ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/physchemau Article

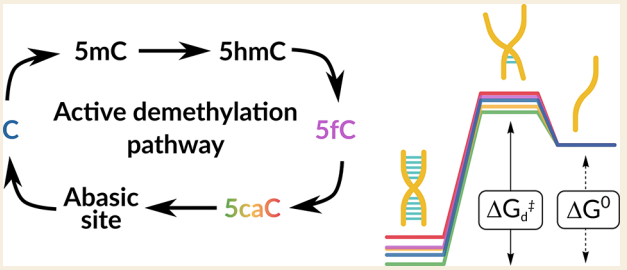
^{1}H NMR Chemical Exchange Techniques Reveal Local and Global Effects of Oxidized Cytosine Derivatives

Romeo C. A. Dubini, Eva Korytiaková, Thea Schinkel,[§] Pia Heinrichs,[§] Thomas Carell, and Petra Rovó*

Cite This: *ACS Phys. Chem Au* 2022, 2, 237−246 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: 5-Carboxycytosine (5caC) is a rare epigenetic modification found in nucleic acids of all domains of life. Despite its sparse genomic abundance, 5caC is presumed to play essential regulatory roles in transcription, maintenance and base-excision processes in DNA. In this work, we utilize nuclear magnetic resonance (NMR) spectroscopy to address the effects of 5caC incorporation into canonical DNA strands at multiple pH and temperature conditions. Our results demonstrate that 5caC has a pH-dependent global destabilizing and a base-pair mobility enhancing local impact on dsDNA, albeit without any detectable influence on the ground-state B-DNA structure. Measurement of hybridization thermodynamics and kinetics of 5caC-bearing DNA duplexes highlighted how acidic environment (pH 5.8 and 4.7) destabilizes the double-stranded structure by ∼10−20 kJ mol^{-1} at 37 °C when compared to the same sample at neutral pH. Protonation of 5caC results in a lower activation energy for the dissociation process and a higher barrier for annealing. Studies on conformational exchange on the microsecond time scale regime revealed a sharply localized base-pair motion involving exclusively the modified site and its immediate surroundings. By direct comparison with canonical and 5-formylcytosine (5fC)-edited strands, we were able to address the impact of the two most oxidized naturally occurring cytosine derivatives in the genome. These insights on 5caC's subtle sensitivity to acidic pH contribute to the long-standing questions of its capacity as a substrate in base excision repair processes and its purpose as an independent, stable epigenetic mark.



KEYWORDS: *DNA, Epigenetic modifications, 5-formylcytosine, 5-carboxycytosine, NMR, Dynamics*

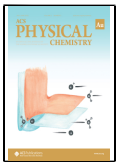
INTRODUCTION

Apart from the four canonical bases, DNA may also contain modified versions of cytosine, thymine, or adenosine nucleosides.[1,2] Such naturally occurring DNA modifications are broadly named epigenetically modified bases, and they constitute an additional regulatory layer by extending genomic complexity and have been shown to play several crucial roles.[3] The discovery of ten-eleven translocation (TET)-induced oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) has been a key catalyst for the exploration and subsequent characterization of both 5-formylcytosine (5fC) and 5-carboxycytosine (5caC).[4]

5fC and 5caC retain a yet undefined number of functional roles, besides being intermediates within the active demethylation pathway,[5,6] as they represent genomically stable, semipermanent modifications with clearly defined tissue distributional patterns.[7−9] Both have been reported as abnormally abundant in prostate, breast and plasma cells cancer.[10−13] Their biological significance is not limited to the initial appearance and progression of diseases, since 5fC and 5caC are transiently accumulated during lineage specification of neural stem cells (NSCs) in culture and *in vivo*,[14] reduce the rate and substrate specificity of RNA polymerase II transcription,[15] or can be selectively recognized by specialized proteins.[16]

Despite extensive research efforts in recent years, it is yet unclear how reader proteins recognize 5fC and 5caC with high specificity and selectivity regardless of their sparse genomic abundance and their chemical similarity to canonical, methylated or hydroxymethylated cytosines. 5fC has been more extensively studied compared to its carboxylated counterpart. Recent research endeavors clarified initial contradictory reports about its impact on structure and stability, leaning toward the notion that 5fC does not affect the B-DNA form.[17−19] As a result of the destabilizing effect that formylation imparts on the C−G base pair, 5fC was found to facilitate melting and hinders annealing, although without affecting the structure to any measurable extent, an aspect which has been investigated via FT-IR and NMR spectroscopy.[20,21]

Received: December 13, 2021
Revised: January 25, 2022
Accepted: January 26, 2022
Published: February 11, 2022



From a structural perspective, it has been recently suggested that 5caC is able to induce a kink in dsDNA, and such geometric alteration has been deemed essential for recognition and enzymatic action by Thymine DNA Glycosylase (TDG).[22] On the other hand, disagreement arises concerning its impact on dsDNA melting and annealing, with studies reporting inconsistent de/stabilization-related properties even when studied under nearly identical conditions.[23−25] While 5fC-induced weakening of the base pair was found to be independent of the mildly basic or acidic conditions naturally occurring within different tissues and/or organisms, with the introduction of the titrable carboxyl group, a pH-dependent behavior emerges for 5caC.[25,26] Protonation/deprotonation events of the carboxyl group induce 5caC to act as an electron-withdrawing group (EWG)/electron-donating group (EDG), respectively. These aspects have been reported to be biologically significant. Indeed, the carboxyl group protonation state has a subtle impact on hydrogen bonding, a behavior that was rationalized via the pK_a values of the two solvent-exposed sites: the nitrogen atom N3 and the carboxyl group itself.[21,26,27] Correspondingly, activity studies on DNA polymerases also concluded that 5caC acts as a base-pair mismatch during DNA replication, signaling that protonated 5caC is a highly destabilizing entity in the context of dsDNA.[28] In addition, structural biology investigations have suggested that the degree of protonation of 5caC's carboxyl group might be a key factor in the mechanism of excision operated by TDG.[26]

Even though other studies have considerably advanced our understanding of epigenetically modified DNA bases within the context of nucleic acids, a number of aspects remain unsettled. In this paper, we seek to shed light on the extent to which 5caC-edited dsDNA's structural features and kinetics-related phenomena deviate from its formylated and canonical equivalents. By employing state-of-the-art methodologies and analytical frameworks in solution-state NMR spectroscopy, we aimed at providing a noninvasive, label-free and site-specific description of the structure and dynamics of these samples using three distinct approaches. First, we compared the impact of the 5caC modification on dsDNA at pH 7.0, 5.8, and 4.7 with respect to 5fC and canonical C by measuring ^{1}H, ^{13}C, and ^{15}N chemical shift perturbations. Second, we applied temperature-dependent ^{1}H chemical exchange saturation transfer (CEST) experiments to assess the impact of 5caC on dsDNA melting and annealing processes providing site-specific parameters for dissociation ($\Delta G^{\ddagger}_{d,37°C}$, $\Delta H^{\ddagger}_{d}$, $\Delta S^{\ddagger}_{d}$) and association kinetics ($\Delta G^{\ddagger}_{a,37°C}$, $\Delta H^{\ddagger}_{a}$, $\Delta S^{\ddagger}_{a}$) and for melting thermodynamics ($\Delta G^{\circ}_{37°C}$, $\Delta H°$, $\Delta S°$).[20,29] Third, we explored the microsecond time scale exchange kinetics via ^{1}H on-resonance $R_{1\rho}$ relaxation dispersion (RD) targeting potential base-specific motions. Watson−Crick (WC) to Hoogsteen (HG) base-pair exchange processes occurring in this regime have been characterized in the context of canonical DNA and protein−DNA complexes,[30−32] but no other exchange phenomena are known to occur in undamaged, non-mismatched DNA helices which are not interacting with a reader protein or enzyme.[33,34]

Collectively, our findings reveal that 5caC's pH-induced chameleonic behavior is not due to any permanent structural changes, opposite to previously reported results.[22] In fact, we observed that the repercussions of 5caC and 5fC incorporation into otherwise canonical dsDNA are uniquely perceptible in the context of conformational dynamics affecting at least two distinct time scales: a slower one (tens of milliseconds) and a faster one (hundreds of microseconds). In addition, we found that carboxycytosine does noticeably affect dsDNA annealing and melting phenomena when exposed to progressively lower pH conditions. Our unified analysis provides a highly comparable and comprehensive overview of the mechanistic details involving both global (such as energetics of DNA strands association and dissociation) and local (site specific motions involving a single base-pair) dynamics phenomena, and sets the stage for future studies of protein−DNA interactions.

RESULTS

We considered the pH-dependent structural features, thermodynamic stability, and site-specific dynamics of a self-complementary 12mer homocarboxylated DNA with the sequence of 5′-GCGATXGATCGC-3′ where X stands for 5caC (Figure 1). The corresponding samples were named

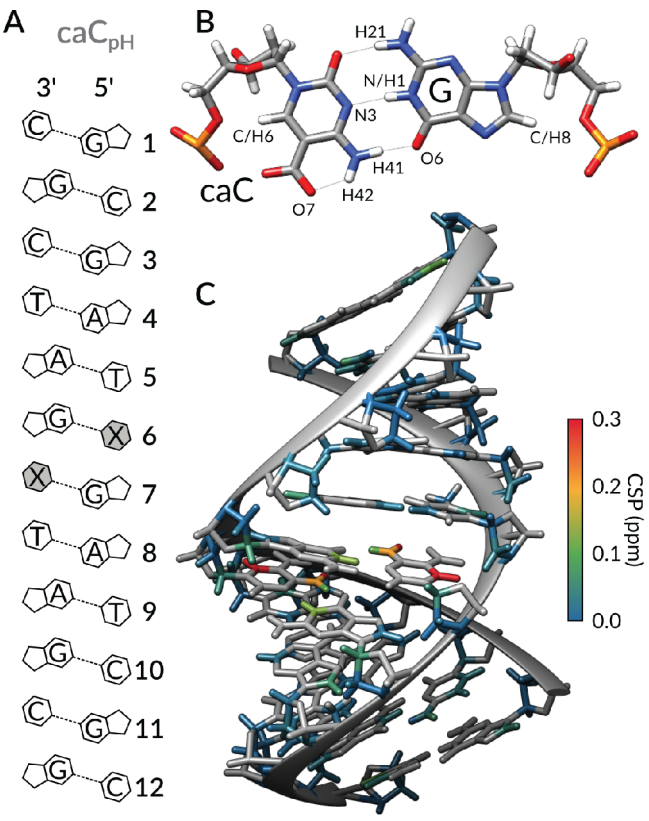


Figure 1. (A) DNA sequence of caC$_{pH}$'s model sequence. The modified cytosine nucleoside is highlighted in gray, marked as X. (B) Structural model showing the expected 5caC−G base pair conformation, consistent with ^{1}H, ^{13}C, and ^{15}N chemical shift values. (C) Absolute CSP values comparing samples at pH 7.0 and 4.7 are schematically displayed onto the B-DNA structural model of caC$_{pH}$.

caC$_{7.0}$, caC$_{5.8}$, and caC$_{4.7}$ indicating the pH at which they were studied. In order to evaluate the influence of cytosine carboxylation in a broader context, we compare the chemical shifts and the NMR-derived thermodynamics, kinetics, and dynamics parameters to analogous values obtained for canonical (C$_{7.0}$) and 5fC-modified (fC$_{7.0}$) samples featuring a sequence identical as what we considered here, which were previously reported in ref 20.

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

Structural Impact by Chemical Shift Analysis

To spot any unusual features induced by inclusion of carboxylated cytosine within otherwise canonical DNA, we recorded a comprehensive set of homo- (^{1}H−^{1}H NOESY) and heteronuclear (^{1}H−^{13}C HSQC, ^{1}H−^{15}N SOFAST HMQC) 2D spectra for resonance assignment and chemical shift analysis purposes.[35]

Previous crystallographic studies have reported contrasting results concerning a potential geometric alteration of the dsDNA helix induced by 5caC.[22,23] Our results, in support of ref 23, indicate that the 5caC nucleobase does not induce any detectable, permanent deviation from the canonical B-DNA structure. At pH 7.0, when compared to the 5fC−G base-pairing interactions, 5caC−G appears to be equally well-tolerated within a canonical double-helix architecture: all of ^{1}H−^{13}C and ^{1}H−^{15}N cross-peaks are superimposable among the canonical, and 5fC and 5caC-containing 12mer DNA constructs, with the sole exception represented by those nuclei in direct proximity to the epigenetic modification (Figures 2, S1, and S2).

An analogous comparison of $caC_{7.0}$ with $caC_{5.8}$ and $caC_{4.7}$ points to no detectable structural change upon acidification. ^{1}H, ^{13}C, and ^{15}N resonances, $^3J_{HH}$ couplings and ^{1}H−^{1}H NOESY cross-peak patterns suggest that carboxylated cytosine leaves the B-DNA structure entirely unperturbed at both neutral and acidic pH values, resulting in no substantial structural difference from the canonical or formylated cytosine-bearing constructs in any studied condition (Figures S3−S6).

Figure 2 compares the perturbations of representative ^{1}H, ^{13}C, and ^{15}N chemical shifts of $fC_{7.0}$, $caC_{7.0}$, $caC_{5.8}$, and $caC_{4.7}$ with respect to the shifts of the canonical 12mer sequence. Figure 2A shows imino protons, which are sensitive reporters of the strength of base-pairing: downfield (upfield) shifted signals, appearing at higher (lower) ppm values, revealing stronger (weaker) intermolecular hydrogen bonds. Interestingly, the only responsive site to pH changes is G7H1, which reports on the X:G base pair, where X is either C, 5fC, or 5caC at the three distinct pH conditions we selected. Substantial chemical shift perturbations are also visible for aromatic ^{1}H sites T5H6 and X6H6, ^{13}C X6C6, and ^{15}N G7N1 (Figures 2B−D and S7−S10). Among those nuclei, the trends suggest that $caC_{4.7}$'s resonances resemble $fC_{7.0}$ the most, while in neutral conditions $caC_{7.0}$ is comparable with canonical $C_{7.0}$.

These pH-induced shifts were most prominent for H-bond forming protons, namely, for caC6H41, H42, and G7H1. The nature of the effect is related to the reshuffling of electron densities around the base-pairing atoms due to altered Pauli repulsion between occupied atomic orbitals.[36] An increased ^{1}H chemical shift is associated with a decreased electron density (^{1}H shielding) around the proton and hence with a stronger H-bond, or shorter H−X distance. As the carboxyl group of 5caC becomes progressively protonated with decreasing pH, the intrabase H-bond between caC6H42 and the carboxyl oxygen O7 gets weaker (caC6H42 $\Delta\delta$ = −0.3 ppm), while the interbase H-bond formed by caC6H41 and G7O6 (Figure 1B) gets stronger (caC6H41 $\Delta\delta$ = +0.1 ppm). Meanwhile, hydrogen bonding between G7H1 and caC6N3 weakens (G7H1 $\Delta\delta$ = −0.15 ppm) due to the decreased electron density at N3 caused by the EWG properties of the protonated carboxyl group, compensating for the increased base-pairing stability gained by the H41−O6 interaction (Figures S7−S10). The fine balance between the strengths of the intra- and interbases hydrogen bonds in the 5caC−G base pair leads to close to optimal base-pairing both at neutral and mildly acidic conditions, leaving the B-DNA structure unperturbed through the entire studied pH range.

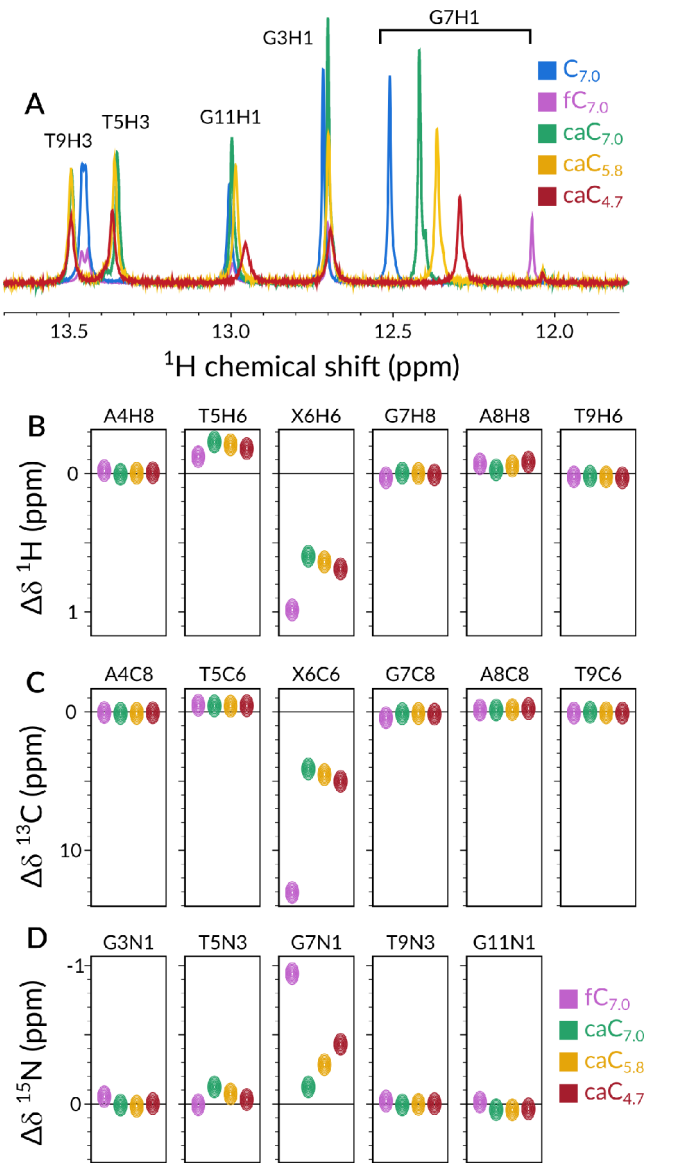


Figure 2. Comparison of chemical shift perturbations $\Delta\delta$ for representative imino ^{1}H (A), aromatic ^{1}H (B), aromatic ^{13}C (C), and imino ^{15}N (D) resonances with respect to the chemical shifts of the canonical 12mer sample (blue spectrum in A, black horizontal line in B−D). Magenta, green, yellow, and red symbols represent $fC_{7.0}$, $caC_{7.0}$, $caC_{5.8}$, and $caC_{4.7}$, respectively. A full comparison of all comparable chemical shifts between the five samples is displayed in Figures S7−S10.

5caC's Influence on DNA Melting and Annealing

A thorough characterization of nucleic acids' global structural rearrangements, such as folding, melting, annealing, and binding should entail a comprehensive analysis of kinetic events occurring on the millisecond to second time scale.[37] CEST experiments have found adoption in modern biomolecular NMR, allowing for quantitative and site-specific determination of population, chemical shift, and exchange kinetics of sparsely populated conformations.[29,38] When measured in a temperature-dependent fashion, the shift in

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

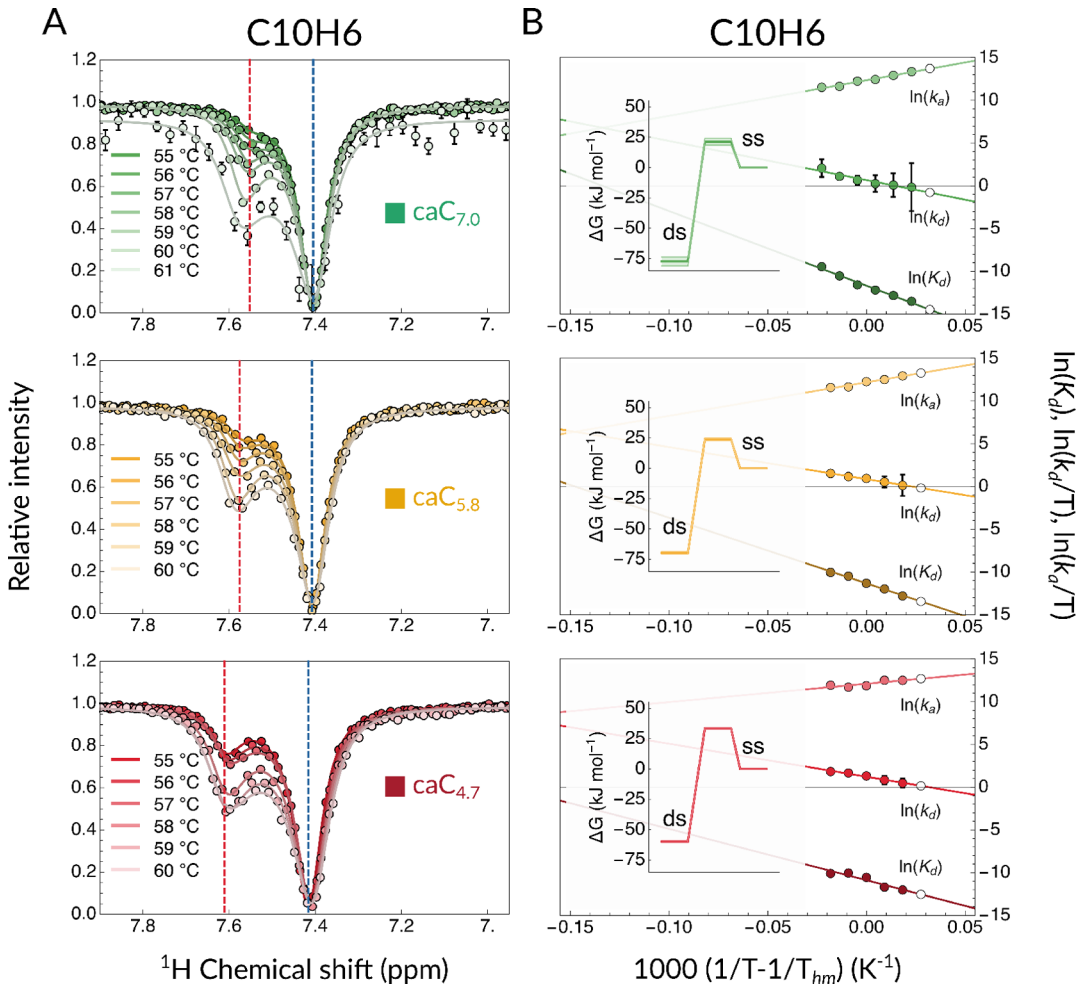


Figure 3. (A) Temperature-dependent CEST melting profiles together with the obtained two-site exchange fits for proton C10H6. $caC_{7.0}$, $caC_{5.8}$, and $caC_{4.7}$ are shown in shades of green, yellow, and red, respectively. Dashed blue and red lines indicate dsDNA and ssDNA chemical shift values. (B) van't Hoff plots relative to the CEST melting profiles. Shades of green, yellow, and red indicate data entries and linear fits of ln k_a, ln k_d, and ln K_d vs $1/T - 1/T_{hm}$, where T_{hm} is the harmonic mean of the measured temperatures, for $caC_{7.0}$, $caC_{5.8}$, and $caC_{4.7}$, respectively. White data points represent back-calculated values for 55 °C that were used to validate the CEST fits. Insets present the relevant Gibbs free energy plots at each pH condition at 37 °C.

exchange parameters can reveal atomistic details about the melting thermodynamics and kinetics of the studied system providing unprecedented insights into the molecular processes. In pursuance of the study of 5caC-induced DNA destabilization, we proceeded by recording CEST profiles for the aromatic protons in all caC_{pH} samples in the 55−61 °C range. As an example, Figure 3A displays the melting CEST profiles for C10H6 at three pH values (profiles for other comparable protons can be found in Figures S11−S18). The appearance of a distinctive secondary dip at increasingly higher temperatures indicates the presence of an alternative conformer, which we identify as the single-stranded conformation (ssDNA) as per comparative chemical shift analysis.

To obtain a more quantitative comparison, we fit the CEST profiles to a two-state exchange model (dsDNA ⇌ 2 ssDNA) that yielded a numerical estimation of populations (p_D for dsDNA, $1 - p_D$ for ssDNA), exchange kinetics (k_{ex}), and chemical shifts of the exchanging states at each of the highest temperatures, while we used the lowest temperature of the ensemble for each sample (55 °C) to compare the predicted back-calculated value from the fits to the experimental data points. ^{1}H longitudinal (R_1) and transverse (R_2) relaxation rates were measured separately at multiple temperatures and used as inputs for the CEST fits assuming that the rate constants of the dsDNA and ssDNA states are the same.

As an example, Figure 3B shows the logarithm of the obtained kinetic rates and equilibrium constants for C10H6. Consistently with results obtained for $fC_{7.0}$ and $C_{7.0}$, two observations can be made: (i) a linear fit could be identified for kinetic rates and equilibrium constants against T^{-1}, suggesting a single transition state is plausible for the melting and annealing phenomena, and (ii) all sites show thermally activated kinetics with a positive dissociation barrier (Arrhenius behavior) and a negative association barrier (anti-Arrhenius).

In addition, CEST measurements run at multiple distinct temperatures allow for the extraction of Gibbs free energy and related parameters by leveraging the temperature-dependent nature of kinetics and thermodynamic phenomena. In order to achieve an estimation of the degree of protonation-induced destabilization to dsDNA, we applied our recently described methodological framework for the decomposition of traditional CEST output into enthalpic and entropic stability and activation changes.[20] Succinctly, we assumed the observed

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

dynamic equilibrium is the reversible melting/annealing process of a dsDNA strand into two single-stranded DNA sequences. If the concentration of DNA is known, then kinetics of association (k_a) and dissociation (k_d) can be extracted. From the temperature dependence of k_a and k_d, activation barriers for both process ($\Delta G^{\ddagger}_{d,37°C}$ and $\Delta G^{\ddagger}_{a,37°C}$) can be derived. Analogously, the ratio between $k_a(T)$ and $k_d(T)$ rates allows for the determination of the equilibrium dissociation constant (K_d) and consequently the quantification of thermodynamic parameters such as $\Delta G^{\circ}_{37°C}$.

This analysis allows us to directly compare $caC_{7.0}$, $caC_{5.8}$, and $caC_{4.7}$ not only between them, but also with otherwise identical canonical and formylated samples ($C_{7.0}$ and $fC_{7.0}$, respectively), which we previously discussed in ref 20. In Table 1, we compare three proton reporters across all five samples

Table 1. Thermodynamic and Kinetic Parameters of the dsDNA Melting Process Obtained from the van't Hoff and Eyring Analysis of the CEST-Derived Exchange Parameters[a]

sample		$\Delta G^{\circ}_{37°C}$ (kJ mol^{-1})	$\Delta G^{\ddagger}_{d,37°C}$ (kJ mol^{-1})	$\Delta G^{\ddagger}_{a,37°C}$ (kJ mol−1)
$C_{7.0}$	C2H6	71.1 ± 1.7	99.2 ± 2.2	28.1 ± 2.1
	T9H6	70.2 ± 1.8	102.2 ± 1.5	31.9 ± 1.5
	C10H6	71.3 ± 1.4	100.3 ± 1.3	29.0 ± 1.3
$caC_{7.0}$	C2H6	73.3 ± 7.4	93.0 ± 4.3	19.7 ± 4.1
	T9H6	75.7 ± 0.8	99.2 ± 0.9	23.4 ± 0.9
	C10H6	77.3 ± 3.6	98.6 ± 3.2	21.3 ± 2.9
$caC_{5.8}$	C2H6	71.1 ± 1.2	94.8 ± 1.4	23.7 ± 1.3
	T9H6	68.0 ± 0.9	96.6 ± 0.9	28.6 ± 1.0
	C10H6	67.4 ± 1.4	92.4 ± 1.2	24.9 ± 1.1
$caC_{4.7}$	C2H6	58.6 ± 0.7	93.1 ± 0.7	34.5 ± 0.9
	T9H6	62.1 ± 0.4	95.2 ± 0.4	33.1 ± 0.5
	C10H6	56.8 ± 0.6	92.9 ± 0.6	36.1 ± 0.6
$fC_{7.0}$	C2H6	74.7 ± 10.4	99.7 ± 5.6	25. ± 5.5
	T9H6	64.6 ± 1.2	97.1 ± 1.0	32.5 ± 1.0
	C10H6	66.4 ± 1.2	97.5 ± 1.0	31.1 ± 1.1

[a]Errors are given as one standard deviation. An extended version of Table 1 can be found in the Supporting Information file (Table S7).

(comprehensive tables including fitting results for all proton reporters can be found in Tables S1−S6, while an extended version of Table 1 is available in Table S7). From a thermodynamic perspective, among the C-modified samples, $caC_{7.0}$ scores as the most stable one, which is highly comparable to $C_{7.0}$, as observed by UV/vis spectroscopy (Figure S19) as well as in other studies.[23,27] Acidification of the buffer to pH 5.8 destabilizes the double-stranded conformer by ∼2−10 kJ mol^{-1} and further by another ∼6−13 kJ mol^{-1} when the pH is decreased from 5.8 to 4.7. Interestingly, $\Delta G^{\circ}_{37°C}$ for $caC_{5.8}$'s proton reporters are very similar to those we obtained for $fC_{7.0}$. Kinetics of dissociation data, as expected, support the notion that $caC_{7.0}$ and $C_{7.0}$ require the most energy for undergoing a dsDNA → 2 ssDNA conformational transition. According to this metric, $fC_{7.0}$ is slower in undergoing the melting process when compared to $caC_{5.8}$ and $caC_{4.7}$, as $\Delta G^{\ddagger}_{d,37°C}$ is higher by ∼3−7 kJ mol^{-1}. Lastly, kinetics of association suggest that $caC_{4.7}$ is the slowest in performing an annealing process, followed by $fC_{7.0}$. Compellingly, $\Delta G^{\ddagger}_{a,37°C}$ data indicate that the rate of association of $caC_{7.0}$ is ∼8 kJ mol^{-1} less energetically demanding when compared to $C_{7.0}$, a behavior that can be rationalized considering the EDG nature of the carboxylate substituent (5caC at neutral pH) when compared to a proton (canonical C).

In Figure 4, we show the correlation between the dissociation and equilibrium free energy changes across all

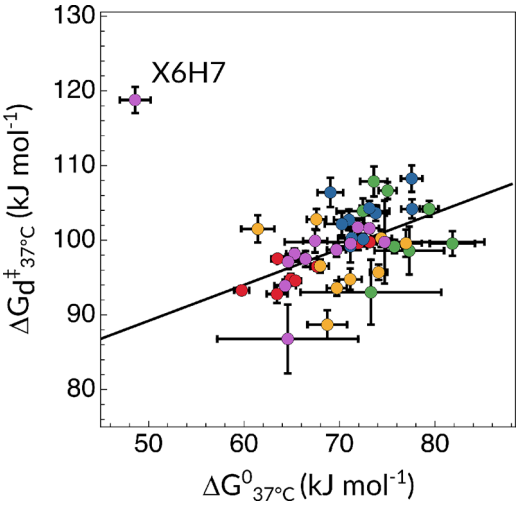


Figure 4. Correlation plot between the changes of the Gibbs free energy of activation for the dissociation process and the equilibrium free energy of previously reported values for $fC_{7.0}$ and $C_{7.0}$ (magenta and blue, respectively) together with 5caC-containing samples ($caC_{7.0}$ as green, $caC_{5.8}$ as yellow, and $caC_{4.7}$ as red) at 37 °C. $fC_{7.0}$ X6H7 (featuring much higher activation free energies and lower equilibrium energies than the rest of the molecule) is an outlier due to its stable intramolecular hydrogen-bond between the formyl O7 atom and the adjacent amino H42 proton.

samples and conditions, for every proton reporter. Here, $\Delta G^{\ddagger}_{d}$ is plotted as a function of $\Delta G°$ for all five samples across all eligible proton reporters. Data points accounting for $caC_{7.0}$ and $C_{7.0}$ tend to cluster at the upper right-hand corner of the plot. Conversely, $\Delta G^{\circ 0}_{37°C}$ and $\Delta G^{\ddagger}_{d,37°C}$ values are substantially decreased whenever $fC_{7.0}$, $caC_{5.8}$, or $caC_{4.7}$ is considered, as elaborated above.

Protonation-Induced Microsecond Dynamics

^{1}H and ^{15}N chemical shift values for G7H1/N1 nuclei, both reporters of the centrally positioned 5caC6-G7 base pair's stability, have evidenced that the protonation state of the exocyclic carboxylic group 5caC has a selective impact on this imino proton resonance (Figure 2A, D), which have long been regarded as key indicators of hydrogen bond strength.[23,39] Because of this observation, we aimed at investigating whether this protonation-induced, localized weakening is accompanied by increased probability of local fast time-scale motions.

Our CEST-based kinetic and thermodynamic analysis has established that $caC_{5.8}$, $caC_{4.7}$, and $fC_{7.0}$ destabilize the double-stranded DNA structure without any apparent static, persistent impact on its helical architecture. In order to investigate the presence of a potentially localized conformational exchange which might contribute to the destabilization, we interrogated the faster, microsecond time scale by applying ^{1}H $R_{1\rho}$ relaxation dispersion (RD) methods.[37,40]

In Figure 5, we show X6H6 (where X = C, 5fC or 5caC, depending on the sample under current consideration) ^{1}H on-resonance $R_{1\rho}$ RD profiles measured at 55 °C, a condition that ensures that the melting process is still rather sparse and

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

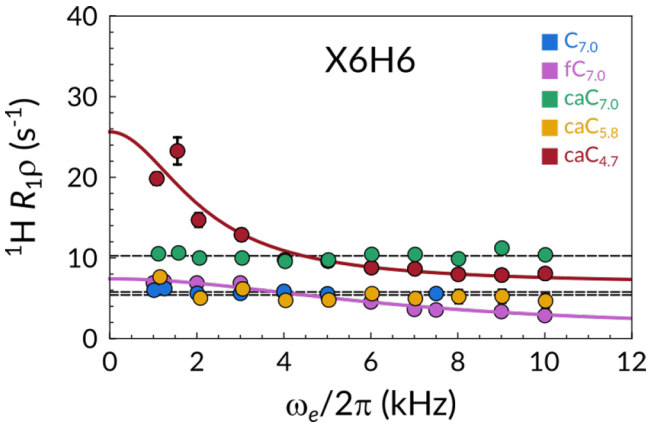


Figure 5. $R_{1\rho}$ relaxation dispersion profiles of $C_{7.0}$ (blue), $fC_{7.0}$ (magenta), $caC_{7.0}$ (green), $caC_{5.8}$ (yellow), and $caC_{4.7}$ (red) recorded at 55 °C. Lines represent best fits to models either accounting for (color-coded) or discounting chemical exchange (black).

infrequent. The data sets recorded for $caC_{7.0}$, $caC_{5.8}$, and $C_{7.0}$ are best fit to a no-exchange model, resulting in flat profiles (black lines). Conversely, profiles for $caC_{4.7}$ and $fC_{7.0}$ fit best to the exchange model, exposing a chemical exchange contribution to its R_2 relaxation rate. Such motions, consistent with a τ_{ex} in the order of hundreds of microseconds, fall within the intermediate exchange regime and are approximately 2 orders of magnitude faster compared to the overall melting process as characterized by our CEST measurements. This result is especially interesting when comparing the H6 proton of the modified base in samples $caC_{4.7}$ and $fC_{7.0}$ to any other available ^{1}H nucleus. No other profile is consistent with an exchange phenomenon on this interval (Figures S20−S24), suggesting that the detected phenomenon is unrelated to the previously reported Watson−Crick to Hoogsteen base-pair exchange process.[30−32] In other words, this motions appears to be sharply localized, leaving all other bases unaffected.

This result can be rationalized by considering our previous speculations on the mechanism by which 5fC (and by extension protonated 5caC) weakens the dsDNA conformer.[20] Deprotonated (or not protonated to a sufficient extent) 5caC proved itself to be either a neutral or even a stabilizing factor in terms of base-pair strength and agrees well both with NOESY-derived chemical shift values and the literature. Instead, whenever the pH of the buffer is sufficiently acidified, the weakening of the 5caC-G hydrogen bond induces a chemical exchange process in the intermediate microsecond time scale. The fact that this motion is sharply localized at the modified site seems to suggest that the protonation of the exocyclic carboxyl moiety and the consequent lowered basicity of the hydrogen-bonded N3 atom could be deemed responsible for generating a sparse and transitory fraying event in the middle of the DNA strand, presumably synergistic with the global destabilization effect on the whole structure.

DISCUSSION

In pursuance of understanding the role of 5caC, we addressed the structural and dynamic features of oligomeric DNA double strands carrying a single version of such modification on each strand. Incorporation of carboxycytosine into DNA is a naturally occurring phenomenon that has been mainly discussed within two biological frameworks: (i) its recently devised (and to date mostly obscure) semipermanent epigenetic role and (ii) as a DNA lesion that undergoes the base excision repair (BER) process.[2−4,41]

Considerations on 5caC Protonation Sites

In their report on divergent mechanisms of enzymatic excision for 5caC and 5fC, Maiti et al. have elaborated how, in a TDG-DNA complex, nitrogen N3 of 5caC (Figure 1B) is likely more basic than the carboxyl group and thus undergoes protonation before the carboxylic exocyclic moiety does.[26] In contrast, despite assigning a pK_{a,COO^-} of 4.7 and $pK_{a,N3}$ of 2.1 for the isolated nucleoside, a distinct infrared spectroscopy and quantum mechanical analysis suggests that, when in the context of a dsDNA strand, the first protonation site is the carboxyl group.[25] Our results support the latter idea. In Figure 2A, we show that no other base pair is affected by the pH change over the entire 7.0−4.7 interval. Indeed, T9H3, T5H3, and G3H1 display negligible chemical shift differences, while G11H1 shifts upfield by ∼0.05 ppm likely due to its proximity to the fraying ends of the oligomeric model system. In sharp contrast, the signal reporting on G7H1 consistently shifts upfield with decreasing pH. While this preference for protonation of the weaker base (i.e., 5caC's carboxylate group over N3) is apparently counterintuitive, we reason that in a dsDNA setting the N3 site is well protected from the solvent environment, both for steric and electrostatic reasons.

5fC and 5caC as Semipermanent Modifications

For a long time, 5fC and 5caC have been mainly considered as transient intermediates within the active demethylation pathway.[42] However, research toward their capacity as standalone epigenetic marks has gained increasingly more traction. For instance, both 5fC and 5caC overlap with H3K4me1 marked regions, associated with active transcription.[43] Also, several developmental and metabolic related genes show 5fC enrichment on promoters before gene upregulation,[44] while 5caC has been reported to transiently accumulate at promoter regions preceding gene expression during lineage specification and differentiation.[45] Lastly, a number of cancerous diseases are correlated with a significant enrichment of such oxidized cytosine epigenetic modifications.[10−13]

From a structural and conformational perspective, both 5fC and 5caC have been reported to induce structural changes localized in the proximity of the modified nucleoside,[22,26] while the pH dependence of 5caC characteristics on dsDNA stability has been previously discussed in the context of short oligomers carrying several clustered modifications.[21,27] In the context of longer (90 bp) DNA strands, Ngo et al. reported a 3-fold enhancement of cyclization rate for 5fC containing strands, while the sample carrying 5caC at pH 8.0 showed no detectable difference when compared to canonical cytosine.[46]

In contrast to refs 22 and 26, our results indicate that, identically to 5fC, 5caC does not induce any permanent structural modification detectable by NMR spectroscopy at any pH condition under consideration (Figures 2 and S2−10). In fact, in analogy to 5fC, 5caC was found to affect dsDNA melting and annealing equilibrium and kinetics, rather than B-DNA average structure. On average, and across all sites, $caC_{5.8}$ and $caC_{4.7}$ resemble $fC_{7.0}$, both thermodynamically and kinetically, when it comes to annealing and melting. In contrast, and consistently with previous cyclization essays and FRET studies, the behavior of deprotonated 5caC is most similar to canonical cytosine.

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

5mC oxidation derivatives have been shown to accumulate and persist in a relatively stable state in certain biological contexts. For this reason, they have been suggested to carry out additional roles apart from being intermediates in biochemical pathways. In light of the foregoing, we contend that our results could correlate with the relative abundance of 5caC and 5fC in cancerous tissues. How protonation of the exocyclic carboxyl moiety affects the thermodynamics and kinetics of melting and annealing *in vitro* has been discussed in this and previous reports.[20,21,25,27] Our results corroborate the protonation-induced destabilization of the double strand, and we speculate that a similar effect might take place in cancerous tissues, triggered by the low pH environment. Cell proliferation is notoriously accelerated in cancer, and the acidic microenvironment where such diseases thrive is widely recognized as a phenotypic trait, making the presence of the two most oxidized cytosine epigenetic derivatives more than circumstantial.[47,48]

5fC- and 5caC-Driven Enzymatic Recognition

In mammals, active DNA demethylation takes place via an enzymatic tandem of TET and TDG, which together govern the initial stages of the BER pathway. In order to successfully complete the removal of 5mC, the presence of either 5fC or 5caC is ultimately necessary, as they are substrates for TDG, which generates the abasic site.[26] Apart from TET and TDG, several additional proteins are able to selectively recognize, bind, and exert their respective enzymatic activity upon 5fC and 5caC.[12,13,49] The mechanism by which different reader enzymes selectively recognize cytosine's epigenetic modifications has long been established as a crucial theme in chemical biology.[6] Across the proposed enzymatic mechanisms, many rely on a specific residue to initiate the base-extrusion process into the active site. For instance, the "pinch−push−pull" mechanism proposed for TDG leans on Arg275 to promote the breakage of an X:G base pair, where X = T, 5fC, or protonated 5caC.[50,51] Alternative studies suggest that partially extruded nucleotide conformations, which are sparse but naturally occurring events, might play a role in recognition and base excision.[52] This second mechanism of action seems to agree with DNA replication studies, which highlighted how 5caC:G base-pair can behave as a DNA lesion.[28]

Although the data hereby presented is not conclusive, we believe our on-resonance $R_{1\rho}$ RD data assist in providing one more piece of evidence in this complex puzzle. By observing site selective and spontaneous (i.e., not triggered by the enzyme) base-flipping of protonated 5caC:G, especially in the context of a weakened dsDNA strand as evidenced in our CEST analysis, we present further experimental evidence that 5fC and protonated 5caC could indeed act as DNA lesions.[33] We hypothesize that, albeit undetected in our study, such kinetic processes could be present at physiological temperatures and decisively impact enzyme recognition and mode of action.

CONCLUSIONS

In this work, we have considered the impact of 5caC incorporation into a model dsDNA oligomer (caC_{pH}) in three pH conditions, namely, at pH 7.0, 5.8, and 4.7. We obtained chemical shift, melting/annealing, and microsecond conformational exchange data which we could reliably compare with our recent study focusing on 5fC and canonical cytosine. Assignment and chemical shift studies on comparable ^{1}H, ^{13}C, and ^{15}N nuclei have shown that there is no evidence of a permanent structural change: all caC_{pH} samples, together with $C_{7.0}$ and $fC_{7.0}$, are compatible with a standard B-DNA helical arrangement. CEST-derived kinetic and thermodynamic data suggested that the reduced cohesion of the X6:G7 base pair, evidenced by chemical shift studies, affects the extent to which nearby bases are able to cooperatively stabilize one another. $caC_{4.7}$ and $fC_{7.0}$ emerged as the most destabilized samples of the cohort, while $caC_{7.0}$ and $C_{7.0}$ showed remarkably similar properties overall.

$caC_{4.7}$ and $fC_{7.0}$ also revealed a detectable chemical exchange process at or in the proximity of the modified nucleoside X6 on the microsecond time scale. The data hereby presented indicate that 5caC's impact on B-DNA is only evident through the lenses of conformational dynamics, as protonation of the exocyclic carboxyl moiety affects the melting-annealing equilibrium as well as induces sparse and localized microsecond time scale base-pair dynamics. We believe our findings are relevant in the context of several open questions concerning this sparse epigenetic mark. Future investigations may consider expanding our initial exploration of the microsecond time scale by recording off-resonance $R_{1\rho}$ RD experiments or studying protein−DNA interactions featuring isotopic labeled 5fC or 5caC nucleosides to unravel the exact mechanistic details of the interaction between TET, TDG (and other enzymes), and cytosine's oxidized derivatives.

MATERIALS AND METHODS

Sample Preparation

The cadC-phosphoramidite (cadC-PA) and subsequently the modified dsDNA samples caC_{pH} were prepared via phosphoramidite chemistry as previously reported.[53] Solid phase synthesis of oligonucleotides containing cadC was performed on an ABI 394 DNA/RNA synthesizer (Applied Biosystems) using standard DNA synthesis conditions with a cartridge scale of 1 μmol. The phosphoramidites Bz-dA, Ac-dC, iBu-dG, and dT as well as the PS carriers were purchased from LinkTechnologies. For the reaction of the cadC-PA a coupling time of 180 s was applied. The terminal DMT protecting group was cleaved after DNA synthesis on the synthesizer. Basic and acidic deprotection of all oligonucleotides was performed according to literature.[53] Purification of the oligonucleotides was achieved with a HPLC system (Agilent 1260 Infinity II 400 bar pump and a Agilent 1260 Infinity II VWD detecting at 260 nm) applying a buffer system of 0.1 M triethylammonium acetate in water (buffer A) and 0.1 M triethylammonium acetate in 80% aqueous MeCN (buffer B), a gradient of 0%−30% buffer B in 45 min and a flow rate of 5.0 mL/min. As stationary phase Nucleodur columns (250/10 mm, C18ec, 5 μm) from Macherey-Nagel were used. Purified oligonucleotides were analyzed by MALDI-TOF (Bruker Autoflex II). Quantification of oligonucleotides was performed via UV/vis spectroscopy with a NanoDrop ND-1000 spectrophotometer at 260 nm. Samples $caC_{7.0}$ and $caC_{5.8}$ were dissolved in aqueous buffers consisting of 15 mM Na_2HPO_4/NaH_2PO_4 (pH 7.0 and 5.8, respectively), 25 mM NaCl in H_2O. Sample $caC_{4.7}$ was prepared by titrating a 1 M HCl solution into the same buffer described above. The thermal stability of the buffer between room temperature and 60 °C was ascertained by pH-meter measurements. Annealing was performed by heating the dsDNA-containing buffer solution to 90 °C for 5 min and slowly cooling it to 5 °C in approximately 90 min, after which it was allowed to return to room temperature. Then, the NMR sample was prepared with the addition of 0.02% NaN_3, 25 μM DSS and 5% D_2O, resulting in final sample concentrations of ∼0.66 mM for all samples, as determined via UV spectrophotometric measurements at 260 nm using the extinction coefficient calculated via the nearest neighbor approximation.

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

UV/Vis Spectroscopy

UV/vis melting profiles of the oligonucleotides were measured at 260 nm with a JASCO V-650 UV/vis spectrophotometer between 20 and 85 °C (scanning rate of 1 °C/min), and each sample was measured four times. Samples were placed into 100 μL cuvettes and diluted with the same Na_2HPO_4/NaH_2PO_4, NaCl aqueous buffer as used in the NMR experiment. Before each measurement, a layer of mineral oil was placed on the surface of the sample in order to prevent water evaporation. $caC_{7.0}$ was measured at four concentrations (1.25, 2.50, 5.00, and 10.00 μM), while $fC_{7.0}$ and $C_{7.0}$ were measured as described in ref 20. All concentration values yielded absorption values within the linear range of the spectrometer.

NMR Spectroscopy

All experiments were performed on Bruker Avance III spectrometer operating at a ^{1}H Larmor frequency of 800 MHz (corresponding to a magnetic field of 18.8 T) equipped with a 5 mm triple-resonance cryogenically cooled TCI probe. Standard 2D NOESY (mixing time 250 ms) spectra were recorded at 37 °C for resonance assignment. Natural abundance ^{1}H−^{13}C and ^{1}H−^{15}N HSQC and HMQC spectra were recorded using standard fast-pulsing pulse sequences.[54] Site-selective spin relaxation measurements were performed following the SELOPE scheme; these included ^{1}H CEST, on-resonance ^{1}H $R_{1\rho}$, recorded either with a single spin-lock strength of 10 kHz or as an entire RD profile ranging from 1 to 10 kHz, and ^{1}H R_1 experiments at temperatures between 37 and 60 °C. The employed on-resonance $R_{1\rho}$ and ^{1}H CEST pulse sequences have been modified from Schlagnitweit et al.[40] CEST profiles and pseudo-2D ^{1}H R_1 and pseudo-3D on-resonance $R_{1\rho}$ experiments were performed, processed, and analyzed as previously described.[20]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphyschemau.1c00050.

^{1}H−^{1}H, ^{1}H−^{13}C, and ^{1}H−^{15}N 2D correlation spectra, schematic comparison of ^{1}H, ^{13}C, and ^{15}N chemical shift perturbations, concentration-dependent UV/vis melting analysis, temperature-dependent CEST profiles, and van't Hoff plots for all available sites, on-resonance $R_{1\rho}$ relaxation dispersion profiles for all samples, CEST-derived dsDNA and ssDNA chemical shifts and relaxation parameters (R_1 and R_2), and CEST-derived thermodynamic and kinetic parameters for all samples (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Petra Rovó − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany; Center for Nanoscience (CeNS), Faculty of Physics, Ludwig-Maximilians-Universität München, 80799 Munich, Germany; Institute of Science and Technology Austria (ISTA), 3400 Klosterneuburg, Austria;* orcid.org/0000-0001-8729-7326; Email: petra.rovo@ist.ac.at

Authors

Romeo C. A. Dubini − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany; Center for Nanoscience (CeNS), Faculty of Physics, Ludwig-Maximilians-Universität München, 80799 Munich, Germany;* orcid.org/0000-0001-6045-271X

Eva Korytiaková − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany*

Thea Schinkel − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany;* orcid.org/0000-0002-1891-2660

Pia Heinrichs − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany;* orcid.org/0000-0002-3128-6676

Thomas Carell − *Faculty of Chemistry and Pharmacy, Department of Chemistry, Ludwig-Maximilians-Universität München, 81377 Munich, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.1c00050

Author Contributions

§T.S. and P.H. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Markus Müller for valued discussions and Felix Xu for assistance in the measurement of UV/vis melting profiles. This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) − SFB 1309-325871075, EU-ITN LightDyNAmics (ID: 765266), the ERC-AG EpiR (ID: 741912), the Center for NanoScience, the Excellence Clusters CIPSM, and the Fonds der Chemischen Industrie. Open access funding provided by Institute of Science and Technology Austria (ISTA).

■ REFERENCES

(1) Kumar, S.; Chinnusamy, V.; Mohapatra, T. Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. *Front. Genet.* **2018**, *9*, 1−14.

(2) Bilyard, M. K.; Becker, S.; Balasubramanian, S. Natural, modified DNA bases. *Curr. Opin. Chem. Biol.* **2020**, *57*, 1−7.

(3) Chen, Y.; Hong, T.; Wang, S.; Mo, J.; Tian, T.; Zhou, X. Epigenetic modification of nucleic acids: from basic studies to medical applications. *Chem. Soc. Rev.* **2017**, *46*, 2844−2872.

(4) Carell, T.; Kurz, M. Q.; Müller, M.; Rossa, M.; Spada, F. Non-canonical Bases in the Genome: The Regulatory Information Layer in DNA Angewandte. *Angew. Chem., Int. Ed.* **2018**, *57*, 4296−4312.

(5) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* **2011**, *333*, 1300−1304.

(6) Wu, X.; Zhang, Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* **2017**, *18*, 517−534.

(7) Bachman, M.; Uribe-Lewis, S.; Yang, X.; Burgess, H. E.; Iurlaro, M.; Reik, W.; Murrell, A.; Balasubramanian, S. 5-Formylcytosine can be a stable DNA modification in mammals. *Nat. Chem. Biol.* **2015**, *11*, 555−557.

(8) Su, M.; Kirchner, A.; Stazzoni, S.; Müller, M.; Wagner, M.; Schröder, A.; Carell, T. 5-Formylcytosine Could Be a Semipermanent Base in Specific Genome Sites. *Angew. Chem., Int. Ed.* **2016**, *55*, 11797−11800.

(9) Koivunen, P.; Laukka, T. The TET enzymes. *Cell. Mol. Life Sci.* **2018**, *75*, 1339−1348.

(10) Guo, M.; Li, X.; Zhang, L.; Liu, D.; Du, W.; Yin, D.; Lyu, N.; Zhao, G.; Guo, C.; Tang, D. Accurate quantification of 5-

Methylcytosine, 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxylcytosine in genomic DNA from breast cancer by chemical derivatization coupled with ultra performance liquid chromatography-electrospray quadrupole time. *Oncotarget* **2017**, *8*, 91248−91257.
(11) Eleftheriou, M.; Pascual, A. J.; Wheldon, L. M.; Perry, C.; Abakir, A.; Arora, A.; Johnson, A. D.; Auer, D. T.; Ellis, I. O.; Madhusudan, S.; et al. 5-Carboxylcytosine levels are elevated in human breast cancers and gliomas. *Clin. Epigenet.* **2015**, *7*, 88.
(12) Hashimoto, H.; Olanrewaju, Y. O.; Zheng, Y.; Wilson, G. G.; Zhang, X.; Cheng, X. Wilms tumor protein recognizes 5-carboxylcytosine within a specific DNA sequence. *Genes Dev.* **2014**, *28*, 2304−2313.
(13) Wang, D.; Hashimoto, H.; Zhang, X.; Barwick, B. G.; Lonial, S.; Boise, L. H.; Vertino, P. M.; Cheng, X. MAX is an epigenetic sensor of 5-carboxylcytosine and is altered in multiple myeloma. *Nucleic Acids Res.* **2017**, *45*, 2396−2407.
(14) Wheldon, L. M.; Abakir, A.; Ferjentsik, Z.; Dudnakova, T.; Strohbuecker, S.; Christie, D.; Dai, N.; Guan, S.; Foster, J. M.; Corrêa, I. R.; et al. Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells. *Cell Rep* **2014**, *7*, 1353−1361.
(15) Kellinger, M. W.; Song, C. X.; Chong, J.; Lu, X. Y.; He, C.; Wang, D. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* **2012**, *19*, 831−833.
(16) Song, J.; Pfeifer, G. P. Are there specific readers of oxidized 5-methylcytosine bases? *BioEssays* **2016**, *38*, 1038−1047.
(17) Raiber, E. A.; Murat, P.; Chirgadze, D. Y.; Beraldi, D.; Luisi, B. F.; Balasubramanian, S. 5-formylcytosine alters the structure of the DNA double helix. *Nat. Struct Mol. Biol.* **2015**, *22*, 44−49.
(18) Wang, S.; Long, Y.; Wang, J.; Ge, Y.; Guo, P.; Liu, Y.; Tian, T.; Zhou, X. Systematic Investigations of Different Cytosine Modifications on CpG Dinucleotide Sequences: The Effects on the B-Z. Transition. *J. Am. Chem. Soc.* **2014**, *136*, 56.
(19) Hardwick, J. S.; Ptchelkine, D.; El-Sagheer, A. H.; Tear, I.; Singleton, D.; Phillips, S. E.; Lane, A. N.; Brown, T. 5-Formylcytosine does not change the global structure of DNA. *Nat. Struct. Mol. Biol.* **2017**, *24*, 544−552.
(20) Dubini, R. C.; Schön, A.; Müller, M.; Carell, T.; Rovó, P. Impact of 5-formylcytosine on the melting kinetics of DNA by 1H NMR chemical exchange. *Nucleic Acids Res.* **2020**, *48*, 8796−8807.
(21) Sanstead, P. J.; Ashwood, B.; Dai, Q.; He, C.; Tokmakoff, A. Oxidized Derivatives of 5-Methylcytosine Alter the Stability and Dehybridization Dynamics of Duplex DNA. *J. Phys. Chem. B* **2020**, *124*, 1160−1174.
(22) Fu, T.; Liu, L.; Yang, Q.-l.; Wang, Y.; Xu, P.; Zhang, L.; Liu, S.; Dai, Q.; Ji, Q.; Xu, G.-l.; et al. Thymine DNA glycosylase recognizes the geometry alteration of minor grooves induced by 5-formylcytosine and 5-carboxylcytosine. *Chem. Sci.* **2019**, *10*, 7407−7417.
(23) Szulik, M. W.; Pallan, P. S.; Nocek, B.; Voehler, M.; Banerjee, S.; Brooks, S.; Joachimiak, A.; Egli, M.; Eichman, B. F.; Stone, M. P. Differential Stabilities and Sequence-Dependent Base Pair Opening Dynamics of Watson-Crick Base Pairs with 5-Hydroxymethylcytosine, 5-Formylcytosine, or 5-Carboxylcytosine. *Biochemistry* **2015**, *54*, 1294.
(24) Sumino, M.; Ohkubo, A.; Taguchi, H.; Seio, K.; Sekine, M. Synthesis and properties of oligodeoxynucleotides containing 5-carboxy-2−deoxycytidines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 274−277.
(25) Dai, Q.; Sanstead, P. J.; Peng, C. S.; Han, D.; He, C.; Tokmakoff, A. Weakened N3 Hydrogen Bonding by 5-Formylcytosine and 5-Carboxylcytosine Reduces Their Base-Pairing Stability. *ACS Chem. Biol.* **2016**, *11*, 470−477.
(26) Maiti, A.; Michelson, A. Z.; Armwood, C. J.; Lee, J. K.; Drohat, A. C. Divergent mechanisms for enzymatic excision of 5-formylcytosine and 5-carboxylcytosine from DNA. *J. Am. Chem. Soc.* **2013**, *135*, 15813−15822.
(27) 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.
(28) Shibutani, T.; Ito, S.; Toda, M.; Kanao, R.; Collins, L. B.; Shibata, M.; Urabe, M.; Koseki, H.; Masuda, Y.; Swenberg, J. A.; et al. Guanine-5-carboxylcytosine base pairs mimic mismatches during DNA replication. *Sci. Rep.* **2015**, *4*, 5220.
(29) Vallurupalli, P.; Bouvignies, G.; Kay, L. E. Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J. Am. Chem. Soc.* **2012**, *134*, 8148−8161.
(30) Liu, B.; Rangadurai, A.; Shi, H.; Al-Hashimi, H. M. Rapid assessment of Watson−Crick to Hoogsteen exchange in unlabeled DNA duplexes using high-power SELOPE imino ^{1}H CEST. *Magn. Reson.* **2021**, *2*, 715−731.
(31) 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.
(32) Zhou, H.; Sathyamoorthy, B.; Stelling, A.; Xu, Y.; Xue, Y.; Pigli, Y. Z.; Case, D. A.; Rice, P. A.; Al-Hashimi, H. M. Characterizing Watson-Crick versus Hoogsteen Base Pairing in a DNA-Protein Complex Using Nuclear Magnetic Resonance and Site-Specifically 13C- and 15N-Labeled DNA. *Biochemistry* **2019**, *58*, 1963−1974.
(33) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E., III On the absence of intrahelical DNA dynamics on the μs to ms timescale. *Nat. Commun.* **2014**, *5*, 5152.
(34) Choi, S.-R.; Kim, N.-H.; Jin, H.-S.; Seo, Y.-J.; Lee, J.; Lee, J.-H. Base-pair Opening Dynamics of Nucleic Acids in Relation to Their Biological Function. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 797−804.
(35) Schanda, P.; Brutscher, B. Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J. Am. Chem. Soc.* **2005**, *127*, 8014−8015.
(36) Zarycz, M. N. C.; Fonseca Guerra, C. NMR ^{1}H-Shielding Constants of Hydrogen-Bond Donor Reflect Manifestation of the Pauli Principle. *J. Phys. Chem. Lett.* **2018**, *9*, 3720−3724.
(37) Marušič, M.; Schlagnitweit, J.; Petzold, K. RNA Dynamics by NMR Spectroscopy. *ChemBioChem.* **2019**, *20*, 2685−2710.
(38) Shi, H.; Liu, B.; Nussbaumer, F.; Rangadurai, A.; Kreutz, C.; Al-Hashimi, H. M. NMR Chemical Exchange Measurements Reveal That N^{6}-Methyladenosine Slows RNA Annealing. *J. Am. Chem. Soc.* **2019**, *141*, 19988−19993.
(39) Wang, Y.; Han, G.; Jiang, X.; Yuwen, T.; Xue, Y. Chemical shift prediction of RNA imino groups: application toward characterizing RNA excited states. *Nat. Commun.* **2021**, *12*, 1595.
(40) Schlagnitweit, J.; Steiner, E.; Karlsson, H.; Petzold, K. Efficient Detection of Structure and Dynamics in Unlabeled RNAs: The SELOPE Approach. *Chem.—Eur. J.* **2018**, *24*, 6067−6070.
(41) Hardwick, J. S.; Lane, A. N.; Brown, T. Epigenetic Modifications of Cytosine: Biophysical Properties, Regulation, and Function in Mammalian DNA. *BioEssays* **2018**, *40*, 1700199.
(42) Zhu, Q.; Stöger, R.; Alberio, R. A Lexicon of DNA Modifications: Their Roles in Embryo Development and the Germline. *Front. Cell Dev. Biol.* **2018**, *6*, 24.
(43) Wu, H.; Wu, X.; Shen, L.; Zhang, Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nat. Biotechnol.* **2014**, *32*, 1231−1240.
(44) Zhu, C.; Gao, Y.; Guo, H.; Xia, B.; Song, J.; Wu, X.; Zeng, H.; Kee, K.; Tang, F.; Yi, C. Single-Cell 5-Formylcytosine Landscapes of Mammalian Early Embryos and ESCs at Single-Base Resolution. *Cell Stem Cell* **2017**, *20*, 720−731.
(45) Lewis, L. C.; Lo, P. C. K.; Foster, J. M.; Dai, N.; Corrêa, I. R.; Durczak, P. M.; Duncan, G.; Ramsawhook, A.; Aithal, G. P.; Denning, C.; et al. Dynamics of 5-carboxylcytosine during hepatic differentiation: Potential general role for active demethylation by DNA repair in lineage specification. *Epigenetics* **2017**, *12*, 277−286.
(46) Ngo, T. T.; Yoo, J.; Dai, Q.; Zhang, Q.; He, C.; Aksimentiev, A.; Ha, T. Effects of cytosine modifications on DNA flexibility and nucleosome mechanical stability. *Nat. Commun.* **2016**, *7*, 10813.

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246

(47) Lee, S. H.; Griffiths, J. R. How and Why Are Cancers Acidic? Carbonic Anhydrase IX and the Homeostatic Control of Tumour Extracellular pH. *Cancers (Basel)* **2020**, *12*, 1616.
(48) Swietach, P.; Vaughan-Jones, R. D.; Harris, A. L.; Hulikova, A. The chemistry, physiology and pathology of pH in cancer. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20130099.
(49) Spruijt, C. G.; Gnerlich, F.; Smits, A. H.; Pfaffeneder, T.; Jansen, P. W.; Bauer, C.; Münzel, M.; Wagner, M.; Müller, M.; Khan, F.; et al. Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell* **2013**, *152*, 1146−1159.
(50) Dodd, T.; Yan, C.; Kossmann, B. R.; Martin, K.; Ivanov, I. Uncovering universal rules governing the selectivity of the archetypal DNA glycosylase TDG. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 5974−5979.
(51) Coey, C. T.; Malik, S. S.; Pidugu, L. S.; Varney, K. M.; Pozharski, E.; Drohat, A. C. Structural basis of damage recognition by thymine DNA glycosylase: Key roles for N-terminal residues. *Nucleic Acids Res.* **2016**, *44*, 10248−10258.
(52) Kanaan, N.; Imhof, P. Interactions of the DNA Repair Enzyme Human Thymine DNA Glycosylase with Cognate and Noncognate DNA. *Biochemistry* **2018**, *57*, 5654−5665.
(53) Schröder, A. S.; Steinbacher, J.; Steigenberger, B.; Gnerlich, F. A.; Schiesser, S.; Pfaffeneder, T.; Carell, T. Synthesis of a DNA Promoter Segment Containing All Four Epigenetic Nucleosides: 5-Methyl-, 5-Hydroxymethyl-, 5-Formyl-, and 5-Carboxy- 2'-Deoxycytidine. *Angew. Chem., Int. Ed.* **2014**, *53*, 315−318.
(54) Favier, A.; Brutscher, B. NMRlib: user-friendly pulse sequence tools for Bruker NMR spectrometers. *J. Biomol. NMR* **2019**, *73*, 199−211.

https://doi.org/10.1021/acsphyschemau.1c00050
ACS Phys. Chem Au 2022, 2, 237−246