PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Measuring thermodynamic preferences to form non-native conformations in nucleic acids using ultraviolet melting

Atul Rangadurai[a,1], Honglue Shi[b,1], Yu Xu[b,1], Bei Liu[a], Hala Abou Assi[a], John D. Boom[b,c], Huiqing Zhou[a,2], Isaac J. Kimsey[a,3], and Hashim M. Al-Hashimi[a,b,4,5]

Edited by Joseph Puglisi, Stanford University School of Medicine, Stanford, CA; received August 8, 2021; accepted April 5, 2022

Thermodynamic preferences to form non-native conformations are crucial for understanding how nucleic acids fold and function. However, they are difficult to measure experimentally because this requires accurately determining the population of minor low-abundance (<10%) conformations in a sea of other conformations. Here, we show that melting experiments enable facile measurements of thermodynamic preferences to adopt nonnative conformations in DNA and RNA. The key to this "delta-melt" approach is to use chemical modifications to render specific minor non-native conformations the major state. The validity and robustness of delta-melt is established for four different non-native conformations under various physiological conditions and sequence contexts through independent measurements of thermodynamic preferences using NMR. Delta-melt is faster relative to NMR, simple, and cost-effective and enables thermodynamic preferences to be measured for exceptionally low-populated conformations. Using delta-melt, we obtained rare insights into conformational cooperativity, obtaining evidence for significant cooperativity (1.0 to 2.5 kcal/mol) when simultaneously forming two adjacent Hoogsteen base pairs. We also measured the thermodynamic preferences to form G-C$^+$ and A-T Hoogsteen and A-T base open states for nearly all 16 trinucleotide sequence contexts and found distinct sequence-specific variations on the order of 2 to 3 kcal/mol. This rich landscape of sequence-specific non-native minor conformations in the DNA double helix may help shape the sequence specificity of DNA biochemistry. Thus, melting experiments can now be used to access thermodynamic information regarding regions of the free energy landscape of biomolecules beyond the native folded and unfolded conformations.

conformational dynamics | melting | relaxation dispersion

Biomolecules do not fold into a single structure but rather form dynamic ensembles of many interconverting conformations (1–3). Low-populated non-native conformational states in the dynamic ensemble can become the major state when a molecule binds a partner molecule and forms a complex (Fig. 1 *A* and *B*). In these cases, energetically favorable intermolecular contacts must cover the energetic cost, or conformational penalty (ΔG°_{conf}), associated with stabilizing the minor non-native conformation relative to the major native state (Fig. 1*A*) (4–6). Therefore, the thermodynamic preferences to form non-native conformations can be essential determinants of binding affinity and specificity and catalytic efficiency.

For example, in the DNA double helix, Watson–Crick G-C and A-T base pairs (bps) exist in dynamic equilibrium with a minor Hoogsteen conformation, with populations on the order of 0.1 to 1.0% (7, 8), corresponding to an energetic cost of ΔG°_{conf} = 2.7 to 4.0 kcal/mol. However, the Hoogsteen bp can become the major conformation when duplex DNA is bound by proteins such as p53 (9) (Fig. 1*B*). Here, energetically favorable intermolecular contacts make up for the conformational penalty to convert the bp from Watson–Crick to the minor Hoogsteen conformation. Indeed, DNA modifications that increase the thermodynamic propensities to form Hoogsteen bps have been shown to improve p53-DNA binding affinity (4, 10). Changes in the structures of RNA and DNA are ubiquitously observed upon binding to proteins and other partners (11, 12). In all cases, the binding energetics will have to make up for the relevant conformational penalties.

Conformational preferences, and therefore conformational penalties, can vary depending on sequence and structural contexts, due to epigenetic and post-transcriptional modifications as well as damage, and due to changes in physiological conditions such as pH, temperature, salt composition, and concentration. This can, in turn, help shape the specificities of biochemical processes.

For example, bps in DNA can spontaneously open (Fig. 1*C*) (13, 14) and when bound to damage repair proteins can adopt extrahelical conformations (15). The thermodynamic preferences to form the base open and extrahelical conformations have been shown to vary with base modifications and sequence context, in turn determining the

Significance

Thermodynamic preferences of nucleic acids to adopt non-native conformations are crucial for understanding how they function but prove difficult to measure experimentally. As a result, little is known about how these thermodynamic preferences vary with sequence and structural contexts, physiological conditions, and chemical modifications. Here, we show that modifications stabilizing non-native conformations and rendering them the major state, in conjunction with melting experiments, enable facile measurements of thermodynamic preferences to form various nonnative conformations in DNA and RNA. Delta-melt provided rare insights into the cooperativity of forming tandem Hoogsteen base pairs and revealed large and distinct sequence-specific preferences to form G-C$^+$ and A-T Hoogsteen and A-T base open conformations in DNA, which may contribute to sequence-specific DNA biochemistry.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]A.R., H.S., and Y.X. contributed equally to this work.

[2]Present address: Department of Chemistry, Boston College, Boston, MA 02467.

[3]Present address: Nymirum, Durham, NC 27713.

[4]To whom correspondence may be addressed. Email: ha2639@cumc.columbia.edu.

[5]Present address: Columbia University Irving Medical Center BSL2 Lab, Columbia University, New York, NY 10032.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2112496119/-/DCSupplemental.

Published June 7, 2022.

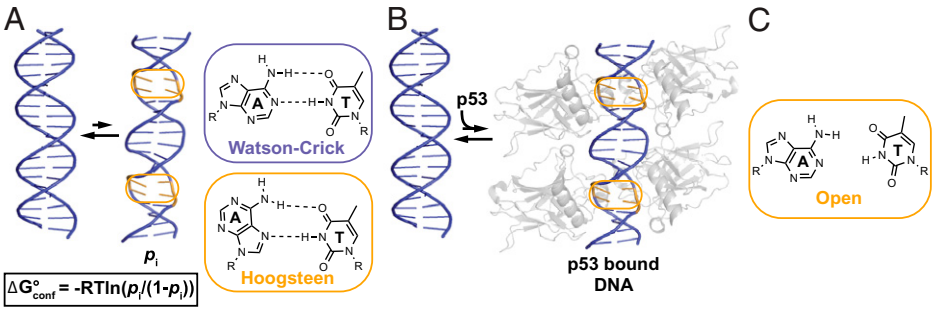

Fig. 1. Thermodynamic preferences to form minor non-native conformations in DNA can influence binding. (*A*) The p53-bound conformation of a DNA duplex with Hoogsteen bps (in orange) exists as a minor conformation with population p_i in the unbound DNA ensemble. The free energy cost to stabilize the minor conformation on binding of p53 depends on p_i and is given by $\Delta G^\circ_{conf} = -RT\ln(p_i/(1-p_i))$, in which R is the universal gas constant (units of kilocalories per mole per degrees Kelvin, and T is the temperature in Kelvin). (*B*) Binding of p53 to duplex DNA (Protein Data Bank ID: 3KZ8) (9) results in a change in the conformation of bps from Watson–Crick (purple) in the free DNA to Hoogsteen (orange) in the DNA–p53 complex. The free DNA structure corresponds to an idealized B-form DNA double helix generated using 3DNA (93). (*C*) Schematic representation of the base open state of A-T bps, which is implicated in DNA–protein recognition and damage repair.

binding affinity (15) and catalytic efficiency (16) of glycosylase enzymes, which excise damaged bases from DNA, as well as the processivity of helicases, which unwind DNA (17). DNA bps near mismatches have been shown to have increased propensities to form Hoogsteen bps, and Hoogsteen bps are often observed in non-canonical regions of DNA in crystal structures of DNA–protein complexes (18, 19). In addition, wobble G•T/U mismatches can form Watson–Crick-like conformations through tautomerization and ionization of the bases (20, 21). The preferences to form Watson–Crick-like conformations can vary with pH and chemical modifications, in turn shaping the probability of misincorporating G•T bps during DNA replication (20) and providing a basis for recoding translation (22–24).

In RNA, sequence-specific preferences to adopt alternative conformations determine the energetics of RNA–RNA tertiary assembly (25, 26) and RNA–protein binding affinities (27). In addition, pH- and temperature-dependent preferences to adopt alternative secondary structures can help control messenger RNA (mRNA) translation (28, 29) and the efficiency of enzymatic processing of micro RNAs (30). By increasing the thermodynamic preferences to form single-stranded RNA, the epitranscriptomic modification N^6-methyladenosine (31) has been shown to improve the binding of single-stranded RNA binding proteins to modified RNAs (32). Aberrant modulation of RNA conformational preferences due to mutations have also been linked to disease (reviewed in ref. 5).

Measuring conformational preferences in DNA and RNA and their variation as a function of sequence, structural and chemical contexts, and physiological conditions remains a significant challenge in biophysics and structural biology. This is because measuring the energetic cost to form non-native conformations requires accurately determining their low abundance in a sea of other conformations in the apo-ensemble. A variety of approaches have been used to measure such conformational equilibria, including NMR (33), electron paramagnetic resonance (34) and infrared spectroscopy (35), fluorescence resonance energy transfer (36), cryogenic electron microscopy and X-ray crystallography (37), and chemical probing (38). However, with the exception of NMR, most of these techniques are not applicable to conformations such as Hoogsteen bps and the base open states, which have populations <1%, are short-lived with lifetimes shorter than a few milliseconds, and which involve localized bp rearrangements. Most of these approaches are also technically demanding, often requiring synthesis of molecules with specialized labels, with NMR in particular requiring large sample quantities. Thus, they do not lend themselves to comprehensively exploring the conformational preferences of nucleic acids. As a result, the thermodynamic preferences to form Hoogsteen or base open states have yet to be measured for all 16 trinucleotide sequence contexts. Therefore, there is a need for alternative approaches that can more readily measure conformational preferences in nucleic acids as well as proteins.

Prior NMR studies have shown that chemical modifications and mutations can render minor non-native conformations (Fig. 2*A*) the major state (population >90%) in proteins (39) and nucleic acids (40), thereby enabling their in-depth structural as well as functional characterization (41). Here, we show that by combining such chemical modifications (Fig. 2*B*), with melting experiments that measure folding energetics, it is feasible to measure the thermodynamic preferences to form minor non-native conformations in RNA and DNA. We used ultraviolet (UV) spectroscopy to measure melting energetics (42, 43) (Fig. 2*C*) given its simplicity, lack of requirement for specific labeling, and reduced sample requirements relative to other approaches (44–46).

The approach termed "delta-melt" builds on prior studies employing differences in melting energetics to estimate the contribution of motifs (44, 47) or specific functional groups (48, 49) to nucleic acid stability. Such studies provided the basis for nearest-neighbor rules and secondary structure prediction algorithms (50, 51). However, in contrast to these prior studies, delta-melt measures the difference in melting energetics between an unmodified (wild-type, WT) molecule and a molecule chemically modified (Mod) to shift the ensemble toward the desired non-native state (Fig. 2 *A* and *B*). Converting differences in melting energetics into differences between the thermodynamic stability of the major and minor conformational states requires calibration against independent measurements of conformational preferences using other biophysical techniques. In this manner, delta-melt enables measurement of thermodynamic preferences using a pair of melting measurements in a facile manner, with fewer sample requirements relative to alternative biophysical techniques such as NMR. In this study, we extensively verified the robustness of delta-melt for four different non-native conformational states under a variety of physiological conditions and sequence contexts through independent measurements of thermodynamic preferences using NMR.

We obtained rare insights into conformational cooperativity using delta-melt, estimating that tandem Hoogsteen bps form cooperatively by 1.0 to 2.5 kcal/mol in DNA. We also measured the thermodynamic preferences to form G-C^+ and A-T Hoogsteen and A-T base open states for nearly all 16 trinucleotide sequence contexts and found distinct sequence-specific variations on the order of 2 to 3 kcal/mol, corresponding to ~10- to 100-fold variation in the population of the minor conformational state. These results suggest that the DNA double helix codes for a rich layer of sequence-specific conformational

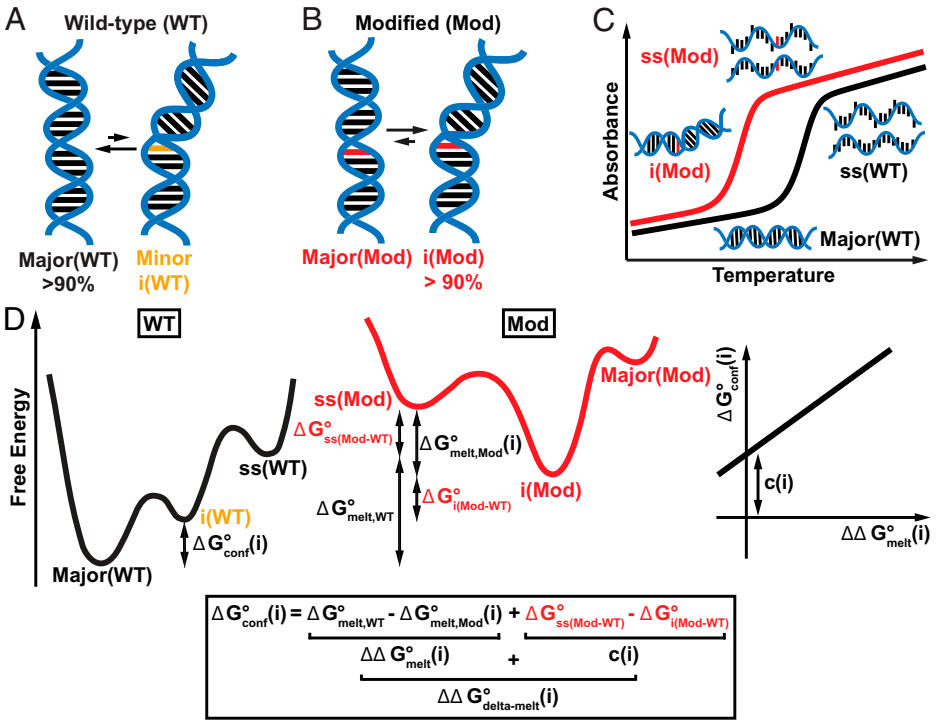

Fig. 2. Conceptual underpinnings of delta-melt for measuring thermodynamic preferences to adopt non-native conformational states in nucleic acids. (*A*) Dynamic equilibrium in the unmodified WT nucleic acid molecule involving a major and the i^{th} minor non-native conformation [i(WT)] involving a change in the conformation of a bp (in orange). (*B*) Chemical modifications (red, Mod) are introduced to bias the conformational ensemble toward the i^{th} non-native conformational state i(WT) of the WT nucleic acid. (*C*) Schematic diagram showing optical melting experiments on unmodified WT (black) and modified Mod (red) nucleic acids. (*D*) Free energy diagram of delta-melt (*Left*), relationship between the free energies in delta-melt (*Bottom*), and expected correlation between $\Delta G^{\circ}_{conf}(i)$ and $\Delta G^{\circ}_{melt}(i)$(*Right*).

preferences, which could in turn tune the specificity of DNA biochemical transactions. Thus, melting experiments can be used to access thermodynamic information regarding regions of the free energy landscape of biomolecules beyond the native folded and unfolded conformations, opening the door to more systematic explorations of how the ensemble varies with sequence and structural contexts, chemical modifications, and physiological conditions.

Results

Conceptual Underpinnings of Delta-Melt. The conceptual underpinnings of delta-melt can be understood by using a free energy diagram (Fig. 2*D*). The desired thermodynamic preference, $\Delta G^{\circ}_{conf}(i)$, in the WT molecule to form the i^{th} non-native conformational state is given by the free energy difference between the i^{th} non-native [i(WT)] and major [Major(WT)] conformations (Fig. 2*A*),

$$\Delta G^{\circ}_{conf}(i) = \Delta G^{\circ}_{conf}\left(i(WT)\right) - \Delta G^{\circ}_{conf}(Major(WT)).$$

In principle, $\Delta G^{\circ}_{conf}(i)$ could be obtained from the difference in the melting energetics (ΔG°_{melt}) between the Major(WT) and i(WT) conformational states in the WT molecule:

$$\Delta G^{\circ}_{conf}(i) = \Delta G^{\circ}_{melt,Major(WT)} - \Delta G^{\circ}_{melt,i(WT)}.$$

In practice however, because the minor i(WT) conformation is lowly populated, its melting energetics are difficult to measure experimentally (52, 53). To address this limitation, we introduce a chemical modification that renders i(WT) the dominant conformation in the modified (Mod) molecule, typically by destabilizing the major state relative to the minor state (Fig. 2*B*). The degree to which the modification will affect the melting energetics relative to the WT molecule will depend on the desired thermodynamic preference to form the minor conformation $\Delta G^{\circ}_{conf}(i)$ and also on how the modification affects the energetics of the minor conformation and single-strand (ss) (Fig. 2*D*). Therefore, the desired $\Delta G^{\circ}_{conf}(i)$ can be obtained from the difference between the melting energetics of WT and Mod molecules, $\Delta G^{\circ}_{melt}(i)$, plus an offset, c(i), which takes into account the difference in free energy to modify ss relative to i(WT) (Fig. 2*D*):

$$\Delta G^{\circ}_{conf}(i) = \Delta G^{\circ}_{melt}(i) + c(i)$$

$$\Delta G^{\circ}_{melt}(i) = \Delta G^{\circ}_{melt,WT} - \Delta G^{\circ}_{melt,Mod}(i)$$

$$c(i) = \Delta G^{\circ}_{ss(Mod-WT)} - \Delta G^{\circ}_{i(Mod-WT)}$$

Thus, $\Delta G^{\circ}_{conf}(i)$ can be estimated using a pair of melting experiments, provided that the offset c(i) is known. The value of c(i) can be determined experimentally provided that it does not vary across different sequence contexts and/or environmental conditions (changes in temperature, pH, and salt) of interest, by calibrating $\Delta G^{\circ}_{melt}(i)$ data against $\Delta G^{\circ}_{conf}(i)$ values independently measured via biophysical techniques such as NMR for a few experimental conditions and/or sequence contexts (Fig. 2*D*). c(i) can be obtained as the intercept from a linear fit of $\Delta G^{\circ}_{melt}(i)$ versus $\Delta G^{\circ}_{conf}(i)$ (Fig. 2*D*). Such calibrations with resource intensive measurements are routinely performed in other contexts such as microarray-based binding measurements, which are calibrated against isothermal titration calorimetry/electrophoretic mobility shift assay experiments to enable measurement of binding affinities in high throughput (4, 54). The calibration enables us to take into consideration the effect of the modification on the energetics of the minor and ss states. This is in contrast to prior studies employing melting experiments to estimate the energetic contribution of specific motifs to nucleic acid stability (44, 47–49, 55), in which it is typically assumed that modifications do not impact the ss energetics (i.e., $\Delta G^{\circ}_{ss(Mod-WT)} \sim 0$). Following calibration to estimate c(i), $\Delta G^{\circ}_{conf}(i)$ for a new sample can be estimated in a facile manner using a pair of melting experiments as the sum ($\Delta G^{\circ}_{delta-melt}(i)$) of the difference in melting energetics $\Delta G^{\circ}_{melt}(i)$ and c(i):

$$\Delta G^{\circ}_{conf}(i) = \Delta G^{\circ}_{delta-melt}(i) = \Delta G^{\circ}_{melt}(i) + c(i).$$

One of the advantages of delta-melt is that in theory there is no limit on the population and lifetime of the minor conformation that can be studied as long as it can be stabilized using a modification. It should be applicable to conformational states that fall outside NMR detection, and we will encounter such an example in this study (see below). Likewise, in theory, there is no size limit provided that the site-specifically modified nucleic acid can be synthesized (56). However, deviations from two-state melting behavior are more likely in large or more complex nucleic acid molecules, and this can complicate analysis of the UV melting data, as observed even for some of the short duplexes studied here (see below). In such cases, the melting experiments have to be supplemented with information from other techniques to understand the structural basis behind changes in absorbance (reviewed in ref. 57). In addition, delta-melt does not provide any information regarding the kinetics of interconversion and, consequently, binding mechanisms (induced fit vs. conformational selection) (58) or structure of the non-native state. It is also assumed that the modification robustly renders the desired i^{th} minor conformation the major state (>90%) across different contexts and experimental conditions. If the modification confers a weaker bias, delta-melt may underestimate the true value of $\Delta G^{\circ}_{conf}(i)$. In the subsequent sections, we apply delta-melt to a wide range of conformational transitions in nucleic acids in order to evaluate the generality and robustness of the approach and test the validity of the above assumptions underpinning the approach.

Testing Delta-Melt on A-T Hoogsteen bps in Duplex DNA. We initially performed a set of experiments to test the underlying assumptions and general utility of delta-melt to measure conformational preferences to form four different minor non-native conformations in DNA and RNA. As a first test, we used delta-melt to measure the preferences to form A-T Hoogsteen bps (7) in duplex DNA. The thermodynamic preferences to form A-T Hoogsteen bps in duplex DNA are of interest because they are proposed to determine DNA–protein binding affinities (4, 10) and the propensity of DNA damage induction (59).

To form a Hoogsteen bp, the purine base in a Watson–Crick bp has to flip 180° about the glycosidic bond to adopt a *syn* conformation, and the DNA backbone has to constrict by ~2 Å to allow hydrogen bonding between the bases (Fig. 3*A*) (60, 61). We used N^1-methylated adenine (m^1A^+), a positively charged naturally occurring form of alkylation damage, to substantially bias (>90%) the conformation of the A-T bp toward Hoogsteen, as shown previously (7, 62). The modification destabilizes the Watson–Crick A-T bp both by disrupting a hydrogen bond and through steric collisions (Fig. 3*A*) (7, 62). Prior NMR studies showed m^1A^+-T to be an excellent structural mimic of the transient A-T Hoogsteen bps in unmodified DNA (63). As m^1A^+-T is a good structural mimic of the transient A-T Hoogsteen bp, it may also serve as a good energetic mimic for delta-melt experiments.

We analyzed sequences shown previously to form stable duplexes in solution and bp positions for which the Watson–Crick to Hoogsteen exchange had previously been characterized using NMR relaxation dispersion (RD) (7, 63, 64) (Fig. 3*B*). This included A-T bps in two different trinucleotide sequences (5′-CAA-3′/5′-AAT-3′) embedded in three different duplexes in two different positional contexts. These data were supplemented with additional NMR RD and corresponding delta-melt experiments under varying buffer and temperature (T) conditions (pH = 4.4 to 6.9, T = 10 to 30 °C, and 25 to 150 mM NaCl) (Fig. 3*B* and *SI Appendix*, Figs. S1, S2, and S4–S6 and Tables S1–S5). The pK_a for ionization of the amino group in the m^1A base is 7.2 (65); thus, under the pH conditions used here, the amino group is likely to be protonated and m^1A is expected to be a good mimic of the Hoogsteen bp, as shown previously (63). Furthermore, prior studies (64) have also shown that under the conditions used for the UV melts the contribution of Dimroth rearrangement of m^1A to N^6-methyladenosine is negligible. Additional measurements were performed when chemically modifying the Hoogsteen bp with ribo-guanosine or its neighbor using N^1-methyl guanosine, both of which are known to perturb the thermodynamic propensity to form Hoogsteen bps (64) (Fig. 3*B*). This yielded a total of 14 data points which could be used to test delta-melt.

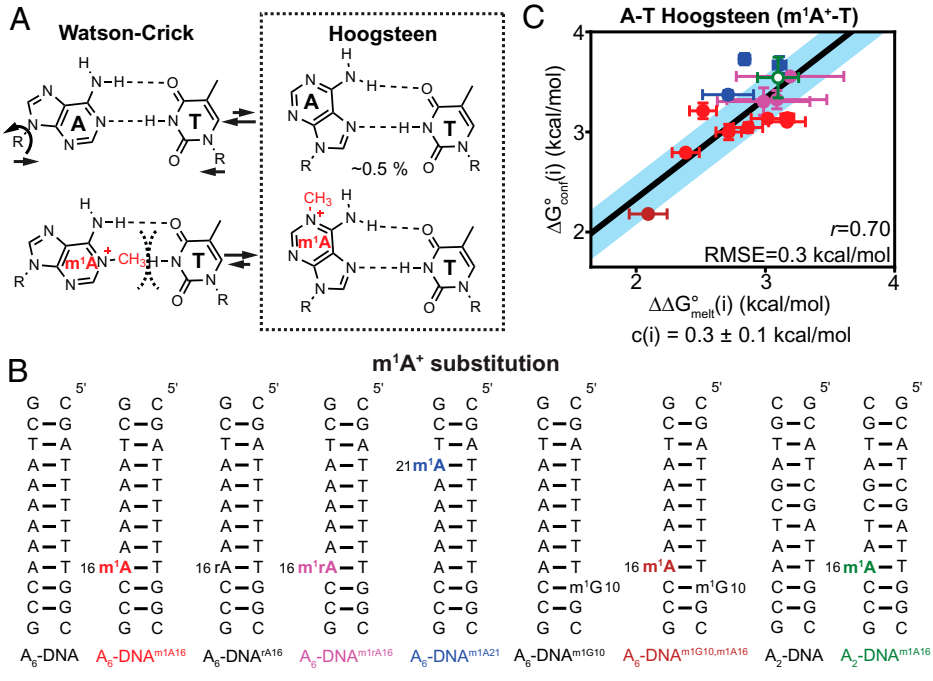

Fig. 3. Testing delta-melt on A-T Hoogsteen bps in DNA using the N^1-methyladenosine mimic. (*A*) In duplex DNA, A-T Watson–Crick bps exist in dynamic equilibrium with minor A-T Hoogsteen bps. Population of the Hoogsteen bp was obtained as described previously (8). N^1-methylation of adenine (m^1A^+, red) biases the equilibrium to favor the Hoogsteen bp. Steric clashes are indicated using dashed lines. (*B*) DNA duplexes used in UV melting experiments with and without m^1A^+ substitution (in bold). rA corresponds to ribo adenosine. (*C*) Correlation plot between $\Delta G^{\circ}_{conf}(i)$ for formation of A-T Hoogsteen bps (*SI Appendix*, Table S5) and $\Delta G^{\circ}_{melt}(i)$ obtained from melting measurements on m^1A^+-T/A-T bp containing DNA (*SI Appendix*, Tables S1 and S2). Points are color-coded according to duplex (in *B*). Error bars for NMR and delta-melt measurements were obtained using a Monte-Carlo scheme (94) and by propagating the uncertainties from UV melts, respectively, as described in *Materials and Methods*. Shown are the Pearson's correlation coefficient (*r*) and RMSE (*Materials and Methods*). The blue shaded region denotes the estimate of the error of linear regression obtained using Monte-Carlo simulations, while open symbols denote data derived from weak RD profiles (*Materials and Methods*).

All delta-melt experiments were performed using the same buffer conditions used in the NMR RD experiments (*Materials and Methods*). The "curve-fitting" approach (66, 67) was used to extract the free energy of melting (ΔG°_{melt}) from the UV melting curves for all samples (66, 67). The approach assumes two-state melting behavior and is commonly used in the literature (66, 67). A more rigorous approach evaluates the two-state assumption using concentration-dependent melting experiments to extract thermodynamic parameters (42, 43, 68). Although concentration-dependent measurements help gauge the two-state behavior and are also required to determine the enthalpy (ΔH°_{melt}) and entropy (ΔS°_{melt}) of melting more accurately, this substantially increases the sample requirements and reduces the throughput of the experiment. In addition, the concentration dependence is a necessary but insufficient condition for two-state melting (44). As we were primarily interested in free energy of melting (ΔG°_{melt}), which can be measured accurately using curve-fitting since the errors in the enthalpy and entropy typically cancel out (42), and since we have already previously verified using NMR that some of the duplexes studied here (A_2 and A_6-DNA) (69) exhibit two-state melting behavior, we favored using the curve-fitting approach in the application of delta-melt. For a more elaborate discussion regarding the UV melting experiments and the assumptions used in the analysis refer to *SI Appendix, Discussion S1*.

Indeed, a reasonable correlation ($r = 0.70$) was observed between $\Delta G^{\circ}_{melt}(i) = \Delta G^{\circ}_{melt,\,A-T} - \Delta G^{\circ}_{melt,m1A+-T}$ (*SI Appendix*, Tables S1 and S2) and $\Delta G^{\circ}_{conf}(i)$ measured using NMR RD (Fig. 3*C* and *SI Appendix*, Table S5). The correlation could be linearly fit ($r = 0.7$; Fig. 3*C*) with a root-mean-square error (RMSE) of 0.3 kcal/mol and with c(i) ~0.3 kcal/mol. The quantitative agreement between the delta-melt energetics and those derived using NMR RD suggest that the offset c(i) corresponding to the difference in energetic cost to methylate adenine-N1 in the ss versus the A-T Hoogsteen bp (*SI Appendix, Discussion S2*) does not vary substantially across the sequence and chemically modified contexts and buffer conditions examined, and that any variations remain small relative to the differences in $\Delta G^{\circ}_{conf}(i)$ (~2 kcal/mol). These results therefore also indicate that delta-melt can be used to measure the thermodynamic preferences to form A-T Hoogsteen bps as a function of experimental conditions.

Nevertheless, the correlation was not ideal and m^1A^+-T is an imperfect mimic of the A-T Hoogsteen bp as the methyl group introduces a positive charge to the adenine base. We therefore also tested neutral modifications such as C-T and T-T mismatches, which were recently shown to mimic the A-T Hoogsteen bp based on NMR chemical shifts analysis (4) (Fig. 4*B*). As in the A-T Hoogsteen bp, the two bases in C-T and T-T mismatches also have to come into closer proximity relative to Watson–Crick bps to hydrogen bond, thus mimicking this key feature of the Hoogsteen conformation (63) (Fig. 4*A*). Mismatches also have the advantage that they can be readily installed when synthesizing DNA. These experiments also assessed the robustness of delta-melt when using different modifications to mimic the minor conformation.

Indeed, an improved correlation ($r = 0.98$ and 0.93) was observed between $\Delta G^{\circ}_{melt}(i) = \Delta G^{\circ}_{melt,\,A-T} - \Delta G^{\circ}_{melt,C-T\text{ or }T-T}$ for both C-T and T-T mismatches (Fig. 4 *C* and *D* and *SI Appendix*, Table S2) and $\Delta G^{\circ}_{conf}(i)$ to form A-T Hoogsteen bps as measured using NMR RD (*SI Appendix*, Table S5). The data could be fit with RMSE = 0.2 kcal/mol for both C-T and T-T (Fig. 4 *C* and *D* and *SI Appendix, Discussion S2*). Nevertheless, the deviation seen for the A16-T9 bp near the m^1G10-C15$^+$ Hoogsteen bp in A_6-DNA^{m1G10} (brown point in Fig. 4 *C* and *D*) indicates that C-T/T-T mismatches may not be good energetic mimics of A-T Hoogsteen bp when neighboring another Hoogsteen bp, possibly because they do not capture unique stacking interactions involving the adjacent *syn* purine base. Indeed, this deviation was not observed when using m^1A^+ (Fig. 3*C*).

Taken together, the above results establish m^1A^+-T, C-T and T-T bps to be good energetic mimics of A-T Hoogsteen bps and support the underlying assumptions used in delta-melt to measure the thermodynamic preferences to form A-T Hoogsteen bps.

In an analogous manner, delta-melt could also be used to measure pH dependent thermodynamic preferences to form G-C$^+$ Hoogsteen base pairs in duplex DNA, by using the naturally occurring N^1-methyl guanosine (m^1G) modification to bias the conformation of the G-C bp towards Hoogsteen (*SI Appendix*, Fig. S7).

Testing Delta-Melt on Base Open A-T States in Duplex DNA. One of the advantages of delta-melt is that it can potentially be applied to low-populated and short-lived states falling below

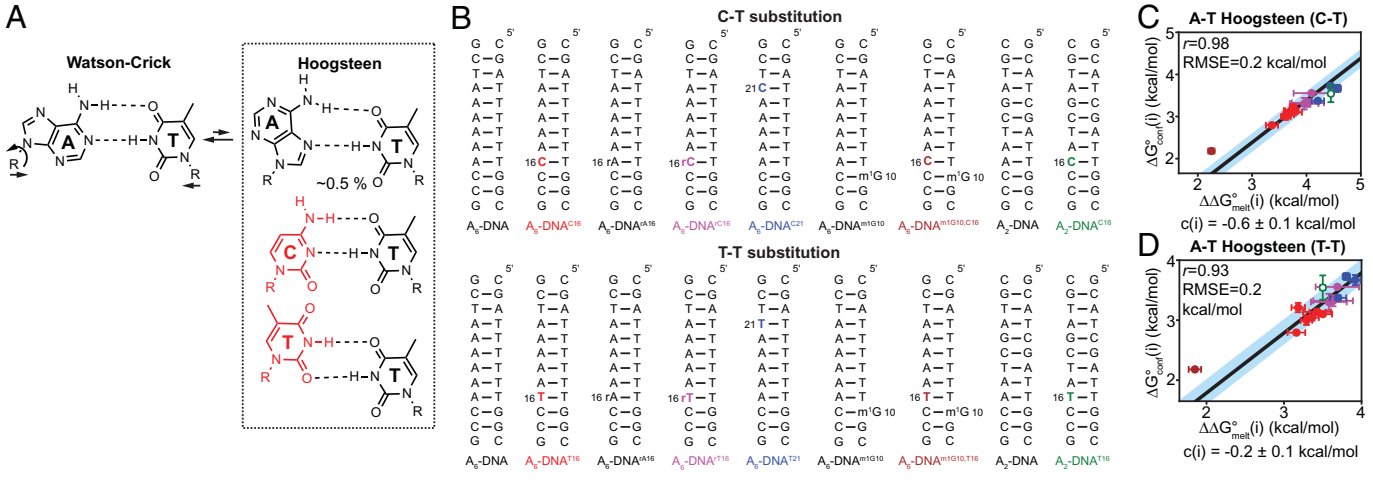

Fig. 4. Testing delta-melt on A-T Hoogsteen bps in DNA using C-T and T-T mismatches as Hoogsteen mimics. (*A*) C-T and T-T mismatches mimic the conformation of the A-T Hoogsteen bp. (*B*) Duplexes used in UV melting experiments with and without C-T and T-T mismatches (in bold). (*C* and *D*) Correlation plot between $\Delta G^{\circ}_{melt}(i)$ obtained from melting experiments on C-T/A-T (*C*) and T-T/A-T (*D*) bp containing DNA (*SI Appendix*, Table S2) and $\Delta G^{\circ}_{conf}(i)$ for formation of A-T Hoogsteen bps as determined using NMR RD (*SI Appendix*, Table S5). Points in *C* and *D* are color-coded according to the corresponding duplex in *B*. Error bars for NMR and delta-melt measurements were obtained using a Monte-Carlo scheme (94) and by propagating the uncertainties from UV melts, respectively, as described in *Materials and Methods*. Shown are the Pearson's correlation coefficient (*r*) and RMSE (*Materials and Methods*). The blue shaded region denotes the estimate of the error of linear regression obtained using Monte-Carlo simulations, while open symbols denote data derived from weak RD profiles (*Materials and Methods*).

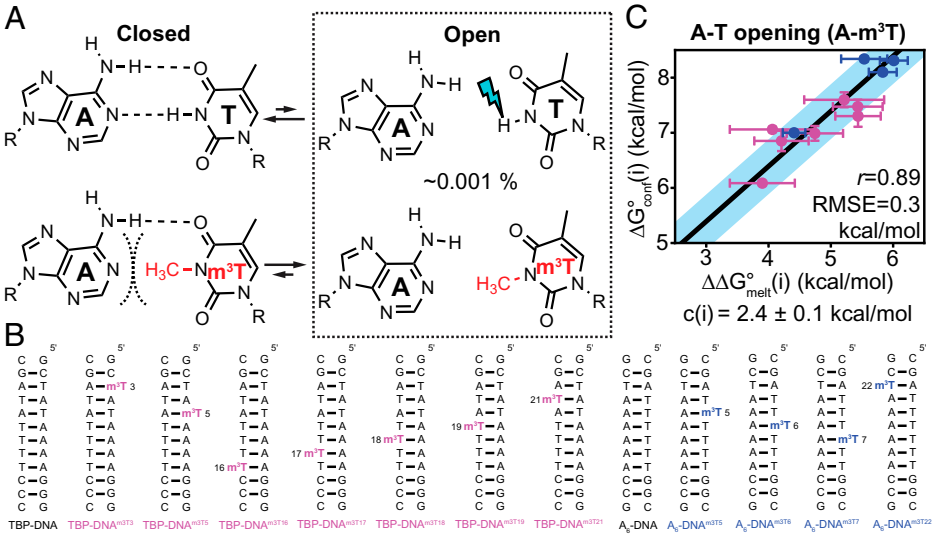

Fig. 5. Testing delta-melt on the open state of A-T bps using the N^3-methylthymine mimic. (*A*) Exchange between closed and open A-T Watson–Crick bps in DNA. The T-H3 imino proton in the open state is susceptible to solvent exchange (blue thunder). m^3T is hypothesized to bias the ensemble to mimic the base open state by disrupting the A(N1)–H3-(N3)T hydrogen bond and by sterically (curved dashed lines) disfavoring the closed bp. (*B*) Duplexes used in UV melting experiments with and without m^3T substitution (in bold). (*C*) Correlation plot between $\Delta G^\circ_{conf}(i)$ for formation of open A-T bps (*SI Appendix*, Table S8) and $\Delta G^\circ_{melt}(i)$ obtained from melting experiments on A-m^3T/A-T bp containing DNA (*SI Appendix*, Tables S1 and S2). Points are color-coded according to the duplex that they correspond to (in *B*). Error bars for NMR and delta-melt measurements were obtained using a Monte-Carlo scheme and by propagating the uncertainties from UV melts, respectively, as described in *Materials and Methods*. Shown are the Pearson's correlation coefficient (*r*) and RMSE (*Materials and Methods*). The blue shaded region denotes the estimate of the error of linear regression obtained using Monte-Carlo simulations (*Materials and Methods*).

detection methods of conventional biophysical methods. We therefore tested its utility on the A-T base open state (Fig. 1*C*), which based on NMR measurements of imino proton exchange rates (14) has an exceptionally low population of ~0.001% and short lifetime on the order of ~0.1 μs (Fig. 5*A*). Compared to G-C^+ and A-T Hoogsteen bps, the base open state is disfavored by ~4 kcal/mol, providing an interesting test for the general applicability of delta-melt on an exceptionally low-populated conformational state that cannot be detected even by NMR RD methods.

Based on imino proton solvent exchange experiments (70), the base open state entails the loss of the A(N1)–H3-(N3)T hydrogen bond. Thus, we hypothesized that substituting T by N^3-methyl thymine (m^3T), so that the thymine imino proton is replaced by a methyl group, should disrupt the A(N1)–H3-(N3)T hydrogen bond and bias the ensemble toward the base open state (70). Indeed, NMR spectra (*SI Appendix*, Fig. S8) revealed that the modification disrupted the A(N1)–H3-(N3)T hydrogen bond in the modified bp with m^3T adopting an intrahelical conformation.

We analyzed sequences shown previously to form stable duplexes in solution and bp positions for which A-T base opening has been previously characterized using NMR (TBP-DNA, Fig. 5*B*) (71). We also performed additional imino proton exchange NMR and corresponding delta-melt experiments for another duplex (A_6-DNA) under different buffer and temperature conditions, yielding in total 11 points for comparison, corresponding to six different trinucleotide sequence contexts (5′-TAG-3′/5′-TAT-3′/5′-AAG-3′/5′-AAA-3′/5′-TAA-3′/5′-GAT-3′) in two different duplexes and a range of buffer conditions (pH 8.0 to 8.8, T = 15 to 25 °C, 100 mM NaCl) (Fig. 5*B* and *SI Appendix*, Figs. S1, S2, S4, and S9, and Tables S1, S2, and S6–S8.

UV melting experiments were used to measure the energetics of duplex melting with and without m^3T substitutions at these positions under the NMR buffer conditions (*SI Appendix*, Figs. S1 and S2 and Tables S1 and S2). Interestingly, a very good correlation ($r = 0.89$) was observed between the difference in melting energetics $\Delta G^\circ_{melt}(i) = \Delta G^\circ_{melt,A-T} - \Delta G^\circ_{melt,A-m3T}$ and $\Delta G^\circ_{conf}(i)$ for base opening measured using NMR imino proton exchange measurements (Fig. 5*C* and *SI Appendix*, Tables S1, S2, and S8) (71). The correlation could be fit with RMSE of 0.3 kcal/mol and c(i) ~2.4 kcal/mol (Fig. 5*C*) (*SI Appendix*, *Discussion S2*). These results demonstrate the utility of delta-melt to measure the thermodynamic preferences for A-T base opening across different sequence contexts and buffer conditions and also establish A-m^3T as a mimic of the base open state which has been recalcitrant to experimental structural characterization.

Testing Delta-Melt on Isomerization of the N^6-Methylamino Group in Duplex RNA. We also examined the utility of delta-melt to address more subtle conformational rearrangements involving rotation of a single methylamino group. N^6-methyladenosine (m^6A) is a highly abundant RNA modification that plays roles in virtually all aspects of mRNA metabolism (31, 72). When paired with uridine, the methylamino group in m^6A isomerizes between *anti* (major) and *syn* (minor) conformations, resulting in the loss of a hydrogen bond (Fig. 6*A*) (73). This conformational change has been proposed to play roles slowing a variety of biochemical processes that involve duplex melting and annealing (69).

Using NMR chemical exchange saturation transfer (CEST) experiments (74), it was recently shown that the minor non-native m^6A(*syn*)-U bp transiently forms with a population of ~1% (73). It was also shown that N^6,N^6-dimethyladenosine ($m^6{}_2A$), a posttranscriptional modification in transfer RNA (tRNA) (75), structurally mimics this non-native m^6A(*syn*)-U bp (73). This modification destabilizes the major m^6A(*anti*)-U conformation both by disrupting the A(N6)-H6–(O4)T hydrogen bond and through steric collisions (curved dashed lines, Fig. 6*A*).

We tested delta-melt on two sequences shown previously to form stable duplexes in solution, at bp positions for which the non-native m^6A(*syn*)-U bp has been previously characterized using NMR and for which the $m^6{}_2A$ modification has been shown to recapitulate the NMR chemical shifts of the non-native m^6A(*syn*)-U bp (73) (*SI Appendix*, Fig. S4). Delta-melt was used to measure temperature-dependent (pH 6.8, T = 37 to 65 °C, 25 mM NaCl) thermodynamic preferences to form the non-native m^6A(*syn*)-U bp at two positions in two duplexes in two different trinucleotide sequence and position contexts (5′-CAA-3′/5′-GAC-3′) (Fig. 6*B*); UV melting experiments were used to measure the energetics of duplex melting with m^6A and $m^6{}_2A$ modifications at these positions under the NMR buffer conditions (*SI Appendix*, Figs. S1 and S4–S6 and Tables S1 and S3–S5).

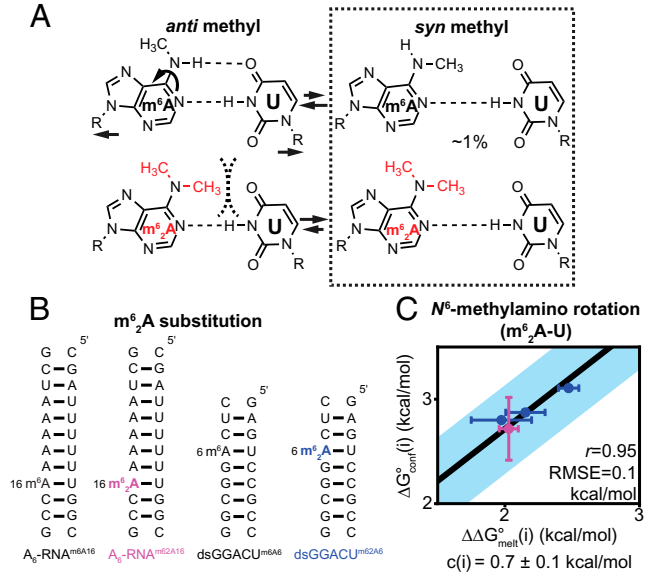

Fig. 6. Testing delta-melt on the isomerization of the N^6-methylamino group in m^{6}A-U bps in RNA. (*A*) (*Top*) Equilibrium between major *anti* and minor *syn* conformations of the N^6-methylamino group in m^{6}A-U bps in RNA. (*Bottom*) m$^6{}_2$A mimics the *syn* methylamino group by sterically disfavoring (black dashes) the conformation with the *anti* methyl. (*B*) Duplexes used in UV melting experiments with m^{6}A and m$^6{}_2$A (indicated in bold) modifications. (*C*) Correlation plot between $\Delta G^\circ_{melt}(i)$ obtained from melting experiments on m^{6}A-U/m$^6{}_2$A-U bp containing RNA (*SI Appendix*, Table S1) and $\Delta G^\circ_{conf}(i)$ measured using NMR RD for formation of m^{6}A(*syn*)-U bps (*SI Appendix*, Table S5). Points are color-coded according to the duplexes in *B*. Error bars for NMR and delta-melt measurements were obtained using a Monte-Carlo scheme (94) and by propagating the uncertainties from UV melts, respectively, as described in *Materials and Methods*. Shown are the Pearson's correlation coefficient (*r*) and RMSE (*Materials and Methods*). The blue shaded region denotes the estimate of the error of linear regression obtained using Monte-Carlo simulations.

An excellent correlation (r = 0.95) was observed between $\Delta G^\circ_{melt}(i) = \Delta G^\circ_{melt,m6A-U} - \Delta G^\circ_{melt,m62A-U}$ and $\Delta G^\circ_{conf}(i)$ for N^6-methylamino isomerization in m^{6}A measured using NMR RD (Fig. 6*C*) (73). The correlation could be fit with an RMSE of 0.1 kcal/mol and c(i) ~0.7 kcal/mol (Fig. 6*C*) (*SI Appendix*, *Discussion S2*). Thus, the energetic destabilization due to incorporation of m$^6{}_2$A is correlated to the energetics of methyl rotation in m^{6}A. These results show that delta-melt can be applied to measure thermodynamic preferences involving isomerization of small chemical groups and broaden its applicability to RNA in addition to DNA.

Using Delta-Melt to Measure the Cooperativity When Forming Tandem Hoogsteen bps in Duplex DNA. Having benchmarked delta-melt and tested its underlying assumptions, we applied the approach to gain new insights into the biophysical properties of duplex DNA, which might not be feasible to obtain using more conventional approaches. In this first application, we used delta-melt to examine the cooperativity of forming two adjacent Hoogsteen bps. Tandem Hoogsteen bps are frequently observed in crystal structures of DNA in complex with proteins and drugs and in the context of DNA lesions that bias the bp conformation to Hoogsteen (reviewed in ref. 76). Furthermore, crystal structures of certain A-T repeat sequences form helices in which all bps are Hoogsteen (77), again pointing to Hoogsteen bps forming cooperatively. Thus, the thermodynamic preference to form a Hoogsteen bp could be modified by the presence of a nearby Hoogsteen bp, in turn impacting conformational penalties integral to recognition of proteins and other molecules. Such cooperativity could extend to other states such as the base open conformation.

Measuring conformational cooperativity is especially challenging as it requires accurately determining the abundance of exceptionally low-populated conformational states in which two or more bp positions are simultaneously in a minor conformation. The population of tandem Hoogsteen bps is expected to be on the order of the product of the probabilities for forming each Hoogsteen bp (~0.01%). Moreover, such conformations must be distinguished from states in which only one bp is Hoogsteen. Owing to these unique challenges, cooperativity has not been measured to date for any bp rearrangement in nucleic acids.

In theory, delta-melt can be used to measure conformational cooperativity for any arbitrary number of conformational transitions by performing experiments employing the appropriate number of modifications in a duplex. We tested this approach and measured the cooperativity of forming tandem Hoogsteen bps using the validated m^{1}A$^+$ and m^{1}G substitutions to induce Hoogsteen bps at specific positions. In what follows, it will be assumed that tandem N^1-methylated bps are good energetic mimics of their unmodified tandem Hoogsteen counterparts and that the offset c(i) for tandem Hoogsteen bps is small, similar to isolated Hoogsteen bps (Fig. 3*C* and *SI Appendix*, Fig. S7*C*) (*SI Appendix*, *Discussion S2*) or is equal to the sum of the offsets for forming the individual Hoogsteen bps (*SI Appendix*, *Appendix S6* and *Appendix S7*).

If tandem Hoogsteen bps do form cooperatively in duplex DNA, the energetic cost to form two Hoogsteen bps simultaneously ($\Delta G^\circ_{conf}(1+2)$) should be smaller than the sum of the energetic cost to form each Hoogsteen bp at each position ($\Delta G^\circ_{conf}(1) + \Delta G^\circ_{conf}(2)$), i.e., $\Delta G^\circ_{conf}(1+2) < \Delta G^\circ_{conf}(1) + \Delta G^\circ_{conf}(2)$. Conversely, $\Delta G^\circ_{conf}(1+2) = \Delta G^\circ_{conf}(1) + \Delta G^\circ_{conf}(2)$ for non-cooperativity and $\Delta G^\circ_{conf}(1+2) > \Delta G^\circ_{conf}(1) + \Delta G^\circ_{conf}(2)$ for the case of negative cooperativity.

Thus, subject to the assumptions outlined above, positive cooperativity predicts that the sum of energetic destabilization of single N^1-methylations ($\Delta G^\circ_{melt}(1) + \Delta G^\circ_{melt}(2)$) is greater than the energetic destabilization due to tandem N^1-methylation ($\Delta G^\circ_{melt}(1+2)$) with the difference between the two quantities ($\Delta G^\circ_{Coop,HG}$) providing a measure of the extent of cooperativity (Fig. 7*A* and *SI Appendix*, Fig. S10),

$$\Delta G^\circ_{melt}(1+2) < \Delta G^\circ_{melt}(1) + \Delta G^\circ_{melt}(2)$$

$$\Delta G^\circ_{Coop,HG} = \Delta G^\circ_{melt}(1) + \Delta G^\circ_{melt}(2) - \Delta G^\circ_{melt}(1+2).$$

We used delta-melt to assess the cooperativity of forming tandem T-A/G-C, C-G/G-C, T-A/T-A, and A-T/T-A Hoogsteen bps corresponding to four different bp steps (5′-TG-3′/5′-CG-3′/5′-TT-3′/5′-AT-3′) embedded within two different duplexes under different buffer conditions (pH 5.4 to 6.8, 25 °C, 150 mM NaCl) (*SI Appendix*, Fig. S11). We used the same DNA sequences employed in the experiments used to measure thermodynamic preferences for forming single Hoogsteen bps (Fig. 3*C* and *SI Appendix*, Fig. S7*C*). Melting experiments were performed on tandem N^1- methylated DNA duplexes in which both bps are expected to be Hoogsteen and their singly methylated counterparts, in which only one bp should be Hoogsteen (Fig. 7*B* and *SI Appendix*, Fig. S11).

We first performed NMR experiments on the tandem methylated duplexes and their singly methylated counterparts to verify the presence of stably formed Hoogsteen bps. As expected, the singly N^1-methylated duplexes showed NMR signatures of Hoogsteen bps (7), including upfield shifted T-H3 and downfield shifted m^{1}A$^+$-H6 protons for A-T Hoogsteen and the downfield shifted C-H4 proton for m^{1}G-C$^+$ Hoogsteen bps, respectively (*SI Appendix*, Fig. S11). The tandem N^1-methylated

duplexes also exhibited the expected downfield shifted purine-C8 resonances corresponding to the N^1-methylated *syn* purine bases (7, 64) and downfield shifted m^1A^+-H6 resonances corresponding to the m^1A^+-T Hoogsteen bps. However, the T-H3 signal was not observable and was likely broadened out of detection due to rapid solvent exchange kinetics.

We then measured the cooperativity of forming Hoogsteen bps using delta-melt experiments on the doubly and singly methylated duplexes (*SI Appendix*, Fig. S1 and Table S1). Indeed, across all four bp steps and different buffer conditions examined, tandem Hoogsteen bps formed cooperatively with sizeable $\Delta G^\circ_{Coop,HG} = 1.0$ to 2.5 kcal/mol (Fig. 7*B*).

If tandem Hoogsteen bps do indeed form cooperatively in DNA as implied by delta-melt, and robustly in a manner independent of the methyl modification, the thermodynamic preferences to form a Hoogsteen bp at an unmodified bp should also increase when the adjacent site is a Hoogsteen bp stabilized by an N^1-methylated purine modification. We tested this prediction using NMR RD experiments to directly measure the thermodynamic preferences to form G10-C15$^+$ and A16-T9 Hoogsteen bps in A_6-DNA when adjacent to pre-formed m^1A16^+-T9 or m^1G10-C15$^+$ Hoogsteen bps, respectively (*SI Appendix*, Fig. S12 *A–D*). Indeed, the thermodynamic preference to form Hoogsteen increased by ~0.2 to 1.0 kcal/mol when the neighboring bp was a modified preformed Hoogsteen. Similar increases were also observed recently when using T-T or C-T mismatches to mimic the Hoogsteen bp (19).

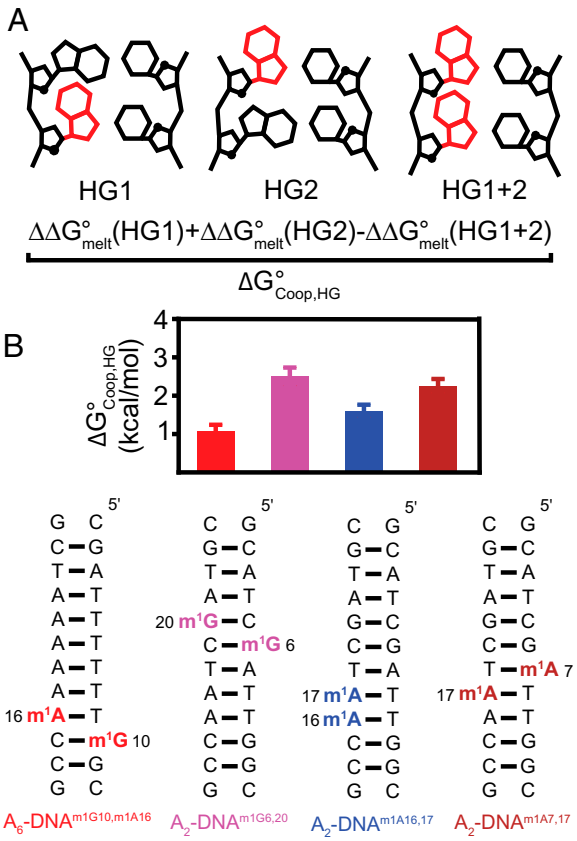

Fig. 7. Measuring cooperativity of Hoogsteen bp formation in DNA. (*A*) Estimating the cooperativity of Hoogsteen bp formation ($\Delta G^\circ_{Coop,HG}$) using delta-melt (*SI Appendix*, *Appendix S7*). Red bases denote modified N^1-methylated nucleotides (m^1A^+ and m^1G) that bias the bp conformation toward Hoogsteen. (*B*) The value of $\Delta G^\circ_{Coop,HG}$ over four different bp steps (*SI Appendix*, *Appendix S7* and Fig. S11). Bp steps are color-coded according to the duplexes shown below. Errors in $\Delta G^\circ_{Coop,HG}$ were determined by propagating the uncertainties from the UV melts as described in *SI Appendix*, *Appendix S7*. Buffer conditions were 15 mM sodium phosphate, 150 mM sodium chloride, and 0.1 mM ethylenediaminetetraacetic acid and pH 5.4 or 6.8 for cooperativity involving G-C and only A-T bps, respectively. All $\Delta G^\circ_{Coop,HG}$ values are reported at 25 °C.

To further test these delta-melt predictions of tandem Hoogsteen cooperativity in an unmodified DNA duplex, we took advantage of the anti-tumor antibiotic echinomycin (78) (*SI Appendix*, Fig. S12 *E–I*). Echinomycin has previously been shown to bind duplex DNA containing two CG steps separated by a CA step, inducing tandem G-C$^+$/A-T Hoogsteen bps at the CA step at low pH 5.3 (79). If the tandem Hoogsteen bps at the CA step form cooperatively, reducing the preference to form a Hoogsteen bp at one position should also reduce the preference to form the Hoogsteen bp at the second position, resulting in the loss of NMR signatures unique to tandem Hoogsteen bps. Indeed, reducing the thermodynamic preference to form G-C$^+$ Hoogsteen either by increasing the pH to 6.9 or by substituting guanine with 7-deaza guanine at pH 5.3 (80) resulted in the loss of NMR signals characteristic of tandem Hoogsteen bps (*SI Appendix*, Fig. S12 *E–I*). Furthermore, the loss of tandem Hoogsteen bp signatures at pH 6.9 could be rescued by using m^1G to stabilize the G-C$^+$ Hoogsteen bp (*SI Appendix*, Fig. S10 *E–I*). The above results suggest that Hoogsteen bps form cooperatively in echinomycin bound DNA, which may explain why multiple intercalating drugs such as echinomycin bind cooperatively to DNA (81).

Taken together, the above results strongly suggest that tandem Hoogsteen bps can form cooperatively in duplex DNA. Future experiments could independently verify tandem Hoogsteen cooperativity, for example by examining the impact of single or double Hoogsteen mimicking substitutions on DNA–protein and DNA–drug binding affinities (4), or through the development of new approaches that can directly detect tandem Hoogsteen bps in unmodified DNA.

Using Delta-Melt to Measure Sequence-Specific Signatures of Hoogsteen and Base Open States in Duplex DNA. Many biochemical processes involving DNA including protein recognition (12) and damage induction (82) and repair (83), as well as mutations including those related to cancer (84), are all strongly dependent on DNA sequence. When biochemical processes include steps that act on minor non-native conformational states of nucleic acids, such as the Hoogsteen and base open state, the sequence-specific thermodynamic preferences to form such conformations can shape the sequence specificities of the biochemical processes as well. However, the "conformational signatures" describing the sequence-specific preferences to form alternative conformational states have not been measured, primarily because of the difficulties in performing such a large number of NMR measurements.

Although the thermodynamic preferences to form minor non-native conformations such as Hoogsteen (8) and base open states (71) have been measured for various bp steps, these prior observations were confounded by additional variation in the global sequence context and position along the helix. We therefore used delta-melt to overcome these barriers and measured conformational signatures for A-T and G-C$^+$ Hoogsteen bps, and the A-T base open state for all 16 trinucleotide sequence contexts, at a given positional context. These measurements were performed on a highly stable scaffold sequence (scaf2, Fig. 8*A*), enabling robust melting measurements across all 16 sequence contexts, including when destabilizing modifications (m^1A^+, m^1G, and m^3T) are present.

In measuring the conformational signatures using delta-melt, our analysis assumed that the offsets c(i) are independent of sequence context. Although this is supported by our benchmarks (Figs. 3–6 and *SI Appendix* Fig. S7), without additional delta-

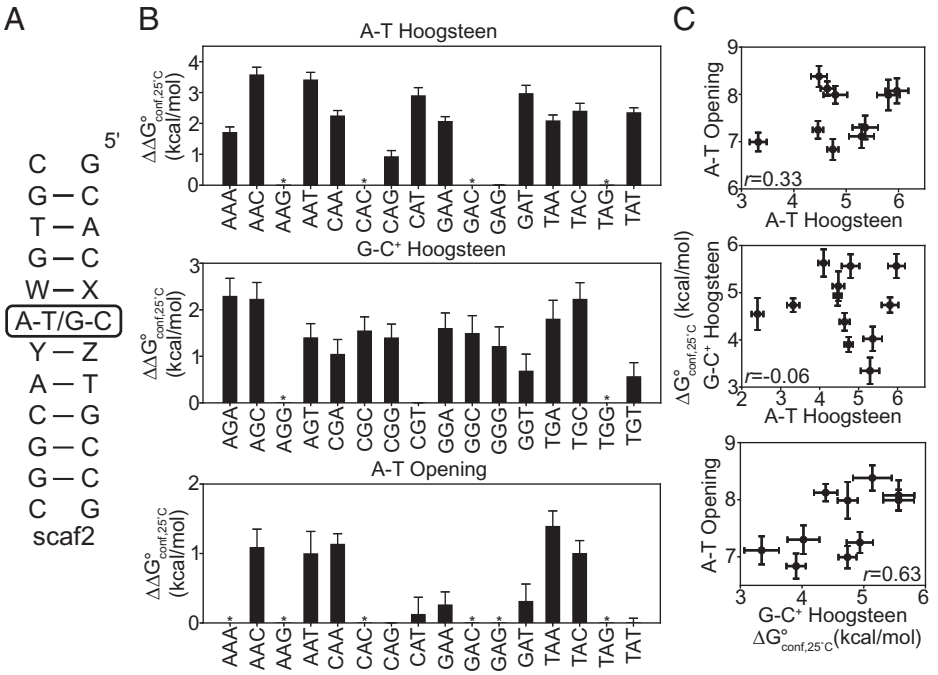

Fig. 8. Measuring sequence-dependent thermodynamic preferences in DNA using delta-melt. (*A*) DNA duplex used to measure sequence-dependent thermodynamic preferences using delta-melt. W-X and Y-Z represent either of the four Watson–Crick bps (A-T/T-A/C-G/G-C). Thermodynamic preferences to form base open or Hoogsteen conformations were measured using delta-melt for the central bp indicated using a box. (*B*) Relative thermodynamic preferences ($\Delta G^{\circ}_{conf,25\,C}$) to form A-T and G-C$^+$ Hoogsteen bps and the A-T base open states measured using delta-melt at 25 °C (*SI Appendix*, Figs. S1 and S2 and Tables S1 and S2). Sequence contexts are given in the 5′ to 3′ direction. Thermodynamic preferences were referenced to the sequence context with the most favorable preference (5′-GAG-3′, 5′-CGT-3′, and 5′-CAG-3′ for A-T Hoogsteen, G-C$^+$ Hoogsteen, and A-T base open states, respectively). Sequences for which reliable thermodynamic preferences could not be obtained due to non-two-state melting behavior are denoted using * (*Materials and Methods*). Buffer conditions for the measurements are described in *Materials and Methods*. (*C*) Correlation plots between the thermodynamic preferences for formation of A-T and G-C$^+$ Hoogsteen bps and open A-T bps at 25 °C. (*Inset*) The Pearson correlation coefficient (*r*). Note that for correlating preferences for A-T and G-C$^+$ Hoogsteen bps, the purine bases were aligned with each other; for example, the sequence context 5′-AAA-3′ for A-T Hoogsteen was correlated with 5′-AGA-3′ for G-C$^+$ Hoogsteen. Error bars in *B* and *C* were obtained by propagating the error from the UV melts and c(i), as described in *Materials and Methods*.

melt and NMR calibration experiments, we cannot rule out different c(i) values are applicable for different sequence contexts. In addition, for some sequence contexts, the melting curves deviated from a simple two-state behavior, and the data could not be used to extract reliable thermodynamic preferences (* symbols in Fig. 8*B*, *SI Appendix*, Fig. S3, and *Materials and Methods*). Additional studies are needed to understand the origins of the non-two-state behavior of the melting curves and how to extract robust thermodynamic parameters from them.

Delta-melt revealed large sequence-specific variations in the thermodynamic preferences to form A-T (~1.7 to 3.5 kcal/mol) and G-C$^+$ (0.5 to 2.2 kcal/mol) Hoogsteen bps, as well as the A-T base open state (0.3 to 1.4 kcal/mol), which are on the order of 2 to 3 kcal/mol, over the sequence contexts which could be measured reliably using delta-melt (Fig. 8*B* and *SI Appendix*, Figs. S1, S2, S13, and S14 and Tables S1 and S2). This corresponds to ~10- to 100-fold sequence-specific variations in the population of the minor conformation in the dynamic ensemble purely due to changes in the sequence of the immediate neighbor.

To verify the conformational preferences obtained using delta-melt, we used NMR RD to independently measure the energetic cost to form G-C$^+$ Hoogsteen bps in the 5′-TGC-3′ and 5′-CGT-3′ sequence contexts (Fig. 9*A*), predicted by delta-melt to have low and high propensities (Fig. 8*B*) to form G-C$^+$ Hoogsteen bps, respectively. Measurements were performed under two different buffer conditions and three different temperatures to test the robustness of the predictions. One of the contexts, 5′-TGC-3′ (data point 6, Fig. 9*B*), with low propensity is predicted by delta-melt to have little to no observable NMR RD even under low pH conditions (pH 5.4, 150 mM NaCl, 15 °C). Such behavior has never been observed previously and therefore represents a strong prediction from the delta-melt approach. On the other side, the 5′-CGT-3′ context (data point 1, Fig. 9*B*) is predicted to have a relatively large population (~0.7%) of the minor G-C$^+$ Hoogsteen bp at low pH (pH 5.4, 25mM NaCl, 30 °C). Prospective comparison of six thermodynamic preferences measured using delta-melt (Fig. 8*B*) and NMR RD (Fig. 9*B*) reveals an average accuracy of ~0.5 kcal/mol, which is small relative to the range of 0.5 to 2.0 kcal/mol measured across different sequence contexts. Delta-melt correctly predicted the lack of observable RD for the 5′-TGC-3′ sequence (data point 6, Fig. 9*B*) and enabled measurements of thermodynamic preferences at this position when it was not feasible using NMR RD. Delta-melt also correctly predicted the 5′-CGT-3′ context having the highest propensity to form G-C$^+$ Hoogsteen (data point 1, Fig. 9*B*). Nevertheless, we also observed a relatively large deviation of ~0.8 kcal/mol between the delta-melt and NMR-derived thermodynamic preferences for G6-C19$^+$ Hoogsteen bp formation in scaf2_TGC^{GC} at pH 5.4 high salt (150mM NaCl) and 30 °C (data point 4, Fig. 9*B*). This sequence context may have a unique c(i) offset (*SI Appendix*, Fig. S7*C*), as only one out of the seven measurements used for establishing calibration curves contains the same 5′-TGC-3′ trinucleotide step (*SI Appendix*, Fig. S7 *B* and *C*). These results support the utility of delta-melt to determine sequence-specific conformational signatures and to identify potential conformational hotspots.

Given the robustness of the delta-melt–derived thermodynamic preferences (Fig. 9*B*), we then proceeded to examine their sequence dependencies. The largest sequence dependence was observed for A-T Hoogsteen ($\Delta G^{\circ}_{conf}(i)$ ~ 1.7 to 3.5 kcal/mol) followed by G-C$^+$ Hoogsteen ($\Delta G^{\circ}_{conf}(i)$ ~ 0.5 to 2.2 kcal/mol) and with the open A-T state showing the weakest sequence-dependence ($\Delta G^{\circ}_{conf}(i)$ ~ 0.3 to 1.4 kcal/mol), possibly because of its relatively increased flexibility. Interestingly, the delta-melt–derived conformational signatures were very different (r ~ 0 to 0.63) for these three conformational states (Fig. 8*C*). Therefore, those trinucleotide sequence contexts showing a preference for forming one state may not also favor the formation of another state. For example, the 5′-GAT-3′/5′-GGT-3′ trinucleotide context is the third most favorable context for forming the A-T base open state and G-C$^+$ Hoogsteen bps but is the third least favorable context for A-T Hoogsteen bps. If different processes act on different conformational states of the DNA, these conformational signatures could translate into different sequence specificities for the distinct DNA biochemical reactions.

The delta-melt data also reveal that the 5′ and 3′ neighboring bps of the purine base of a given bp influence thermodynamic

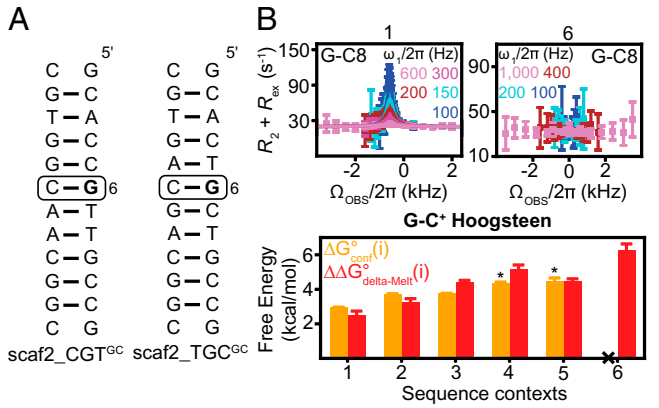

Fig. 9. Testing the accuracy of delta-melt–derived thermodynamic preferences to form G-C$^+$ Hoogsteen bps using NMR. (*A*) Duplexes used to test delta-melt derived thermodynamic preferences. $^{13}C/^{15}N$ isotopically labeled residues are in bold. All remaining residues are unlabeled. (*B*, *Bottom*) Comparison of measured thermodynamic preferences to form G-C$^+$ Hoogsteen bps using NMR (orange, *SI Appendix*, Table S5) and values measured using delta-melt (red, *SI Appendix*, Table S1). Numbers 1 to 6 correspond to combinations of sequences and buffer/temperature conditions for the measurements and are given in *SI Appendix*, Table S9. Lower-confidence data points (*Materials and Methods*) obtained from weak $R_{1\rho}$ RD profiles are denoted using *. Data point 6 with a flat $R_{1\rho}$ RD profile is denoted using "x," indicating that the minor G-C$^+$ Hoogsteen bp in this case falls outside the NMR RD detection limit. Error bars for free energies were obtained by propagating the uncertainty from UV melts (*Materials and Methods*). (*B*, *Top*) Representative examples of $R_{1\rho}$ RD profiles showing the measured effective transverse relaxation rate ($R_2 + R_{ex}$) as a function of offset ($\Omega_{OBS}/2\pi$) and color-coded spin-lock power ($\omega_1/2\pi$) (*Materials and Methods*). Errors bars were obtained by propagating the uncertainty in $R_{1\rho}$ as described in *Materials and Methods*.

preferences to form the three non-native conformational states to different extents. The 3′ neighboring bp dominates the preferences for A-T and G-C$^+$ Hoogsteen bps, whereas for the A-T base open state both neighbors contribute equally (*SI Appendix*, Fig. S15). In addition, substantial variations (~1.5 to 2.0 kcal/mol) in the sequence-specific thermodynamic preferences due to the third bp are seen for a given dinucleotide step, indicating that conformational signatures must at minimum be described by a triplet code (*SI Appendix*, Fig. S15).

Another advantage of delta-melt is that it provides unique insights into the origins of these different sequence-specific conformational signatures. For example, the sequence-specific variations to form A-T Hoogsteen bps appear to be dominated by the sequence-specific variations in the melting stability of the Watson–Crick A-T bp (*SI Appendix*, Fig. S15). Therefore, those sequences that are least stable in the unmodified DNA have the highest propensities to form A-T Hoogsteen bps. This is in line with a prior study showing that A-T Hoogsteen bps have high propensities to form in CA and TA steps, which are the least stable in the Watson–Crick form (8).

Interestingly, in contrast, the sequence-specific variations to form G-C$^+$ Hoogsteen bps appears to be dominated by the sequence-specific variations in the melting stability of the m^{1}G-C$^+$ Hoogsteen bp (*SI Appendix*, Fig. S15 and *Discussion S3*). These variations could arise due to large sequence-specific variations in the pK$_a$ of the cytosine imino proton, possibly mediated by sequence-specific stacking. Finally, in the case of the base open state, the sequence-specific preferences are determined by both the stability of the A-T Watson–Crick and base open (or specifically A-m^{3}T) states (*SI Appendix*, Fig. S15).

Discussion

Melting experiments have been used with great success to understand factors contributing to the stability of the major folded conformations of nucleic acids relative to the unfolded state, providing the basis for the nearest-neighbor rules, which are widely used to predict nucleic acid secondary structure and other properties of interest (44, 47). These prior studies took advantage of the relatively high throughput and low sample requirements of melting experiments, which permit broad explorations of how folding thermodynamics varies across different sequence and structural motifs and experimental conditions. In this study, we showed that this well-established approach (43, 44, 47–49) can be extended to the thermodynamic stabilities of minor non-native conformational states in the ensemble, whose contribution to melting experiments are typically negligible (53).

The key to delta-melt is to use chemical modifications to render minor non-native conformational states in the ensemble the major state. This then makes it possible to access information regarding thermodynamic stabilities of minor conformations. It is assumed that the modification biases the ensemble toward the major state so that its population exceeds 90% robustly across the different sequence contexts and experimental conditions examined. We have validated this assumption using NMR for all of the modifications presented here under a variety of sequence and experimental conditions (7, 61) (*SI Appendix*, Figs. S8, S11, and S14). However, it may be necessary to test this assumption particularly for cases which give rise to outlier behavior. Thus, in practice, we recommend independent validation of the structure of the modified molecules using biophysical techniques such as NMR spectroscopy or X-ray crystallography for use with delta-melt. Furthermore, a fundamental assumption in delta-melt is that the relative energetic cost for modifying the minor conformation and single strands does not vary with sequence contexts and conditions. Under this assumption, calibration with independent measurements of conformational preferences over different sequence contexts and conditions as done here using NMR provides a basis for exploring the energetics of other conformations in the ensemble using melting experiments. While our study supports this assumption for the conformational states studied here, it needs to be further verified, especially for different sequence contexts through additional measurements of sequence-specific thermodynamic preferences using NMR. This might reveal subgroups of sequence contexts that should be calibrated independently, thus improving the accuracy of the approach.

Prior studies established the utility of modifications as structural mimics for a variety of minor non-native conformations in nucleic acids (reviewed in ref. 40). Our results further establish chemical modifications as energetic mimics of minor conformational states. Thus, one can measure thermodynamic preferences to form non-native states in a facile manner using melting experiments, in contrast to NMR RD based approaches, which are comparatively more costly and sample intensive. In addition, our study introduced m^{3}T as a mimic of the A-T base open state. Solving high-resolution structures and dynamic ensembles (63) for m^{3}T modified DNA duplexes could shed light on the structure of this elusive state, in a manner analogous to how m^{1}A$^+$-T provided structural insights into transient Hoogsteen bps (61).

Delta-melt made a number of unique applications possible. It allowed us to gain rare insights into the cooperativity of forming non-native bp conformations in nucleic acids, revealing that tandem Hoogsteen bps form cooperatively in duplex DNA. This implies that N^1-methylation and possibly other forms of DNA damage can increase the Hoogsteen preferences at neighboring sites, and this could in turn be exploited by damage-repair enzymes such as glycosylases during recognition and repair. Delta-melt could be used in the future to more broadly examine how Hoogsteen preferences are

altered with other perturbations such as mismatches, nicks, and other lesions (19). Delta-melt can also potentially be used to measure cooperativity involving different types of bp conformational changes such Hoogsteen and base open states as well as to assess whether cooperativity can arise between bps that are not immediately adjacent to one another (85).

Delta-melt enabled us to measure the preferences to form non-native conformations in DNA as a function of the trinucleotide sequence context, revealing ∼10- to 100-fold differences in the probabilities of forming A-T and G-C^+ Hoogsteen and A-T base open states due to changes in sequence context alone. We observe rich behavior in the sequence-dependent preferences to form these three conformational states. Not only did the scale of the sequence dependencies vary across the three states, being weakest for the base open conformation, we also observed differences in the extent to which these sequence preferences are driven by the stability of the major versus minor state. These results strongly suggest that hidden within the DNA double helix is a rich landscape of sequence-specific minor conformations. Future studies should verify these predictions from delta-melt and examine their biological implications.

In the future, delta-melt could be used to study how changes in DNA sequence beyond the trinucleotide context affect conformational preferences and examine how the conformational preferences are shaped by epigenetic modifications such as 5-methylcytosine. The sequence-dependent thermodynamic propensities obtained from delta-melt can also be cross-referenced with sequence-dependent biological processes, including over 60 mutational signatures linked to cancer (86), to potentially identify minor conformations that drive them (*SI Appendix*, Fig. S18 and *Discussion S4*). For example, we observe a strong correlation between the sequence-dependent probabilities for T > A mutations in mutational signature 17A and the thermodynamic propensities for A-T Hoogsteen bp formation obtained using delta-melt, suggesting potential biological roles for Hoogsteen bps that require future investigation (*SI Appendix*, *Discussion S4*).

It should be possible to apply delta-melt to other minor conformational states not examined here. For example, O^6-methylguanine, a naturally occurring mutagen (87), could be used to measure thermodynamic preferences to form Watson–Crick like G•T/U mismatches in DNA, RNA, and RNA–DNA hybrids, which have been proposed to evade fidelity checkpoints, contributing to misincorporation errors during DNA replication, translation, and transcription (20). Furthermore, N^3-methylcytosine could be used to measure thermodynamic preferences for opening G-C bps, and 8-oxo guanine and adenine, which are commonly occurring form of oxidative damage (88), could be used to measure the thermodynamic preferences to form alternative conformations of A-G mismatches such as A(*anti*)$^+$-G(*syn*) and A(*syn*)-G(*anti*). Delta-melt could also be used to study the impact of experimental conditions on forming non-native conformations involving secondary structural rearrangements in RNA (89). It may be possible to extend delta-melt to proteins, for which mutants stabilizing minor conformations have been reported (90, 91). The throughput of delta-melt could also be increased in the future using high-throughput melting experiments (92). Through these applications, it may be feasible to extend the nearest-neighbor model to predict the energetics of various alternative nucleic acid conformational states, an important step toward predicting the conformational ensembles of nucleic acids.

Materials and Methods

Sample Preparation.

Buffer preparation. Sodium phosphate based buffers used for the NMR and UV experiments were prepared as described in *SI Appendix, Appendix S1*.

Annealing and buffer exchange. Duplex DNA and RNA samples were prepared by mixing the two complementary strands (∼1 mM) in a 1:1 ratio in water, heating to 95 °C for 5–10 min, followed by slow annealing at room temperature. Hairpin DNA and RNA samples were prepared by diluting the samples in water to concentrations < 100 μM, heating to 95 °C for 5–10 min, followed by rapid annealing on ice. Additional procedures for sample preparation and buffer exchange are described in *SI Appendix, Appendix S2*.

Oligonucleotides. Unlabeled and labeled oligonucleotides were purchased or synthesized in-house as described in *SI Appendix, Appendix S3*.

NMR experiments. The imino ^{1}H exchange experiments were carried out on a Bruker Avance III 700 MHz spectrometer equipped with a HCN room temperature probe while the remaining NMR data were collected on Bruker Avance III 600 MHz or 700 MHz NMR spectrometers equipped with HCPN and HCN cryogenic probes, respectively. Procedures for resonance assignment and analysis of $R_{1\rho}$, CEST, and imino proton exchange data are given in *SI Appendix, Appendix S4*.

UV Melting.

Experiments and sample conditions. Optical melting experiments were carried out on a PerkinElmer Lambda 25 UV/VIS spectrometer with an RTP 6 Peltier Temperature Programmer and a PCB 1500 Water Peltier System, and a Cary-100 UV-Vis spectrophotometer. At least three measurements were carried out for each DNA and RNA duplex using a sample volume of 400 μL in a Teflon-stoppered 1 cm path length quartz cell. The absorbance at 260 nm was monitored while the temperature was varied at a ramp rate of 1 °C/min. Procedures for data analysis are given in *SI Appendix, Appendix S5*.

Comparison of thermodynamic preferences between NMR experiments and delta-melt. Thermodynamic preferences for conformational changes as measured by NMR were compared to melting energetics as measured by UV melting experiments, as outlined in Fig. 2*C* and described in detail in *SI Appendix, Appendix S6*.

Data Availability. All the raw data and scripts used for the analysis are available on GitHub at https://github.com/alhashimilab/delta-melt-data.

ACKNOWLEDGMENTS. We thank all members of the H.M.A.-H. laboratory for critical input and Prof. Terry Oas for discussions with regard to UV melting experiments. We also thank Dr. Richard Brennan (Duke University) for providing access to the UV spectrophotometer for the UV melts performed in this study and Dr. Christopher Kreutz (University of Innsbruck, Austria) for providing us site-specifically labeled RNA phosphoramidites. The research in this study was funded by NIH Grant R01GM089846, US National Institute for General Medical Sciences Grant 1R01GM132899, and a Mathers Foundation grant to H.M.A.-H.

Author affiliations: [a]Department of Biochemistry, Duke University School of Medicine, Durham, NC 27710; [b]Department of Chemistry, Duke University, Durham, NC 27710; and [c]Department of Biomedical Engineering, Duke University, Durham, NC 27708

Author contributions: A.R., H.S., Y.X., and H.M.A.-H. designed research; A.R., H.S., Y.X., B.L., H.A.A., J.D.B., H.Z., and I.J.K. performed research; A.R., H.S., and Y.X. contributed new reagents/analytic tools; A.R., H.S., Y.X., B.L., J.D.B., H.Z., and I.J.K. analyzed data; and A.R., H.S., Y.X., B.L., H.A.A., and H.M.A.-H. wrote the paper.

1. H. J. Dyson, P. E. Wright, Perspective: The essential role of NMR in the discovery and characterization of intrinsically disordered proteins. *J. Biomol. NMR* **73**, 651–659 (2019).
2. M. R. Jensen, R. W. Ruigrok, M. Blackledge, Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr. Opin. Struct. Biol.* **23**, 426–435 (2013).
3. F. A. Mulder, A. Mittermaier, B. Hon, F. W. Dahlquist, L. E. Kay, Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 932–935 (2001).
4. A. Afek *et al.*, DNA mismatches reveal conformational penalties in protein-DNA recognition. *Nature* **587**, 291–296 (2020).
5. L. R. Ganser, M. L. Kelly, D. Herschlag, H. M. Al-Hashimi, The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **20**, 474–489 (2019).
6. M. Fischer, R. G. Coleman, J. S. Fraser, B. K. Shoichet, Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. *Nat. Chem.* **6**, 575–583 (2014).
7. E. N. Nikolova *et al.*, Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498–502 (2011).
8. H. S. Alvey, F. L. Gottardo, E. N. Nikolova, H. M. Al-Hashimi, Widespread transient Hoogsteen base pairs in canonical duplex DNA with variable energetics. *Nat. Commun.* **5**, 4786 (2014).
9. M. Kitayner *et al.*, Structural basis of DNA recognition by p53 tetramers. *Mol. Cell* **22**, 741–753 (2006).
10. D. Golovenko *et al.*, New insights into the role of DNA shape on its recognition by p53 proteins. *Structure* **26**, 1237–1250.e6 (2018).

11. D. E. Draper, Themes in RNA-protein recognition. *J. Mol. Biol.* **293**, 255–270 (1999).
12. R. Rohs *et al.*, Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **79**, 233–269 (2010).
13. J. L. Leroy, D. Broseta, M. Guéron, Proton exchange and base-pair kinetics of poly(rA).poly(rU) and poly(rI).poly(rC). *J. Mol. Biol.* **184**, 165–178 (1985).
14. M. Guéron, J. L. Leroy, Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **261**, 383–413 (1995).
15. J. T. Stivers, Extrahelical damaged base recognition by DNA glycosylase enzymes. *Chemistry* **14**, 786–793 (2008).
16. B. J. Dow, S. S. Malik, A. C. Drohat, Defining the role of nucleotide flipping in enzyme specificity using ^{19}F NMR. *J. Am. Chem. Soc.* **141**, 4952–4962 (2019).
17. F. Colizzi *et al.*, Asymmetric base-pair opening drives helicase unwinding dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22471–22477 (2019).
18. H. Zhou *et al.*, New insights into Hoogsteen base pairs in DNA duplexes from a structure-based survey. *Nucleic Acids Res.* **43**, 3420–3433 (2015).
19. H. Shi *et al.*, Revealing A-T and G-C Hoogsteen base pairs in stressed protein-bound duplex DNA. *Nucleic Acids Res.***49**, 12540–12555 (2021).
20. I. J. Kimsey *et al.*, Dynamic basis for dG•dT misincorporation via tautomerization and ionization. *Nature* **554**, 195–201 (2018).
21. I. J. Kimsey, K. Petzold, B. Sathyamoorthy, Z. W. Stein, H. M. Al-Hashimi, Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature* **519**, 315–320 (2015).
22. E. Strebitzer *et al.*, 5-Oxyacetic acid modification destabilizes double helical stem structures and favors anionic Watson-crick like cmo^5 U-G base pairs. *Chemistry* **24**, 18903–18906 (2018).
23. W. A. Cantara, F. V. Murphy, 4th, H. Demirci, P. F. Agris, Expanded use of sense codons is regulated by modified cytidines in tRNA. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10964–10969 (2013).
24. P. F. Agris *et al.*, Celebrating wobble decoding: Half a century and still much is new. *RNA Biol.* **15**, 537–553 (2018).
25. S. K. Denny *et al.*, High-throughput investigation of diverse junction elements in RNA tertiary folding. *Cell* **174**, 377–390.e20 (2018).
26. J. D. Yesselman *et al.*, Sequence-dependent RNA helix conformational preferences predictably impact tertiary structure formation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16847–16855 (2019).
27. W. R. Becker *et al.*, Quantitative high-throughput tests of ubiquitous RNA secondary structure prediction algorithms via RNA/protein binding. *biorXiv* [Preprint] (2019). https://www.biorxiv.org/content/10.1101/571588v1 (Accessed 22 February 2022).
28. B. Houck-Loomis *et al.*, An equilibrium-dependent retroviral mRNA switch regulates translational recoding. *Nature* **480**, 561–564 (2011).
29. A. Reining *et al.*, Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature* **499**, 355–359 (2013).
30. J. T. Baisden, J. A. Boyer, B. Zhao, S. M. Hammond, Q. Zhang, Visualizing a protonated RNA state that modulates microRNA-21 maturation. *Nat. Chem. Biol.* **17**, 80–88 (2021).
31. S. Zaccara, R. J. Ries, S. R. Jaffrey, Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
32. N. Liu *et al.*, N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
33. A. G. Palmer III, Chemical exchange in biomacromolecules: Past, present, and future. *J. Magn. Reson.* **241**, 3–17 (2014).
34. G. Jeschke, DEER distance measurements on proteins. *Annu. Rev. Phys. Chem.* **63**, 419–446 (2012).
35. N. T. Hunt, 2D-IR spectroscopy: Ultrafast insights into biomolecule structure and function. *Chem. Soc. Rev.* **38**, 1837–1848 (2009).
36. B. Schuler, Single-molecule FRET of protein structure and dynamics–A primer. *J. Nanobiotechnology* **11** (suppl. 1), S2 (2013).
37. M. Bonomi, M. Vendruscolo, Determination of protein structural ensembles using cryo-electron microscopy. *Curr. Opin. Struct. Biol.* **56**, 37–45 (2019).
38. A. M. Mustoe, N. N. Lama, P. S. Irving, S. W. Olson, K. M. Weeks, RNA base-pairing complexity in living cells visualized by correlated chemical probing. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24574–24582 (2019).
39. D. M. Korzhnev, V. Y. Orekhov, L. E. Kay, R. Off-resonance, Off-resonance R(1rho) NMR studies of exchange dynamics in proteins with low spin-lock fields: An application to a Fyn SH3 domain. *J. Am. Chem. Soc.* **127**, 713–721 (2005).
40. A. Rangadurai, E. S. Szymaski, I. J. Kimsey, H. Shi, H. M. Al-Hashimi, Characterizing micro-to-millisecond chemical exchange in nucleic acids using off-resonance $R_{1\rho}$ relaxation dispersion. *Prog. Nucl. Magn. Reson. Spectrosc.* **112-113**, 55–102 (2019).
41. L. R. Ganser *et al.*, Probing RNA conformational equilibria within the functional cellular context. *Cell Rep.* **30**, 2472–2480.e4 (2020).
42. S. J. Schroeder, D. H. Turner, Optical melting measurements of nucleic acid thermodynamics. *Methods Enzymol.* **468**, 371–387 (2009).
43. J. D. Puglisi, I. Tinoco Jr., Absorbance melting curves of RNA. *Methods Enzymol.* **180**, 304–325 (1989).
44. J. SantaLucia Jr., D. H. Turner, Measuring the thermodynamics of RNA secondary structure formation. *Biopolymers* **44**, 309–319 (1997).
45. G. Bonnet, O. Krichevsky, A. Libchaber, Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8602–8606 (1998).
46. K. J. Breslauer, J. M. Sturtevant, I. Tinoco Jr., Calorimetric and spectroscopic investigation of the helix-to-coil transition of a ribo-oligonucleotide: rA7U7. *J. Mol. Biol.* **99**, 549–565 (1975).
47. P. N. Borer, B. Dengler, I. Tinoco Jr., O. C. Uhlenbeck, Stability of ribonucleic acid double-stranded helices. *J. Mol. Biol.* **86**, 843–853 (1974).
48. D. H. Turner, N. Sugimoto, R. Kierzek, S. D. Dreiker, Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.* **109**, 3783–3785 (1987).
49. J. L. SantaLucia, R. Kierzek, D. H. Turner, Functional group substitutions as probes of hydrogen bonding between GA mismatches in RNA internal loops. *J. Am. Chem. Soc.* **113**, 4313–4322 (1991).
50. D. H. Mathews, J. Sabina, M. Zuker, D. H. Turner, Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940 (1999).
51. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
52. C. E. Longfellow, R. Kierzek, D. H. Turner, Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry* **29**, 278–285 (1990).
53. H. Abou Assi *et al.*, 2′-O-Methylation can increase the abundance and lifetime of alternative RNA conformational states. *Nucleic Acids Res.* **48**, 12365–12379 (2020).
54. M. F. Berger *et al.*, Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.* **24**, 1429–1435 (2006).
55. D. J. Williams, K. B. Hall, Experimental and theoretical studies of the effects of deoxyribose substitutions on the stability of the UUCG tetraloop. *J. Mol. Biol.* **297**, 251–265 (2000).
56. Y. Liu *et al.*, Applications of PLOR in labeling large RNAs at specific sites. *Methods* **103**, 4–10 (2016).
57. P. Dumas, E. Ennifar, F. Disdier, P. Walter, "UV melting studies with RNA" in *Handbook of RNA Biochemistry: Second, Completely Revised and Enlarged Edition*, R. K. Hartmann, A. Bindereif, A. Schon, E. Westhof, Eds. (Wiley-VCH, Weinheim, 2014), pp. 445–480.
58. G. G. Hammes, Y. C. Chang, T. G. Oas, Conformational selection or induced fit: A flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13737–13741 (2009).
59. Y. Xu *et al.*, Hoogsteen base pairs increase the susceptibility of double-stranded DNA to cytotoxic damage. *J. Biol. Chem.* **295**, 15933–15947 (2020).
60. K. Hoogsteen, The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **12**, 822–823 (1959).
61. B. Sathyamoorthy *et al.*, 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.* **45**, 5586–5601 (2017).
62. H. Yang, Y. Zhan, D. Fenn, L. M. Chi, S. L. Lam, Effect of 1-methyladenine on double-helical DNA structures. *FEBS Lett.* **582**, 1629–1633 (2008).
63. H. Shi *et al.*, 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* **70**, 229–244 (2018).
64. H. Zhou *et al.*, m(1)A and m(1)G disrupt A-RNA structure through the intrinsic instability of Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **23**, 803–810 (2016).
65. P. Brooks, P. D. Lawley, The methylation of adenosine and adenylic acid. *J. Chem. Soc.*539–545 (1960).
66. M. Petersheim, D. H. Turner, Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* **22**, 256–263 (1983).
67. J. A. McDowell, D. H. Turner, Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAGGUCUC)2 by two-dimensional NMR and simulated annealing. *Biochemistry* **35**, 14077–14089 (1996).
68. F. H. Martin, O. C. Uhlenbeck, P. Doty, Self-complementary oligoribonucleotides: Adenylic acid-uridylic acid block copolymers. *J. Mol. Biol.* **57**, 201–215 (1971).
69. H. Shi *et al.*, NMR chemical exchange measurements reveal that N^6-methyladenosine slows RNA annealing. *J. Am. Chem. Soc.* **141**, 19988–19993 (2019).
70. U. D. Priyakumar, A. D. MacKerell Jr., Computational approaches for investigating base flipping in oligonucleotides. *Chem. Rev.* **106**, 489–505 (2006).
71. C. Chen, I. M. Russu, Sequence-dependence of the energetics of opening of at basepairs in DNA. *Biophys. J.* **87**, 2545–2551 (2004).
72. P. Ji, X. Wang, N. Xie, Y. Li, N6-Methyladenosine in RNA and DNA: An epitranscriptomic and epigenetic player implicated in determination of stem cell fate. *Stem Cells Int.* **2018**, 3256524 (2018).
73. B. Liu *et al.*, A quantitative model predicts how m6A reshapes the kinetic landscape of nucleic acid hybridization and conformational transitions. *Nat Commun.* **12**, 5201 (2021).
74. P. Vallurupalli, A. Sekhar, T. Yuwen, L. E. Kay, Probing conformational dynamics in biomolecules via chemical exchange saturation transfer: A primer. *J. Biomol. NMR* **67**, 243–271 (2017).
75. D. Su *et al.*, Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nat. Protoc.* **9**, 828–841 (2014).
76. E. N. Nikolova *et al.*, A historical account of Hoogsteen base-pairs in duplex DNA. *Biopolymers* **99**, 955–968 (2013).
77. N. G. Abrescia, A. Thompson, T. Huynh-Dinh, J. A. Subirana, Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2806–2811 (2002).
78. X. L. Gao, D. J. Patel, 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* **27**, 1744–1751 (1988).
79. Y. Xu, J. McSally, I. Andricioaei, H. M. Al-Hashimi, Modulation of Hoogsteen dynamics on DNA recognition. *Nat. Commun.* **9**, 1473 (2018).
80. E. N. Nikolova, F. L. Gottardo, H. M. Al-Hashimi, Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J. Am. Chem. Soc.* **134**, 3667–3670 (2012).
81. C. Bailly, F. Hamy, M. J. Waring, Cooperativity in the binding of echinomycin to DNA fragments containing closely spaced CpG sites. *Biochemistry* **35**, 1150–1161 (1996).
82. B. D. Thrall, D. B. Mann, M. J. Smerdon, D. L. Springer, DNA polymerase, RNA polymerase and exonuclease activities on a DNA sequence modified by benzo[a]pyrene diolepoxide. *Carcinogenesis* **13**, 1529–1534 (1992).
83. A. M. Sanchez, D. E. Volk, D. G. Gorenstein, R. S. Lloyd, Initiation of repair of A/G mismatches is modulated by sequence context. *DNA Repair (Amst.)* **2**, 863–878 (2003).
84. L. B. Alexandrov *et al.*; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
85. S. Kim *et al.*, Probing allostery through DNA. *Science* **339**, 816–819 (2013).
86. S. Bamford *et al.*, The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br. J. Cancer* **91**, 355–358 (2004).
87. J. J. Warren, L. J. Forsberg, L. S. Beese, The structural basis for the mutagenicity of O(6)-methyl-guanine lesions. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19701–19706 (2006).
88. M. S. Cooke, M. D. Evans, M. Dizdaroglu, J. Lunec, Oxidative DNA damage: Mechanisms, mutation, and disease. *FASEB J.* **17**, 1195–1214 (2003).
89. E. A. Dethoff, K. Petzold, J. Chugh, A. Casiano-Negroni, H. M. Al-Hashimi, Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).
90. D. M. Korzhnev, T. L. Religa, W. Banachewicz, A. R. Fersht, L. E. Kay, A transient and low-populated protein-folding intermediate at atomic resolution. *Science* **329**, 1312–1316 (2010).
91. G. Bouvignies *et al.*, Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111–114 (2011).
92. J. H. Bae, J. Z. Fang, D. Y. Zhang, High-throughput methods for measuring DNA thermodynamics. *Nucleic Acids Res.* **48**, e89 (2020).
93. X. J. Lu, W. K. Olson, 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* **3**, 1213–1227 (2008).
94. J. R. Bothe, Z. W. Stein, H. M. Al-Hashimi, Evaluating the uncertainty in exchange parameters determined from off-resonance R1rho relaxation dispersion for systems in fast exchange. *J. Magn. Reson.* **244**, 18–29 (2014).