

Photochemistry of DNA Fragments via Semiclassical Nonadiabatic Dynamics

Anastassia N. Alexandrova* and John C. Tully

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Giovanni Granucci

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, v. Risorgimento 35, I-56126, Pisa, Italy

Received: April 13, 2010; Revised Manuscript Received: August 6, 2010

Forming upon absorption of a UV photon, excited states of DNA are subject to nonadiabatic evolution, via either internal conversion (IC) back to the ground state or mutagenesis. Nonadiabatic processes following the formation of the first singlet excited states, S1, in 10 different small DNA fragments—4 single 4'H-nucleosides, 2 Watson–Crick base pairs, and 4 nucleotide quartets—have been investigated. Simulations were done via the nonadiabatic direct trajectory surface hopping semiclassical dynamics. The electronic wave function was obtained with configuration interaction, based on the semiempirical AM1 and PM3 Hamiltonians with fractional orbital occupation numbers. The evolution of the electronic wave function was governed by the time-dependent Schrödinger equation with a locally diabatic representation, intrinsically stable near surface crossings. The nuclei evolved on adiabatic potential energy surfaces, as prescribed by classical Newtonian dynamics, with sudden hops between potential energy surfaces to account for nonadiabatic transitions. The “fewest switches” surface hopping algorithm coupled the quantum and classical parts of the system. The dynamics simulations revealed several routes of nonadiabatic relaxation in these systems, which were not reported previously, and also recovered known routes of IC.

Introduction

Upon absorption of UV light in the UV-B and UV-C regions of the solar spectrum, DNA undergoes electronic excitations. The excited states formed can further submit to an ultrafast nonradiative internal conversion (IC) back to the ground electronic state, S0. Alternatively, excited electronic states can evolve toward new minima on the ground potential energy surface (PES), which constitutes mutagenesis. Damaged DNA requires intracellular repair, while the accumulation of mutagenic products in the genome may ultimately lead to cancer.

The dynamics of small DNA fragments following photoexcitation have been subjected to femtosecond pump–probe experimental and theoretical investigations. Kang et al.¹ reported experimental results indicating that IC for A, C, T, and U consists of a short step (<15 fs) and a long step (750–1020 fs). Canuel et al.² also found that the decay for all bases occurs in two steps; however, from their experiments, a short step should take 100–160 fs and the long step should take 1–5 ps, depending on the base. Pecourt et al.³ performed femtosecond pump–probe spectroscopy on nucleosides in water. They found that the excited state lifetimes of Ado, Guo, Cyd, and Thd were 290, 460, 720, and 540 fs, respectively.³ However, full dynamical information about the mechanisms of IC and mutagenesis is hardly ever accessible experimentally, since various dark states are known to be involved in the dynamics.

A variety of conical intersections were proposed for IC in pyridine bases, most of which proceed via one of the out-of-plane bending modes of the aromatic system.^{1,2,4–15} Ullrich et al.⁹ reported femtosecond time-resolved photoelectron spectra of nucleobases and, for adenine, proposed the predominant IC mechanism, in which the initial $\pi\pi^*$ excited state first crossed

to the $n\pi^*$ state, which in turn crossed to S0. This mechanistic proposal came in accord with theoretical studies by Blancfort,¹⁰ Perun et al.,¹¹ and Serrano-Andrés et al.,¹² who also reported the major role of the $\pi\pi^*$ (1L_b) state in the dynamics: at CASPT2//CASSCF, this state decays either directly through a conical intersection with S0 or indirectly via a relay $n\pi^*$ state, either with or without a barrier. The semiempirical OM2/MR-CI dynamics simulations by Fabiano and Thiel¹³ also yielded the two paths for IC—direct and indirect—and the indirect $\pi\pi^*$ to $n\pi^*$ to S0 path was found to be predominant. At the same time, Coni et al.¹⁴ claimed the $\pi\pi^*$ (1L_b) state to be solely responsible for the dynamics, both at short and at long time scales, whereas $n\pi^*$, $\pi\pi^*$ (1L_a), and $\pi\sigma^*$ do not play a significant role in IC. Controversially, Sobolewski and Domcke¹⁵ put forward a mechanism of IC via the population of the N–H dissociative $\pi\sigma^*$ state. The decay path was also found to be dependent upon the initial excitation energies;⁹ for example, the $\pi\sigma^*$ IC channel was shown to play a role in 267 nm excitations. Finally, Barbatti and Lischka,⁵ in their dynamics study performed at the MR-CIS level and supported by CASPT2 calculations on conical intersections, reported many different conical intersections for 9H-adenine, including several ring puckering structures, and one rare structure with an open C8–N9 bond. The authors found no dynamics evidence for the direct $\pi\pi^*$ to S0 IC mechanism and proposed that the short component of the decay is due to S3–S1 relaxation and the long component is due to the decay from S1 to S0. They also emphasized that dynamics simulations are required to assess photochemical processes and that their mechanisms are virtually impossible to deduce from static simulations, mainly because of coupling between many involved internal coordinates.

Chen and Li,¹⁶ Yamazaki et al.,¹⁷ and Serrano-Andrés et al.¹⁸ performed CASPT2//CASSCF studies on 9H-guanine and

* To whom correspondence should be addressed. E-mail: anastassia.alexandrova@yale.edu.

proposed the IC mechanism consisting of a direct $\pi\pi^*$ (${}^1\text{L}_\alpha$) to S0 decay. This IC path is barrierless at CASPT2//CASSCF, but at DFT, there is a quasiplanar minimum of the $\pi\pi^*$ surface.^{19a} The $n\pi^*$ state could also decay via first converging to $\pi\pi^*$ and then to S0.¹⁶ The N–H dissociation via the $\pi\sigma^*$ state constitutes the second possible mechanism of decay, from CASPT2//CASSCF¹⁷ and Car–Parrinello molecular dynamics studies.^{19b} Serrano-Andrés et al.¹⁸ also proposed that the $n_0\pi^*$ state should be involved and responsible for the slower decay via the $n_0\pi^*/S0$ conical intersection. However, Chen and Li¹⁶ reported the involvement of $\pi\sigma^*$ and $n_0\pi^*$ states to be energetically unfavorable, and thus unlikely. In the dynamics study by Lan et al.,²⁰ 9*H*-guanine was shown to undergo IC via two channels, both involving bending of the base: a faster process via C2 going out of the molecular plane and a slower process via distortion of the NH₂ group. A Car–Parrinello molecular dynamics study of 9*Me*-guanine showed that this species also decays via the NH₂ group going out of the molecular plane.²¹ Unlike in 9*H*-guanine, however, this deformation on S1 is much stronger, and optical activity of 9*Me*-guanine in the Franck–Condon region is much smaller.²¹ This may be an indication that 9*H*-guanine nucleobases are not the best models of the biologically relevant systems.

For pyrimidine bases, cytosine and thymine, generally, two alternative paths for IC were proposed: from the $\pi\pi^*$ state, directly via the conical intersection between $\pi\pi^*$ and S0, and indirectly through a minimum on $\pi\pi^*$, to $n\pi^*$, to S0.^{3,22–28} T is the base that takes the longest to relax from its UV-excited state.^{29b} Another experimental work showed that in fact there are three components for IC in T: <50 fs, 490 fs, and 6.4 ps.^{9b} The long lifetime was attributed by different authors to a trapping $\pi\pi^*$ minimum on S2,^{29,30} a $n_0\pi^*$ minimum on S2,²⁷ a $\pi\pi^*$ minimum on S1,³¹ and a $n\pi^*$ minimum on S1.^{27,28,32} Interestingly, the dynamics CASSCF simulations paired with CASPT2 calculations²⁶ revealed that CASSCF tends to overestimate the rate of the $n\pi^*$ decay for T, thereby questioning all quantitative timing information obtained from on-the-fly CASSCF dynamics. Perun et al. found three conical intersections between S1 and S0 in T, at CASPT2//CASSCF and CC2 levels of theory.²⁸ All these structures are characterized by out-of-plane deformation of the base. The most energetically accessible one involves the CH₃ group going strongly out of the molecular plane and corresponds to a crossing between $\pi\pi^*$ and S0. A comprehensive ab initio study by Zechmann and Barbatti³² revealed a total of eight extremes on the crossing seam for thymine, characterized by puckering of the ring in different ways, and bond-length shortening and elongation.

In a few studies of IC in C, the direct IC path from $\pi\pi^*$ to S0 was proposed to be preferred, without the involvement of $n_0\pi^*$ and $n_N\pi^*$ states, in contrast with that for T.^{22,33–35} For the direct $\pi\pi^*$ to S0 IC, the deformation of C on S1 involves bending of the base. In the CAS-MRCI//CASSCF study by Kistler and Matsika,³⁵ the $\pi\pi^*$ state had two channels for IC: via N3 going out of plane and via twisting of the C5–C6 bond. The C5–C6 twisting mechanism was confirmed at CASPT2//CASSCF.³⁶ The authors³⁶ also showed that both $n\pi^*$ states might be involved in IC, but they have minima that slow the decay down, and that may explain its biexponential character. The conical intersection between $n_N\pi^*$ and S0 was claimed to be much more energetically accessible than the $n_0\pi^*$ –S0 one.³⁶ However, in the study by Ismail et al.,³⁷ the $\pi\pi^*$ state first converted to $n_0\pi^*$, and only then decayed to S0, whereas the $n_N\pi^*$ state was not involved. Finally, Blancafort and Robb³⁸ reported a three-state conical intersection between S0, $\pi\pi^*$, and

$n_O\pi^*$, and named it dominant in the decay of the first singlet excited state of C. With the three-state conical intersection, different combinations of bond inversion and pyramidalization modes may be involved in the dynamics.³⁸ The on-the-fly SA-CASSCF spawning Gaussian dynamics study by Hudock and Martínez³⁹ revealed the complexity of IC in C, and showed that multiple IC paths can be simultaneously operational. All $\pi\pi^*$, $n_N\pi^*$, and $n_0\pi^*$ states appeared to be involved in IC at different time scales. The associated reaction coordinates were various base bending and bond stretching and shortening modes. This study dismissed the proposal about a three-state conical intersection, however. Thus, there is a variety of opinions in the literature about IC in C.

IC processes for a few systems containing more than one nucleobase have been documented. IC from the first excited state in the cytosine–guanine base pair in the gas phase, in solution, and in the context of a DNA double helix is known to occur via proton transfer from N1 of guanine to N3 of cytosine and a skeletal deformation of the bases, a downhill process on the S1 PES, which leads to the conical intersection with S0.^{40–44} The corresponding excited state involved is G to C charge transfer (CT). For the AT pair, the analogous mechanism, involving CT, followed by H-transfer, was proposed, based on CC2 calculations.⁴⁵ In general, IC in the AT pair is largely understudied. Interestingly and controversially, in ref 46, nearly all low-energy excitations in CG and AT base pairs were proposed to be localized on just one base in the base pair, thus raising a concern that CT states and their decay may not be as relevant to IC.

When two pyrimidine bases (thymine or cytosine) neighbor each other in a DNA strand, upon UV excitation, the system can either undergo IC via a local excitation on a single base⁴⁷ or it can undergo an extensively studied ultrafast dimerization.^{47–58} The crystal structures of DNA containing a TT dimer lesion alone⁵⁸ and bound to repair enzymes⁵⁷ were also recently obtained.

Most of the theoretical works so far explored single minimum energy paths (MEPs) and minimum energy conical intersections (MECIs), which can be insufficient to describe the photochemistry, as was stated more than once.^{5,33,39} Additional uncertainties arise from the levels of theory used in both static and dynamic studies. Static simulations usually rely on CASSCF geometry optimizations, followed by single point energy calculations at CASPT2. CASSCF, along with semiempirical methods, is also the most affordable method for the dynamics on-the-fly. However, CASSCF lacks dynamic electron correlation and suffers from ambiguity due to the choice of CI active space. Also, CASSCF dynamics simulations can be afforded only with small basis sets. The number and locations of extremes on PESs and MEPs predicted by CASSCF may thus be inaccurate. Due to these methodological problems, there are controversies in the literature, regarding the mechanisms of IC in most of the systems considered so far, calling for further dynamics and high level ab initio studies. Semiempirical methods are less accurate but considerably cheaper, which at least allows for longer dynamics simulations, on a larger number of trajectories, and potentially for larger systems. A large number of different trajectories also seems to be essential to sample a variety of starting conditions for the dynamics, because it was shown that the destiny of the excited state depends upon the excess energy in this state.⁹

Additionally, system-wise, no comprehensive dynamics simulations were accomplished for systems in which more than two nucleobases were assessed quantum mechanically. Furthermore,

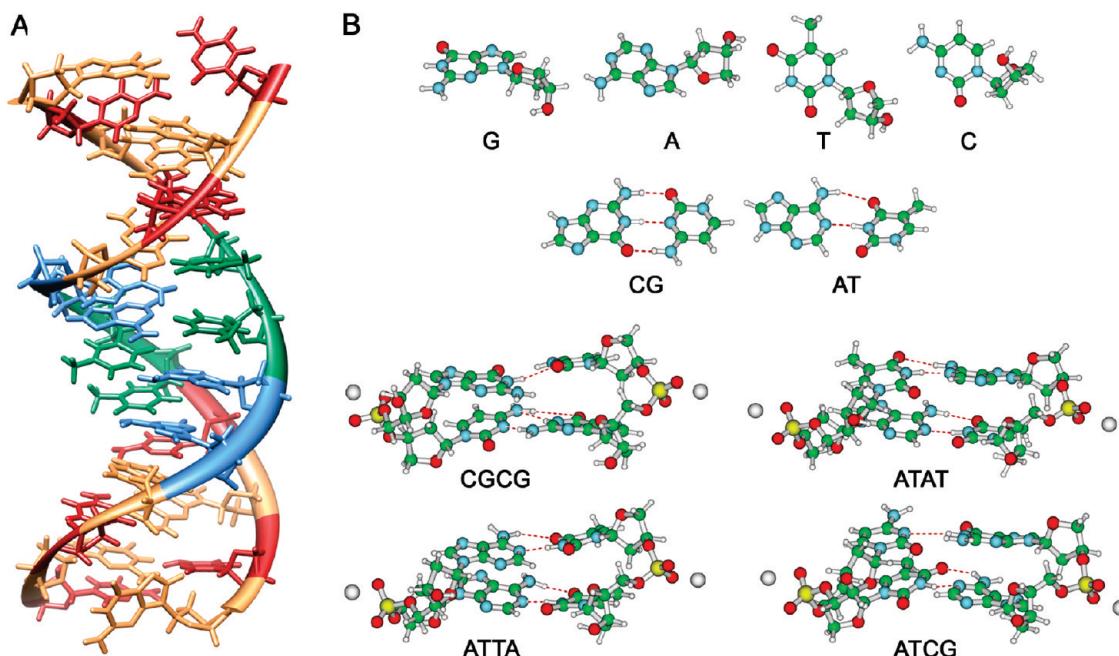


Figure 1. (A) Representative structure of the DNA double helix from the end of the MC run: G - yellow, A - purple, C - red, T - green, “imperfect” base pairs may be seen. (B) Fragments cut from the structure in part A and considered in this work.

even for single bases, the role of the sugar was not addressed, whereas, for example, 9*H*-guanine and 9*Me*-guanine were shown to have differences in IC,²¹ suggesting the potential importance of the substituent at N9. In the present study, semiclassical nonadiabatic dynamics simulations with semiempirical Hamiltonians are conducted for 10 small DNA fragments: 4 4*H*-nucleosides (base bound to its sugar), 2 base pairs, and 4 nucleotide quartets, including the backbone. Most of these fragments were not considered previously. The study features an extensive sampling of initial conditions in the dynamics, and numerous trajectories for each system, as well as fast semiclassical dynamics simulations. The aim is to identify the likely IC and mutagenic paths from the first singlet excited state, in a qualitative fashion, and using a unified methodology. We present a few new routes of IC and mutagenesis that are likely to be involved in the photochemistry of DNA.

Methods

The starting point for the present study is the crystal structure of a DNA double helix fragment (Protein Data Bank code 1N4E).⁵⁸ The 1N4E structure is a dimer, and only a 20-base monomer without a lesion was retained for simulations (Figure 1A). Crystallographic water molecules were removed. The system was relaxed using Monte Carlo (MC) simulation in an NTP ensemble at 25 °C and 1 atm, as implemented in the MCPRO 2.00 package.⁵⁹ The OPLS-AA force field was used.⁶⁰ The solvent was represented with an implicit generalized Born/surface area (GB/SA) solvation model,^{61–63} with an ionic strength of 0.5 imitating a 1 M Na⁺ solution, for the compensation of the negative charge on the DNA backbone. For the nucleotide bases, all valence angles and torsions were sampled in MC, whereas the bond lengths and the DNA backbone were kept fixed. The simulation consisted of a total of 30 × 10⁶ configurations. Noticeably, all DNA structures resulting from the MC run contained several “imperfect” (twisted and shifted) Watson–Crick pairs (Figure 1A). This is quite expected, since the DNA structure is rather flexible, but it is also interesting for the present investigation, since it was shown previously that

the fate and nature of excited states in DNA is dependent upon the nucleotide sequence, and the higher-order structure, i.e., the geometry of the base-stacking.^{64,65,56,66–70} In view of this, fragments taken from the relaxed double helix should present a more realistic model than the ground state geometries of the isolated fragments in the gas phase. Five representative low-energy configurations from the end of the MC simulation were taken as starting points for further investigation. DNA fragments of interest were cut out of these relaxed structures (Figure 1B). For the nucleotide quartets, two Na⁺ cations were manually added to compensate the negative charge on the backbone, and their positions were optimized with AM1.⁷¹

Semiclassical dynamics simulations were conducted using the direct trajectory surface hopping method by Granucci et al.⁷² Nuclei moved classically on a single adiabatic potential energy surface (PES), as prescribed by Newton’s equations of motion, and are allowed to hop to other PESs at any time, according to transition probabilities computed with the “fewest switches” algorithm.⁷³ The evolution of the electronic wave function was governed by the time-dependent Schrödinger equation, with a locally diabatic representation.⁷² The latter makes the solution stable in situations of near-degeneracy of the adiabatic states, for example, near conical intersections. The electronic wave functions were built with a full configuration interaction within a chosen active space (CAS-CI). In the present work, a CAS-CI space of two electrons in two orbitals, CI(2,2), was used, and always two electronic states were involved in the dynamics: S0 and S1. The reference function for the CI was computed with the semiempirical molecular orbital (MO) methods, mainly AM1,⁷¹ and additionally PM3,⁷⁴ with floating orbital occupation numbers (FON),⁷² using MOPAC 2000.⁷⁵ Unless otherwise indicated, the reported results were obtained with AM1 CI(2,2). The Gaussian width was set at 0.1. The choice for the electronic Hamiltonian was justified by the smallest number of trajectories running into the problem of SCF nonconvergence and energy nonconservation (the problem mainly concerned adenine 4' *H*-nucleoside). Other tested Hamiltonians, however, included CI(4,4), (6,6), and (10,10), with AM1 or PM3 as a reference

function, and the Gaussian width of 0.3. The dynamics started from a vertical excitation from S0 to S1. For each atom, the norm of the velocity (v) was assigned on the basis of the equipartition principle, at a temperature of 300 K. The direction of the velocity was assigned at random, and the system was allowed to evolve. Velocities were different for each trajectory.

Each semiclassical dynamics trajectory entailed 10 000 steps of 0.1 fs, or a total of 1 ps. In the case of the adenine–thymine Watson–Crick pair, simulations were extended up to 5 ps, to better capture the mechanism of IC in this system, which appeared to be slow. For each system, each of the five representative configurations resulting from the MC run gave rise to 100 trajectories, i.e., a total of 500 unique trajectories. Occasionally, the simulations still ran into the energy nonconservation problem, for reasons that we could not deduce. Such trajectories were discarded. All results reported herein are averaged over only good and complete trajectories, whose number was unavoidably smaller than 500 in each case. The sampling of starting structures was our way to account for the sensitivity⁹ of the dynamics in DNA fragments to the excess energy in the vertically excited species.

The performance of semiempirical MO methods for excited states has been validated for a number of classes of compounds.^{76–78} They are reliable in providing at least a qualitative description of the topology of excited state PESs. The results of the present work, including the general trends in nuclear rearrangements during the evolution of the excited states, the relative prevalence of different routes in the dynamics, the locations of conical intersections, and the rates on nonadiabatic processes, are, therefore, qualitative in nature. However, semiempirical methods are also fast, which allows for performing the on-the-fly dynamics. Using *ab initio* methods for the electronic part of the problem would be nearly computationally prohibitive, especially for the larger systems considered here. Once the routes of the dynamics are qualitatively mapped out, and the relevant reaction coordinates are revealed, more accurate theoretical treatments could be carried out.

In order to unveil the nature of the observed electronic excitations, and the evolution of the electronic structure of the systems as they progress toward a conical intersection, the electronic states were monitored via molecular orbital (MO) analysis. MOs were obtained at the same level of theory as the corresponding dynamics runs. Some additional DFT calculations were done using Gaussian 03.⁷⁹

Results and Discussion

Single 4'H-Nucleosides. Figure 1B shows representative starting ground state structures of G, A, T, and C bound to their sugars, as resulting from the MC relaxation of the DNA double helix. The total charge of each species is zero.

Guanine. The starting structures taken from the DNA double helix differ from those for the optimized nucleoside in the gas phase: the relative positions of the base and sugar are significantly different for the two contexts, and also there are small out-of-plane distortions in the base taken from DNA. The energy difference between our starting structures and the optimized one varies from 1.5 to 2.4 eV, again, majorly due to the position of the sugar. This destabilization of the ground state, however, does not directly correlate with the S0–S1 energy difference in the Franck–Condon region.

Out of 500 trajectories run for guanine 4'H-nucleoside, 498 finished without running into SCF convergence or energy nonconservation problems. The results of these dynamics simulations are shown in Figure 2. A part of a typical trajectory

from the dynamics simulations is shown in Figure 2A. In the beginning of the simulations, G is vertically excited to the first singlet excited state S1 and it has the ground starting geometry \mathbf{G}_v (vertical red arrow in Figure 2A). The excitation energy in our simulations is 4.11 ± 0.32 eV, depending on the starting structure (compare 4.92 eV, the experimental value for the isolated 4'H-nucleoside, as measured at the center of the absorption band).⁸⁰ The red points in the trajectory (Figure 2A) belong to the S1 PES, and the blue points belong to S0. The changes in the intramolecular degrees of freedom in G are monitored throughout the simulations. It appeared that the out-of-plane skeletal deformation in the base, corresponding to simultaneous pyramidalization at N1 and C2, is the internal coordinate most relevant to the dynamics. We chose to plot the trajectories against the (N1–C2–N3–C4) torsion, as one always relevant reaction coordinate. Bending of the base leads to a conical intersection between S1 and S0, in accord with refs 16–21. A representative hopping point, \mathbf{G}_{hop} , is shown in Figure 2B. The presence of the sugar does not qualitatively change the dynamics in the case of G; IC occurs solely via changes in the internal coordinates of the base.

From Figure 2A, it is evident that there are two almost equivalent conical intersections relevant to bending of G: that in the direction of the positive (N1–C2–N3–C4) dihedral angles and that in the direction of negative (N1–C2–N3–C4) dihedral angles. The hops occurred at $\text{dih}(\text{N1–C2–N3–C4}) = \pm 16.9 \pm 8.9^\circ$. The representative trajectory in Figure 2A also illustrates that the S1 PES is relatively flat with respect to this motion of the molecule. As a result, the molecule easily moves in both directions along the reaction coordinate while on S1. Furthermore, sometimes in the same trajectory, hops occur both at positive and at negative values of $\text{dih}(\text{N1–C2–N3–C4})$. Figure 2D shows the density plots of values of the relevant torsions versus the corresponding difference in energy between S1 and S0. In this case, these plots just show that hops for G may occur at a wide range of values for these torsions, in both positive and negative directions, and greater values are characteristic of $\text{dih}(\text{N1–C2–N3–C4})$ rather than $\text{dih}(\text{C2–N3–C4–N1})$.

Figure 2C shows the fractions of trajectories evolving on S1 as a function of time. It can be seen that the half-life of the S1 state, i.e., the time at which ~50% of the trajectories switched from S1 to S0, is ca. 240 fs. Also, a tiny fraction of trajectories decayed within 120 fs, which might be the fast component of the decay known from the literature. 240 fs is a bit faster than the experimentally obtained timing of 360 fs.² Partly, this is probably due to the uncertainties associated with the semiempirical electronic Hamiltonian. It is also likely that the timing depends on the presence or absence of the solvent, and the sugar bound to the nucleobase. By the end of 1 ps, 74% of trajectories experience IC. Trajectories that did not end up on S0 by 1 ps corresponded to situations where a rehop from S0 to S1 happened near 1 ps, and G continued to flop on S1 along the same reaction coordinate. Rehops were frequent in the dynamics. The mechanism of IC was always the same, and we did not observe trapping in any minima on S1 such that the system would fail to decay.

The results were controlled via running additional dynamics simulations with AM1 CI(6,6), PM3 CI(2,2), and PM3 CI(4,4) Hamiltonians. The trend in nuclear rearrangements and the timing for IC were found to be qualitatively the same. For comparison, in Figure 2C (blue curve), we show the fraction of trajectories on S1 as a function of time, obtained with PM3 CI(2,2).

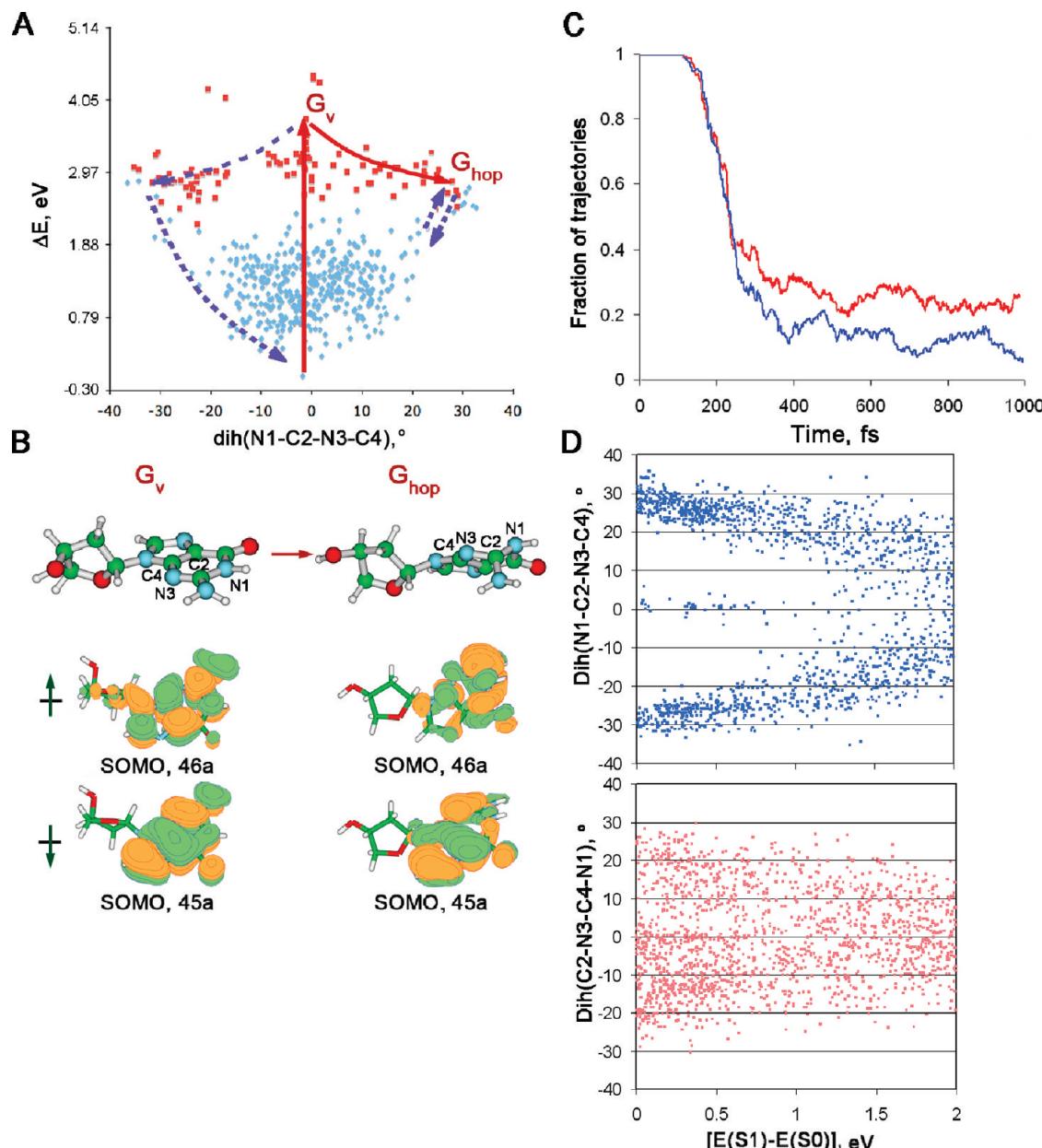


Figure 2. Nonadiabatic dynamics in the G 4'H-nucleoside: (A) a representative trajectory: electronic energy relative to the starting ground state structure is shown on the vertical axis; points are registered every fifth configuration, red points belong to the S1 PES, blue points belong to S0; the vertical red arrow indicates the initial S0–S1 excitation to \mathbf{G}_v ; the system then evolves to a hopping point, \mathbf{G}_{hop} (red arrow pointing to the right), and hops from S1 to S0; multiple rehops from S0 to S1 and back, including those through bending in the direction of negative dih(N1–C2–N3–C4) (e.g., purple dashed arrows), happen in this trajectory; the system ends up near the minimum on S0, where dih(N1–C2–N3–C4) $\sim 0^\circ$. (B) Structures of representative \mathbf{G}_v and \mathbf{G}_{hop} and the comparison between their MOs. (C) The fraction of trajectories evolving on S1 as a function of time: red - at the AM1 CI(2,2) level; blue - at PM3 CI(2,2). (D) The density plot showing the values of relevant internal coordinates at hopping points versus the corresponding energy difference between S1 and S0.

To support and illustrate the dynamics results, we analyzed the changes in the electronic structure of the system by monitoring the molecular orbitals (MOs), obtained with the same AM1 CI(2,2) method (Figure 2B). The first electronic excitation in G 4'H-nucleoside corresponds to the promotion of an electron from the HOMO (45a) to the LUMO (46a), yielding two singly occupied molecular orbitals (SOMOs). The S0–S1 excitation is almost completely localized on the base and does not involve the sugar or the glycosidic bond. The forming state is $\pi\pi^*$ (L_a). The redistribution of the π -electron density in the base manifests itself in the out-of-plane motion on S1. In the vertically excited structure, \mathbf{G}_v , both SOMOs are mixes of π -MOs on the aromatic system and the lone pair (LP) on the amino group, in the framework of MO-LCAO. The SOMOs of the representative

hopping point, \mathbf{G}_{hop} , obviously resemble those in \mathbf{G}_v , except now the SOMO (45a) is mostly a π -MO on the base and the SOMO (46a) contains more electron density on the LP of the amino group. The nature of the S1 state does not change, however; it is still $\pi\pi^*$ (L_a). Hence, our simulations predict a straightforward decay from the $\pi\pi^*$ S1 state to S0, in agreement with earlier reports.^{16–19a} A few trajectories that underwent IC on a shorter time scale corresponded to a smaller deviation from the planarity in the base. The SOMO (46a) in the hopping geometries in these trajectories more closely resembles the SOMO (46a) in \mathbf{G}_v , with a smaller shift of electron density from the base to the amino group.

Adenine. Figure 3 presents results for adenine. The starting structures for A were taken from a particular context in DNA

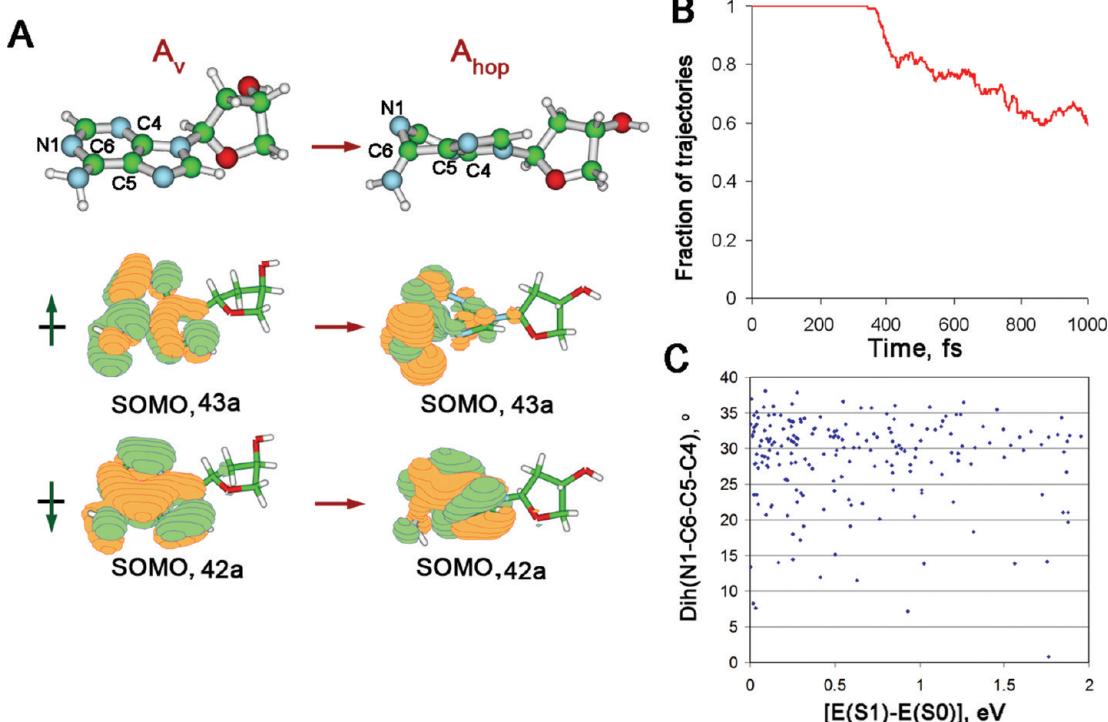


Figure 3. Nonadiabatic semiclassical dynamics in A 4'H-nucleoside: (A) A representative starting vertically excited structure, \mathbf{A}_v , and a hopping point, \mathbf{A}_{hop} , and the comparison between their SOMOs. (B) Averaged over 422 trajectories, the fraction of trajectories evolving on S1, as a function of time. (C) The density plot showing the values of $\text{dih}(\text{N1}-\text{C6}-\text{C5}-\text{C4})$ at hopping points versus the corresponding energy difference between S1 and S0.

that happened to be quite far from the optimized ground state structure (2.2–2.8 eV), due to both the position of the sugar and a slight bend in the base. The vertical excitation energies were 3.51 ± 0.39 eV. The experimental value for the isolated nucleoside is 4.78 eV, at the absorption band maximum,⁸⁰ though, again, the direct comparison between these values might be misleading due to structural differences. The motion on the excited state leading to IC is similar to that observed in guanine: it is the bending in the aromatic system of the base, which is consistent with previous reports.^{5,9–14} The most appropriate reaction coordinate is the ($\text{N1}-\text{C6}-\text{C5}-\text{C4}$) torsion, as 96.0% of the hops occurred via N1 going out of the molecular plane and the remaining 4.0% of the hops involved C6 going out of the molecular plane. A representative hopping point, \mathbf{A}_{hop} , is shown in Figure 3A. As averaged over 422 trajectories, at the hopping point, $\text{dih}(\text{N1}-\text{C6}-\text{C5}-\text{C4}) = 30.15 \pm 17.70^\circ$ (Figure 3C), and the electronic energy is -109.172334 ± 0.047551 au, which is ca. 1.2 eV below \mathbf{A}_v . There is a qualitative consensus between results obtained with AM1 CI(2,2), AM1 CI(6,6), PM3 CI(6,6), and PM3 CI(10,10), although we had difficulties obtaining a significant number of trajectories with these Hamiltonians due to energy nonconservation and SCF nonconvergence problems.

Noticeably, in our simulations, IC in adenine is also slower than that in guanine. Figure 3B shows the fraction of trajectories evolving on S1 as a function of time. It illustrates how inefficient IC is: on average, switching from S1 to S0 does not begin until ca. 400 fs into the dynamics. Figure 3C also illustrates how scarce S1–S0 transitions were in the dynamics for A (compare to Figure 2D). Unlike for guanine, we obtained no fast decaying trajectories. By the end of 1 ps, the population of S1 is still slightly greater than that of S0. It looks like our dynamics simulations do not capture the fast component of the decay that was detected experimentally, and we only see the slow

component, which is known to take ca. 1.1 ps.² However, Barbatti and Lischka in their dynamics study⁵ proposed that A can live on S1 ($\pi\pi^*$) in the neighborhood of the structure similar to our \mathbf{A}_{hop} , which they called ${}^2\text{E}$. In their simulations, this trapped structure slowly decayed via the conical intersection with S1, whereas the fast component was due to the preceding relaxation from S3 to S1. If this mechanism is right, we do not see the fast component of the decay because we start the dynamics from S1. Hence, our results indirectly confirm this IC mechanism by Barbatti and Lischka.

The S0–S1 excitation in the starting geometry of A is of the $\pi\pi^*$ (L_b) character (Figure 3A), which is different from that in G. The L_b character of the S1 state in A was also predicted by ab initio studies (for example, ref 8). In Figure 3A, the SOMOs in \mathbf{A}_v are matched against those in the hopping point, \mathbf{A}_{hop} . The SOMO (43a) in \mathbf{A}_{hop} has the largest component of the electron density localized on the amino group of the base, just like in \mathbf{G}_{hop} (Figure 2). The electron density on the base is significantly perturbed due to ring puckering: the nodal structure of this SOMO changes, and also, the contributing p_z -AOs on C6 and N3 orient at large angles to those on the remaining atoms in the base. Thus, the $\pi\pi^*$ (L_b) state adiabatically changes in the dynamics on S1, and it is no longer possible to characterize it as pure L_b at \mathbf{A}_{hop} . The major role of the L_b state and its evolution in the dynamics is in agreement with ab initio results obtained earlier.^{10–14}

There is a significant amount of controversy in the literature regarding IC in **pyrimidine bases**. Mainly, the argument focuses on whether the $\pi\pi^*$ excited state decays directly to S0 or proceeds through a $n\pi^*$ state or states along the way. It is evident that the $\pi\pi^*$ and $n\pi^*$ excited states are closely spaced in the Franck–Condon region, at least for the isolated pyrimidine bases. Considering this fact, and the innate weaknesses of

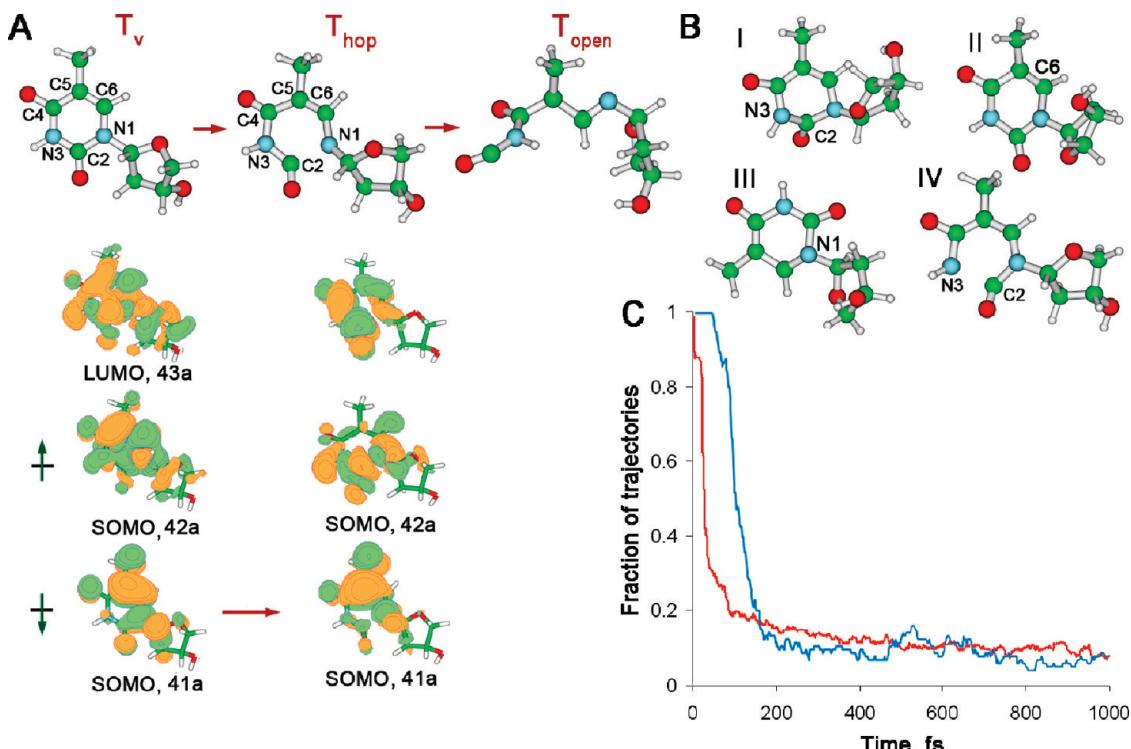


Figure 4. Nonadiabatic semiclassical dynamics in T 4'H-nucleoside: (A) The starting vertically excited structure, T_v , and a representative hopping point, T_{hop} , with cleaved N1–C2 bond, and the comparison between the MOs. (B) Other types of hopping points, with marked pyramidalization sites (I–III), and the C2–N3 stretch (IV). (C) Averaged over 396 trajectories, the fraction of trajectories evolving on S1, as a function of time, at AM1 CI(2,2) (red curve); the blue curve represents the PM3 CI(2,2) result for comparison.

semiempirical methods, we hereby proceed with caution in interpreting results for T and C.

Thymine (Figure 4). The starting structures for T were 1.3–1.9 eV above the optimized structure, and the vertical excitation happens to take 4.19 ± 0.17 eV (the experimental value is 4.65 eV for the center of the absorption band).⁸⁰ The vertical excitation is $\pi\pi^*$. Interestingly, the electron density of the sugar moiety is involved in both SOMOs (Figure 4A), and thus is likely to influence the dynamics, which was never addressed previously.

The most prominent paths in the dynamics were reported to involve ring puckering of the base.^{27–32} However, in our simulations, only 18.98% of the trajectories involve solo bending of the base via pyramidalization at various sites: N1, C2, N3, and C6, sometimes with a slight N2–C3 stretch (Figure 4B). In these trajectories, the S1 state retained its $\pi\pi^*$ character at the hopping point. Surprisingly, the majority of the trajectories (76.2%) proceeded via stretching the N1–C2 bond to ca. 2.02 Å, accompanied by its ring puckering via N3 going out of the molecular plane (Figure 4A). A typical hopping point occurred at 0.28 eV below T_v . In T_{hop} , the SOMO (42a) is a $n_0\pi^*$ -MO that puts antibonding character on the N1–C2 bond, which lead to the N1–C2 stretch. 31.5% of these trajectories returned back to the ground state structure. 44.7% proceeded to the complete cleavage of the N1–C2 bond (T_{open} , Figure 4A). We confirmed the existence of this minimum with B3LYP/6-311G*. The mechanism was also found with PM3 CI(2,2) and PM3 CI(4,4).

The lifetime of S1 in our simulations is ca. 30 fs at AM1 CI(2,2) and ca. 110 fs at PM3 CI(2,2), whereas the experimental value for the short IC component, for example, by Ullrich et al.^{9b} is less than 50 fs. By 1 ps, 90% of trajectories are on S0. The trajectories that failed to converge to the ground state by the end of 1 ps did not get trapped on S1, but rehops occurred from S0 to S1 toward the end of the run, and mostly that

happened with the structures already containing open N1–C2 bonds. We did not capture the long components of the decay (490 fs and 6.4 ps).^{9b}

In order to elaborate on the origin of the discrepancy between our results and those obtained with ab initio methods, we ran an additional 500 trajectories of dynamics simulations with AM1 CI(2,2) for 1H-thymine, e.g., the base without the sugar. The five representative geometries of T were taken from those used for the nucleoside. The vertical excitation energy in this case was 4.50 ± 0.19 eV. We found that the isolated base never undergoes the ring opening during IC, and only in 4.2% of the trajectories is the N1–C2 bond slightly stretched to ca. 1.6 Å at hopping points. Instead, the system indeed adopts one of the ring puckering paths during IC. The most prevalent route, adopted in 95.1% of trajectories, entails C4 going out of the molecular plane. Hence, the mechanism of IC in the nucleoside, as opposed to the isolated base, is predicted by AM1 CI(2,2) to be altered by the presence of the sugar. This correlates with the involvement of the electron density from the sugar in the SOMOs. There is still a possibility that, for the nucleoside, the unusual IC path via the N1–C2 stretch could be a result of incorrect order of excited states predicted by semiempirical methods. The mechanism requires further confirmation at higher levels of theory. It is also anticipated that, in the context of the WC pair, this mechanism would be hindered by H-bonding to the opposing base.

Cytosine (Figure 5). Initial structures of the nucleoside were 0.9–2.2 eV above the ground state, and the vertical excitation energies were 4.17 ± 0.49 eV (experimental value is 4.52 eV).⁸⁰ The S1 state is $\pi\pi^*$ in nature, in agreement with previous reports.^{22,33–35} However, both SOMOs also include the electron density from the sugar moiety, which suggests the possible importance of sugar for the dynamics.

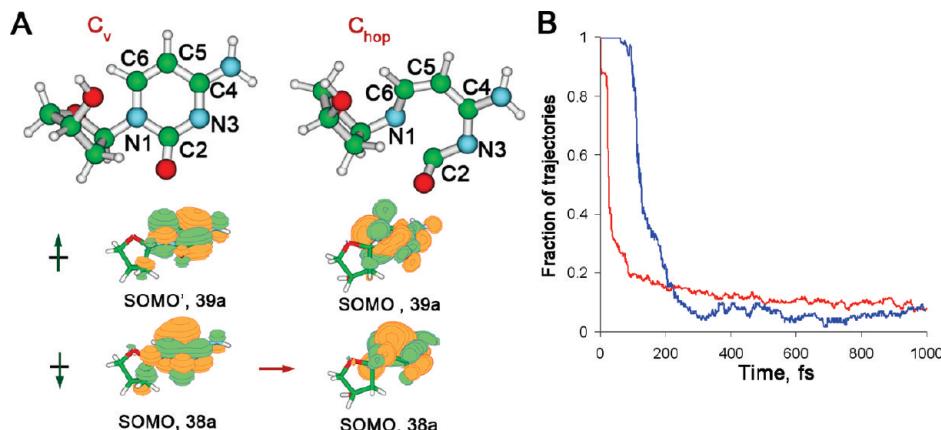


Figure 5. Nonadiabatic semiclassical dynamics in C4'H-nucleoside: (A) The starting vertically excited structure, C_v , and a representative hopping point, C_{hop} , and the comparison between the MOs. (B) Averaged over 361 trajectories, the fraction of trajectories evolving on S1, as a function of time, at AM1 CI(2,2) (red curve); the blue curve represents the PM3 CI(2,2) result for comparison.

22.8% of trajectories involved IC via bending of the base, which corresponded to the direct $\pi\pi^* \rightarrow S0$ decay, in agreement with earlier studies on the isolated C base (for example, ref 39). However, again, the prevailing IC route occurred via stretching the N1–C2 bond accompanied by ring puckering (Figure 5A). 27.2% of such trajectories returned back to the equilibrium geometry of S0, and 50.0% proceeded toward the new minimum on S0, corresponding to the open N1–C2 bond. We confirmed the existence of this minimum at B3LYP/6-311G*. In C_{hop} , $R(N1-C2) = 1.74 \pm 0.25 \text{ \AA}$, i.e., smaller than in T_{hop} . The N1–C2 stretch is due to the population of the SOMO (39a), which is a mixed π^* - and $n_O\pi^*$ -MO that puts a σ^* antibonding character in the region of the N1–C2 bond. The population of $n_O\pi^*$ is in agreement with earlier reports,^{37–39} though the N1–C2 stretch was not found previously. We suspect that this path of IC for C and T may be hindered in the CG WC pair.

The half-life of S1 is ca. 60 fs at AM1 CI(2,2) and ca. 120 fs at PM3 CI(2,2), which might correspond to the fast component of the decay. By the end of 1 ps, 96% of all trajectories convert to S0. Trajectories that failed to return to the ground state by the end of 1 ps corresponded to rehops via very similar nuclear rearrangements: base puckering and stretching N1–C2 bond. No trajectories remained on S1 for the entire duration of the simulation. We again fail to see the slow component of IC.

Whether the found mechanism originates from the poor performance of semiempirical methods remains to be explored. However, we performed additional simulations for the isolated C (500 trajectories at AM1 CI(2,2)) and found that for this system we recover the IC routes reported in the literature. Specifically, 80.6% of the trajectories evolved via C4 going out of the molecular plane, and other trajectories involved other modes of ring puckering. The N1–C2 opening was never observed for the isolated base, though its stretch to ca. 1.6 \AA was found in 30.4% of the hopping geometries. The vertical excitation energy in 1H-cytosine was $4.50 \pm 0.32 \text{ eV}$. Hence, in our simulations, the N1–C2 stretch is due to the presence of the sugar.

We would like to emphasize again that the $\pi\pi^*$ and $n\pi^*$ states are close in energy, at least in isolated T and C, and as predicted by higher levels of theory. Hence, there is a risk that semiempirical methods predict the wrong order of excited states, and take the systems along the wrong PESs in IC. We do recover IC paths known for isolated bases, and that adds certain credibility to our methodology. However, still, the new prevail-

ing IC mechanism needs to be confirmed at higher levels of theory. We hope our simulations at least bring up a warning that the sugar might need to be included when studying IC in pyrimidine nucleotides.

Watson–Crick Base Pairs. For two or more bases in the system, several different possibilities arise for the vertical electronic excitations. There can be local excitations on the same base, or there can be interbase CT excitations. The cross sections for the latter are known to be smaller. We find that the energy ordering of low-energy excited states depends on the geometry of the pairs.

Cytosine–Guanine. The nonadiabatic dynamics of the CG WC pair has been extensively studied, and therefore presents a good benchmark case for our investigation. In our simulations, the system consisted of 9H-guanine and 1H-cytosine (Figure 1B). A harmonic constraint was introduced between N1_G and N3_C, holding the two atoms within the range from 2.5 to 3.3 \AA apart (at equilibrium, $R(N1_G-N3_C) = 2.9 \text{ \AA}$), in order to keep the two bases paired. Without the constraint, the system always fell apart during the dynamics, as the hydrogen bonding within the Watson–Crick pair was insufficient to keep the pair bound.

Figure 6A shows the results. The vertical excitation energy was $5.37 \pm 0.33 \text{ eV}$. At AM1 (CI2,2), 98.5% of the trajectories indeed underwent IC via the N1–H stretch, or complete N1_G–N3_C proton transfer. The structure with completely transferred proton is not a minimum on the S0 PES, however. Among those, 67.5% of trajectories also involved a skeletal deformation, mainly due to pyramidalization at the N3_C site. CG_{hop} is a characteristic hopping point (Figure 6A). A typical CG_{hop} occurred at $-131.49973 \pm 0.02263 \text{ au}$, which is ca. 0.13 eV below the averaged vertically excited structure (averaged over 442 successful trajectories). The remaining 1.5% of the trajectories involved skeletal deformation alone. IC in CG is very efficient. S1 decays in the time scale of 30 fs, and by the end of 1 ps, 96% of the trajectories end up on S0.

For all studied geometries, the first excited state in CG_v corresponds to charge transfer from G to C. It is illustrated in Figure 6A: the SOMO (49a) is fully localized on G, whereas the newly populated SOMO (50a) is localized solely on C. This structure of MOs is retained in CG_{hop}. The negative charge on C attracts the shuttling proton from G, and this motion straightforwardly leads to the S0–S1 conical intersection.

It is also known from previous reports⁴³ and from our B3LYP/6-311G* calculations that there is a minimum corresponding to a double proton transfer in CG: between N1_G and N3_C and between the amino group in C and the carbonyl in G. However,

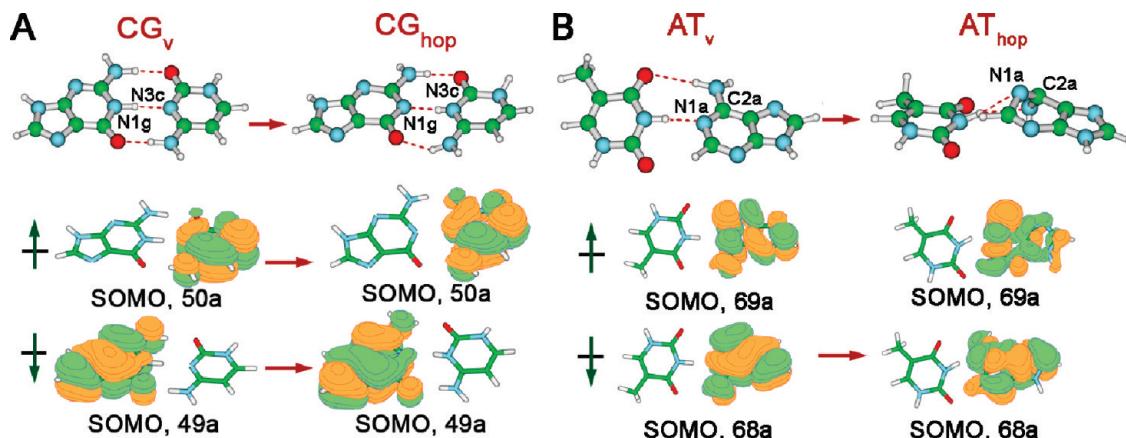


Figure 6. (A) Nonadiabatic semiclassical dynamics in CG: the starting vertically excited structure, CG_v , and a hopping point, CG_{hop} , and the comparison between their MOs. (B) Same information for the AT pair.

the apparently low-probability event of its formation was never observed in our simulations, perhaps due to insufficient statistics. Overall, the mechanism of IC in CG obtained with our semiclassical dynamics simulations appears to be in agreement with previously reported, more accurate, *ab initio* calculations.^{40–44}

Adenine–Thymine. The nature and order of excited states and dynamics in the $(\text{A})_n \cdot (\text{T})_n$ systems have been the subject of some studies,^{18,36,37,40} though IC in this system is less understood than that in CG. Perun et al.³⁷ predicted the involvement of a $\pi\pi^*$ CT state in the IC of the AT WC pair, which occurs via H-transfer. The authors also mentioned the possibility for the dynamics to occur via bending of A or T, when the corresponding π^* -MOs get populated. However, they omitted further exploration of these pathways, and left it to be the subject of future studies. On the other hand, Crespo-Hernandez et al.³⁶ argued that for A and T base stacking is responsible for nonradiative decay, rather than WC pairing.

We found IC in the AT pair to be inefficient. In 1 ps, only 13.7% of the trajectories converted from S1 to S0. For 20 trajectories, the simulations were then extended to 5 ps, during which 55.5% of the S1 population decayed to S0. All 500 trajectories for AT successfully completed, without having SCF convergence problems.

The vertical excitation energy was 4.52 ± 0.23 eV. The main IC route observed 89.4% of the time (in 1 ps trajectories) was characterized only by the out-of-plane motion of the N1_A atom on adenine. The averaged value of dih(N1–C6–C6–C2)_A in A at AT_{hop} was $\pm 41.3 \pm 15.9^\circ$. Figure 6B shows a representative vertically excited structure (AT_v) and a hopping point (AT_{hop}), which was found ca. 1.15 eV below AT_v . 12.5% of the trajectories involved both N1_A going out of plane and an additional smaller out-of-plane deformation of T. However, in those trajectories, the averaged value of dih(N3–C2–C1–C6)_T in T was $\pm 20.7 \pm 4.1^\circ$, suggesting that the deformation of A has a greater relevance to the dynamics. When the system was tested with other Hamiltonians and CI spaces, PM3 CI(2,2), PM3 CI(4,4), the trends in geometrical changes during IC were the same.

At the AM1 level, the S0–S1 excitation in AT is a $\pi\pi^*$ (L_b) excitation localized on A (Figure 6B), and this electronic structure is retained all the way to AT_{hop} , just like in the unpaired A (Figure 3). Consistently, it leads to the observed out-of-plane distortion of A. Thus, the S1–S0 IC path in our simulations involves only A. However, its timing is slowed down by H-bonding with T, as compared to the case of isolated A. We did not find any H-transfer involvement in the dynamics, which was expected, since in our simulations S1 is not a CT state.

Nucleotide Quartets. The larger DNA fragments investigated herein are nucleotide quartets: two stacked Watson–Crick base pairs with the backbone phosphate groups and counterions, 2Na^+ . The presence of the backbone was essential not only for a more realistic representation of the system but also to keep the integrity of the quartets during the dynamics. There are a total of eight possible base arrangements in a quartet; here, we consider only four: CGCG, ATAT, ATTA, and ATCG (Figure 1B). Four geometric constraints were imposed on each system: the distance between the N3 atom of C or T and the N1 atom of G or A, respectively, was kept restrained between 2.5 and 4.0 Å; the stacked bases were kept close via constraining the distance between the N1 atoms of pyrimidine bases and the N9 atoms of the purine bases to fall between 4.0 and 5.5 Å. The semiclassical dynamics simulations for these systems are quite demanding, even at the semiempirical level of theory. For each system, only 20 trajectories were run, 10 at AM1 CI(2,2) and 10 at PM3 CI(2,2), using 5 starting structures. This might be insufficient to be statistically significant. Hence, the results presented herein are qualitative, and no statistics are reported. For nucleotide quartets, AM1 exhibited more SCF convergence problems than did PM3, so we focus on the PM3 CI(2,2) results. We found that, at AM1, the 3s atomic orbitals on Na^+ cations can lay as low in energy as the LUMO of quartets, which is not physical. These orbitals can get populated in the dynamics, and lead to erroneous results.

CGCG Quartet. At both AM1 and PM3, the S1 PES in CGCG exhibited a repulsive character along the out-of-plane bending mode of one of the G bases, labeled “G” in Figure 7. No other noticeable geometric changes were observed at the hopping points. According to MO analysis, at a typical hopping point, CGCG_{hop} , the structural deformation is associated with the promotion of an electron from the π -HOMO to the LUMO mostly consisting of the lone pair on N of NH_2 and some π^* -character of the base. Both resulting SOMOs are localized on the deforming G. However, at the starting geometry shown in Figure 7, the S0 to S1 excitation is different: it is CT from G to the C π -stacked to it in the strand. The SOMO (285a) switches with the LUMO (286a) in the dynamics. It is remarkable that the MOs near the Fermi level in CGCG_v are closely spaced in energy, their order depends on the starting geometry, and it is different in the initial structures used here but never the same as in CGCG_{hop} . This contrasts what was observed for the CG WC pair, where the S0–S1 excitation is always CT within the pair. Hence, our results suggest that the IC process in DNA may be more complicated than that in an isolated CG pair. First of all, the nature of the first excited state

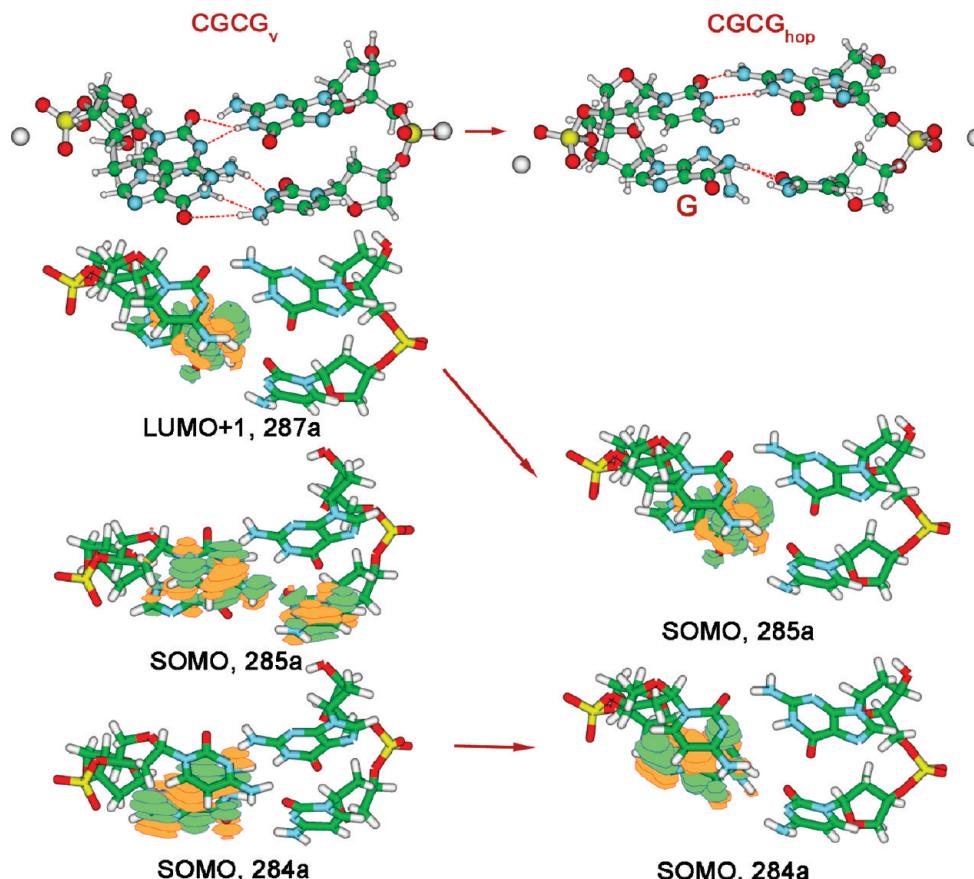


Figure 7. Characteristic structures and MOs involved in the S1–S0 nonadiabatic dynamics in the CGCG nucleotide quartet.

in the vertically excited structure is not always the CT found for CG. Second, the most accessible conical intersection is not the same as that in CG. Our simulations suggest that the IC path localized on a single G base becomes the most prominent in the dynamics, when more than two bases become involved.

Despite much effort, we could not obtain noteworthy results for the **ATAT quartet**. The system exhibited no S1–S0 hops, and lived on the S1 state for the entire duration of the simulation. Most likely, this is another manifestation of the same inefficient IC mechanism as in the AT pair. It is possible that ATAT cannot undergo IC without involvement of another base. It is also possible that the simulations were not long enough to observe IC. However, at least we observed consistent results at AM1 and PM3, with CI(2,2).

For the **ATTA quartet**, no S1–S0 transitions were observed at AM1 CI(2,2). The system again remained on S1 for the entire duration of the simulations. At PM3 CI(2,2), all trajectories involved bending in both of the T bases during IC. A representative hopping point structure is **ATTA_{hop}** (Figure 8). This structure near the S1–S0 conical intersection is reminiscent of that reported, for example, by Boggio-Pasqua et al.,²⁶ who studied photoinduced [2 + 2] cycloaddition T–T, using CASPT2/cc-PVDZ//CASSCF/6-31G* to construct S1 and S0 PESs. The simulations, therefore, seem to recover the general trend in the nonadiabatic dynamics of two stacked T bases. However, we did not observe the formation of the mutant (product of [2 + 2] cycloaddition), which, again, can be due to insufficient statistics, or a distorted topology of the PES as represented in semiempirical methods.

The bending in T is due to the population of its π^* -SOMO in **ATTA_{hop}**. The S0–S1 excitation in this structure is fully localized on one of the deforming T bases (Figure 8). The MOs

of **ATTA_v** and **ATTA_{hop}** are also compared in Figure 8. One may see that the vertical excitation in the **ATTA_v** has a different origin than that in **ATTA_{hop}**: both the SOMO (285a) and the SOMO (284a) are localized on purine bases. The SOMO (285a) in **ATTA_{hop}** is most reminiscent of the LUMO+1 (287a) in **ATTA_v**, and the **ATTA_v** MO most similar to the SOMO (284a) in **ATTA_{hop}** is a mix between π -MOs on both T bases (the HOMO-2). We would like to point out that, in the context of this quartet, the S0–S1 excitation localized on T does not lead to the most prominent IC route that we found for isolated T (that via C1–N12 bond stretch). This might be an indication that that path is hindered in DNA.

The S1–S0 internal conversion in the **ATCG quartet** proceeded via the bending of G, and, to a smaller extent, purine base A, whereas T and C remain stacked against each other, at their near-equilibrium geometry (structure **ATCG_{hop}**, Figure 9). We did not observe the tendency to form a [2 + 2] T–C dimer in our simulations.

At **ATCG_{hop}**, the two SOMOs are π and π^* , localized on bending bases. In the vertically excited structure, **ATCG_v**, however, the S0–S1 excitation entails the transfer of an electron from the π -MO on G to the π^* -MO on C. The AO contributions to the SOMOs of **ATCG_{hop}** are spread among several different MOs in **ATCG_v**, and no direct correspondence can therefore be reported. Hence, S1 changes its origin multiple times before the conical intersection with S0 and decay. Noticeably, in various systems containing G, IC is very prone to proceeding via bending of G. Whether this is realistically one of the most facile IC mechanisms in DNA or just an artifact of the methodology remains to be seen.

It is worth mentioning again that the uncertainty associated with the dynamics of the quartets is greater, because of the

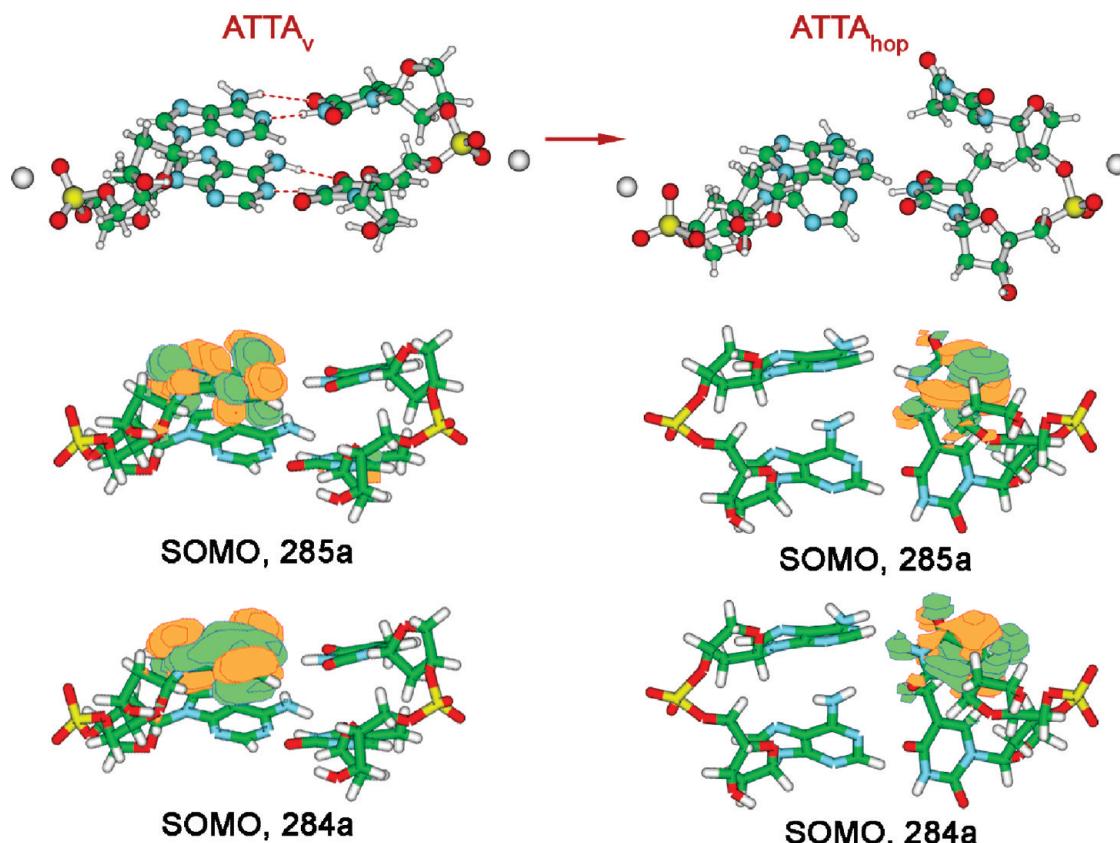


Figure 8. Characteristic structures and MOs involved in the S1–S0 nonadiabatic dynamics in the ATTA nucleotide quartet.

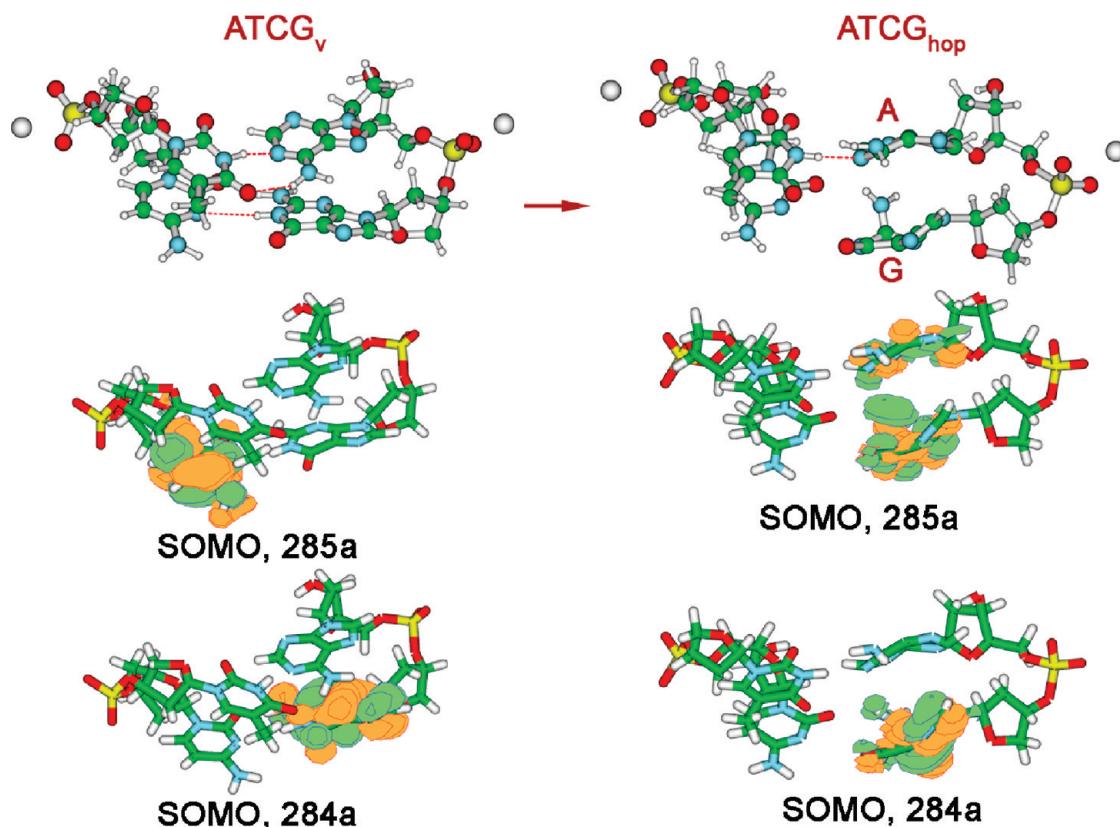


Figure 9. Characteristic structures and MOs involved in the S1–S0 nonadiabatic dynamics of the ATCG nucleotide quartet.

limited statistics, multiple imposed constraints, and a larger number of closely spaced excited states that might not be well-described by semiempirical methods. Therefore, these results should be taken with caution. Nevertheless, we recovered known

routes of IC for ATTA, suggested new IC mechanisms for CGCG and ATCG, and found that IC in ATAT is very inefficient. We also would like to challenge the use of just an isolated CG pair as a model for IC in real DNA. We hope our

results will inspire further exploration of nonadiabatic processes in these and larger and more realistic DNA fragments.

Conclusions

The photochemistry of small DNA fragments, one to four bases, was investigated herein via semiclassical nonadiabatic dynamics, with semiempirical Hamiltonians, and in conjunction with the fewest switches surface hopping algorithm. Sample starting structures of these fragments were taken from Monte Carlo simulations, with the intention to represent them in a more realistic way. The structures did not correspond to the optimized global minima of the fragments but instead to those occurring in a flexible DNA double helix. The use of semiempirical electronic structure methods allowed for simulating 400–500 unique 1 ps trajectories for each system, thus partly accounting for the known sensitivity of the dynamics in DNA to the excess energy in the vertically excited structures. To date, every theoretical method used for assessing the dynamics in DNA has limitations, ours not being an exception. However, the advantage of our approach is considerable sampling and fairly long trajectories. The qualitative trends in the evolution of the first singlet excited state, S1, were revealed. In most cases, our results showed a qualitative agreement with results obtained at higher levels of theory, in cases when those were available from previous works. We also found a few new potential routes of IC and mutagenesis in these fragments, which need to be subjected to studies at higher levels of theory.

(1) The S1 state in single purine 4'H-nucleosides, G and A, evolved via the ring puckering of the base. The difference is that in G IC is more facile than that in A.

(2) In T and C, the IC was found to mostly proceed via the stretch of the N1–C2 bond, accompanied by the skeletal bending of the base, whereas the solo base puckering routes play a lesser role. There are also minima of S0 corresponding to the open N1–C2 bond. The N1–C2 stretch would probably be hindered in the context of DNA, or even a WC pair. This new route of decay is in disagreement with previous findings, and we do encourage skepticism and further studies at higher levels of theory. At the same time, our method does recover the known IC paths in the isolated T and C bases: the mechanism involves various ring puckering modes, and never a N1–C2 bond opening. Hence, the new IC route seems to be adopted due to the presence of the sugar. Furthermore, we found that the sugar moiety contributes to the MOs relevant to the dynamics in the nucleoside, and that correlates with its active role in the dynamics. The sugar was never included in any earlier studies, and we could like to bring up a warning that the sugar might actually be important.

(3) The CG base pair undergoes IC via H⁺-shuttling between the bases accompanied by skeletal deformations. This is in agreement with extensive studies on this system reported earlier.

(4) The IC in AT is found to be much less efficient, and very slow. It involves bending of A, and the process is completely localized on A, at least when only S1 and S0 are considered in the dynamics.

(5) In CGCG, the IC is fast and proceeds via bending of one of the G bases. Bending of G in general seems to be one of the most efficient mechanisms of decay of an electronically excited state in DNA. Because the quartet adopts this IC path, different from that for the CG pair, we suspect that the isolated CG pair is not the best model for studying IC in real DNA. In ATAT, no IC was observed in the course of the simulations, potentially indicating the inability of this system to undergo IC without CT to another base. In ATTA, IC involved bending of the two

T bases, which is known to lead to the formation of the pyrimidine dimer. We did not observe the formation of this mutagenic product, but the dynamics on S1 seems to be qualitatively correct. In ATCG, IC also involved bending of purine bases, and mostly that of G.

For all processes, the electronic nature of all the initial S0–S1 excitations and the electronic structure at hopping points were successfully elucidated via the MO analysis. The orbital populations were consistent with the nuclear rearrangements in the dynamics and shed light on the electronic origin of the observed IC and mutagenesis processes.

The on-the-fly dynamics reported herein was only possible due to the use of semiempirical Hamiltonians. Such calculations are computationally inexpensive but have a limited accuracy. The results of the present survey are suggestive in nature. However, once the reaction coordinates relevant to the dynamics in each case are known, more accurate *ab initio* calculations can be done, without doing the full dynamics on the fly. Several new routes on S1–S0 dynamics in DNA fragments found in this work have initiated further exploration at higher levels of theory, which will be the subject of future publications.

Acknowledgment. Gratitude is expressed to the American Cancer Society (PF-08-165-01GMC) and to Yale University.

References and Notes

- (1) Kang, H.; Jung, B.; Kim, S. K. *J. Chem. Phys.* **2003**, *118*, 6117.
- (2) Canuel, C.; Mons, M.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.
- (3) (a) Pecourt, J.-M. L.; Peon, J.; Kohler, B. *J. Am. Chem. Soc.* **2000**, *122*, 9348. (b) Pecourt, J.-M. L.; Peon, J.; Kohler, B. *J. Am. Chem. Soc.* **2001**, *123*, 10370.
- (4) Crespo-Hernandez, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
- (5) Barbatti, M.; Lischka, H. *J. Am. Chem. Soc.* **2008**, *130*, 6831.
- (6) Yamazaki, S.; Kato, S. *J. Am. Chem. Soc.* **2007**, *129*, 2901.
- (7) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.
- (8) Chen, H.; Li, S. H. *J. Phys. Chem. A* **2005**, *109*, 8443.
- (9) (a) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *J. Am. Chem. Soc.* **2004**, *126*, 2262. (b) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796.
- (10) Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210.
- (11) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
- (12) (a) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Chem.—Eur. J.* **2006**, *12*, 6559. (b) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
- (13) Febiano, E.; Thiel, W. *J. Phys. Chem. A* **2008**, *112*, 6859.
- (14) Conti, I.; Garavelli, M.; Orlandi, G. *J. Am. Chem. Soc.* **2009**, *131*, 16108.
- (15) Sobolewski, A. L.; Domcke, W. *Eur. Phys. J. D* **2002**, *20*, 369.
- (16) Chen, H.; Li, S. *J. Chem. Phys.* **2006**, *124*, 154315.
- (17) Yamazaki, S.; Domcke, W.; Sobolewski, A. L. *J. Phys. Chem. A* **2008**, *112*, 11965.
- (18) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473.
- (19) (a) Kanger, H.; Doltsinis, N. L. *Chem. Phys. Phys. Chem.* **2003**, *5*, 4516. (b) Langer, H.; Doltsinis, N. L.; Marx, D. *ChemPhysChem* **2005**, *6*, 1734.
- (20) Lan, Z.; Fabiano, E.; Thiel, W. *ChemPhysChem* **2009**, *10*, 1225.
- (21) Kanger, H.; Doltsinis, N. L. *Chem. Phys. Phys. Chem.* **2004**, *6*, 2742.
- (22) Merchan, M.; Serrano-Andres, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
- (23) Hare, P. M.; Crespo-Hernandez, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 435.
- (24) Markwick, P. R. L.; Doltsinis, N. L. *J. Chem. Phys.* **2007**, *126*, 175102.
- (25) Markwick, P. R. L.; Doltsinis, N. L.; Schlitter, J. *J. Chem. Phys.* **2007**, *126*, 045104.
- (26) Asturiol, D.; Lasorne, B.; Robb, M. A.; Blancafort, L. *J. Phys. Chem. A* **2009**, *113*, 10211.
- (27) Szymczak, J. J.; Barbatti, M.; Hoo, J. T. S.; Adkins, J. A.; Windus, T. L.; Nachtigallová, D.; Lischka, H. *J. Phys. Chem. A* **2009**, *113*, 12686.

- (28) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 13238.
- (29) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martínez, T. J. *J. Phys. Chem. A* **2007**, *111*, 8500.
- (30) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martínez, T. J. *J. Phys. Chem. A* **2007**, *111*, 8500.
- (31) Merchan, M.; Gonzalez-Luque, R.; Clement, T.; Serrano-Andres, L.; Rodriguez, E.; Reguero, M.; Pelaez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
- (32) Zechmann, C.; Barbatti, M. *J. Phys. Chem. A* **2008**, *112*, 8273.
- (33) Lan, Z.; Fabiano, E.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 3548.
- (34) Zgierski, M. Z.; Patchkovskii, S. *J. Chem. Phys.* **2005**, *123*, 081101.
- (35) Kistler, K. A.; Matsika, S. *J. Phys. Chem. A* **2007**, *111*, 2650.
- (36) Blancafort, L. *Photochem. Photobiol.* **2007**, *83*, 603.
- (37) Ismail, N.; Blancafort, L.; Olivucci, M.; Kohler, B.; Robb, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 6818.
- (38) Blancafort, L.; Robb, M. A. *J. Phys. Chem. A* **2004**, *108*, 10609.
- (39) Hudock, H. R.; Martinez, T. J. *ChemPhysChem* **2008**, *9*, 2486.
- (40) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
- (41) Schultz, T.; Samoylova, E.; Radloff, W. I. V. H.; Sobolewski, A. L.; Domcke, W. *Science* **2004**, *206*, 1765.
- (42) Villani, G. *J. Chem. Phys.* **2008**, *128*, 114306.
- (43) Groenhof, G.; Schafer, L. V.; Boggio-Pasqua, M.; Goette, M.; Grubmuller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 6812.
- (44) Sobolewski, A. L.; Domcke, W.; Hattig, C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17903.
- (45) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 9031.
- (46) Shukla, M. K.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 4709.
- (47) Kwok, W.-M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 5131.
- (48) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernandez, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W. B.; Kohler, B. *Science* **2007**, *315*, 625.
- (49) Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L.; Stewart, J. J.; Callis, P. R. *Photochem. Photobiol.* **1994**, *59*.
- (50) Min, J.-H.; Pavletich, N. P. *Nature* **2007**, *449*, 570.
- (51) Masson, F.; Laino, T.; Tavernelli, I.; Rothlisberger, U.; Hutter, J. *J. Am. Chem. Soc.* **2008**, *130*, 3443.
- (52) Serrano-Perez, J. J.; Gonzalez-Ramirez, I.; Coto, P. B.; Merchan, M.; Serrano-Andres, L. *J. Phys. Chem. B* **2008**, *112*, 14096.
- (53) Roca-Sanjuan, D.; Olaso-Gonzales, G.; Gonzales-Ramirez, I.; Serrano-Andres, L.; Mecham, M. *J. Am. Chem. Soc.* **2008**, *130*.
- (54) Boggio-Pasqua, M.; Groenhof, G.; Schefer, L. V.; Grubmuller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996.
- (55) Tachikawa, H.; Kawabata, H. *Chem. Phys. Lett.* **2008**, *462*, 321.
- (56) Hariharan, M.; Lewis, F. D. *J. Am. Chem. Soc.* **2008**, *130*, 11870.
- (57) Mees, A.; Klar, T.; Gnau, P.; Hennecke, U.; Eker, A. P. M.; Carell, T.; Essen, L.-O. *Science* **2004**, *206*, 1789.
- (58) Park, H.; Zhang, K.; Ren, Y.; Nadji, S.; Sinha, N.; Taylor, J.-S.; Kang, G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15965.
- (59) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689.
- (60) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (61) Still, W. C.; Tempezyk, A. L.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
- (62) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.
- (63) Jorgensen, W. L.; Ulmschneider, J. P.; Tirado-Rives, J. *J. Phys. Chem. B* **2004**, *108*, 16264.
- (64) Czader, A.; Bittner, E. R. *J. Chem. Phys.* **2008**, *128*, 035101.
- (65) Emanuele, E.; Zakrzewska, K.; Markovitsi, D.; Lavery, R.; Millié, P. *J. Phys. Chem. B* **2005**, *109*, 16109.
- (66) Crespo-Hernandez, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141.
- (67) Jean, J. M.; Hall, K. B. *Biochemistry* **2002**, *41*, 13152.
- (68) Schwalb, N. K.; Temps, F. *Science* **2008**, *322*, 243.
- (69) Lange, A. W.; Herbert, J. M. *J. Am. Chem. Soc.* **2009**, *131*, 3913.
- (70) Emanuele, E.; Markovitsi, D.; Millié, P.; Zakrzewska, K. *ChemPhysChem* **2005**, *6*, 1387.
- (71) Dewar, M. J. S.; Zoebisch, E. F.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (72) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.
- (73) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061.
- (74) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221.
- (75) Stewart, J. J. P. *MOPAC 2000*; Tokyo, Japan, 1999.
- (76) Ridly, J. E.; Zerner, M. C. *Theor. Chim. Acta* **1976**, *42*, 223.
- (77) Baraldi, I.; Carnevali, A.; Momicchioli, F.; Ponterini, G. *Spectrochim. Acta, Part A* **1993**, *49A*, 471.
- (78) Germain, A.; Millié, P. *Chem. Phys.* **1997**, *219*, 265.
- (79) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Ciosowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
- (80) (a) Voet, D.; Gratzer, W. B.; Cox, R. A.; Doty, P. *Biopolymers* **1963**, *1*, 193. (b) Onidas, D.; Markovitsi, D.; Marguet, S.; Sharonov, A.; Gustavsson, T. *J. Phys. Chem. B* **2002**, *106*, 11367.

JP103322C