. However, even in this most favorable case, the magnitude of electrostatic preorganization is too small to account for the cooperativity of DNA annealing, which amounts to −2 to −3 kcal/mol per base-pair step.[70]

The van der Waals components of ΔG_{ins} show a large dependence on the reference state (Table 3). For example, the insertion of C opposite G yields a ΔG_{ins}^{vdW} of −4.8 and −2.6 kcal/mol for the dN and dNpN reference states, respectively. Because the calculated electrostatic and preorganization components of ΔG_{ins} are independent of the reference state, the model-dependent differences in ΔG_{ins} mirror those in the van der Waals term. Since there are no stacking interactions in the dN model, these differences indicate that stacking interactions in the 5′-GC, GT, and GA dinucleoside monophosphates are −2.2, −3.1, and −2.8 kcal/mol, respectively. Of course, accounting for stacking in the reference system is fundamentally important, which makes the results calculated with the dNpN model more reliable.

The order of stacking free energies is consistent with the differences in the structural flexibility of the dinucleoside monophosphates obtained by the analysis of histograms of N9−N1 or N9−N9 distances (cf. Figures 1S, 3S, and 6S of the Supporting Information). Using these histograms, stacking free energies can be estimated by counting the number of configurations with N9−N1 (or N9−N9) distances below and above a reasonable cutoff that separates stacked and unstacked conformations.[71] Using a 6 Å cutoff, the stacking free energies of GC, GT, and GA dinucleoside monophosphates were −1.0, −2.3, and −1.3 kcal/mol, respectively. Interestingly, the unstacked configurations of GT mostly correspond to T-shaped complexes (Figure 4S). The histogram analysis also showed that uncharging one of the bases does not significantly affect the stacking propensity of dinucleoside monophosphates (Figures 2S, 5S, and 7S of the Supporting Information).

Since the differences in ΔG_{ins}^{vdW} for dN and dNpN reference states are larger (in absolute value) than stacking free energies determined directly by counting stacked and unstacked configurations, the coefficient α used for scaling vdW interactions in the LRA/α approximation should be decreased to better represent stacking interactions. The least-squares fit of LRA/α parameters on experimental data provided the lowest rmsd for $\beta = 0.34$ and $\alpha = 0.43$ (Figure 6). Our refined LRA/α parameter $\beta = 0.34$ is consistent with the value of this parameter ($\beta = 0.43$) recommended for use in LIE calculations of protein−ligand binding free energies[50] and DNA destabilization by base substitutions[34] because a larger β compensates for the lack of the electrostatic preorganization term in the LIE method. In addition, the refined α parameter of 0.43 decreases the magnitude of ΔG_{ins}^{vdW} so that the LRA/α stacking free energy estimates for dinucleoside monophosphates become closer to those obtained by direct counting of stacked and unstacked configurations or by quantum mechanical calculations combined with the Langevin dipoles solvation model.[4]

Although the refined LRA/α parameters are important to properly describe stacking interactions, their effect on the magnitude of the preorganization contributions to base insertion free energies is within the error bars of the values presented in Table 3, which were obtained with the generic LRA/α parameters. Similarly, the error cancellation renders the use of perfect LRA/α parameters less critical for obtaining reasonably accurate predictions of the relative insertion free energies (eq 7, Table 4). Thus, the overall DNA duplex stabilization free energies and their components presented in Table 4 are also relatively unaffected by the choice of LRA/α parameters or models describing the single-stranded DNA.

The relative insertion free energies presented in Table 4 represent the magnitude of DNA stabilization due to replacing

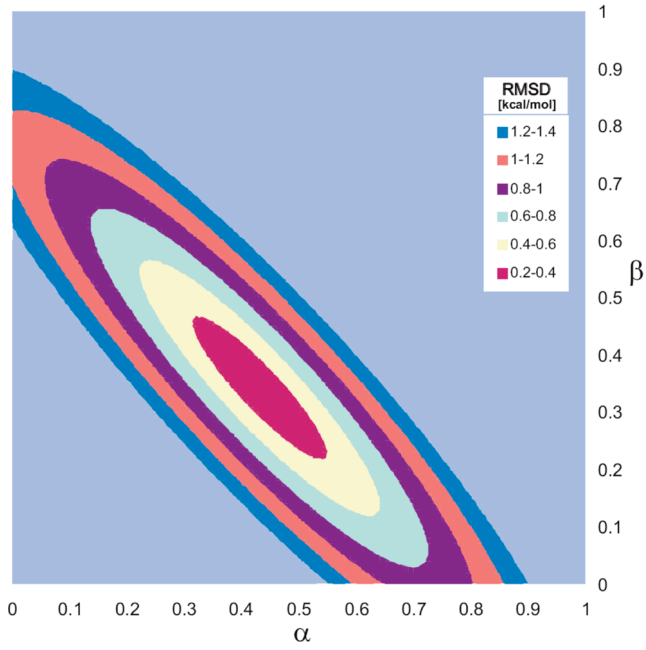

Figure 6. Root mean-square deviation of the calculated and experimental insertion, ΔG_{ins}(B), as well as $\Delta\Delta G_{A\rightarrow C}$ and $\Delta\Delta G_{T\rightarrow C}$ DNA duplex stabilization free energies as a function of the LRA/α empirical parameters α and β. The calculated data were obtained using the dNpN reference state.

TABLE 4: The calculated DNA Duplex Stabilization Free Energies and Their Components

SS model[a]	X[b]	free energy (kcal/mol)[c]				
		$\Delta\Delta G_{X\rightarrow C}^{ES}$	$\Delta\Delta G_{X\rightarrow C}^{vdW}$	$\Delta\Delta G_{X\rightarrow C}^{PRE}$	$\Delta\Delta G_{X\rightarrow C}$	experiment
dN	T	−1.8	0.4	−0.1	−1.5	−1.8[d], −2.0[e]
	A	−1.0	−0.6	0.0	−1.6	−1.4[e]
dNpN	T	−1.4	−0.5	−0.6	−2.5	−1.8[d], −2.0[e]
	A	−0.6	−1.2	0.1	−1.7	−1.4[e]

[a] Model used to calculate ΔG_{solv}^{SS}(B) (see Methods). [b] Probe in the mispair. [c] $\Delta\Delta G_{X\rightarrow C}$, eq 7, $\alpha = 0.56$, $\beta = 0.5$. [d] DSC, this work. [e] Determined using nearest-neighbor empirical parameters[2] for terminal DNA mismatches in 0.1 M NaCl and at 25 °C via program HYTHER.[74]

wobble G·T or G·A base pairs by a Watson−Crick G·C base pair at the end of DNA a double helix. This stabilization is of fundamental importance for assessing the intrinsic contribution of DNA polymerases to both their dNTP insertion fidelity and proofreading proficiency. In the case of the G·dCTP → G·dATP mismatch, the electrostatic preorganization was calculated to contribute 2.7 kcal/mol of the total 4.6 kcal/mol binding contribution to the fidelity of DNA replication by T7 DNA polymerase.[72] Similarly, earlier LRA/α binding studies of fidelity of DNA polymerase β have shown significant preorganization effects,[15] which were attributed by the free energy decomposition to the interactions between the templating base and the base of the incoming dNTP substrate. However, since these studies did not consider dNTP binding in the absence of the protein, it was not possible to convincingly attribute this preorganization to the steric and electrostatic restrictions placed upon the positioning of the template and dNTP bases by the enzyme active site. In fact, it could be argued that the template and dNTP base positioning could originate from stacking interactions with the DNA, the restrictions imposed by the presence of the phosphodiester backbone, or by strong binding of the triphosphate moiety. These possibilities have been excluded by the present calculations, which show that in the absence of the polymerase, the preorganization component of the G·C → G·A mismatch destabilization actually favors the mismatch by 0.1 kcal/mol (Table 4). Although the preorganization term does destabilize the G·T mismatch, the amount of this destabilization is relatively small in the absence of DNA polymerase. These results reinforce the notion that the electrostatic preorganization of the DNA polymerase active site represents an important contribution to DNA replication fidelity.[17,73]

Acknowledgment. This work was supported by the National Institutes of Health Grant 1U19CA10501 and by the Slovenian Research Agency through Grants Nos. P1-0201, P1-0002, P1-0294, J1-6653, and Z1-2000 as well as by the WFS scholarship. V.M. would like to thank Professor Marie Stiborová (Charles University, Prague, Czech Republic) for her continuous support over many years.

Supporting Information Available: Histograms of N9−N1 distances in solvated dinucleoside monophosphates and a T-shaped geometry of the GT dinucleoside monophosphate. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books: Sausalito, CA, 2000.
(2) SantaLucia, J., Jr. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415.
(3) Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **2006**, *34*, 564.
(4) Florián, J.; Sponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884.
(5) Sponer, J.; Florián, J.; Ng, H. L.; Sponer, J. E.; Spackova, N. *Nucleic Acids Res.* **2000**, *28*, 4893.
(6) Poland, D. *Cooperative Equilibria in Physical Biochemistry*; Clarendon Press: Oxford, UK, 1978.
(7) Luo, R.; Gilson, H. S. R.; Potter, M. J.; Gilson, M. K. *Biophys. J.* **2001**, *80*, 140.
(8) Norberg, J.; Nilsson, L. *Acc. Chem. Res.* **2002**, *35*, 465.
(9) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250.
(10) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629.
(11) Warshel, A. *Acc. Chem. Res.* **1981**, *14*, 284.
(12) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035.
(13) Warshel, A.; Florián, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950.
(14) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257.
(15) Florián, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 5739.
(16) Florián, J.; Warshel, A.; Goodman, M. F. *J. Phys. Chem. B* **2002**, *106*, 5754.
(17) Florián, J.; Goodman, M. F.; Warshel, A. *Biopolymers* **2003**, *68*, 286.
(18) Lee, F. S.; Chu, Z. T.; Bolger, M. B.; Warshel, A. *Protein Eng.* **1992**, *5*, 215.
(19) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. *Proteins: Struct. Funct. Genet.* **2000**, *39*, 393.
(20) Åqvist, J.; Medina, C.; Samuelson, J. E. *Protein Eng.* **1994**, *7*, 385.
(21) Gelfand, C. A.; Plum, G. E.; Grollman, A. P.; Johnson, F.; Breslauer, K. J. *Biochemistry* **1998**, *37*, 7321.
(22) Vesnaver, G.; Chang, C.; Eisenberg, M.; Grollman, A. P.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3614.
(23) Gelfand, C. A.; Plum, G. E.; Grollman, A. P.; Johnson, F.; Breslauer, K. J. *Biopolymers* **1996**, *38*, 439.
(24) Cuniasse, P.; Sowers, L. C.; Eritja, R.; Kaplan, B.; Goodman, M. F.; Cognet, J. A. H.; Le Bret, M.; Guschlbauer, W.; Fazakerley, G. V. *Biochemistry* **1989**, *28*, 2018.
(25) Kouchakdjian, M.; Eisenberg, M.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1991**, *30*, 3262.
(26) Goljer, I.; Kumar, S.; Bolton, P. H. *J. Biol. Chem.* **1995**, *270*, 22980.
(27) Lah, J.; Drobnak, I.; Dolinar, M.; Vesnaver, G. *Nucleic Acids Res.* **2008**, *36*, 897.
(28) Drobnak, I.; Seručnik, M.; Lah, J.; Vesnaver, G. *Acta Chim. Slov.* **2007**, *54*, 445.
(29) Tikhomirova, A.; Beletskaya, I. V.; Chalikian, T. V. *Biochemistry* **2006**, *45*, 10563.
(30) Mikulecky, P. J.; Feig, A. L. *Biochemistry* **2006**, *45*, 604.
(31) Rouzina, I.; Bloomfield, V. A. *Biophys. J.* **1999**, *77*, 3242.
(32) Tikhomirova, A.; Taulier, N.; Chalikian, T. V. *J. Am. Chem. Soc.* **2004**, *126*, 16387.
(33) Ladbury, J. E.; Chowdhry, B. Z. *Biocalorimetry: Applications of Calorimetry in the Biological Sciences*; John Wiley: Chichester, 1998.
(34) Bren, U.; Martinek, V.; Florián, J. *J. Phys. Chem. B* **2006**, *110*, 10557.
(35) Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.* **1991**, *60*, 477.
(36) Bren, U.; Martinek, V.; Florián, J. *J. Phys. Chem. B* **2006**, *110*, 12782.
(37) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.
(38) Cantor, C. R.; Warshaw, M. M.; Shapiro, H. *Biopolymers* **1970**, *9*, 1059.
(39) Randall, S. K.; Eritja, R.; Kaplan, B. E.; Petruska, J.; Goodman, M. F. *J. Biol. Chem.* **1987**, *262*, 6864.
(40) Takeshita, M.; Chang, C.-N.; Johnson, F.; Will, S.; Grollman, A. P. *J. Biol. Chem.* **1987**, *262*, 10171.
(41) Job, P. *Ann. Chim. (Paris)* **1928**, *9*, 113.
(42) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601.
(43) Lah, J.; Bester-Rogac, M.; Perger, T.-M.; Vesnaver, G. *J. Phys. Chem. B* **2006**, *110*, 23279.
(44) Lah, J.; Prislan, I.; Krzan, B.; Salobir, M.; Francky, A.; Vesnaver, G. *Biochemistry* **2005**, *44*, 13883.
(45) Lah, J.; Simic, M.; Vesnaver, G.; Marianovsky, I.; Glaser, G.; Engelberg-Kulka, H.; Loris, R. *J. Biol. Chem.* **2005**, *280*, 17397.
(46) Lah, N.; Lah, J.; Zegers, I.; Wyns, L.; Messens, J. *J. Biol. Chem.* **2003**, *278*, 24673.
(47) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipies*; Cambridge University Press: Oxford, 1992.
(48) Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97*, 3100.
(49) Levy, R. M.; Belhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.
(50) Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27.
(51) Kubo, R. *J. Phys. Soc. Jpn.* **1957**, *12*, 570.
(52) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210.
(53) Born, M. *Z. Phys.* **1920**, *1*, 45.
(54) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283.
(55) Roux, B.; Yu, H. A.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683.
(56) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
(57) Florián, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B* **2000**, *104*, 10092.
(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(59) Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. *J. Mol. Graphics Modell.* **1999**, *16*, 213.

(60) King, G.; Warshel, A. *J. Chem. Phys.* **1989**, *91*, 3647.
(61) Sham, Y. Y.; Warshel, A. *J. Chem. Phys.* **1998**, *109*, 7940.
(62) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205.
(63) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(64) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14.1*, 33.
(65) Mikulecky, P. J.; Feig, A. L. *Biopolymers* **2006**, *81*, 38.
(66) Dill, K. A.; Truskett, T. M.; Vlachy, V.; Hribar-Lee, B. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 173.
(67) Johnson, R. C.; Power, T. D.; Holt, J. S.; Immaraporn, B.; Monat, J. E.; Sissoko, A. A.; Yanik, M. M.; Zagorodny, A. V.; Cybulski, S. M. *J. Phys. Chem.* **1996**, *100*, 18875.
(68) Sievers, A.; Beringer, M.; Rodnina, M. V.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7897.
(69) Cerny, J.; Kabelac, M.; Hobza, P. *J. Am. Chem. Soc.* **2008**, *130*, 16055.
(70) Porschke, D. *Mol. Biol. Biochem. Biophys.* **1977**, *24*, 191.
(71) Norberg, J.; Nilsson, L. J. A. *Chem. Soc.* **1995**, *117*, 10832.
(72) Florián, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6819.
(73) Martínek, V.; Bren, U.; Goodman, M. F.; Warshel, A.; Florián, J. *FEBS Lett.* **2007**, *581*, 775.
(74) Peyret, N.; SantaLucia, J., Jr. *HyTher, version 1.0*; Wayne State University: Detroit, MI, 2002.

JP9064246