THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Probing the Hydrogen-Bonding Environment of Individual Bases in DNA Duplexes with Isotope-Edited Infrared Spectroscopy

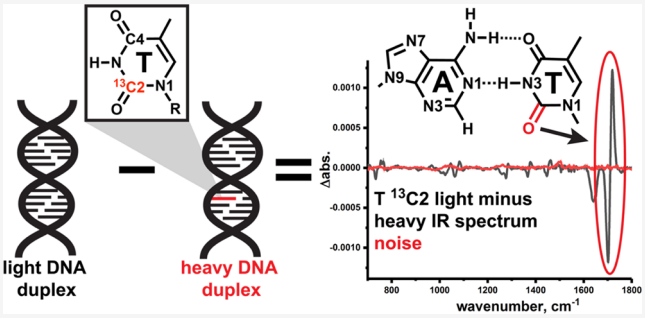
Robert J. Fick, Amy Y. Liu, Felix Nussbaumer, Christoph Kreutz, Atul Rangadurai, Yu Xu, Roger D. Sommer, Honglue Shi, Steve Scheiner, and Allison L. Stelling*

Cite This: *J. Phys. Chem. B* 2021, 125, 7613–7627

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Measuring the strength of the hydrogen bonds between DNA base pairs is of vital importance for understanding how our genetic code is physically accessed and recognized in cells, particularly during replication and transcription. Therefore, it is important to develop probes for these key hydrogen bonds (H-bonds) that dictate events critical to cellular function, such as the localized melting of DNA. The vibrations of carbonyl bonds are well-known probes of their H-bonding environment, and their signals can be observed with infrared (IR) spectroscopy. Yet, pinpointing a single bond of interest in the complex IR spectrum of DNA is challenging due to the large number of carbonyl signals that overlap with each other. Here, we develop a method using isotope editing and infrared (IR) spectroscopy to isolate IR signals from the thymine (T) C2═O carbonyl. We use solvatochromatic studies to show that the TC2═O signal's position in the IR spectrum is sensitive to the H-bonding capacity of the solvent. Our results indicate that C2═O of a single T base within DNA duplexes experiences weak H-bonding interactions. This finding is consistent with the existence of a third, noncanonical CH···O H-bond between adenine and thymine in both Watson–Crick and Hoogsteen base pairs in DNA.



1. INTRODUCTION

Determining the hydrogen bonding (H-bonding) status of individual bonds in nucleic acids is key to understanding their structure, dynamics, and biological functions. This includes recognition by proteins, localized melting required for transcription and replication of DNA, and determining how specific sequences can be targeted with drugs. H-bonds are also critical for determining the complex tertiary structures formed by RNA. Understanding how key H-bonding interactions impact the three-dimensional structure of DNA is essential for determining how our chemical code is recognized by cellular machinery on a molecular level. This knowledge is critical for unraveling processes key to vitality, including replication and transcription, particularly, as it is the dysregulation of these processes that drives genetic diseases such as cancer. Yet, measuring the strength of individual hydrogen bonds (H-bonds) in DNA duplexes between the adenine (A)–thymine (T) and guanine–cytosine base pairs proposed by Watson and Crick[1] (Figure 1A) in the 1950s remains a challenge, particularly in the native solution environment.

While the Protein Data Bank (PDB) contains thousands of DNA and DNA/protein structures at angstrom-level resolution, these structures can be profoundly impacted by the crystal packing. For example, there are a number of AT-rich DNA duplexes that crystallize as all-Hoogsteen base pairs[2,3]

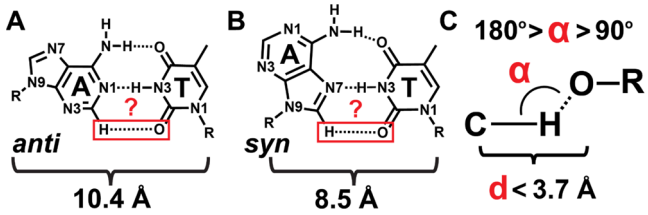


Figure 1. CH···O H-bonding in A–T DNA base pairs. (A) Potential, third CH···O H-bond in Watson–Crick and (B) Hoogsteen A–T DNA base pairs. (C) Typically used geometric definition for CH···O H-bonds.

rather than the canonical Watson–Crick pairs. In Hoogsteen base pairs[4,5] (Figure 1B), the A or G base is flipped 180° from its anti configuration to syn, forming a completely new set of H-bonds with the T or C and exposing the Watson–Crick face of the purine base. However, many of these all-Hoogsteen

Received: February 13, 2021
Revised: May 19, 2021
Published: July 8, 2021

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613–7627

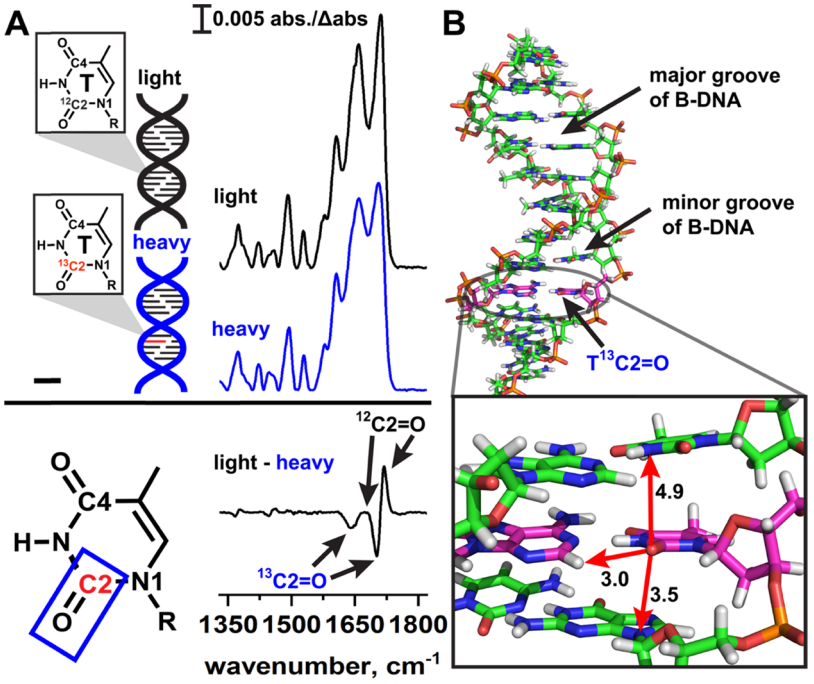


Figure 2. Isotope editing and spectral subtraction isolate isotope signals from individual T bases in DNA duplexes. (A) In isotope editing experiments, an isotopically labeled base is site-specifically incorporated into a DNA duplex (heavy, blue) and the heavy IR spectrum is subtracted from the IR spectrum of a light duplex (black), which has an identical sequence. In the "light−heavy" subtraction spectrum, signals due to the light $^{12}C{=}O$ appear as positive bands, and signals due to the heavy $^{13}C{=}O$ appear as negative bands. Signals that arise from vibrations in which the C2 atom does not participate are canceled out, thus reducing spectral congestion so that C=O stretching of a specific bond of an individual DNA base can be observed in the duplex environment. (B) B-form DNA duplex illustrating the local environment of a TC2=O bond. Arrows indicate nearby functional groups on neighboring bases that may impact the stretching frequency. Distances are in angstrom.

duplexes are in fact Watson−Crick when examined in solution NMR,[6] highlighting a need for methods sensitive to the molecular structure of individual DNA bases that can be performed in the native aqueous environment.

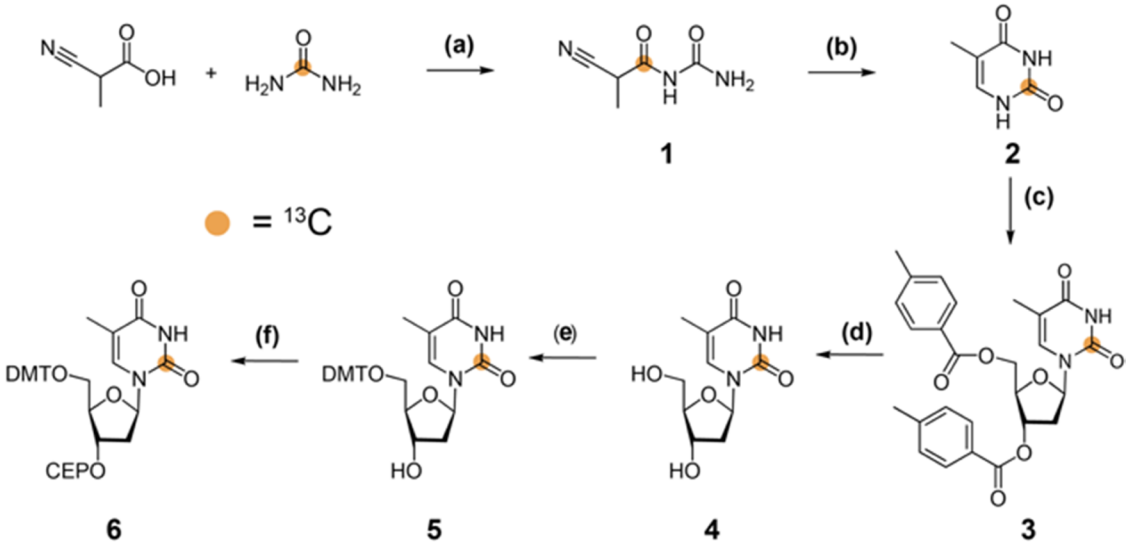
Two of the predominantly used molecularly resolved methods for detecting H-bonding in solution are NMR and vibrational spectroscopies. While NMR is exquisitely sensitive to the structure and electrostatic environment of individual bases, size limitations and chemical exchange render detection and measurement of H-bonds in large DNA/protein complexes challenging. Infrared (IR) spectroscopy is one of the few solution-state methods sensitive enough to changes in H-bonding and molecular structure that it can, for example, distinguish between Watson−Crick and Hoogsteen base pairs, even in large DNA/protein complexes.[7] The frequencies of carbonyl stretching vibrations in particular are well-known[8−11] to change when in strong vs weak H-bonds. C=O stretches can serve as sensitive probes for their immediate electrostatic environment[12−14] and tend to have strong IR intensities that fall in a spectral window clear of signals due to the other molecular bonds. Carbonyl stretches have a long history of use for probing important H-bonding interactions in biological systems using vibrational spectroscopy, particularly as changes in C=O positions can be directly related to the enthalpy of H-bond formation,[10,15−18] provided suitable calibrations are performed. For example, the C=O bond of *S*-acetyl-CoA was used as a probe to measure the degree of H-bonding to the acyl group when the cofactor was bound to α-chymotrypsin.[16]

DNA has been extensively studied by both one-[19−22] and two-dimensional[23−28] IR spectroscopy to probe the ground states and time-resolved methods[29−31] to probe excited electronic states. However, three out of four of the bases contain at least one C=O bond, resulting in a highly overlapped region between 1600 and 1700 cm^{-1} in nucleic acid IR spectra. This renders detection of a particular C=O signal from any single base for use as a molecular probe of the base's local electrostatic environment quite challenging. Here, we overcome this challenge using site-specifically incorporated, isotopically labeled bases and spectral subtraction to dramatically reduce spectral overlap. Isotope labeling is a well-established method for assigning signals to vibrations from particular atomic groups in congested IR and Raman spectra of polyatomic molecules. A multitude of studies have incorporated $^{13}C{=}O$ or $^{13}C{=}^{18}O$ labels into the amide backbone of proteins[32,33] and peptides,[34−38] or of small-molecule enzymatic substrates,[15,39,40] to analyze the hydrogen bonding status of these groups. A recent study used fully ^{15}N- and ^{13}C-labeled DNA bases to examine coupling in G-quadruplexes with ultrafast 2DIR spectroscopy.[41]

A specific application that exploits isotope labeling is isotope editing experiments in which the IR (or Raman) spectrum of the heavy isotope is subtracted from its light counterpart (see the schematic in Figure 2A), canceling out overlapping signals from modes due to other bonds in the molecule. It is more straightforward to ascribe the resulting difference signals to specific functional groups in large molecules, particularly those vibrations which are well-localized to two or three atoms. In nucleic acids, previous studies examined the conformation of adenine bases within DNA duplexes[42] and DNA/protein complexes[43] using deuterium-labeled adenine bases. More recent studies employed isotope editing experiments using commercially available, fully ^{15}N- and ^{13}C-labeled RNA bases

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

Scheme 1. Synthesis of 2-^{13}C-Labeled Thymidine Phosphoramidite[a]



[a](a) Acetic anhydride, 90 °C, 1 h, and 81%; (b) Pd/$BaSO_4$ 5%, H_2, in AcOH/H_2O, 16 h, rt, and 67%; (c) **1**. HMDS, TMS-Cl, 120 °C, 16 h, then Hoffer's α-chlorosugar, in $CHCl_3$, 4 h, 40 °C, and 56%; (d) CH_3NH_2 in EtOH 8 M, 16 h, RT, and 88%; (e) DMT-Cl, DMAP, in pyridine, 3 h, rt, and 71%; and (f) CEP-Cl, DIPEA, in CH_2Cl_2, RT, 3 h, and 74%. DMT, 4,4′-dimethoxytrityl and CEP, 3′-*O*-(*N*,*N*-diisopropyl phosphoramidite).

to track incorporation of a single labeled base using Raman microscopy during an RNA polymerase reaction.[44−46] However, to the best of our knowledge, to date, no isotope editing studies have been performed that label an individual C═O in nucleic acids.

Here, we develop an IR-based isotope editing method to report the first detection of the TC2═O frequency in individual DNA bases when in duplexes containing two different types of A−T base pairs: Watson−Crick[1] and Hoogsteen.[5,47] For these initial experiments, we focused on the C2 atom of T as the C2═O stretch as, based on earlier[48] and more recent[28,49] experimental studies, we expected it to be localized to the C and O atoms, making it a suitable reporter for its immediate chemical environment. The TC2═O bond is located in the minor groove of B-form DNA (Figure 2B) and is therefore expected to be sensitive to interactions with adenine C2−H, any solvent waters or ions that may bind to the minor groove, and stacking interactions from the adjacent bases immediately above or below it.

Here, we present the "light minus heavy" isotope subtraction spectra (Figure 2A) for a series of DNA duplexes, each 12 base pairs in length. These specific sequences were chosen for two main reasons. First, they have been extensively examined with solution NMR methods and, second, two out of the three sequences can form A−T Hoogsteen base pairs when either bound to a drug (echinomycin) or when chemically modified to enforce Hoogsteen pairing, thus allowing us to determine if the stretching is sensitive to changes in the sequence context, in base pairing, and in the minor groove environment. Our third sequence serves as a control for the drug-bound A−T Hoogsteen base pairs as this sequence binds to the drug but maintains Watson−Crick A−T base pairing when bound. Thus, this initial series of duplexes allows us to test the sensitivity of the C2═O bond to two different types of base pairing, Watson−Crick and Hoogsteen, in systems that have been validated to form each under solution conditions by NMR spectroscopy.

We compare the position of the TC2═O stretching frequency in these DNA duplexes to that from the thymine base in solvents with different H-bonding capacities. Collectively, our results indicate that TC2═O in DNA duplexes experiences a weak H-bonding interaction. This may be due to the presence of a weak CH···O H-bond (Figure 1C) in A−T Watson−Crick and Hoogsteen DNA base pairs, which is suggested to exist in computational studies[50,51] but has been vigorously disputed.[52−54] Alternatively, the interaction may arise due to a weakly bound water molecule in the minor groove of the DNA duplexes, which has been observed in crystal structures,[55−63] solution NMR[64,65] and 2DIR[66−69] studies, and molecular dynamics caluclations.[70,71]

Our results are a powerful proof of concept of how isotope editing with labeled DNA bases is a robust, solution-state detection method capable of pinpointing individual chemical bonds encased in very large macromolecular complexes. As carbonyls are present in every nucleic acid save for adenine, we anticipate that this method can be applied to study the electrostatic environment of individual bases in a broad range of DNA- and RNA-containing systems, including their complexes with drugs and proteins.

2. MATERIALS AND METHODS

2.1. Chemicals. Unless otherwise indicated, all chemicals for buffers, thymine base small-molecule mimics, echinomycin, and solvents were purchased from Sigma-Aldrich. Natural-abundance DNA duplexes were purchased from Integrated DNA Technologies. Fully labeled TA-DNA was purchased from Yale-Keck.

2.2. Synthesis of ^{13}C2 -Labeled Thymine Phosphoramidite. Scheme 1 shows the chemical structure of the synthetic intermediates and an overview of the route. See the Supporting Information Methods section for experimental details, yields, and spectroscopic characterization for compounds **1**−**6**.

2.3. IR Experiments. *2.3.1. General.* All DNA IR measurements were performed at DNA concentrations

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

between 5 and 10 mM in a sodium phosphate buffer (25 mM sodium phosphate, 15 mM NaCl) at pH 6.9. IR measurements were performed with a PerkinElmer Frontier FTIR spectrometer using an NB MCT detector, a PIKE MIRacle diamond ATR attachment, and a CaF_2 coverslip as previously described.[7,72] In general, 128 scans at 4 cm^{-1} resolution were taken for each measurement and the average of minimum three measurements was used for averaging in OriginPro. All DNA duplex spectra were baselined with the Spectrum multipoint baselining function and normalized to the phosphate band at ~1080 cm^{-1}. Isotope editing subtractions were performed in the Spectrum (PerkinElmer) software using the subtraction function and the average of at least three independent replicates was exported and used for averaging in OriginPro.

2.3.2. Oligonucleotide Synthesis. Single strands containing the T $^{13}C2$ isotope were synthesized on a MerMade synthesizer using standard protocols.[73] The ^{13}C thymine C2-labeled single strands were synthesized in-house using a MerMade 6 Oligo synthesizer. ^{13}C C2-labeled thymine phosphoramidites, standard DNA phosphoramidites (*n*-ibu-G, *n*-bz-A, *n*-ac-C, and T Chemgenes), and columns (1000 Å from Bioautomation) were used with a coupling time of 1 min, with the final 5′-dimethoxytrityl (DMT) group retained during synthesis. The oligonucleotides were cleaved from the supports (1 μmol) using ~1 mL of AMA (1:1 ratio of ammonium hydroxide and methylamine) for 30 min and deprotected at room temperature for 2 h. The single strands were then purified using Glen-Pak DNA cartridges, with ethanol precipitated and resuspended in water for subsequent annealing.

All unlabeled DNA duplexes were annealed as previously described.[72] Briefly, complementary single-stranded DNA were combined in 0.5 mL of ddH_2O or buffer, allowed to anneal at 95 °C for ~5 min, and cooled at ambient temperature for at least 45 min. The samples were then concentrated and exchanged at least three times into the sodium phosphate buffer using 0.5 mL Amicon (Millipore) with a MW cutoff of 10 kDa. All IR measurements used the flowthrough from the previous exchange as the background spectrum.

2.3.3. Echinomycin Binding. Echinomycin reactions with both TA- and AT-DNA duplexes were performed as previously described[74−80] with minor modifications. Echinomycin was dissolved in methanol (~5 mL methanol per 1 echinomycin) and the concentration was determined via UV−vis spectroscopy using the extinction coefficient at 325 nm (11.5 mM^{-1} cm^{-1}).[81] A three-fold molar excess of echinomycin was used for binding. To ensure binding, DNA was diluted such that it was at an equal volume with the methanol/echinomycin solution and shaken at room temperature for at least 2 h. The mixture was then divided into 1.5 mL Eppendorf tubes and allowed to slowly air-dry overnight. Once dry, the reaction was resuspended using ddH_2O in a volume equal to that of the buffer and concentrated in 0.5 mL Amicon to a volume of ~30 to 50 μL. This was then exchanged into the sodium phosphate buffer at least three times before measurement with IR. The extent of binding was determined by inspection of the symmetric phosphate stretching band at 1085 cm^{-1} in the IR spectra of complexes, which shows a characteristic loss of the shoulders at ~1020 and 1050 cm^{-1} and gain at ~1040 cm^{-1} upon binding. Additional markers for drug binding in the IR include an increase in the intensity in the C═O stretching region at ~1650 cm^{-1} and the red shift to the antisymmetric phosphate stretching by ~6 cm^{-1}.

2.3.4. Solvatochromatic IR Experiments with Thymine Bases. For experiments in water, both dTTP (~100 mM) and a single-stranded poly-T 12-mer (~3 mM) were used to determine the C2═O positions. For dimethylformamide and d_6-DMSO measurements, 1-methylthymine was used to increase solubility. Briefly, 2−3 mg of 1-methylthymine was dissolved by heating at ~95 °C for less than 10 min, and 16 scans were used to avoid sample evaporation. The highest frequency C═O signal, which usually lies between 1650 and 1800 cm^{-1}, was then plotted against the solvent acceptor number and a linear regression was performed in Excel.

2.4. Crystallization of 9-Methyladenine and 1-Methylthymine. Crystals of the dimer were grown by combining equal amounts (~1 mL) of 0.1 M solutions of 9-methyladenine and 1-methylthymine in a 5 dram vial. The volume was reduced by slow evaporation to approximately one-half. The vial was then covered and the cap loosened to allow for further evaporation to 1/4 volume. This process yielded large rod-shaped crystals. Typical crystals had a height and a width of 0.1 and 0.2 mm, respectively, and varied in length from 0.5 to 1.2 mm. A representative crystal (0.096 × 0.212 × 0.572 mm^3) was selected and mounted on a 0.1 mm MiTeGen mount.

Diffraction data were collected using a Bruker-Nonius X8 Kappa APEX II diffractometer and Mo Kα radiation (λ = 0.71073 Å) from a fine-focus sealed tube and graphite monochromator. Corrections were applied using SADABS (Bruker-AXS Inc. (2019)), Madison Wisconsin. The structure was solved using direct methods and refined using least-squares refinement of F^2 and the SHELXTL software package (Bruker-AXS Inc. (2019), Madison Wisconsin). All nonhydrogen atoms were refined anisotropically. Most H atoms were visible in the difference map in the later stages of refinement but were placed at standard calculated positions. Alkyl and aryl H atoms were allowed to ride on the parent atom position with isotropic displacement parameters set to 1.2 and 1.5 times that of the parent atom, respectively. Positions for hydrogen atoms involved in close contacts of interest for this study were allowed to be refined while maintaining the isotropic displacement parameter restraint.

The dimer crystallizes in space group $P2_1/m$. General information is tabulated in Table S1. The molecule sits on the mirror plane. Close contacts between A and T are consistent with those of earlier structures[47,82−84] with negligible variation in intermolecular contacts. For this work, the rotational disorder of the methyl groups was modeled across the mirror plane. Structure determination confirms that the crystals used for this work are indeed consistent with the structure used to analyze the spectroscopic and electronic properties in this study.

2.5. Ab Initio Calculations. Ab initio calculations made use of the MP2 treatment of electron correlation along with the polarized 6-31+G* basis set, within the context of the Gaussian-09 set of codes.[85] Heavy atoms were frozen in their experimentally determined positions, while H atom positions were fully optimized. Interaction energies were evaluated as the difference in energy between each base pair and the energy sum of the individual monomers and then corrected for the basis set superposition error by the standard counterpoise protocol.[86] The densities of the H-bond critical points were evaluated by the AIMALL program.[87]

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

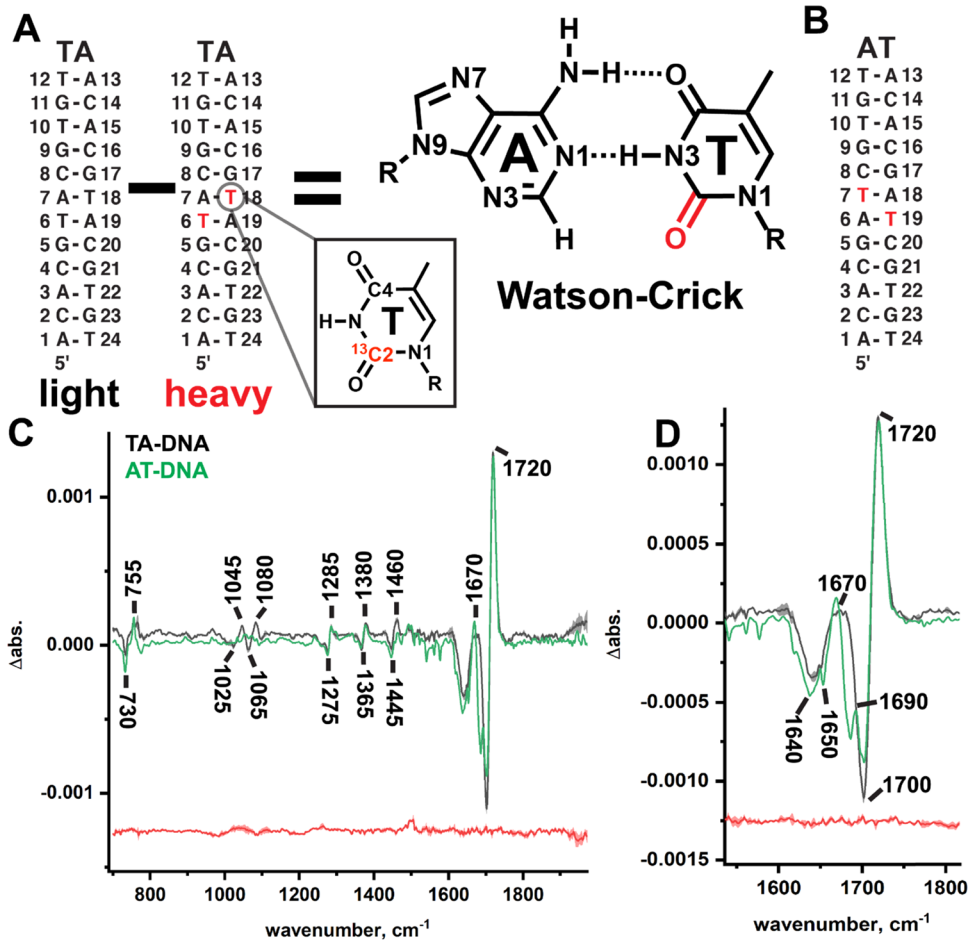


Figure 3. Isotope editing experiments pinpoint TC2═O IR signals in Watson−Crick base pairs in DNA duplexes. The sequence of the palindromes TA-DNA (A) and AT-DNA (B). Labeled T bases are indicated in red. (C) Light minus heavy IR difference spectrum of ^{13}C2-labeled TA-DNA (black) and AT-DNA (green). Positive bands from ^{12}C2═O and negative bands from ^{13}C2═O are labeled. (D) Enhanced view of the carbonyl region in (C) showing splitting in the minor signal of the AT-DNA duplex. A representative noise line (replicate minus replicate for TA-DNA) is shown in red. The mean of four independent measurements is shown in bold, and the standard deviation in shade.

3. RESULTS

3.1. Isotope Editing Pinpoints TC2═O Signals from Individual Bases in Large DNA Duplexes and DNA/Drug Complexe. *3.1.1. Detection of Watson−Crick TC2═O Signals in Two DNA Duplexes with Isotope Editing.* For performing isotope editing experiments, there is a choice between two different T isotopes. The commercially available, fully ^{13}C- and ^{15}N-labeled nucleic acid bases used in previous IR and Raman studies are labeled at every carbon and nitrogen atom, including the sugar and backbone carbons. Isotopically substituting many atoms results in a heavier mass change, which may result in a stronger difference spectrum and thus provide higher contrast in large DNA/protein complexes. However, the difference signal obtained from multiple isotope incorporation may be more complex than that of a duplex containing only one or two labeled atoms due to the multiple mass changes. This mass effect might make ascribing physical interpretations to specific atomic groups, such as an individual C═O bond, very challenging as it would be difficult to determine the extent to which the observed signal is due to a vibration that is localized to the carbonyl of interest. We therefore first sought to compare the difference spectra taken between light and heavy DNA duplexes when using the fully labeled T and a T with only one ^{13}C label inserted at the C2 position, which was chosen as the C2═O stretching vibration is more likely to be localized to the C and O atoms[88] and therefore be a faithful reporter of the H-bonding environment of the minor groove.

We therefore site-specifically installed fully ^{13}C- and ^{15}N-labeled and single-atom ^{13}C2-labeled thymine bases into a TA-DNA duplex (Figure S1). As this sequence is a palindrome (Figure 3A), each duplex now contained two labeled T bases. The spectrum of the unlabeled (light) TA-DNA duplex was subtracted from that of the labeled (heavy) one (Figure S1A) to produce a difference spectrum for each type of label (Figure S1B) in which positive signals are due to the light isotope vibrations, and negative signals from the heavy ones. The difference signal in the fully labeled T is more pronounced but significantly more complex than that of the single-atom ^{13}C- and ^{15}N-labeled T due to the multiple mass changes, rendering any potential physical interpretation of the data challenging. To circumvent this complexity, we decided to develop singly ^{13}C2-labeled T as a probe for the H-bonding environment of A−T base pairs in DNA duplexes using two well-studied[74] palindromes, TA-DNA (Figure 3A) and AT-DNA (Figure 3B). The major positive signal in both ^{13}C2-labeled TA-DNA and AT-DNA (Figures 3C,D and S1) is an intense band at ∼1720 cm^{-1} that we preliminarily assign to the localized stretching of

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

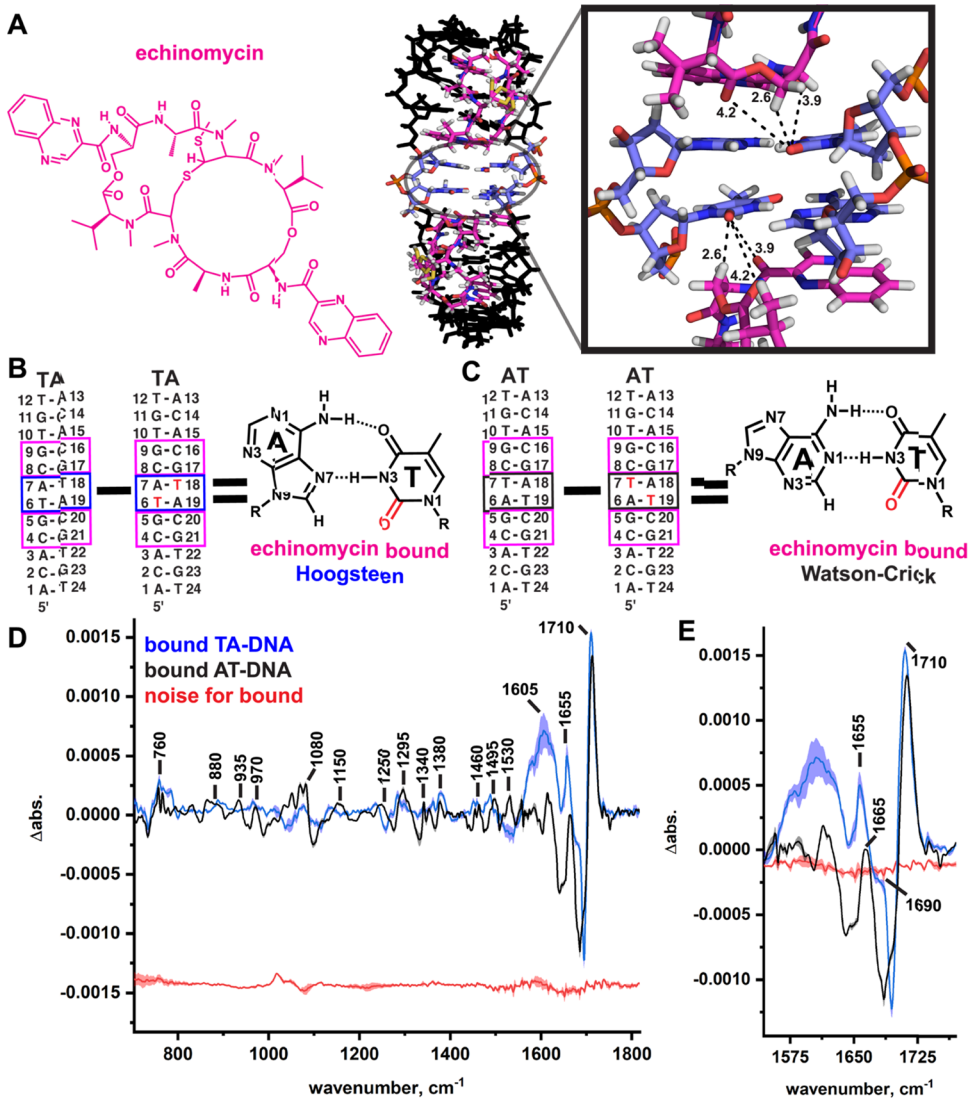


Figure 4. Isolation of the TC2═O stretching signal in a DNA/drug complex with Hoogsteen base pairs. (A) Chemical structure of echinomycin (magenta) and a crystal structure of the bound TA-DNA (PDBID: 1XVN),[91] which induces the formation of two Hoogsteen base pairs (blue). Dashes indicate functional groups from echinomycin close to the labeled carbonyls (distances in angstrom). (B) and (C) GC-step binding sites of echinomycin in TA-DNA are boxed (magenta). Echinomycin was bound to both the light and heavy TA-DNA duplexes and their difference spectra calculated to isolate signals from the two A−T Hoogsteen (T6/T18, blue box) base pairs' TC2═O stretching signals (see the chemical structure; TC2═O is in red). (D) and (E) Light minus heavy IR difference spectrum of the ^{13}C2-labeled DNA/echinomycin complex with the positive signals from the light isotope labeled. Representative noise lines (replicate minus replicate) are shown in red. The mean of three to four independent measurements is shown in bold, and the standard deviation in shade.

TC2═O in a B-form duplex based on values in the experimental literature for isotopically labeled T bases[28,48,49,89] and theoretical calculations.[88] The negative signal at ∼1700 cm^{-1} is downshifted by the expected amount (∼20 cm^{-1}) for a ^{12}C to ^{13}C mass shift,[37,38] which is well in excess of our instrument's resolution (set at 4 cm^{-1}).

We also observe a minor positive signal in each duplex centered at ∼1670 cm^{-1}, whose intensity is roughly one quarter that of the major signal in both duplexes. Interestingly, the two different sequences show differences in this minor signal: in TA-DNA, the signals are broad and in AT-DNA, the signals appear to be splitting into two separate components (Figure 3D). We preliminarily assign the minor signal to delocalized C4═O/C5═C6 stretching with amplitude from the labeled ^{13}C2 atom based on experimental isotope shifts[48] for the isolated T base and theoretical calculations of the A−T base pair.[88] This is consistent with 2DIR studies[24,28] of AT-rich Watson−Crick duplexes that observe four bands in this region due to stretching from both bases, three of which are thought to be due to vibrations localized to the T base and one to adenine. In deuterated buffer, the three from the T base are located at 1690 (C2═O stretching, T_{2R}, which we observe in water as our "major signal" at 1720 cm^{-1}), 1660 (C4═O stretching, T_{4R}), 1640 (T ring stretching, T_R), and 1630 cm^{-1} (adenine ring stretching, A_{R1}). It may be noted that we observe T C5═C6 stretching with our T ^{13}C2 label, resulting in three signals contributing to our difference spectrum. The C═C stretching mode in the T mononucleotide[49] (T_R) was found to be fairly localized to the C5 and C6 atoms, but its character may change upon entry into the duplex environment, allowing for it to be observed with our ^{13}C2 isotope. Thus, our "minor signal" in both duplexes may be a composite band containing

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

two overlapping signals: one from T_R and the other from T_{4S}. This assignment, and the dependence of the splitting of this signal on the sequence context, will be investigated more deeply in future studies using T ^{13}C4-labeled bases.

The T_R stretching mode is also thought to be more sensitive to interstrand base pairing than to intrastrand stacking interactions with its flanking base pairs, as this mode shifts down to 1630 cm^{-1} upon heating past the duplex melting temperature.[28] Yet, as mentioned, our isotope-edited spectra for these two fully base-paired duplexes show a few differences, the most prominent of which can be seen in the minor positive signal centered at 1670 cm^{-1}. In TA-DNA, this signal is quite broad with a band width at half maximum of ∼30 cm^{-1}. In AT-DNA, the minor ∼1670 cm^{-1} signal is split into two components: a dominant one at ∼1665 cm^{-1} and a weak shoulder at ∼1690 cm^{-1} (Figure 3D).

There are a few physical models that might account for the splitting observed in the minor signal. One is that the minor signal splitting in the T base in AT-DNA experiences two separate H-bonding environments at its C2═O, giving rise to two populations of strongly and weakly bound carbonyls with two different band positions. A previous study using NMR chemical shifts and N−H coupling constants[90] has indicated that the hydrogen bond strength of A−T base pairs may be sequence-dependent and thus our AT-DNA sequence may have weak base pair H-bonding, which in turn can result in the breaking of the Watson−Crick H-bond between C4═O and the adenine −NH_2. This results in a small signal due to a population of C4═O that is blue-shifted owing to formation of a nonbonded C4═O. Another possibility that may account for the splitting might arise from the two different sequence contexts experienced by TC4═O in each duplex due to the different sequence contexts for the T base. In TA-DNA, both C2-labeled T bases are adjacent to guanines and are in proximity (about 4 Å) to the guanine C6═O. In AT-DNA, the T bases are flanked with cytosines and are at a much greater distance from the cytosine's C4═O (about 6 Å). If the T_R mode is sensitive to stacking and base pairing, then, different sequence contexts would be expected to shift this mode down, revealing the T_{4S} mode at 1690 cm^{-1} in AT-DNA.

A third potential model is that the sequence context impacts the character of base modes in DNA duplexes, and thus changes in through-bond coupling between T base modes cannot be rejected completely as the potential source of the splitting. Meanwhile, through-bond coupling due to the frequency shift upon mass labeling is unlikely to be the origin for the splitting, as if this was the case, then the splitting ought to be independent of the sequence context. Future studies using a range of sequence contexts and isotope labels, abasic sites, temperature variance, and 2DIR will aid in discriminating between these physical models.

The edited spectra of TA-DNA and AT-DNA overlay quite well below ∼1600 cm^{-1}, although one additional difference is observed: TA-DNA has two weak signals at 1080 and 1045 cm^{-1} that are not present in the spectrum of AT-DNA. Signals at this frequency have been attributed to delocalized T ring motions,[48,89] and it is possible that the sequence context has altered the mode's character such that the C2 atom now contributes a higher percentage of amplitude to the mode in TA-DNA.

3.1.2. Detection of TC2═O IR Signals in Drug-Bound Hoogsteen and Watson−Crick A−T Base Pairs. To test if isotope editing at the T C2 position can isolate signals from atoms in a large DNA/drug complex, and to determine if Hoogsteen base pair formation of A−T base pairs impacts the TC2═O signal, we examined two DNA duplexes when bound to the drug echinomycin (Figure 4A),[74,78,79,91] a peptide inhibitor[92] that possesses anticancer, antiviral, and antibacterial activities. In the TA-DNA duplex, two echinomycin molecules bind to two GC sites (Figure 4A) and induce the formation of tandem A−T Hoogsteen base pairs (Figure 4B) that are sandwiched between the drug-binding sites. In contrast, when echinomycin is bound to the AT-DNA duplex, these tandem AT base pairs maintain Watson−Crick pairing[76,77,79] (Figure 4C). Thus, these two duplexes allow us to test the sensitivity of our edited IR signal to two different types of DNA base pairs, Watson−Crick and Hoogsteen, while simultaneously controlling the presence of the drug as it binds each sequence in an identical manner. Additionally, as echinomycin is an intercalator, binding ought to also remove contributions from stacking: the tandem A−T base pairs are effectively isolated, yet still in a DNA duplex, and this allows us to determine the impact of stacking on the splitting observed in the free duplexes by comparison of the AT-DNA free and drug-bound spectra.

Labeled and unlabeled TA-DNA and AT-DNA duplexes were bound to echinomycin and the light spectrum subtracted from the T ^{13}C2-labeled spectrum. The extent of binding was determined by examining changes in the IR signals from the DNA backbone and C═O bonds from both echinomycin and DNA (Figure S2) that are characteristic of echinomycin binding,[72] specifically, a decrease in the shoulder at ∼1050 cm^{-1}, an increase at ∼1040 cm^{-1}, an increase at 1740 cm^{-1}, and large increases at ∼1700 cm^{-1} and between 1650 and 1610 cm^{-1}. Our results (Figures 3C and S3A) show that the difference signal is strong in both the edited spectra, even in the large DNA/drug complex, confirming that our method can be used to isolate C═O signals from specific bases in large, macromolecular complexes.

The major, most intense positive signal observed in the isotope-edited spectra of both drug-bound duplexes are red-shifted by ∼10 cm^{-1} from their unbound signals from 1720 cm^{-1} down to ∼1710 cm^{-1} (Figures 4D,E and S3A,B). This red shift may be due to duplex unwinding upon drug intercalation, allowing water or buffer salts to more deeply penetrate the minor groove, or due to interactions with the drug itself. As can be seen from the crystal structure, there are two carbonyls from the drug relatively close (∼4 Å) to each TC2═O in TA-DNA (Figure 4A), as well as a −CH_2 group (∼2.6 Å). The bond dipoles of the drug's C═O groups are orientated opposite to those of the TC2═Os. Thus, if they interacted with each other, then a blue shift and not a red shift would be expected.[93] We therefore conclude that the source of the ∼10 cm^{-1} red shift in both drug-bound duplexes relative to the free DNA likely arises from increased H-bonding from solvent waters to C2═O as the minor groove is widened by echinomycin. This red shift may also arise from destacking due to drug intercalation, which effectively "deletes" interactions between the T base and the neighboring base; however, shifts due to stacking are not anticipated to result in such a strong red shift.[94] These hypotheses will be tested further using D_2O exchange in future studies to determine if there is an observable isotope shift and inserting abasic sites flanking the isotopically labeled base, which ought to remove any stacking.

Interestingly, the minor signal observed at ∼1670 cm^{-1} in the free TA-DNA duplex dramatically splits when echinomycin

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

is bound and Hoogsteen base pairs are formed into two components at ~1690 and ~1655 cm^{-1} (see Figure S3A for an overlay of the drug-bound and free TA-DNA). Meanwhile the AT-DNA/echinomycin complex, which is bound to the drug in an identical fashion to TA-DNA but preserves the Watson–Crick base pairs when bound, exhibits a single minor signal at ~1665 cm^{-1} (see Figure S3B for an overlay of the drug-bound and free AT-DNA). As with the free duplexes, the physical origins of this minor-band splitting in the drug-induced Hoogsteen base pairs may arise from different mechanisms (assuming that this mode arises from the T_{4S} and T_R modes, with minor participation from the ^{13}C2-labeled atom). If we ascribe the ~1690 cm^{-1} signal once again to T_{4S}, this would imply that the T_R mode downshifts by ~15 cm^{-1} in the drug-induced Hoogsteen base pair formed in the TA-DNA/drug complex. This is in line with a previous study,[28] indicating that this T ring mode is sensitive to base pairing.

Many 2DIR studies[23,25,26] have illustrated that coupling between vibrational modes is a major determinant of why IR spectroscopy is sensitive to DNA structure and base paring. Splitting in duplexes may arise from both interactions with modes located within flanking bases and interactions with modes located on its base-pairing partner.[25,94] A recent theoretical study[94] indicates that in WC A–T pairs, T modes are coupled to ring C=C bond vibrations in the partnering adenine, and this has also been observed using 2DIR spectroscopy.[24,28,95−97] This coupling would likely be impacted by the flipping of the adenine base from anti to syn in the Hoogsteen base pair, which switches the orientation of the adenine ring. Future experiments on these duplexes using ^{13}C4-labeled T will aid not only in the interpretation of this splitting but also in the validation of its suitability for use as a marker band for A–T Hoogsteen base pairs.

Finally, it is worth noting another signal centered at ~1605 cm^{-1} (broad in TA-DNA and narrow in AT-DNA) that is present in both bound duplexes and absent in the free duplexes, the origins of which are not clear. This may be due to adenine $-NH_2$ scissoring, as computational studies indicate that base modes in general can become highly delocalized upon base pairing[88] and modes can contain contributions from both bases. Thus, our labeled T ^{13}C2 atom may participate in adenine vibrations (such as amino scissoring) that would be expected to be impacted by drug binding. Future studies with ^{13}C4=O isotope and isotopes on the adenine and T ring atoms and complimentary 2DIR studies will aid in the assignment of these minor signals and provide physical interpretations for the observed splitting.

3.1.3. Detection of TC2=O Signals in Singly ^{13}C-Labeled Watson–Crick and m^{1}A-Hoogsteen Base Pairs. As mentioned, both the TA- and AT-DNA duplexes are palindromic and so contain two ^{13}C-labeled carbon atoms. To test if the IR difference signal arising from a single ^{13}C-labeled atom in a large 12-mer DNA duplex is detectable above the noise for the measurement and to examine if the observed frequency is altered when changing the sequence context, we incorporated a ^{13}C2 T base into the A6-DNA duplex (Figure 5A). To determine if the splitting observed in the drug-induced Hoogsteens in TA-DNA arises from the drug or from Hoogsteen formation, we installed an N1-methyladenine (m^{1}A, Figure 5B), which is a known form of DNA damage,[98] opposite the labeled T base. The addition of the methyl on the N1 of adenine (Figure 5B) results in a positive charge being

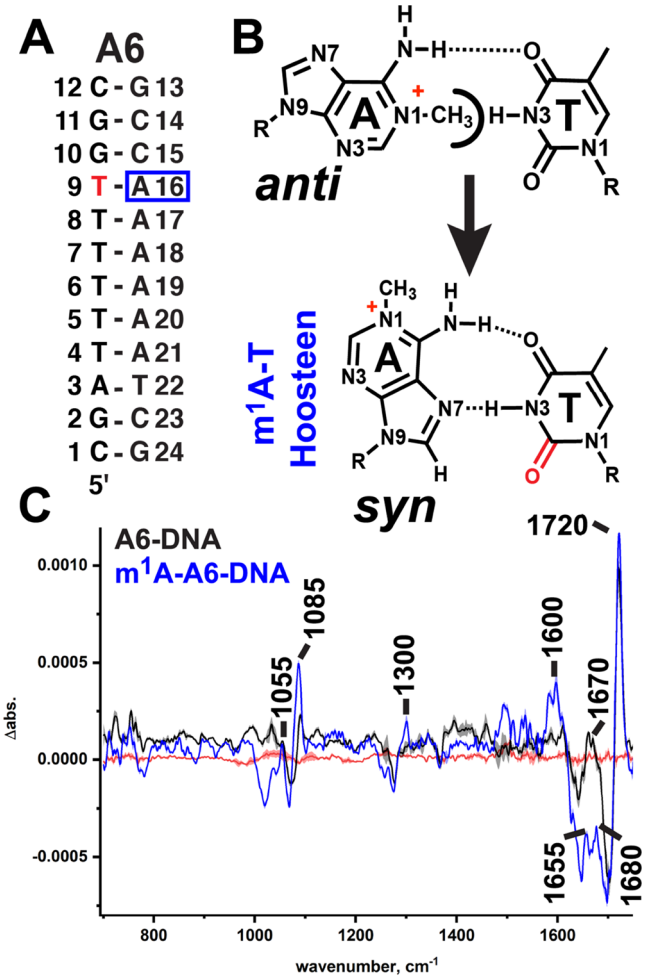


Figure 5. Isotope editing pinpoints a single T atom in two DNA duplexes. (A) Sequence of A6-DNA with the labeled T (T9) in red. (B) Chemical structure of N1-methyladenine (m^{1}A) incorporated at A16, opposite the ^{13}C2-labeled T (red). Methyl group on the adenine N1 has a steric clash with the T N3–H, resulting in the purine flipping over into its Hoogsteen form. (C) Light minus heavy difference spectra of A6-DNA (black) and m^{1}A-A6-DNA (blue).

placed on the base, and a steric clash with the T N3–H forces it to flip 180° and form Hoogsteen H-bonds.[99−106]

In the isotope-edited spectra of T ^{13}C2-labeled A6-DNA, the signal is ~50% that of the doubly labeled T ^{13}C2s in TA and AT-DNA but quite detectable above the noise for the measurement (Figure 5C). The A6-DNA difference spectrum is almost identical to that of TA-DNA (see Figure S4A for overlays of the A6-DNA, TA-DNA, and AT-DNA spectra), with a major signal at 1720 cm^{-1} and a broad minor signal at ~1670 cm^{-1}. In m^{1}A-A6-DNA (Figure 5C), the major signal maintains its position at 1720 cm^{-1}. As in the drug-bound, Hoogsteen-forming TA-DNA duplex, the minor signal at 1670 cm^{-1} is split into two bands (see Figure S4B for an overlay of m^{1}A-A6-DNA with the bound TA-DNA), with one now at ~1680 cm^{-1} and the other at ~1655 cm^{-1}. The ~1680 cm^{-1} band is ~10 cm^{-1} red-shifted from the signal observed in the nonchemically enforced Hoogsteen base pairs formed in the TA-DNA/echinomycin complex. This may be due to the positive charge on the chemically modified m^{1}A base, which could increase the strength of the H-bond between TC4=O

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

and the adenine $-NH_2$ and in turn lower the frequency of the T_{4S} stretch.

Taking together the results from the m[1]A-A6-DNA duplex and from the free TA-DNA and AT-DNA duplexes, we conclude that Hoogsteen H-bonding does not alter the frequency of the major TC2=O stretching signal, whose position is thus far independent of the sequence context and base pairing but does appear to be sensitive to the H-bonding environment in the minor groove. Meanwhile, the strong red shift down to ~1655 cm^{-1} observed in the minor signal in both drug-induced and chemically enforced A−T Hoogsteen base pairs (Figure S4B) is not observed in either drug-bound or free WC base pairs (see Figure S4A for overlays) and may arise solely from either a different coupling from T base modes to adenine ring vibrations in a syn vs an anti adenine or due to increased H-bonding at TC4=O in a chemical-enforced Hoogsteen base pair. Either way, this splitting may therefore be suitable for use as a marker band for detecting protein-bound A−T Hoogsteen pairs in solution. We note that our short duplexes have about equal AT-to-GC content (six A−T pairs in TA- and AT-DNA and seven A−T pairs in A6-DNA), and so studies that dramatically change the AT-to-GC ratios of the sequence along with studies on much longer duplexes must be performed to determine the limits of detection in our isotope-edited spectra.

3.2. Sensitivity of TC2=O IR Signals to the Solvent Environment. *3.2.1. Solvatochromatic Studies of TC2=O and Its Sensitivity to Solvent H-Bonding.* The isotope editing experiments provide a means to measure the frequency of TC2=O in solution-state DNA duplexes, but calibration is required to determine if this band position changes as a function of the strength of H-bonding to C2=O. To provide an initial interpretation for the relative strength of interactions at TC2=O in a duplex environment, we measured IR spectra of the T base in solvents with different H-bonding abilities. We also crystallized the 9-methyladenine/1-methylthymine dimer from Hoogsteen's original paper[5,47] and measured its IR spectrum. (The IR spectrum of Hoogsteen A−T crystals was examined in earlier studies;[107,108] however, the sample was either too thick for a transmission measurement in the 1800−1300 cm^{-1} region[109] or only the amide stretching region was examined.)[108] If this CH···O H-bond exists in Watson−Crick and/or Hoogsteen base pairs, then the frequency of C2=O ought to be greater than 1720 cm^{-1} in solvents with decreased H-bonding power, as measured by Gutmann's H-bond acceptor number (AN)[110] for the given solvent. Our solvatochromatic results are in line with those of previous studies using Raman spectroscopy,[111] reporting that both T C=O stretching signals are sensitive to solvent H-bond strength.

The highest frequency signal of the T base (when not in a duplex) has been shown via isotope labeling studies to arise from localized C2=O stretching in its Raman and IR spectra.[48,89] We plotted the highest positions (Figure S5A,B) vs the H-bond acceptor number of the solvent in which T was dissolved (Figure 6). We used the equation obtained from a linear regression of these data to calculate the AN values for the band positions that we observed via T ^{13}C2 isotope editing in the DNA duplexes, the 9-methyladenine/1-methylthymine Hoogsteen crystal, and the isolated N1-methylthymine's observed IR position in an argon matrix (from ref[112] 112). We find that in solvents with diminished H-bonding power such as d_6-DMSO (AN = 21), the C2=O signal lies at ~1750 cm^{-1}, while in strong H-bonding solvents like water (AN = 54), the signal is shifted down to ~1695 cm^{-1}.

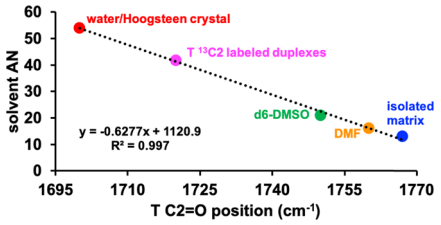


Figure 6. Acceptor numbers for A−T base pairs in DNA duplexes. A plot of the solvent's Gutmann H-bond acceptor number (AN) vs the observed TC2=O stretching. Values for ^{13}C2-labeled duplexes, the Hoogsteen crystal, and the isolated matrix N1-methylthymine were calculated based on the regression line derived from the observed values for the T base in water, dimethylformamide (DMF), and d_6-dimethyl sulfoxide (d_6-DMSO). The isolated matrix value is from N1-methylthymine in the study by Person and co-workers.[112]

In the 9-methyladenine/1-methylthymine dimer crystal, which crystallizes with Hoogsteen H-bonding (Figure S6A,B),[5,47] the IR position of TC2=O is about the same as that observed in water (Figure S6C), which might indicate a very strong CH···O H-bond. However, in our structure, the measured distance between the carbon and oxygen atoms is 3.664 Å, which is just under the typical cutoff distance. Upon inspection of the crystal packing (Figure S6B), we observed another H-bonding interaction between TC2=O and an adenine $-NH_2$ group from an adjacent unit cell, which is likely the source of the strong red shift in the Hoogsteen crystal and supports our hypothesis that TC2=O stretching is a robust probe of the H-bonding environment.

In summary, the frequency of C2=O is shifted down in energy by ~30 cm^{-1} when in a Watson−Crick or a Hoogsteen H-bond in a DNA duplex from its position in d_6-DMSO, indicating that C2=O experiences a H-bond interaction in DNA duplexes in both WC and Hoogsteen base pairs whose strength is less than that of water. This red shift in the paired duplex from the d_6-DMSO and the isolation matrix values is consistent with either a CH···O interaction with the adenine C2/C8−H (Figure 6), which has been observed in nucleic acid crystal structures,[113] or a weakly bound water molecule forming the spine of hydration[55−57,97] in the minor groove of the duplexes.

3.3. Calculations of the CH···O Bond Strength for A−T Base Pairs. *3.3.1. Ab Initio Calculations of the CH···O Bond in A−T Base Pairs.* To investigate the energetics of the CH···O interaction, we performed MP2 correlated calculations using several different A−T base pair coordinates. One set used the PDB coordinates (PDBID: 2ADW) for a DNA sequence bound to echinomycin that contained two duplexes in the unit cell, one of which had WC A−T pairs and the other had Hoogsteen pairs (Figure S7A), so that energies could be effectively compared when in drug-bound DNA. The other set used NMR coordinates[101,114] for the A16−T9 base pair in the A6-DNA (PDBID: 5UZF) and m[1]A-A6-DNA (PDBID: 5UZI, methyl group has been removed) (Figure S7B) duplexes. Interaction energies were assessed as the difference between the energy of each base pair and the sum of the two monomers.

We evaluated the interaction energy of the Watson−Crick and Hoogsteen dimers (Table 1), which were then compared to two controls to estimate the strength of the CH···O bonds.

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

Table 1. Interaction Energies (E_{int}, kcal/mol) for Watson−Crick and Hoogsteen A−T Base Pairs for Echinomycin-Bound DNA (PDBID: 2ADW) and the NMR Structures for A6-DNA (PDBID: 5UZF) and m^{1}A6-DNA (PDBID: 5UZI), along with AIM Electron Densities (au) for the H-Bond Path Critical Points

	E_{int} (kcal/mol)				ρ_{BCP} (au)		
	A−T pair	C−H to N[a]	change	90° rot[b]	A−NH···O−T	T−NH···N−A	A−CH···O−T
WC (2ADW)	13.06	10.51	−2.55	0.64	0.0308	0.0375	0.0047
Hoogsteen (2ADW)	13.79	10.46	−3.33	1.45	0.0242	0.0350	0.0045
WC (5UZF)	12.96	10.72	−2.24	−0.10	0.0281	0.0412	0.0036
Hoogsteen (5UZI)	13.54	9.71	−3.83	0.28	0.0246	0.0427	0.0042

[a]The change of C2/C8−H of adenine to N removes possible CH···O H-bonds. [b]Rotation of T around φ(A−CH···OC−T) to 90° so as to destroy NH···O and NH···N HBs and retain only CH···O.

In one control, the C2/8H···O interaction is abrogated by substitution of the C2 or C8 with a N atom (Figure S7C). This loss of the CH···O H-bond reduces the interaction energy by 2.2−3.8 kcal/mol. In the other control, we rotated the T base by 90° out of the joint plane, around the CH···OC axis, so that the two canonical H-bonds are deleted (Figure S7D) and the CH···O interaction is the only remaining H-bond in the dimer system.

In addition to removing the canonical H-bonds, this rotation also takes the two bases out of their coplanar arrangement and so has an additional destabilizing impact on the system's energy. Nonetheless, the interaction energy remains attractive by as much as 1.5 kcal/mol, which must be attributed to the remaining CH···O H-bond, and is in agreement with similar analyses[115,116] performed on idealized A−T base pairs to determine the relative energy contributions of this bond. Altogether, the interaction energies from both the crystal and NMR (Table 1) coordinates are consistent with the existence of a weak CH···O H-bond. This is consistent with the frequency we observe in the duplexes (1720 cm^{-1}) being greater than the value in water (1700 cm^{-1}) and lower than that in d_6-DMSO (∼1750 cm^{-1}).

An atoms-in-molecules (AIM) analysis of the electron density at various H-bond critical points allows for an independent assessment of H-bond strengths, along with the H-bond lengths. As seen in Table 1, these densities are in the range of 0.025−0.043 au for the pair of canonical H-bonds, with NH···N slightly stronger than NH···O. But of the greatest importance here is that there is a bond path in all four systems linking the CH group with the O, consistent with its characterization as a H-bond. In keeping with the hypothesis that this bond is weaker than its canonical analogues, its critical point density is smaller but nonetheless falls within the range expected for a weak H-bond.

From these results, we conclude that there is a weak CH···O H-bond present in A−T WC and Hoogsteen base pairs. A few DFT studies on A−T WC base pairs have strongly disputed[52−54] the existence of a CH2/8···O H-bond, but later studies have indicated that it is present[117,118] although weak. Our experimental and theoretical results are in line with those of previous theoretical studies that support the existence of a third H-bond in A−T base pairs.

4. DISCUSSION

Here, we describe a powerful method for isolating carbonyl vibrational frequencies from the highly congested IR spectra of DNA duplexes and DNA/drug complexes in aqueous solution. This in conjunction with plots relating the change in band position to the H-bonding strength of a solvent and allowing us to make qualitative statements about the H-bonding status of TC2═O in DNA duplexes in solution. A previous study[119] incorporating modified adenines containing a N atom at the C2 position, which is incapable of H-bonding with TC2═O, reported a decrease in duplex melting temperature by 5 °C per modified base incorporated. The authors state that this may be due to interruption of the hydration spine of water present in duplexes or the repulsive interactions between the N and O atom lone pairs. Their observed decrease in the duplex melting temperature may also be due to interruption of a CH···O H-bond, the loss of which starts to destabilize the duplex. They noted that destabilization appeared to be independent of the sequence, which is consistent with our present results indicating that the position of the major TC2═O signal is unchanged with the sequence context. Base pair melting is an important parameter for cellular processes that require genomic DNA to be single-stranded such as replication, and thus this potential CH···O H-bond may supply a weak, sequence-independent energetic barrier for localized duplex melting.

Additionally, 2DIR studies also indicate that the TC2═O stretching motion may be a good reporter for minor groove interactions (such as with the spine of hydration) in general,[97] and our results showing a red shift in this signal in our drug-bound duplexes support this conclusion. Every TC2═O stretching frequency in naked DNA duplexes that we have measured thus far has a frequency of 1720 cm^{-1} and so the "acceptor number" for TC2═O in a duplex environment (∼40) lies midway between those of DMSO (AN = 21) and water (AN = 54). While the adenine C2/H8 is not as strong as one of the most powerful H-bonders, liquid water, it is significantly stronger than DMSO. We therefore conclude that our observations thus far are consistent with a CH···O H-bond in A−T base pairs which is weak, but present, and could contribute to base-pair stability. Our DFT calculations support this conclusion and indicate that the energy of this bond (as compared to the unpaired A and T monomers) is ∼2 kcal/mol, supporting the conclusion that the bond is weak but certainly present and likely to contribute to the base pair's stability.

It must be noted, however, that this is a qualitative interpretation as calculated T vibrational modes are well-known to change dramatically (particularly in the degree of their delocalization over multiple atoms)[88] when going from a T monomer to the base-paired, duplex environment. Furthermore, at present, we cannot rule out H-bonding interactions with TC2═O from the spine of hydration[55−63] observed in the minor groove in duplex crystal structures whose existence in solution is supported by a number of solution NMR[64,65] and 2DIR[66−69] experiments and molecular dynamics simulations.[70,71] Thus, the position of the C2═O signal in our isotope editing IR experiments may arise from

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

interaction with a weakly bound water molecule in the minor grove. However, the positions of the waters in the spine of hydration are thought to be sequence-dependent.[120−122] If a water molecule from the spine of hydration were to impact the position of the IR signal, then we would have expected increases or decreases in wavenumber as a function of sequence as TC2═O experiences increases or decreases in H-bonding. We do not observe such changes, supporting the conclusion that there is a weak CH···O interaction between TC2═O and adenine.

5. CONCLUSIONS

In summary, we have demonstrated that ^{13}C isotope editing at TC2═O can be used to isolate the stretching frequency of this carbonyl from the congested IR spectrum of DNA duplexes and DNA/drug complexes. Our solvatochromatic experiments combined with our isotope editing results are thus far consistent with a weak CH···O interaction with the TC2═O bond in both Watson−Crick and Hoogsteen A−T base pairs. Future work will employ labeling with isotopes at different positions to assign the major signal to the TC2═O signal in a DNA duplex and variable temperature studies to quantify the relationship between H-bond enthalpy and carbonyl stretches in DNA bases. D_2O exchange studies will be conducted to probe the exposure of TC2═O to solvent waters. These may be combined with isotope editing experiments utilizing C2/8-D labeled adenine to observe mass shifts and determine the extent to which this TC2═O bond is H-bonded to water vs the adenine C2/8H.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c01351.

Synthesis of ^{13}C2-labeled thymine phosphoramidite; crystallographic information for the 9-methyladenine/1-methylthymine dimer; full IR spectrum of unlabeled TA-DNA and the fully ^{15}N, ^{13}C labeled and ^{13}C2 labaled T isotopes; full IR spectra of free and bound AT-DNA; overlays of the isotope-edited spectra for all DNA duplexes examined in this work; observed density map for crystals of the N9-methyladenine and N1-methylthymine Hoogsteen dimers; and models and controls used in the ab initio calculations for Watson−Crick and Hoogsteen A−T base pairs (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Allison L. Stelling − *Department of Chemistry and Biochemistry, The University of Texas at Dallas, Richardson, Texas 75080, United States; Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710, United States;* orcid.org/0000-0002-0368-0852; Phone: (972) 883-6718; Email: stelling@utdallas.edu

Authors

Robert J. Fick − *Department of Chemistry and Biochemistry, The University of Texas at Dallas, Richardson, Texas 75080, United States*

Amy Y. Liu − *Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710, United States*

Felix Nussbaumer − *Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck 6020, Austria*

Christoph Kreutz − *Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck 6020, Austria;* orcid.org/0000-0002-7018-9326

Atul Rangadurai − *Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710, United States*

Yu Xu − *Department of Chemistry, Duke University, Durham, North Carolina 27710, United States*

Roger D. Sommer − *Molecular Education, Technology, and Research Innovation Center, North Carolina State University, Raleigh, North Carolina 27695, United States;* orcid.org/0000-0003-1422-5967

Honglue Shi − *Department of Chemistry, Duke University, Durham, North Carolina 27710, United States;* orcid.org/0000-0003-3847-1652

Steve Scheiner − *Department of Chemistry and Biochemistry, Utah State University, Logan, Utah 84322, United States;* orcid.org/0000-0003-0793-0369

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c01351

Author Contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This work was supported by the National Institutes of Health [1F32GM125213-01 and 5F32GM125213-02 to A.L.S.; A.R., H.S., and Y.X. were supported by R01GM089846], and C.K. was supported by the Austrian Science Fund (FWF, projects P30370 and P32773) and Austrian Research Promotion Agency FFG (West Austrian BioNMR, 858017).

Notes

The authors declare no competing financial interest.

Atomic coordinates for the 9-methyladenine/1-methylthymine crystal have been deposited in the Cambridge Crystallographic Data Centre under deposition number 2053806.

■ ACKNOWLEDGMENTS

We are grateful for critical commentary and many useful discussions with Prof. Al-Hashimi and members of the Al-Hashimi laboratory at Duke University, Prof. Trievel at the University of Michigan (Ann Arbor), and Prof. Horowitz at the University of Denver. We additionally thank Prof. Moran at Southern Illinois University at Carbondale for many useful discussions concerning the impact of vibrational coupling in nucleic acids. This work was performed in part by the Molecular Education, Technology and Research Innovation Center (METRIC) at NC State University, which is supported by the State of North Carolina. A.L.S. thanks the Al-Hashimi laboratory, the Department of Biochemistry at Duke University Medical Center, and the Department of Chemistry and Biochemistry at the University of Texas at Dallas for support.

■ ABBREVIATIONS

T, thymine; A, adenine; C, cytosine; G, guanine; WC, Watson−Crick

REFERENCES

(1) Watson, J. D.; Crick, F. H. C. A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737−738.
(2) Acosta-Reyes, F. J.; Alechaga, E.; Subirana, J. A.; Campos, J. L. Structure of the DNA Duplex d(ATTAAT)2 with Hoogsteen Hydrogen Bonds. *PLoS One* **2015**, *10*, No. e0120241.
(3) Abrescia, N. G. A.; Thompson, A.; Huynh-Dinh, T.; Subirana, J. A. Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2806−2811.
(4) Hoogsteen, K. R. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1963**, *16*, 907−916.
(5) Hoogsteen, K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1959**, *12*, 822.
(6) Abrescia, N. G. A.; Gonzalez, C.; Gouyette, C.; Subirana, J. A. X-ray and NMR Studies of the DNA Oligomer d(ATATAT): Hoogsteen Base Pairing in Duplex DNA. *Biochemistry* **2004**, *43*, 4092−4100.
(7) Stelling, A. L.; Liu, A. Y.; Zeng, W.; Salinas, R.; Schumacher, M. A.; Al-Hashimi, H. M. Infrared Spectroscopic Observation of a G−C+ Hoogsteen Base Pair in the DNA:TATA-Box Binding Protein Complex Under Solution Conditions. *Angew. Chem., Int. Ed.* **2019**, *58*, 12010−12013.
(8) Thijs, R.; Zeegers-Huyskens, T. Infrared and Raman studies of hydrogen bonded complexes involving acetone, acetophenone and benzophenone—II. Raman intensity of the νC═O band. *Spectrochim. Acta, Part A* **1984**, *40*, 1057−1061.
(9) Lewell, X. Q.; Hillier, I. H.; Field, M. J.; Morris, J. J.; Taylor, P. J. Theoretical studies of vibrational frequency shifts upon hydrogen bonding. The carbonyl stretching mode in complexes of formaldehyde. *J. Chem. Soc., Faraday Trans. 2* **1988**, *84*, 893−898.
(10) Rodríguez Ortega, P. G.; Montejo, M.; Valera, M. S.; López González, J. J. Studying the Effect of Temperature on the Formation of Hydrogen Bond Dimers: A FTIR and Computational Chemistry Lab for Undergraduate Students. *J. Chem. Educ.* **2019**, *96*, 1760−1766.
(11) Chen, J.-S.; Wu, C.-C.; Kao, D.-Y. New approach to IR study of monomer−dimer self-association: 2,2-dimethyl-3-ethyl-3-pentanol in tetrachloroethylene as an example. *Spectrochim. Acta, Part A* **2004**, *60*, 2287−2293.
(12) Pazos, I. M.; Ghosh, A.; Tucker, M. J.; Gai, F. Ester Carbonyl Vibration as a Sensitive Probe of Protein Local Electric Field. *Angew. Chem., Int. Ed.* **2014**, *53*, 6080−6084.
(13) Deng, H. Chapter Five - Enzyme Active Site Interactions by Raman/FTIR, NMR, and Ab Initio Calculations. In *Advances in Protein Chemistry and Structural Biology*, Christov, C. Z., Ed.; Academic Press, 2013; Vol. *93*, pp 153−182.
(14) Schneider, S. H.; Boxer, S. G. Vibrational Stark Effects of Carbonyl Probes Applied to Reinterpret IR and Raman Data for Enzyme Inhibitors in Terms of Electric Fields at the Active Site. *J. Phys. Chem. B* **2016**, *120*, 9672−9684.
(15) Tonge, P. J.; Carey, P. R. Length of the acyl carbonyl bond in acyl-serine proteases correlates with reactivity. *Biochemistry* **1990**, *29*, 10723−10727.
(16) Tonge, P. J.; Carey, P. R. Forces, bond lengths, and reactivity: fundamental insight into the mechanism of enzyme catalysis. *Biochemistry* **1992**, *31*, 9122−9125.
(17) Tonge, P. J.; Fausto, R.; Carey, P. R. FTIR studies of hydrogen bonding between α,β-unsaturated esters and alcohols. *J. Mol. Struct.* **1996**, *379*, 135−142.
(18) Latajka, Z.; Scheiner, S. Correlation between interaction energy and shift of the carbonyl stretching frequency. *Chem. Phys. Lett.* **1990**, *174*, 179−184.
(19) Banyay, M.; Sarkar, M.; Gräslund, A. A library of IR bands of nucleic acids in solution. *Biophys. Chem.* **2003**, *104*, 477−488.
(20) Menssen, R. J.; Tokmakoff, A. Length-Dependent Melting Kinetics of Short DNA Oligonucleotides Using Temperature-Jump IR Spectroscopy. *J. Phys. Chem. B* **2019**, *123*, 756−767.
(21) Ashwood, B.; Sanstead, P. J.; Dai, Q.; He, C.; Tokmakoff, A. 5-Carboxylcytosine and Cytosine Protonation Distinctly Alter the Stability and Dehybridization Dynamics of the DNA Duplex. *J. Phys. Chem. B* **2020**, *124*, 627−640.
(22) Wood, B. R. The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues. *Chem. Soc. Rev.* **2016**, *45*, 1980−1998.
(23) Hithell, G.; Ramakers, L. A. I.; Burley, G. A.; Hunt, N. T. Applications of 2D-IR Spectroscopy to Probe the Structural Dynamics of DNA. In *Frontiers and Advances in Molecular Spectroscopy*, Laane, J., Ed.; Elsevier, 2018; Chapter 3, pp 77−100.
(24) Hithell, G.; Shaw, D. J.; Donaldson, P. M.; Greetham, G. M.; Towrie, M.; Burley, G. A.; Parker, A. W.; Hunt, N. T. Long-Range Vibrational Dynamics Are Directed by Watson−Crick Base Pairing in Duplex DNA. *J. Phys. Chem. B* **2016**, *120*, 4009−4018.
(25) Krummel, A. T.; Zanni, M. T. DNA Vibrational Coupling Revealed with Two-Dimensional Infrared Spectroscopy: Insight into Why Vibrational Spectroscopy Is Sensitive to DNA Structure. *J. Phys. Chem. B* **2006**, *110*, 13991−14000.
(26) Krummel, A. T.; Mukherjee, P.; Zanni, M. T. Inter and Intrastrand Vibrational Coupling in DNA Studied with Heterodyned 2D-IR Spectroscopy. *J. Phys. Chem. B* **2003**, *107*, 9165−9169.
(27) Zhang, X.-X.; Brantley, S. L.; Corcelli, S. A.; Tokmakoff, A. DNA minor-groove binder Hoechst 33258 destabilizes base-pairing adjacent to its binding site. *Commun. Biol.* **2020**, *3*, No. 525.
(28) Hithell, G.; Donaldson, P. M.; Greetham, G. M.; Towrie, M.; Parker, A. W.; Burley, G. A.; Hunt, N. T. Effect of oligomer length on vibrational coupling and energy relaxation in double-stranded DNA. *Chem. Phys.* **2018**, *512*, 154−164.
(29) Zhang, Y.; de La Harpe, K.; Beckstead, A. A.; Martínez-Fernández, L.; Improta, R.; Kohler, B. Excited-State Dynamics of DNA Duplexes with Different H-Bonding Motifs. *J. Phys. Chem. Lett.* **2016**, *7*, 950−954.
(30) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Thymine Dimerization in DNA Is an Ultrafast Photoreaction. *Science* **2007**, *315*, 625.
(31) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. DNA Excited-State Dynamics: From Single Bases to the Double Helix. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217−239.
(32) Torres, J.; Kukol, A.; Goodman, J. M.; Arkin, I. T. Site-specific examination of secondary structure and orientation determination in membrane proteins: The peptidic 13C═18O group as a novel infrared probe. *Biopolymers* **2001**, *59*, 396−401.
(33) Decatur, S. M. Elucidation of Residue-Level Structure and Dynamics of Polypeptides via Isotope-Edited Infrared Spectroscopy. *Acc. Chem. Res.* **2006**, *39*, 169−175.
(34) Brielle, E. S.; Arkin, I. T. Quantitative Analysis of Multiplex H-Bonds. *J. Am. Chem. Soc.* **2020**, *142*, 14150−14157.
(35) Scheerer, D.; Chi, H.; McElheny, D.; Keiderling, T. A.; Hauser, K. Isotopically Site-Selected Dynamics of a Three-Stranded β-Sheet Peptide Detected with Temperature-Jump Infrared-Spectroscopy. *J. Phys. Chem. B* **2018**, *122*, 10445−10454.
(36) Setnička, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. IR Study of Cross-Strand Coupling in a β-Hairpin Peptide Using Isotopic Labels. *J. Am. Chem. Soc.* **2005**, *127*, 4992−4993.
(37) Barber-Armstrong, W.; Donaldson, T.; Wijesooriya, H.; Silva, R. A. G. D.; Decatur, S. M. Empirical Relationships between Isotope-Edited IR Spectra and Helix Geometry in Model Peptides. *J. Am. Chem. Soc.* **2004**, *126*, 2339−2345.
(38) Decatur, S. M.; Antonic, J. Isotope-edited infrared spectroscopy of helical peptides. *J. Am. Chem. Soc.* **1999**, *121*, 11914−11915.
(39) Deng, H.; Vedad, J.; Desamero, R. Z. B.; Callender, R. Difference FTIR Studies of Substrate Distribution in Triosephosphate Isomerase. *J. Phys. Chem. B* **2017**, *121*, 10036−10045.

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

(40) Tonge, P. J.; Carey, P. R. Direct observation of the titration of substrate carbonyl groups in the active site of alpha-chymotrypsin by resonance Raman spectroscopy. *Biochemistry* **1989**, *28*, 6701.
(41) Price, D. A.; Kartje, Z. J.; Hughes, J. A.; Hill, T. D.; Loth, T. M.; Watts, J. K.; Gagnon, K. T.; Moran, S. D. Infrared Spectroscopy Reveals the Preferred Motif Size and Local Disorder in Parallel Stranded DNA G-Quadruplexes. *ChemBioChem* **2020**, *19*, 2792−2804.
(42) Toyama, A.; Fujimoto, N.; Hanada, N.; Ono, J.; Yoshimitsu, E.; Matsubuchi, A.; Takeuchi, H. Assignments and hydrogen bond sensitivities of UV resonance Raman bands of the C8-deuterated guanine ring. *J. Raman Spectrosc.* **2002**, *33*, 699−708.
(43) Toyama, A.; Matsubuchi, A.; Fujimoto, N.; Takeuchi, H. Isotope-edited UV Raman spectroscopy of protein−DNA interactions: binding modes of cyclic AMP receptor protein to a natural DNA recognition site. *J. Raman Spectrosc.* **2005**, *36*, 300−306.
(44) Chen, Y.; Eldho, N. V.; Dayie, T. K.; Carey, P. R. Probing Adenine Rings and Backbone Linkages Using Base Specific Isotope-Edited Raman Spectroscopy: Application to Group II Intron Ribozyme Domain V. *Biochemistry* **2010**, *49*, 3427−3435.
(45) Antonopoulos, I. H.; Warner, B. A.; Carey, P. R. Concerted Protein and Nucleic Acid Conformational Changes Observed Prior to Nucleotide Incorporation in a Bacterial RNA Polymerase: Raman Crystallographic Evidence. *Biochemistry* **2015**, *54*, 5297−5305.
(46) Antonopoulos, I. H.; Murayama, Y.; Warner, B. A.; Sekine, S.-i.; Yokoyama, S.; Carey, P. R. Time-Resolved Raman and Polyacrylamide Gel Electrophoresis Observations of Nucleotide Incorporation and Misincorporation in RNA within a Bacterial RNA Polymerase Crystal. *Biochemistry* **2015**, *54*, 652−665.
(47) Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **1963**, *16*, 907−916.
(48) Zhang, S. L.; Michaelian, K. H.; Loppnow, G. R. Vibrational Spectra and Experimental Assignments of Thymine and Nine of Its Isotopomers. *J. Phys. Chem. A* **1998**, *102*, 461−470.
(49) Peng, C. S.; Jones, K. C.; Tokmakoff, A. Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy. *J. Am. Chem. Soc.* **2011**, *133*, 15650−15660.
(50) Leonard, G. A.; McAuleyhecht, K.; Brown, T.; Hunter, W. N. Do C-H...O Hydrogen-Bonds Contribute To The Stability Of Nucleic-Acid Base-Pairs. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 136−139.
(51) Ghosh, A.; Bansal, M. C−H···O hydrogen bonds in minor groove of A-tracts in DNA double helices. *J. Mol. Biol.* **1999**, *294*, 1149−1158.
(52) Srinivasadesikan, V.; Sahu, P. K.; Lee, S.-L. Spectroscopic probe on N−H···N, N−H···O and controversial C−H···O contact in A−T base pair: A DFT study. *Spectrochim. Acta, Part A* **2014**, *120*, 542−547.
(53) Šponer, J.; Leszczynski, J.; Hobza, P. Structures and Energies of Hydrogen-Bonded DNA Base Pairs. A Nonempirical Study with Inclusion of Electron Correlation. *J. Phys. Chem. A* **1996**, *100*, 1965−1974.
(54) Shishkin, O. V.; Sponer, J.; Hobza, P. Intramolecular flexibility of DNA bases in adenine-thymine and guanine-cytosine Watson-Crick base pairs. *J. Mol. Struct.* **1999**, *477*, 15−21.
(55) Shui, X.; Sines, C. C.; McFail-Isom, L.; VanDerveer, D.; Williams, L. D. Structure of the Potassium Form of CGCGAATTCGCG: DNA Deformation by Electrostatic Collapse around Inorganic Cations. *Biochemistry* **1998**, *37*, 16877−16887.
(56) Woods, K. K.; Lan, T.; McLaughlin, L. W.; Williams, L. D. The role of minor groove functional groups in DNA hydration. *Nucleic Acids Res.* **2003**, *31*, 1536−1540.
(57) Schneider, B.; Berman, H. M. Hydration of the DNA bases is local. *Biophys. J.* **1995**, *69*, 2661−2669.
(58) Shakked, Z.; Guzikevich-Guerstein, G.; Frolow, F.; Rabinovich, D.; Joachimiak, A.; Sigler, P. B. Determinants of repressor/operator recognition from the structure of the trp operator binding site. *Nature* **1994**, *368*, 469−473.
(59) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179.
(60) Valls, N.; Wright, G.; Steiner, R. A.; Murshudov, G. N.; Subirana, J. A. DNA variability in five crystal structures of d(CGCAATTGCG). *Spectrochim. Acta, Part A* **2004**, *60*, 680−685.
(61) Edwards, K. J.; Brown, D. G.; Spink, N.; Skelly, J. V.; Neidle, S. Molecular structure of the B-DNA dodecamer d-(CGCAAATTTGCG)2 An examination of propeller twist and minor-groove water structure at 2·2Åresolution. *J. Mol. Biol.* **1992**, *226*, 1161−1173.
(62) Larsen, T. A.; Kopka, M. L.; Dickerson, R. E. Crystal structure analysis of the B-DNA dodecamer CGTGAATTCACG. *Biochemistry* **1991**, *30*, 4443−9.
(63) Vlieghe, D.; Turkenburg, J. P.; Van Meervelt, L. B-DNA at atomic resolution reveals extended hydration patterns. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1495−502.
(64) Liepinsh, E.; Otting, G.; Wüthrich, K. NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution. *Nucleic Acids Res.* **1992**, *20*, 6549−6553.
(65) Halle, B.; Denisov, V. P. Water and monovalent ions in the minor groove of B-DNA oligonucleotides as seen by NMR. *Biopolymers* **1998**, *48*, 210−233.
(66) Fingerhut, B. P.; Elsaesser, T. Noncovalent Interactions of Hydrated DNA and RNA Mapped by 2D-IR Spectroscopy. In *Springer Series in Optical Sciences*, Cho, M., Ed.; Springer Singapore: Singapore, 2019; Vol. *226*, pp 171−195.
(67) Guchhait, B.; Liu, Y.; Siebert, T.; Elsaesser, T. Ultrafast vibrational dynamics of the DNA backbone at different hydration levels mapped by two-dimensional infrared spectroscopy. *Struct. Dyn.* **2015**, *3*, No. 043202.
(68) Siebert, T.; Guchhait, B.; Liu, Y.; Costard, R.; Elsaesser, T. Anharmonic Backbone Vibrations in Ultrafast Processes at the DNA−Water Interface. *J. Phys. Chem. B* **2015**, *119*, 9670−9677.
(69) Yang, M.; Szyc, Ł.; Elsaesser, T. Decelerated Water Dynamics and Vibrational Couplings of Hydrated DNA Mapped by Two-Dimensional Infrared Spectroscopy. *J. Phys. Chem. B* **2011**, *115*, 13093−13100.
(70) Feig, M.; Pettitt, B. M. Modeling high-resolution hydration patterns in correlation with DNA sequence and conformation11Edited by B. Honig. *J. Mol. Biol.* **1999**, *286*, 1075−1095.
(71) Young, M. A.; Jayaram, B.; Beveridge, D. L. Intrusion of Counterions into the Spine of Hydration in the Minor Groove of B-DNA: Fractional Occupancy of Electronegative Pockets. *J. Am. Chem. Soc.* **1997**, *119*, 59−69.
(72) Stelling, A. L.; Xu, Y.; Zhou, H.; Choi, S. H.; Clay, M. C.; Merriman, D. K.; Al-Hashimi, H. M. Robust IR-based detection of stable and fractionally populated G-C+ and A-T Hoogsteen base pairs in duplex DNA. *FEBS Lett.* **2017**, *591*, 1770−1784.
(73) Damha, M. J.; Ogilvie, K. K. Oligoribonucleotide Synthesis. In *Protocols for Oligonucleotides and Analogs*; Springer, 1993; Vol. *20*, pp 81−114.
(74) Xu, Y.; McSally, J.; Andricioaei, I.; Al-Hashimi, H. M. Modulation of Hoogsteen dynamics on DNA recognition. *Nat. Commun.* **2018**, *9*, No. 1473.
(75) Gilbert, D. E.; Feigon, J. Proton NMR study of the [d(ACGTATACGT)]2-2echinomycin complex: conformational changes between echinomycin binding sites. *Nucleic Acids Res.* **1992**, *20*, 2411−2420.
(76) Gilbert, D. E.; Feigon, J. The DNA sequence at echinomycin binding sites determines the structural changes induced by drug binding: NMR studies of echinomycin binding to [d (ACGTACGT)] 2 and [d (TCGATCGA)] 2. *Biochemistry* **1991**, *30*, 2483−2494.
(77) Gao, X.; Patel, D. J. Antitumour drug-DNA interactions: NMR studies of echinomycin and chromomycin complexes. *Q. Rev. Biophys.* **1989**, *22*, 93−138.

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

(78) Gilbert, D. E.; van der Marel, G. A.; van Boom, J. H.; Feigon, J. Unstable Hoogsteen base pairs adjacent to echinomycin binding sites within a DNA duplex. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3006−3010.
(79) Gao, X.; Patel, D. J. NMR studies of echinomycin bisintercalation complexes with d (A1-C2-G3-T4) and d (T1-C2-G3-A4) duplexes in aqueous solution: sequence-dependent formation of Hoogsteen A1-T4 and Watson-Crick T1-A4 base pairs flanking the bisintercalation site. *Biochemistry* **1988**, *27*, 1744−1751.
(80) Xu, Y.; Manghrani, A.; Liu, B.; Shi, H.; Pham, U.; Liu, A.; Al-Hashimi, H. M. Hoogsteen base pairs increase the susceptibility of double-stranded DNA to cytotoxic damage. *J. Biol. Chem.* **2020**, *295*, 15933−15947.
(81) Leng, F.; Chaires, J. B.; Waring, M. J. Energetics of echinomycin binding to DNA. *Nucleic Acids Res.* **2003**, *31*, 6191−6197.
(82) Frey, M. N.; Koetzle, T. F.; Lehmann, M. S.; Hamilton, W. C. Precision neutron diffraction structure determination of protein and nucleic acid components. XII. A study of hydrogen bonding in the purine-pyrimidine base pair 9-methyladenine · 1-methylthymine. *J. Chem. Phys.* **1973**, *59*, 915−924.
(83) King, M. D.; Ouellette, W.; Korter, T. M. Noncovalent Interactions in Paired DNA Nucleobases Investigated by Terahertz Spectroscopy and Solid-State Density Functional Theory. *J. Phys. Chem. A* **2011**, *115*, 9467−9478.
(84) Jarzembska, K. N.; Goral, A. M.; Gajda, R.; Dominiak, P. M. Hoogsteen−Watson−Crick 9-Methyladenine:1-Methylthymine Complex: Charge Density Study in the Context of Crystal Engineering and Nucleic Acid Base Pairing. *Cryst. Growth Des.* **2013**, *13*, 239−254.
(85) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision B.01, Wallingford, CT, 2009.
(86) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(87) Keith, T. A. *AIMALL*, TK Gristmill Software, Overland Park KS, 2013.
(88) Lee, C.; Park, K.-H.; Cho, M. Vibrational dynamics of DNA. I. Vibrational basis modes and couplings. *J. Chem. Phys.* **2006**, *125*, No. 114508.
(89) Yarasi, S.; Billinghurst, B. E.; Loppnow, G. R. Vibrational properties of thymine, uracil and their isotopomers. *J. Raman Spectrosc.* **2007**, *38*, 1117−1126.
(90) Manalo, M. N.; Kong, X.; LiWang, A. JNH Values Show that N1···N3 Hydrogen Bonds Are Stronger in dsRNA A:U than dsDNA A:T Base Pairs. *J. Am. Chem. Soc.* **2005**, *127*, 17974−17975.
(91) Cuesta-Seijo, J. A.; Sheldrick, G. M. Structures of complexes between echinomycin and duplex DNA. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 442−448.
(92) Kong, D.; Park, E. J.; Stephen, A. G.; Calvani, M.; Cardellina, J. H.; Monks, A.; Fisher, R. J.; Shoemaker, R. H.; Melillo, G. Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity. *Cancer Res.* **2005**, *65*, 9047−55.
(93) Schneider, S. H.; Kratochvil, H. T.; Zanni, M. T.; Boxer, S. G. Solvent-Independent Anharmonicity for Carbonyl Oscillators. *J. Phys. Chem. B* **2017**, *121*, 2331−2338.
(94) Jiang, Y.; Wang, L. Modeling the vibrational couplings of nucleobases. *J. Chem. Phys.* **2020**, *152*, No. 084114.
(95) Fritzsch, R.; Greetham, G. M.; Clark, I. P.; Minnes, L.; Towrie, M.; Parker, A. W.; Hunt, N. T. Monitoring Base-Specific Dynamics during Melting of DNA−Ligand Complexes Using Temperature-Jump Time-Resolved Infrared Spectroscopy. *J. Phys. Chem. B* **2019**, *123*, 6188−6199.
(96) Ramakers, L. A. I.; Hithell, G.; May, J. J.; Greetham, G. M.; Donaldson, P. M.; Towrie, M.; Parker, A. W.; Burley, G. A.; Hunt, N. T. 2D-IR Spectroscopy Shows that Optimized DNA Minor Groove Binding of Hoechst33258 Follows an Induced Fit Model. *J. Phys. Chem. B* **2017**, *121*, 1295−1303.
(97) Hithell, G.; González-Jiménez, M.; Greetham, G. M.; Donaldson, P. M.; Towrie, M.; Parker, A. W.; Burley, G. A.; Wynne, K.; Hunt, N. T. Ultrafast 2D-IR and optical Kerr effect spectroscopy reveal the impact of duplex melting on the structural dynamics of DNA. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10333−10342.
(98) Delaney, J. C.; Essigmann, J. M. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14051.
(99) Yang, H.; Zhan, Y.; Fenn, D.; Chi, L. M.; Lam, S. L. Effect of 1-methyladenine on double-helical DNA structures. *FEBS Lett.* **2008**, *582*, 1629−1633.
(100) Yang, H.; Lam, S. L. Effect of 1-methyladenine on thermodynamic stabilities of double-helical DNA structures. *FEBS Lett.* **2009**, *583*, 1548−1553.
(101) Sathyamoorthy, B.; Shi, H.; Zhou, H.; Xue, Y.; Rangadurai, A.; Merriman, D. K.; Al-Hashimi, H. M. Insights into Watson−Crick/Hoogsteen breathing dynamics and damage repair from the solution structure and dynamic ensemble of DNA duplexes containing m1A. *Nucleic Acids Res.* **2017**, *45*, 5586−5601.
(102) Zhou, H.; Kimsey, I. J.; Nikolova, E. N.; Sathyamoorthy, B.; Grazioli, G.; McSally, J.; Bai, T.; Wunderlich, C. H.; Kreutz, C.; Andricioaei, I.; Al-Hashimi, H. M. m(1)A and m(1)G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **2016**, *23*, 803−810.
(103) Zhou, H.; Hintze, B. J.; Kimsey, I. J.; Sathyamoorthy, B.; Yang, S.; Richardson, J. S.; Al-Hashimi, H. M. New insights into Hoogsteen base pairs in DNA duplexes from a structure-based survey. *Nucleic Acids Res.* **2015**, *43*, 3420−3433.
(104) Alvey, H. S.; Gottardo, F. L.; Nikolova, E. N.; Al-Hashimi, H. M. Widespread Transient Hoogsteen Base-Pairs in Canonical Duplex DNA with Variable Energetics. *Nat. Commun.* **2014**, *5*, No. 4786.
(105) Nikolova, E. N.; Gottardo, F. L.; Al-Hashimi, H. M. Probing Transient Hoogsteen Hydrogen Bonds in Canonical Duplex DNA Using NMR Relaxation Dispersion and Single-Atom Substitution. *J. Am. Chem. Soc.* **2012**, *134*, 3667−3670.
(106) Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **2011**, *470*, 498−502.
(107) Kyogoku, Y.; Higuchi, S.; Tsuboi, M. Intra-red absorption spectra of the single crystals of 1-methyl-thymine, 9-methyladenine and their 1:1 complex. *Spectrochim. Acta, Part A* **1967**, *23*, 969−983.
(108) Etter, M. C.; Reutzel, S. M.; Choo, C. G. Self-organization of adenine and thymine in the solid state. *J. Am. Chem. Soc.* **1993**, *115*, 4411−4412.
(109) Kyogoku, Y.; Higuchi, S.; Tsuboi, M. Infrared absorption spectra of the single crystals of 1-methyl-thymine, 9-methyladenine and their 1:1 complex. *Spectrochim. Acta, Part A* **1967**, *23*, 969−983.
(110) Mayer, U.; Gutmann, V.; Gerger, W. The acceptor number — A quantitative empirical parameter for the electrophilic properties of solvents. *Monatsh. Chem.* **1975**, *106*, 1235−1257.
(111) Beyere, L.; Arboleda, P.; Monga, V.; Loppnow, G. R. The dependence of thymine and thymidine Raman spectra on solvent. *Can. J. Chem.* **2004**, *82*, 1092−1101.
(112) Szczepaniak, K.; Szczesniak, M. M.; Person, W. B. Raman and Infrared Spectra of Thymine. A Matrix Isolation and DFT Study. *J. Phys. Chem. A* **2000**, *104*, 3852−3863.

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627

(113) Horowitz, S.; Trievel, R. C. Carbon-Oxygen Hydrogen Bonding in Biological Structure and Function. *J. Biol. Chem.* **2012**, *287*, 41576−41582.
(114) Shi, H.; Clay, M. C.; Rangadurai, A.; Sathyamoorthy, B.; Case, D. A.; Al-Hashimi, H. M. Atomic structures of excited state A−T Hoogsteen base pairs in duplex DNA by combining NMR relaxation dispersion, mutagenesis, and chemical shift calculations. *J. Biomol. NMR* **2018**, *70*, 229−244.
(115) Quinn, J. R.; Zimmerman, S. C.; Del Bene, J. E.; Shavitt, I. Does the A·T or G·C Base-Pair Possess Enhanced Stability? Quantifying the Effects of CH···O Interactions and Secondary Interactions on Base-Pair Stability Using a Phenomenological Analysis and ab Initio Calculations. *J. Am. Chem. Soc.* **2007**, *129*, 934−941.
(116) Yurenko, Y. P.; Zhurakivsky, R. O.; Samijlenko, S. P.; Hovorun, D. M. Intramolecular CH...O Hydrogen Bonds in the AI and BI DNA-like Conformers of Canonical Nucleosides and their Watson-Crick Pairs. Quantum Chemical and AIM Analysis. *J. Biomol. Struct. Dyn.* **2011**, *29*, 51−65.
(117) Souri, M.; Mohammadi, A. K. Investigation of solvent effect on adenine-thymine base pair interaction. *J. Mol. Liq.* **2017**, *230*, 169−174.
(118) Gatti, C.; Macetti, G.; Boyd, R. J.; Matta, C. F. An Electron Density Source-Function Study of DNA Base Pairs in Their Neutral and Ionized Ground States†. *J. Comput. Chem.* **2018**, *39*, 1112−1128.
(119) Sugiyama, T.; Schweinberger, E.; Kazimierczuk, Z.; Ramzaeva, N.; Rosemeyer, H.; Seela, F. 2-Aza-2′-deoxyadenosine: Synthesis, Base-Pairing Selectivity, and Stacking Properties of Oligonucleotides. *Chem. - Eur. J.* **2000**, *6*, 369−378.
(120) Nguyen, B.; Neidle, S.; Wilson, W. D. A Role for Water Molecules in DNA−Ligand Minor Groove Recognition. *Acc. Chem. Res.* **2009**, *42*, 11−21.
(121) Jóhannesson, H.; Halle, B. Minor Groove Hydration of DNA in Solution: Dependence on Base Composition and Sequence. *J. Am. Chem. Soc.* **1998**, *120*, 6859−6870.
(122) Verma, S. D.; Pal, N.; Singh, M. K.; Sen, S. Sequence-Dependent Solvation Dynamics of Minor-Groove Bound Ligand Inside Duplex-DNA. *J. Phys. Chem. B* **2015**, *119*, 11019−11029.

https://doi.org/10.1021/acs.jpcb.1c01351
J. Phys. Chem. B 2021, 125, 7613−7627