BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 3225–3236 3225
DOI: 10.1021/bi901920g

Driving Forces for Nucleic Acid pK_a Shifting in an $A^+ \cdot C$ Wobble: Effects of Helix Position, Temperature, and Ionic Strength†

Nathan A. Siegfried,‡ Bernie O'Hare, and Philip C. Bevilacqua*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802. ‡*Present address: Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599.*

Received November 8, 2009; Revised Manuscript Received March 5, 2010

ABSTRACT: Secondary structure plays critical roles in nucleic acid function. Mismatches in DNA can lead to mutation and disease, and some mismatches involve a protonated base. Among protonated mismatches, $A^+ \cdot C$ wobble pairs form near physiological pH and have relatively minor effects on helix geometry, making them especially important in biology. Herein, we investigate effects of helix position, temperature, and ionic strength on pK_a shifting in $A^+ \cdot C$ wobble pairs in DNA. We observe that pK_a shifting is favored by internal $A^+ \cdot C$ wobbles, which have low cooperativities of folding and make large contributions to stability, and disfavored by external $A^+ \cdot C$ wobbles, which have high folding cooperativities but make small contributions to stability. An inverse relationship between pK_a shifting and temperature is also found, which supports a model in which protonation is enthalpically favored overall and entropically correlated with cooperativity of folding. We also observe greater pK_a shifts as the ionic strength decreases, consistent with anticooperativity between proton binding and counterion-condensed monovalent cation. Under the most favorable temperature and ionic strength conditions tested, a pK_a of 8.0 is observed for the $A^+ \cdot C$ wobble pair, which represents an especially large shift (~4.5 pK_a units) from the unperturbed pK_a value of adenosine. This study suggests that protonated $A^+ \cdot C$ wobble pairs exist in DNA under biologically relevant conditions, where they can drive conformational changes and affect replication and transcription fidelity.

In RNA and DNA, the nucleobases typically exist as neutral species at biological pH. The pK_a values for the imino nitrogens are ~3.5 and 4.2 for single-stranded A and C, respectively, and ~9.3 for single-stranded G and U/T (*1*, *2*). Nevertheless, protonated bases play important roles in biology (*3*−*6*), with pK_a values shifted toward neutrality (*7*−*11*). Understanding the driving forces behind pK_a shifting would therefore provide insight into numerous biological phenomena and provide a means of predicting behavior in uncharacterized systems.

On one hand, ionization of imino functionalities interferes with Watson−Crick base pairing and is therefore typically disfavored, resulting in pK_a values shifted further from neutrality (*12*). On the other hand, ionization of functional groups promotes certain non-Watson−Crick base pairs, resulting in pK_a values shifted toward neutrality. For instance, protonation of C at N3 promotes $C \cdot C^+$ mismatch formation (*13*), while protonation of A at N1 promotes GA (GN7-AN1$^+$) pairing (*14*). In both instances, it is the gain of hydrogen bonding, stacking, and electrostatic interactions upon protonation that drives folded-state pK_a values toward or away from 7 (*15*, *16*).

Some bases require complex tertiary structures to protonate, and in these cases, electrostatic focusing often plays an important role (*4*, *16*, *17*). Examples of protonated tertiary bases include cytosines and adenines that play structural and catalytic roles in ribozymes and viral RNAs (*8*, *11*, *18*−*24*). In catalytic RNAs, bases with pK_a values of neutrality can optimally facilitate general acid−base catalysis, while bases with pK_a values of >7 are possible oxyanion holes. Such catalytic functions parallel roles of histidine and lysine/arginine in protein enzymes, effectively increasing the functional diversity of RNA.

Certain protonated bases require just a simple helical context to form, making them amenable to model studies. The most common protonated base pair of this type is the $A^+ \cdot C$ wobble, which is isosteric with a GT wobble and has a pK_a shifted toward 7 (Figure 1). Protonated $A^+ \cdot C$ wobbles are common in RNA, where they play structural and functional roles (*4*, *7*, *25*, *26*). For example, N1 adenine pK_a values shifted toward neutrality have been reported for $A^+ \cdot C$ wobbles in the lead-dependent RNA enzyme (*12*, *25*), the U6 loop of the splicesome (*9*), and tobacco ringspot virus (*7*). Similarly, a pK_a near 7 has been identified in or near the peptidyl transferase center in 23S rRNA (*27*) and likely arises from an $A^+ \cdot C$ wobble, although assignment to a particular nucleobase has been challenging (*28*).

Protonated $A^+ \cdot C$ wobbles also occur in DNA where they can be mutagenic and carcinogenic (*29*, *30*). For example, $A^+ \cdot C$ wobbles created during repair of oxidative DNA damage can be incorrectly converted to a GC base pair (rather than an AT pair), thereby "immortalizing" a mutation (*29*). In addition, the highly carcinogenic aflatoxin B_1-DNA adduct often causes an $A^+ \cdot C$ wobble to be incorporated upstream of the modification, an event that causes a GC-to-AT mutation (*30*).

In this study, effects of helix position, temperature, and ionic strength on the folded-state pK_a of an $A^+ \cdot C$ wobble are studied. Results are interpreted in terms of a thermodynamic framework that couples protonation and DNA folding. Thermodynamic driving forces revealed herein lead to pK_a values at and beyond physiological pH, with implications for biological function.

†Supported by National Science Foundation Grant MCB-0527102.
*To whom correspondence should be addressed. Phone: (814) 863-3812. Fax: (814) 865-2927. E-mail: pcb@chem.psu.edu.

Published on Web 03/25/2010

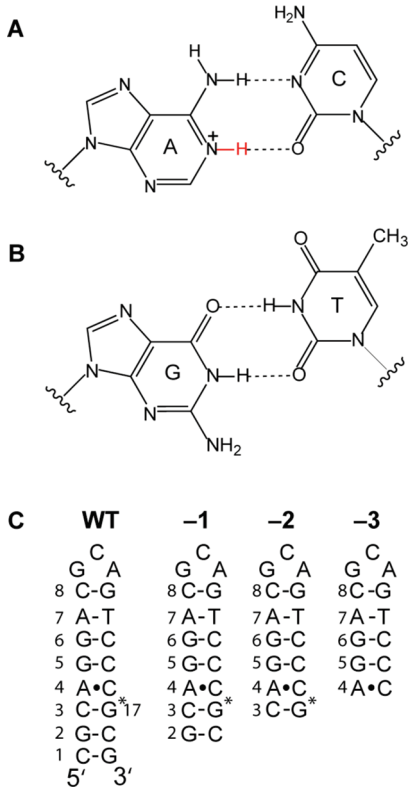

FIGURE 1: Wobble base pairs in dsDNA. (A) $A^+ \cdot C$ wobble base pair. Protonation of N1 of adenine (red) is required for formation. (B) $G \cdot T$ wobble base pair. The base pairs in panels A and B have a common wobble geometry in which both nucleosides are anti and have two hydrogen bonds and the pyrimidine 4-position is in the major groove. (C) Full-length (WT) DNA hairpin with the three constructs (−1, −2, and −3) resulting from terminal base pair deletions. The asterisk denotes the location of the position of the phosphorothioate label (G17), which is constant across these constructs. Numbering of base pairs is provided and maintained in deletion constructs.

MATERIALS AND METHODS

Preparation of DNA. The standard DNA hairpin used in this study has an $A^+ \cdot C$ wobble located in an internal helical context, a phosphorothioate at G17, and is termed WT[1,2] (Figure 1C). To modulate the position of the $A^+ \cdot C$ wobble in the helix while maintaining nearest neighbors, we deleted terminal base pairs in a stepwise manner up until the mismatch at the base of the helix, which resulted in three sequences: −1, −2, and −3 (Figure 1C). These four constructs provide a platform for studying the relationship between pK_a shifting and helical context for an $A^+ \cdot C$ wobble in double-stranded DNA (dsDNA).

To favor a single cationic counterion species to the phosphate backbone, synthetic DNA oligonucleotides (IDT DNA, Coralville, IA) were subjected to salt exchange followed by desalting using the following two-step procedure: (1) dialysis against 100 mM KCl for 6 h and (2) dialysis against sterile deionized water for 12−16 h (the dialysis method is described in the next paragraph). Measurements by atomic absorption confirmed that this length of time was sufficient for both steps to come to equilibrium (see the Supporting Information). After dialysis and prior to each experiment, additional K^+ was added from a concentrated stock solution to provide the desired final free K^+ concentration. Dialysis was found to be important because samples prepared in KCl without the first step of dialysis against 100 mM KCl exhibited smaller (~0.5 unit) pK_a shifts; moreover, analysis by MALDI-TOF mass spectrometry revealed that samples dialyzed without the first step had a mixture of Na^+ and K^+ counterions present, while those dialyzed with this two-step procedure had K^+ counterions only (data not shown).

All dialysis steps were conducted in an eight-well microdialysis system (Gibco-BRL Life Technologies, Rockville, MD) attached to a peristaltic pump set at a flow rate of ~15 mL/min; thus, in the course of a 6 h dialysis run, ~5.5 L of buffer flowed against the sample. The final DNA concentration for NMR studies was 0.6−1.7 mM in 10% ^2H_2O (added after dialysis). Because it would interfere with pH titrations, no buffer was present during dialysis. The pH of the DNA sample before and after each step of dialysis was found to be ~5.5. Cation counterion quantities were determined by atomic absorption.

Atomic Absorption. Analysis of cationic species was conducted on a Varian 220 FS atomic absorption spectrophotometer using a Varian potassium hollow cathode lamp to detect K^+ (at 766.5 nm) ions. Calibration of the instrument was performed prior to analysis using KCl standards prepared by serial dilution (see the Supporting Information). Lamp output was monitored to ensure stability and maximized signal following a minimal 30 min warm-up.

NMR Titrations. The oligonucleotide concentrations were 0.6−1.7 mM in 90% H_2O and 10% D_2O. To favor hairpin formation, oligonucleotides were renatured before each experiment by being incubated at 90 °C for 2 min, followed by being cooled to room temperature on the bench for 15 min. Standard conditions are defined herein as 100 mM KCl and 10 °C. Salt concentrations of greater than ~750 mM were not tested because of NMR experimental constraints, although a pK_a value at 1 M salt could be obtained by extrapolation.

Proton-decoupled ^{31}P NMR was used to monitor a phosphorothioate-labeled position (G17 herein), as described previously (*31*). pH titrations were conducted in the NMR tube, and the pH of the solution was measured using an ISFET micro pH probe (IQ Scientific Instruments) at the temperature of the titration. We previously reported that ion-dependent effects for this probe result in errors in measurement of pH for certain solutions (*32*). The effects of potassium ions on pH meter readings were therefore examined. A series of dilute HCl solutions at known concentrations between 10^{-2} and 10^{-4} M were prepared in 100 mM KCl, as was a series of dilute NaOH solutions at known concentrations between 10^{-2} and 10^{-3} M. Observed meter readings were compared to solution pH values, which were calculated using the relationship $pH = -\log a_{H^+} = -\log c_{H^+} - \log f_{H^+}$, with $-\log f_{H^+}$ determined from the Debye−Hückel equation, as previously discussed (*32*). No significant differences between meter readings and solution pH values were observed for KCl; as such, meter readings are used synonymously with pH herein. ^{31}P NMR titrations were conducted on a Bruker Avance III 500 MHz spectrometer, utilizing a 5 mm broadband observe probe, as well as a Bruker Avance 360 MHz spectrometer with a 5 mm quadranuclear probe. Titration data were sigmoidal and fit well to eq 1

$$\delta = \delta_P + \frac{\Delta\delta}{1 + 10^{n(pK_a - pH)}} \quad (1)$$

[1]Abbreviations: EDTA, ethylenediaminetetraacetic acid; T_M, temperature at which the folded and unfolded populations of an oligonucleotide are equal; WC, Watson−Crick; WT, wild-type; −1, −2, and −3, hairpins from which one, two, and three terminal base pairs, respectively, have been deleted, resulting in progressively shorter hairpins.

[2]This particular sequence has been previously studied using ^{31}P NMR and was termed DNA2 (*31*). We opt to label this sequence as "WT" herein because variants were designed on the basis of this core sequence. We used a similar WT nomenclature in another report (*36*).

where $\Delta\delta = \delta_U - \delta_P$ (δ_U and δ_P are defined as the chemical shifts of the unprotonated and protonated folded DNA states, respectively) and n is a Hill constant related to the number of proton binding sites. Hill constants were close to unity, consistent with uptake of a single proton binding, or binding to several independent sites.

All one-dimensional (1D) imino proton (^{1}H) spectra were recorded on a Bruker Avance-III 600 MHz spectrometer equipped with a CTI cryoprobe, operated at 274 K to slow chemical exchange. Water suppression was achieved for all ^{1}H-detected NMR experiments using excitation sculpting techniques with pulsed field gradients as implemented in standard Bruker pulse programs (*33*).

Temperature and Salt Dependence of NMR Data. (*i*) *Temperature-Dependent Data*. Titrations were performed at temperatures ranging from 10 to 42 °C, and pK_a values were plotted versus $1/T$. The standard van't Hoff equation can be rearranged to afford

$$pK_a = \frac{\Delta H^\circ}{2.303RT} - \frac{\Delta S^\circ}{2.303R} \quad (2)$$

where

$$\Delta G^\circ = 2.303RT \times pK_a \quad (3)$$

Plots of pK_a versus $1/T$ were fit to eq 2 using linear least-squares regression to obtain ΔH° and ΔS°. Error analysis is presented in the Supporting Information. In discussing proton association phenomena we reversed the algebraic signs, since pK_a is for proton dissociation.

(*ii*) *Salt-Dependent Data*. To aid in the interpretation of salt-dependent data, the following molecular model was used, wherein m K^+ ions replace a proton (Scheme 1).

$$FH^+ + mK^+ \rightleftharpoons F\cdot K^+_m + H^+ \quad \text{(Scheme 1)}$$

The overall equilibrium expression for this scheme, K_{tot}, is provided by eq 4.1

$$K_{tot} = \frac{[F\cdot K^+_m][H^+]}{[FH^+][K^+]^m} \quad (4.1)$$

while the acid dissociation constant, in this case K_a apparent ($K_{a,app}$), is provided by eq 4.2

$$K_{tot} = \frac{K_{a,app}}{[K^+]^m} \quad (4.2)$$

Taking the negative base 10 logarithm of eq 4.2 results in eq 5

$$pK_{a,app} = pK_{tot} - m\log[K^+] \quad (5)$$

where p$K_{a,app}$ equals pK_{tot} at 1 M KCl.

Thermal Denaturation Experiments. Thermodynamic parameters were obtained via UV-monitored thermal denaturation (melting) experiments, conducted in 100 mM KCl to maintain a constant and physiologically relevant ionic strength (*34*) and to match our standard experimental salt conditions for NMR. Melts were conducted at either pH 5.7 or 8.0, maintained in 25 mM MES or HEPES, respectively. These buffers have a shallow temperature dependence to their pK_a values, with $\Delta pK_a/°C$ values of −0.011 and −0.014 for MES and HEPES, respectively (*35*). Solution conditions near the transition midpoint have the greatest effect on ΔH° and ΔS°; buffers were therefore prepared such that they would be at the desired pH value near the T_M, similar to the method previously reported (*15*).

Melting experiments were conducted at 280 nm on a Gilford Response II spectrophotometer equipped with a temperature-controlled microcuvette assembly, and data were collected every 0.5 °C. To favor hairpin formation, oligonucleotides were renatured in the same manner as in the NMR experiments. For each experiment, temperature scans over the range of 5−95 °C were performed in both heating and cooling directions; overlapping spectra were obtained, consistent with the reversibility of the folding transition. Data were repeated three or more times, and representative melts are provided in the text.

Simulations of Protonation and pH Denaturation. In this study, two types of folding events are presented: global and local. Global folding is represented by transitions between unfolded (U) and folded (F) states and assesses stability either at pH values far removed from measured pK_a values or at high temperatures (Scheme 2).

$$U \rightleftharpoons F \text{ (or } FH^+) \quad \text{(Scheme 2)}$$

Local folding, on the other hand, is represented by the transition between the folded protonated (FH^+) state with the $A^+\cdot C$ wobble pair and the folded unprotonated (F) state and assesses stability at temperatures low enough that WT remains globally folded (Scheme 3).

$$FH^+ \rightleftharpoons F + H^+ \quad \text{(Scheme 3)}$$

Schemes 2 and 3 can be combined to provide a global view of all actions of proton binding on RNA folding (Scheme 4).

$$\{U^+\} \rightleftharpoons FH^+ \rightleftharpoons F + H^+ \rightleftharpoons \{U^-\} \quad \text{(Scheme 4)}$$

For Scheme 4, an ensemble of unfolded protonated (i.e., U, U^+, ^+U, $^+U^+$, etc.) and unfolded deprotonated (U, U^-, ^-U, $^-U^-$, etc.) states are represented at the left and right ends of the expression, and represent results of acidic and alkaline denaturation, respectively.

Experimental data of chemical shift (δ) versus pH were simulated according to equations derived from Scheme 4 using linked equilibria (see the Supporting Information for the derivation). Doing so produces a partition function Q, provided in eq 6

$$Q = 10^{n(pK_{aF} - pH)} + 1 + K_U(1+10^{pK_{aA} - pH})^{n^A}(1+10^{pK_{aC} - pH})^{n^C} + K_U(1+10^{pH - pK_{aG}})^{n^G}(1+10^{pH - pK_{aT}})^{n^T} - K_U \quad (6)$$

where n^A, n^C, n^G, and n^T are the number of A, C, G, and T residues, respectively, in the DNA oligonucleotide, K_U is [U]/[F], and n is the number of folded-state proton binding sites. The fractional population of a given state is that state's term in Q divided by Q itself. The observed spectroscopic signal, δ_{obs}, is then given by the population-weighted average of all spectroscopic signals (δ_{U^+}, δ_U, δ_{U^-}, δ_F, and δ_{FH^+}), affording eq 7

$$\delta_{obs} = [\delta_{FH^+} \times 10^{n(pK_{aF} - pH)} + \delta_F + \delta_{U^+}K_U(1+10^{pK_{aA} - pH})^{n^A}(1+10^{pK_{aC} - pH})^{n^C} + \delta_{U^-}K_U(1+10^{pH - pK_{aG}})^{n^G}(1+10^{pH - pK_{aT}})^{n^T} - \delta_U K_U]/Q \quad (7)$$

In practice, the final term in Q carries essentially no weight and δ_U is not known, so this term was not included in the simulation. Also, δ_{U^-} appears to be linearly dependent on pH and so was treated as a positively sloping baseline. The mathematical model uses

as inputs the folded-state pK_a determined in this study as well as the unfolded-state pK_a values determined in a previous study (*15*).

RESULTS

Overview of Results. The first three sections describe the effects of helix position, temperature, and ionic strength on folded-state $A^+\cdot C$ wobble pK_a values in dsDNA, which were obtained from ^{31}P NMR titrations using pH values in the near-neutral range of ~5−9. Helix position studies establish effects of mismatch location on pK_a shifting, while temperature and ionic strength studies ascertain effects of potentially biologically relevant conditions on pK_a shifting and allow determination of thermodynamic parameters. The fourth section describes pH titration over a wider range of pH values (~3.5−12) and relates pH to both local and global stability; in this section, protonation-coupled folding data are compared to a mathematical model. Lastly, the fifth section describes structural and thermodynamic properties of the two states present in the near-neutral pH range, using ^1H NMR and UV-monitored thermal denaturations.

Effect of Helix Position on pK_a Shifting. We previously reported folding cooperativity in RNA and DNA helices using double mutant cycles (*36*). Mutations were introduced into internal or external helical contexts, maintaining the same nearest neighbors, and folding cooperativity was assessed using UV melting. It was determined that for DNA, external sequence changes had significantly more cooperativity than internal changes (by a minimum of ~0.5 kcal/mol). Internal sequence mutations, however, provided greater destabilization than external changes (by ~1 kcal/mol in $\Delta G°_{37}$); this latter trend also held for RNA (*36*, *37*). These findings indicate that in simple helical systems the most cooperative regions do not correspond to the most stable regions. Given that cooperativity and base pairing strength are modulated by helix position, it seemed possible that the pK_a of an $A^+\cdot C$ wobble pair would also be affected by helix position.

In all constructs, a phosphorothioate was installed on G17, which is immediately downstream of the C of the $A^+\cdot C$ wobble (Figure 1C). This modification generates a mixture of diastereomers, which was intentionally left mixed to provide two spectrally isolated labels for ^{31}P NMR titrations. A previous study gave pK_a values for the two diastereomers that were in good agreement with each other and with pK_a values for phosphorothioates at other positions in the DNA, suggesting that these labels are nonperturbing (*31*).

Titration curves were collected for WT, −1, and −2 constructs, and the ^{31}P chemical shift of the phosphorus of G17 was plotted as a function of pH (Figure 2). To ensure that the equilibrium observed is between folded (F) and folded protonated (FH^+) states as shown in Scheme 3, titrations were performed at 10 °C, which is well below the melting temperature (T_m) of WT, −1, and −2 (see below). For both WT and −1, the data have a distinct sigmoidal shape that can be fit well to eq 1; however, the data for −2 fit only to a straight line, suggesting that this construct does not ionize appreciably over the pH range of 5.5−9.5.

The pK_a value is largest (and above biological pH) for WT, at 7.66 ± 0.07 (Table 1).[3] The pK_a is slightly smaller in −1, 7.13 ± 0.085, and is missing altogether in −2, wherein the $A^+\cdot C$ wobble is penultimate to the base of the helix.[4] Also, for both WT and −1, the Hill coefficient is unity within error, consistent with a single protonation event. Notably, the chemical shift of −2 at all pH values is similar to the chemical shift of a diastereomer of −1 at higher pH, 52.03 and 51.88 ppm, respectively, consistent with −2 being deprotonated between pH 5.5 and 9.5 (Figure 2). The similarity of the chemical shift of −2 to the high-pH chemical shift of −1 rather than that of WT is consistent with high pH inducing complete unfolding of the lower stem of −1 but not of WT, as described below. It should also be noted that the total change in parts per million observed for WT is smaller than that observed for −1 [$\Delta\delta_{avg}$ = 0.20 and 0.38 ppm, respectively (Table 1)], supporting a greater change in chemical environment of the phosphorothioate label in the −1 construct. In summary, the $A^+\cdot C$ mismatch requires at least two base pairs at its base to form, with more base pairing leading to greater pK_a shifting and less change in chemical shift. This observation is in keeping with greater stability correlating with lower cooperativity (*36*), which in these systems provides greater pK_a shifting (see Discussion).

Effect of Temperature on pK_a Shifting. Protonation and coupled DNA structure formation may result in observable

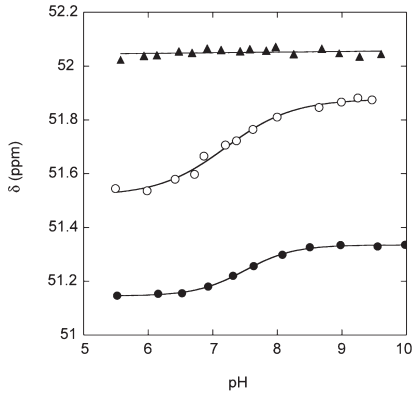

FIGURE 2: Context dependence of pK_a titrations. The ^{31}P chemical shift is plotted vs pH for WT (●), −1 (○), and −2 (▲) constructs. Traces for peak 1 are provided, where peak 1 is defined as the furthest downfield-shifted peak. pK_a values for peak 1 from fitting to eq 1 are 7.67 ± 0.05, 7.11 ± 0.05, and < 5, respectively. pK_a values for individual peaks and their averages are provided in Table 1. Data for −2 were fit to a straight line whose slope turned out to be ~0. Experiments were conducted under standard conditions (100 mM KCl and 10 °C).

Table 1: Context Dependence NMR Titration Data[a]

sequence	peak[b]	Δδ (ppm)	n[c]	pK_a[d]	$\Delta G°_{10}$ (kcal/mol)[e]
WT	1	0.28	1.32 ± 0.18	7.67 ± 0.05	−9.93
	2	0.12	1.31 ± 0.21	7.65 ± 0.05	−9.90
	av	0.20	1.32 ± 0.27	7.66 ± 0.07	−9.92
−1	1	0.39	0.91 ± 0.09	7.11 ± 0.05	−9.21
	2	0.36	0.83 ± 0.12	7.15 ± 0.07	−9.26
	av	0.38	0.87 ± 0.15	7.13 ± 0.085	−9.23
−2	1	—	—	< 5	—
	2	—	—	< 5	—
	av	—	—	< 5	—

[a]Experiments were conducted under standard conditions of 100 mM KCl at 10 °C. Values are calculated from eqs 1 and 3. [b]Peak 1 is defined as the furthest downfield-shifted peak. [c]Errors are from data fitting. [d]Errors are from data fitting; values are shown as error bars in plots. Averages are used in the text. [e]Values are for protonation and so were calculated according to the negative of eq 3.

[3]The average pK_a value of the two diastereomers is provided in the text.

[4]On the basis of this observation and the lack of pH dependence of melting and ^1H NMR spectra for the −3 construct (see below and the Supporting Information), ^{31}P NMR titrations were not attempted for −3.

Table 2: Temperature Dependence NMR Titration Data[a]

temp (°C)	peak	Δδ (ppm)	*n*	pK_a	ΔG° (kcal/mol)[b]
			WT		
10[c]	1	0.28	1.32 ± 0.18	7.67 ± 0.05	−9.93
	2	0.12	1.31 ± 0.21	7.65 ± 0.05	−9.90
	av	0.20	1.32 ± 0.27	7.66 ± 0.07	−9.92
15	1	0.31	0.98 ± 0.10	7.65 ± 0.04	−10.08
	2	0.16	1.04 ± 0.16	7.45 ± 0.08	−9.82
	av	0.24	1.01 ± 0.19	7.55 ± 0.09	−9.95
20	1	0.33	0.81 ± 0.12	7.20 ± 0.07	−9.65
	2	0.23	0.80 ± 0.15	7.27 ± 0.10	−9.74
	av	0.28	0.80 ± 0.19	7.23 ± 0.12	−9.70
25	1	0.31	0.99 ± 0.15	7.49 ± 0.06	−10.21
	2	0.20	1.12 ± 0.24	7.49 ± 0.08	−10.22
	av	0.26	1.05 ± 0.28	7.49 ± 0.10	−10.22
30	1	0.26	0.91 ± 0.20	7.05 ± 0.09	−9.77
	2	0.20	1.09 ± 0.27	7.06 ± 0.09	−9.79
	av	0.23	1.00 ± 0.33	7.05 ± 0.12	−9.78
37	1	0.33	0.78 ± 0.10	7.15 ± 0.06	−10.14
	2	0.21	1.21 ± 0.12	7.20 ± 0.04	−10.21
	av	0.27	0.99 ± 0.16	7.18 ± 0.07	−10.19
42	1	0.21	1.04 ± 0.53	6.76 ± 0.18	−9.74
	2	0.17	1.12 ± 0.24	6.92 ± 0.07	−9.98
	av	0.19	1.08 ± 0.59	6.84 ± 0.19	−9.86
			−1		
10	1	0.39	0.91 ± 0.091	7.11 ± 0.05	−9.21
	2	0.36	0.83 ± 0.12	7.15 ± 0.07	−9.26
	av	0.38	0.87 ± 0.15	7.13 ± 0.09	−9.24
15	1	0.43	0.84 ± 0.12	7.04 ± 0.07	−9.27
	2	0.37	0.88 ± 0.13	7.06 ± 0.07	−9.31
	av	0.40	0.86 ± 0.18	7.05 ± 0.10	−9.29
20	1	0.35	1.03 ± 0.14	7.13 ± 0.05	−9.55
	2	0.36	0.97 ± 0.14	7.15 ± 0.06	−9.59
	av	0.36	1.00 ± 0.22	7.14 ± 0.08	−9.57
20	1	0.32	0.96 ± 0.21	7.22 ± 0.09	−9.67
	2	0.36	0.97 ± 0.14	7.19 ± 0.06	−9.63
	av	0.34	1.00 ± 0.22	7.20 ± 0.09	−9.65
25	1	0.26	1.63 ± 0.28	7.07 ± 0.05	−9.64
	2	0.36	1.08 ± 0.13	6.93 ± 0.06	−9.45
	av	0.30	1.36 ± 0.31	7.00 ± 0.08	−9.55
30	1	0.29	0.97 ± 0.12	6.94 ± 0.06	−9.61
	2	0.40	0.97 ± 0.08	6.96 ± 0.04	−9.64
	av	—	0.97 ± 0.14	6.95 ± 0.07	−9.63
37	1	0.24	1.14 ± 0.38	7.00 ± 0.13	−9.92
	2	0.51	0.71 ± 0.11	6.61 ± 0.09	−9.37
	av	—	0.92 ± 0.40	6.80 ± 0.16	−9.65
42	1	—	—	—	—
	2	—	—	—	—
	av		—	—	—

[a]Experiments were conducted under standard conditions of 100 mM KCl. All other details are as described in Table 1. [b]Values are reported at the temperature at which the titration was conducted and are for protonation and so were calculated according to the negative of eq 3. [c]An additional titration was conducted at 10 °C in the presence of low EDTA that gave a pK_a value of 7.86 ± 0.16 (Figure 3).

temperature dependence for pK_a (Table 2). The effect of temperature on pK_a was therefore examined for WT and −1 over the range of 10−42 °C (Figure 3). Titrations were fit to eq 1, and resulting pK_a values were plotted versus inverse temperature to allow van't Hoff analysis.

As shown in Figure 3, slopes are positive, indicating that lower temperatures lead to larger pK_a values for both WT and −1. Data were fit by linear regression providing slopes of 2300 ± 440 and 1100 ± 340 K and *y*-intercepts of −0.26 ± 1.5 and 3.4 ± 1.2 for

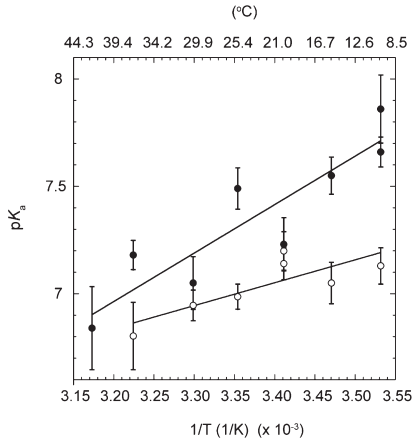

FIGURE 3: Temperature dependence of pK_a titrations. WT (●) and −1 construct (○) pK_a values are plotted as a function of inverse temperature. As the temperature decreases from 42 to 10 °C, pK_a values increase from 6.84 ± 0.1 to 7.86 ± 0.1. The −1 construct did not display a pK_a transition at 42 °C, so this point is missing. Data were fit with linear regression to eq 2, with R^2 values of 0.81 and 0.67 for WT and −1, respectively. Fits produced slopes of 2300 and 1100 K and *y*-intercepts of −0.26 and 3.4 pK_a units for WT and −1, respectively. This led to ΔH° and ΔS° values for protonation of −10 ± 2 kcal/mol and −1.2 ± 7 eu for WT, and −4.9 ± 1.6 kcal/mol and 15 ± 5 eu for the −1 construct, respectively. Experiments were conducted under standard conditions of 100 mM KCl. Error analysis is presented in the Supporting Information.

WT and −1, respectively. According to eq 2, the enthalpic and entropic contributions to the overall energy for protonation of WT are −10 ± 2 kcal/mol and −1.2 ± 7 eu, respectively. In contrast, the enthalpic contribution for protonation of −1 is half that of WT (−4.9 ± 1.6 kcal/mol), but the entropic contribution is a favorable +15 ± 5 eu, which is equivalent to a $-T\Delta S°$ contribution to free energy of −4.65 kcal/mol at 37 °C. Thus, WT protonation is favored enthalpically, while −1 protonation is favored both enthalpically and entropically.

Effect of Ionic Strength on pK_a Shifting. The presence of negatively charged phosphates in nucleic acids results in unique physicochemical properties. For example, unfavorable interactions between backbone charges rigidify the DNA helix and provide a strong dependence on salt interaction for secondary and tertiary structure formation (*38*, *39*). In a similar manner, a positively charged $A^+ \cdot C$ mismatch may interact favorably with nearby negative backbone charges, altering the salt dependence of structure formation and providing an additional driving force for pK_a shifting.

The effect of monovalent salt on WT pK_a was examined over a range of ~10−750 mM free KCl; fitting data for each phosphorothioate peak are listed in Table 3, and average pK_a values are plotted in Figure 4. Titrations were fit to eq 1, and resulting pK_a values are plotted versus the logarithm of potassium ion concentration. As shown in Figure 4, the slope is negative, indicating that the pK_a decreases with salt concentration and therefore that a monovalent cation interacts anticooperatively with a proton. Data were fit by linear regression and provided a slope of −0.34 ± 0.04 and a *y*-intercept of 7.38 ± 0.05 pK_a units, which is the pK_a extrapolated to 1 M KCl. According to Scheme 1 and eqs 4 and 5, these data indicate that ~0.34 potassium ion is taken up for every proton released.

Effects of pH on Local and Global Structure. Next, we consider the extent of structure change induced by pH. As mentioned, oligonucleotides used throughout this study have a phosphorothioate at G17, which has previously been shown to

Table 3: Ionic Strength NMR Titration Data[a]

$[K^+]$ (mM)	peak	$\Delta\delta$ (ppm)	n	pK_a	$\Delta G°_{10}$ (kcal/mol)
10	1	0.28	0.83 ± 0.10	8.09 ± 0.06	−10.47
	2	0.12	1.23 ± 0.37	7.99 ± 0.12	−10.35
	av	0.20	1.03 ± 0.38	8.04 ± 0.13	−10.41
30	1	0.28	1.03 ± 0.10	7.91 ± 0.04	−10.25
	2	0.18	0.99 ± 0.20	7.92 ± 0.09	−10.26
	av	0.23	1.01 ± 0.22	7.92 ± 0.10	−10.26
100	1	0.27	1.32 ± 0.18	7.67 ± 0.05	−9.93
	2	0.15	1.31 ± 0.21	7.65 ± 0.05	−9.90
	av	0.21	1.32 ± 0.27	7.66 ± 0.07	−9.91
316	1	0.26	1.12 ± 0.18	7.60 ± 0.07	−9.84
	2	0.19	1.51 ± 0.39	7.64 ± 0.09	−9.89
	av	0.23	1.32 ± 0.43	7.62 ± 0.11	−9.86
750	1	0.32	0.89 ± 0.07	7.38 ± 0.05	−9.55
	2	0.20	0.91 ± 0.11	7.39 ± 0.07	−9.57
	av	0.26	0.90 ± 0.13	7.38 ± 0.08	−9.56

[a]All experiments were conducted in KCl at standard conditions of 10 °C on WT. Other details are as described in Table 1.

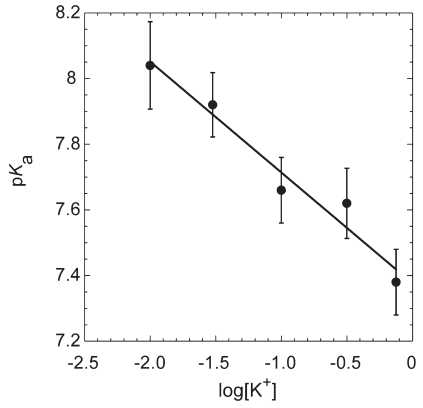

FIGURE 4: Ionic strength dependence of pK_a titrations. Plot of pK_a values for WT (●) as a function of the logarithm of K^+ concentration. These pK_a values are the average of the values for each phosphorothioate and are listed in Table 3. As ionic strength increases, pK_a values decrease from 8.04 ± 0.1 to 7.38 ± 0.1. Data were well fit by linear regression with an R^2 value of 0.96. Error bars in individual pK_a values are from data fitting, increased to a minimum of 0.1 to allow for potential systematic errors. The slope and y-intercept are −0.34 ± 0.04 and 7.38 ± 0.05 pK_a units, respectively. Experiments were conducted under standard conditions at 10 °C.

respond well to local structural changes as depicted in Scheme 3 (*31*). However, the capacity of this labeled oligonucleotide to report global structural changes, such as unfolding shown at the far left and right of Scheme 4, has not been previously assessed.

We performed a titration on WT over a wide pH range in which acid and alkaline denaturation occur because of protonation and deprotonation of the imino nitrogens on the Watson−Crick face. These data are provided in Figure 5 and have five distinct features. First, from pH ~3.8 to 4.5, a negative slope occurs, which reflects acid denaturation of the global structure. Second, a relatively flat region occurs between pH 4.5 and 6, which reflects the folded protonated state that is not ionizing. Third, a positive slope occurs between pH 6 and 8, which reflects local structural changes in response to ionization of the $A^+ \cdot C$ wobble. Fourth, another flat section occurs between pH 8 and 10, which reflects the folded deprotonated state that is not ionizing. Lastly, at pH 10, the data take on a fifth feature, a positive slope that continues through pH 12, which reflects alkaline denaturation of the global structure. Notably, the second through fourth

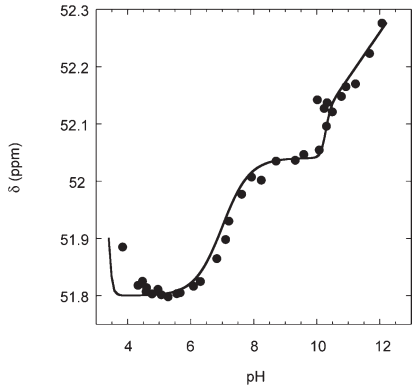

FIGURE 5: Wide-range pH titration of WT (●). NMR data reflect changes in the environment of the phosphorothioate label. Data were collected under standard conditions of 100 mM KCl at 10 °C. Chemical shift signal changes are observed in response to both local and global structural changes and are superimposed by a mathematical model (—) expressed in eqs 6 and 7 that incorporates protonation near neutrality, as well as acidic and alkaline denaturation events.

features comprise the data fit to obtain pK_a values for local $A^+ \cdot C$ wobble protonation; they also represent the states depicted in Scheme 3 in which the protonation of primary interest takes place.

The pH data in Figure 5 were overlayed with output from eq 7, which was developed with a partition function (eq 6) that accounts for all actions of proton binding on oligonucleotide folding. As seen in Figure 5, simulated data agree well with all five features of the experimental data. In these equations, the observed signal is the weighted average of each population; the populations are treated as being in fast exchange because monitored phosphorus chemical shift changes smoothly as a function of pH, rather than splitting into multiple populations (*40*).

Probing the Structure and Thermodynamics of Various States. In an effort to characterize the low- and high-pH folded states described in Scheme 3, NMR imino proton ^{1}H data were collected for WT, −1, −2, and −3 constructs, as well as two controls, at pH 5.7 and 8.0, at 1 °C (Figure 6 and Supporting Information). First, we consider low-pH ^{1}H NMR spectra. A persistent peak was present at ~13.3 ppm in all spectra, while the remaining peaks resonated between 12 and 13 ppm. The most upfield-shifted peak, found at 10.6 ppm at pH 5.7 only, was determined to result from protonation of a loop base (see spectra on Duplex Control in the Supporting Information) and is not considered further, except to note that its presence supports formation of the upper portion of the stem at low pH in all constructs. Given that GC and AT Watson−Crick base pairs resonate at 12−13 and 13−15 ppm, respectively (*41*), the peak at 13.3 ppm was assigned to the single AT (base pair 7 in WT). We also note that, with the exception of the AT Base Pair Control (see Supporting Information), there were no additional resonances found for the various constructs upon opening the window up to 15 ppm.

Assignments of the remaining peaks were made by considering low-pH spectra. Comparing the pH 5.7 spectra of WT and −1 (Figure 6A,B) reveals a small resonance at 12.81 ppm that is absent in the −1 spectrum, suggesting that this resonance corresponds to the terminal base pair in WT (base pair 1). Moreover, the peak at 12.81 ppm is broadened upon increasing the temperature to 10 °C (data not shown), as expected for a terminal base pair.

The low-pH WT spectrum exhibits six peaks: 13.29, 12.81, 12.76, 12.65, 12.45, and 10.58 ppm. The 13.29, 12.81, and 10.58 ppm resonances are assigned as described above. Similarly, the

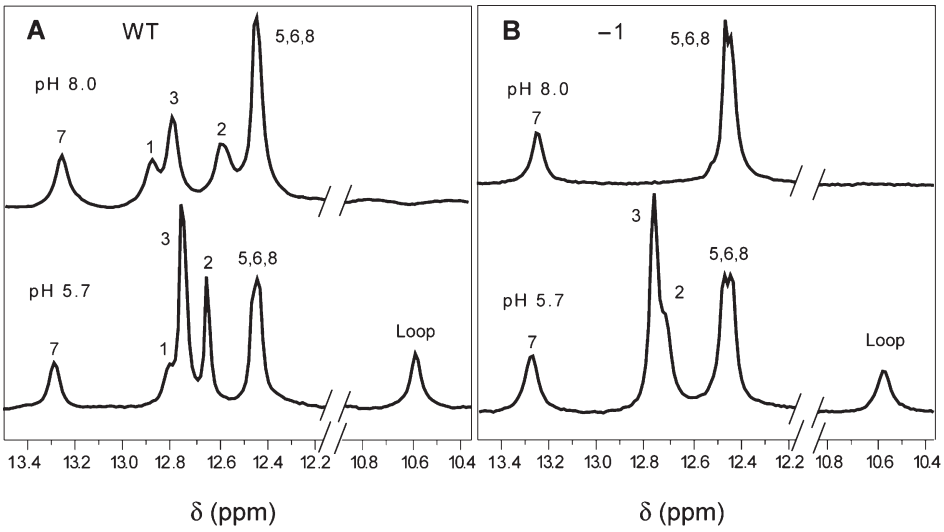

FIGURE 6: ^{1}H NMR spectra of WT and −1 constructs. Spectra were recorded at 1 °C for (A) WT and (B) −1. Data are presented at pH 8.0 (top) and 5.7 (bottom). Resonances are assigned using the base numbering in Figure 1C.

peak at 12.65 ppm can be assigned to base pair 2, as it is the terminal base pair in the −1 construct, where it has shifted only slightly downfield, to 12.72 ppm, which is partly due to its overlap with the nearby large resonance at 12.76 ppm. This large peak at 12.76 and 12.77 ppm in the WT and −1 spectra, respectively, can be assigned to base pair 3: it is distinctly sharper in WT at low pH and disappears in −1 at high pH. Sharpening of this resonance in WT at low pH supports nearby $A^+\cdot C$ wobble formation (*13*), while absence of this resonance at high pH in −1 suggests that base pair 3 breaks with the AC wobble. Additionally, the resonance for base pair 3 is effectively absent from the spectra of the −2 and −3 constructs at both low and high pH (Supporting Information), which supports the earlier finding that −2 does not exhibit an observable pK_a (Figure 2), as well as our earlier studies in which a terminal base pair did not form in the absence of a penultimate base pair (*36*).

The remaining broad feature centered near 12.45 ppm in the pH 5.7 spectrum of WT is likely a set of poorly resolved peaks corresponding to GC base pairs 5, 6, and 8. The upper portion of the stem in WT, which contains these base pairs, appears to be isolated from the local dynamics of the $A^+\cdot C$ wobble. In the constructs with shortened helices, the effective helical terminus is closer to base pairs 5, 6, and 8, and partial dispersion of these features occurs (Figure 6 and Supporting Information). Additional spectral analysis would be necessary to assign base pairs 5, 6, and 8; however, as they have no pH dependence and thus do not have an impact on analysis, this was not pursued. The WT spectrum at pH 8.0 displays the same number of peaks as the low-pH spectrum, with the exception of the loop peak, but the peaks between 12.5 and 13 ppm are downfield-shifted and broadened, consistent with increased exposure to and subsequent chemical exchange with solvent.

The imino proton spectrum of −1 exhibits a more pronounced dependence on pH than the WT spectrum. At low pH, the feature corresponding to base pairs 5, 6, and 8 is seen centered at 12.46 ppm and, as discussed above, is slightly more resolved. The base pair 7 resonance is present at 13.28 ppm, and the sharp base pair 3 resonance appears at 12.77 ppm, nearly obscuring the base pair 2 peak at 12.72 ppm. Remarkably, at pH 8.0, only peaks corresponding to base pairs between the $A^+\cdot C$ wobble site and the hairpin loop are present, suggesting that deprotonation and subsequent disruption of this wobble induce the neighboring base pairs (2 and 3) to break in a cooperative fashion. This observation also supports the greater $\Delta\delta$ observed for −1 versus that for WT in the pH titrations as presented above (Figure 2 and Table 1).

Additional constructs were also analyzed by ^{1}H NMR, and their spectra appear in the Supporting Information. Included is a hairpin identical to WT except that the $A^+\cdot C$ wobble is replaced with a Watson−Crick AT base pair (AT Control). As expected, imino proton NMR spectra of this construct, which had an additional downfield peak for the new AT, were nearly identical at low (5.7) and high (8.0) pH and did not show the peak broadening observed in the pH 8.0 WT spectrum, which supports the presence of a protonated base in WT at pH 5.7.

Next, thermodynamic properties of the various DNA constructs were determined by UV-monitored thermal denaturation at pH 5.7 and 8.0 (see Materials and Methods for details of the temperature dependence of pK_a). Buffer-corrected and normalized data at pH 5.7 reveal that WT and −1 unfold with significantly greater hyperchromicity (~23%) than −2 and −3 (~17%) (Figure 7A). This trend is consistent with the imino NMR experiments, which revealed that at low pH, WT and −1 form more base pairs than shorter constructs.

At pH 5.7, the melting transition for −1 was significantly broader than for WT (Figure 7A). Indeed, the first-derivative plot (Figure 7B) reveals two maxima for −1, supporting non-cooperative, three-state melting at low pH. At pH 8.0, the three-state behavior of −1 is much less pronounced; moreover, WT now has greater hyperchromicity than −1, suggesting that some base pairs do not form in −1 at high pH. Overall, the melting data support a model in which the lower portion of the stem of −1 melts independently at low pH and does not form at high pH. This finding is consistent with the imino proton NMR data (Figure 6) and the phosphorus NMR titrations (Figure 2).

DISCUSSION

Protonation of the nucleobases plays significant functional roles in biology and chemistry. Protonated bases have been implicated in mutagenesis, viral frame shifting, and catalysis, for example (*6*, *8*, *20*, *29*, *30*, *42*). It is therefore important to understand driving forces for nucleobase protonation. In this study, we examined the effects of three factors on pK_a shifting:

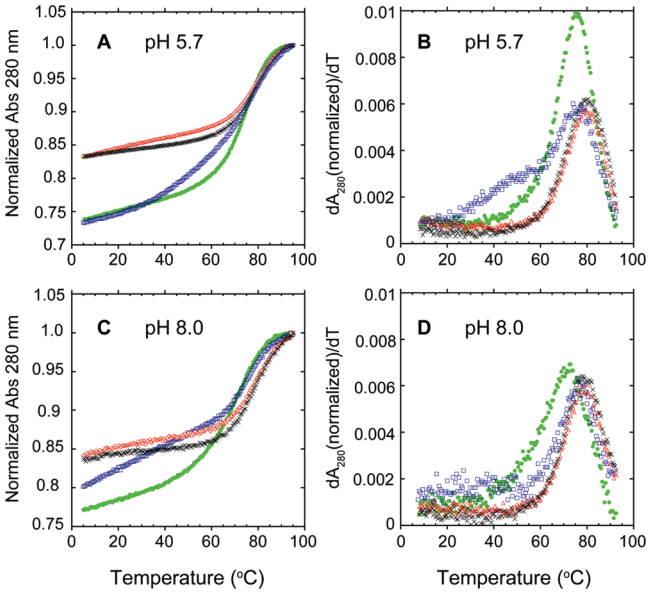

FIGURE 7: UV-monitored thermal denaturation of WT, −1, −2, and −3 constructs. Melts were conducted at 280 nm at (A) pH 5.7 and (C) pH 8.0. First derivatives of data are also provided for (B) pH 5.7 and (D) pH 8.0. The −2 (red ◇) and −3 (black ×) constructs have similar melting profiles in all panels, consistent with similar secondary structures. The WT (green ○) and −1 (blue □) constructs melt similarly at low pH (A), although panel B reveals that −1 has an additional transition that is absent at pH 8.0 (D), where it behaves more like −2 and −3.

helix position, temperature, and ionic strength. All three strongly influence the pK_a of an A^+·C mismatch, leading to average pK_a values as high as 8.0 and as low as 6.8. In the Discussion, we consider possible molecular explanations for pK_a shifting, as well as biological relevance.

Large Folding Cooperativity Is Not Tantamount to Large pK_a Shifting. The pK_a shift associated with A^+·C mismatch formation depended on its location in the helix. The most internally positioned A^+·C wobble, in WT, which is 3 bp from the helix terminus, displayed a pK_a of 7.66 at 10 °C. This value corresponds to a remarkable shift of ~4.2 units from the unfolded-state pK_a of adenosine (*15*) and according to eq 3 to an additional $\Delta G°_{10}$ for protonation of −5.4 kcal/mol. The −1 construct, which places the A^+·C mismatch 1 bp closer to the helix terminus, also has a shifted, albeit slightly smaller, pK_a of 7.13 (~3.6 pK_a unit shift), corresponding to an additional $\Delta G°_{10}$ for protonation of −4.7 kcal/mol. Finally, the −2 construct, which places the A^+·C mismatch penultimate to the base of the stem, did not give an observable change in chemical shift down to pH 5.5 (Figure 2), indicating insufficient thermodynamic driving force to form the A^+·C mismatch.

We are interested in discerning driving forces for protonation. One useful path for considering A^+·C wobble formation is loading of the proton onto an adenine with an unshifted pK_a followed by favorable structure formation, as depicted in Figure 8.[5] We begin by considering how data for −2 fit this mechanism. Failure of the A^+·C mismatch to form in −2 as revealed by ^{31}P NMR (Figure 1) indicates that the energetics of formation of an A^+·C mismatch and one terminal CG base pair are insufficient to overcome the energy required to protonate adenine at pH >5.5. Given that the free energy for step 2 in Figure 8, upper scheme, is given by $\Delta G°_{10} = -1.30\Delta pK_a$ (*15*), we infer that an A^+·C mismatch and a single terminal CG base pair contribute less than −2.6 kcal/mol to stability at 10 °C [= 1.30(5.5 − 3.5)]. This conclusion is strengthened by the absence of observable base pairing below the A^+·C mismatch at low pH according to ^{1}H NMR (Supporting Information).

In the −1 construct, titration data support A^+·C mismatch formation, which indicates that a second GC base pair below an A^+·C mismatch provides sufficient driving force for protonation. We therefore conclude that the two base pairs below the A^+·C mismatch in −1 provide approximately −4.7 kcal/mol in $\Delta G°_{10}$, while the three base pairs below the A^+·C mismatch in WT provide approximately −5.4 kcal/mol (see above). The ^{1}H NMR data for −1 indicate that the additional base pairing in the lower portion of the stem forms cooperatively, while ^{1}H NMR data for WT support additional base pairing forming noncooperatively (compare low- and high-pH NMR data for WT and −1 in Figure 6). Higher cooperativity occurring nearer the terminus of the helix agrees with previous studies on single mismatches (*36*, *37*). Thus, the high cooperativity of secondary structure formation is not tantamount to large pK_a shifting.

In our previous study on single mismatches, an inverse correlation between cooperativity and stability was found: DNA hairpins with mismatches near the terminus of the helix displayed high cooperativity but low stability, while mismatches located internally had low cooperativity but high stability (*36*). For example, a terminal GC base pair did not form in the absence of a penultimate AT base pair (strong cooperative coupling), and disruption of the AT base pair resulted in a destabilization of just 3.24 kcal/mol in $\Delta G°_{10}$. In contrast, the same disruption in the middle of a helix only slightly weakened neighboring base pairs (a less cooperative result) but resulted in a large energetic disruption of 4.14 kcal/mol in $\Delta G°_{10}$. The pK_a shift observed for WT herein is associated with a stabilization of −5.4 kcal/mol in $\Delta G°_{10}$, suggesting that the A^+·C wobble contributes approximately −1.3 kcal/mol (= −5.4 − −4.1) beyond AT formation in the simple helix. This extra stabilization likely arises from favorable electrostatic interaction between protonated A and neighboring phosphates. In summary, A^+·C wobble formation in dsDNA follows free energy rules similar to those for helix formation: formation of the wobble is more stabilizing in internal rather than external portions of a helix.

The Thermodynamics of Protonation Depend on Folding. An inverse relationship was observed between oligonucleotide pK_a and temperature (Figure 3). Moreover, the pK_a of WT has a steeper dependence on temperature than the pK_a of −1, providing an enthalpy change of −10 kcal/mol for WT and just −4.9 kcal/mol for −1, and an entropy change of −1.2 eu for WT and 15.5 eu for −1. The overall thermodynamics can be considered in two steps: base protonation and concomitant structure formation (Figure 8). Protonation of adenine has been reported to have a favorable enthalpy change of −4.8 kcal/mol and a favorable entropy change of 1.9 eu at 25 °C (close to 37 °C used here) from studies on dADP at 100 mM ionic strength, which is our standard state (*1*). Given the overall thermodynamic values and the adenine values, we thus infer that the structure formation step in WT is associated with a favorable enthalpy change of −5.2 kcal/mol and a near-zero entropy change of −3 eu (a $-T\Delta S°$ at 37 °C of 1.0 kcal/mol). The structure formation step in −1, on the other hand, is associated with a near-zero enthalpy

[5]Note that Figure 8 provides values for $\Delta G°_{37}$, which are discussed in the next section, while this section presents $\Delta G°_{10}$ values; this is done because the experiments in this section were conducted at 10 °C. Differences between these free energy values are small and do not affect the general trends.

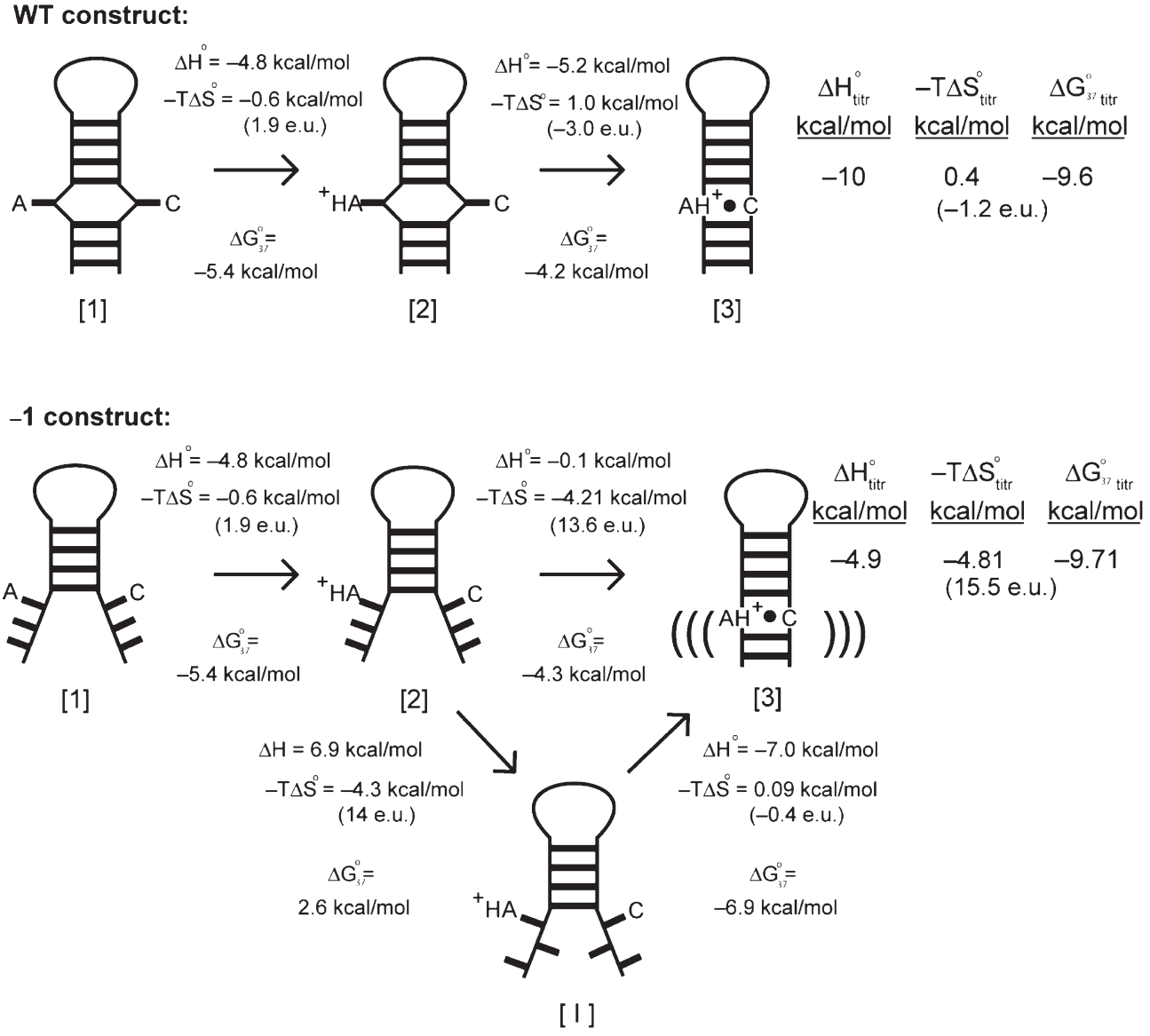

FIGURE 8: Dissection of thermodynamic parameters for WT and −1 . Titration of each construct consists of at least two steps, with the first step depicted as protonation of A prior to structure formation and the second step as structure formation. States are numbered for reference purposes. $\Delta H°$ and $\Delta S°$ values in the right-hand columns were determined from the van't Hoff analysis in Figure 3, and $\Delta G°_{37}$ is calculated from these terms; agreement with the experimental $\Delta G°$ determined by titration at 37 °C (Table 2) is good. Thermodynamic parameters for protonation of A (step 1) are from ref *1* (see text for details), while parameters for structure formation (step 2) are calculated from step 1 and the overall experimental parameters. The −1 construct has two additional steps, which are to and from an intermediate state [I] with unstacked ends. The thermodynamic parameters from [2] to [I] are the sum of parameters for d(AAA) unstacking from the 5′- and 3′-ends of a DNA hairpin (*43*), while the thermodynamic parameters for [I] back to [3] are calculated by difference. The lower helix in state [3] in −1 is likely dynamic, as discussed in the text, and is depicted as such. Steps are depicted with a forward arrow to denote the directionality for thermodynamic parameters and are really in thermodynamic equilibrium.

change of −0.1 kcal/mol but a large favorable entropy change of 13.6 eu (a $-T\Delta S°$ at 37 °C of −4.21 kcal/mol) (Figure 8).

The simplest model that accounts for the near-zero entropy change of WT is a largely preorganized binding site for protonation, which would require little change in conformational entropy to fold into the final protonated state, although small structural accommodations and compensations are possible. This scenario is consistent with similar NMR spectra for WT at low and high pH (Figure 6A), as well as with the reduced folding cooperativity characteristic of internally located mismatches (*36*). The favorable enthalpy change parameter for WT of −5.2 kcal/mol could be due to an increase in the number of hydrogen bonding or stacking interactions upon base protonation, as well as associated structural accommodations.

To account for the structure formation step of −1, we invoke a pathway involving an intermediate ([I] in Figure 8) in which the three dangling nucleotides on each of the helix ends unstack prior to base pair formation. From the work of Sugimoto and co-workers on d(AAA) dangling ends on a DNA hairpin (*43*), unstacking the 3′-end contributes a $\Delta H°$ of 6.6 kcal/mol and a $\Delta S°$ of 17 eu, while unstacking the 5′-end contributes a $\Delta H°$ of 0.3 kcal/mol and a $\Delta S°$ of −3 eu. Summing these parameters provides a $\Delta H°$ of 6.9 kcal/mol and a $\Delta S°$ of 14 eu (a $-T\Delta S°$ at 37 °C of −4.3 kcal/mol) for the [2] to [I] step depicted in Figure 8, bearing in mind that these numbers are estimates since the dangling ends used herein have different sequences. These considerations support the protonated intermediate [2] for the −1 construct having highly ordered dangling ends. The thermodynamic parameters for the

step from [I] to [3] are calculated by difference to consist of a favorable $\Delta H°$ of −7.0 kcal/mol and a slightly unfavorable $\Delta S°$ of −0.4 eu (a $-T\Delta S°$ at 37 °C of 0.09 kcal/mol). The entropy change of [I] to [3] may be ∼0 because [3] is dynamic as suggested by NMR and UV melting studies. The free energy for the [2] to [I] step is unfavorable (2.6 kcal/mol), as expected for unstacking, while the free energy for the [I] to [3] step is quite favorable (−6.9 kcal/mol), as expected for formation of a dynamic group of 3 bp from an unstacked state. Overall, we conclude that the gain in entropy for the structure at the base of the helix upon going from protonated state [2] to final state [3] is due to the protonated intermediate [2] having stacked and highly ordered dangling ends and the final state [3] having this stacking removed into a presumably dynamic lower helix.

With regard to the relationship between base pair protonation and DNA folding, one more consideration is in order. Theory and experiments from Record and co-workers indicate that there are likely DNA length-dependent extents of couterion condensation for oligonucleotides (*44*, *45*). For double-stranded oligonucleotides, they suggest that a Coulombic end effect is expected for the first 35 bp of a long duplex. One can then expect that in short DNA oligonucleotides, like those used here, and in most cellular RNAs, where double-stranded stretches have typically fewer than 10 bp, such end effects will be important. The work of Record and co-workers also indicates that sites near the end of the DNA have weaker cation affinity than interior sites and that this difference is exacerbated by salt concentrations below ∼0.3 M. In addition, theoretical studies by Pack and co-workers suggest the presence of proton gradients around DNA, with high concentrations of protons in the minor and major grooves, and these regions of high local proton density can give rise to upward pK_a shifts of 2 units (*46*). Experimental measurements of pK_a values on amine adducts in the minor groove of the DNA by these investigators further supported these calculations (*47*). Because the relative contributions of Coulombic end effects to proton and cation binding to $A^+\cdot C$ wobble pairs very close to the end of the DNA are unknown at present, it is unclear how much of a difference end effects make to the competition between ions and protons. Nonetheless, it is possible that such effects, which likely vary when the construct is changed from WT to −1, −2, and −3, alter the competition between monovalent cations and protons (see next section) and are important for the pK_a values measured here. This would provide additional complexity to the competition between these species depicted in Scheme 1. Further study will be needed to address these more complex models.

The Magnitude of pK_a Shifting Is Dampened by an Increasing Ionic Strength. Several models of the interaction between a protonated nucleobase (A or C) and a cation can be considered, each of increasing complexity. Ionic strength has no significant effect on the pK_a value for N3 of cytidine or N1 of adenine (*1*), as expected for a protonation event that forms a cationic species (*48*). Similarly, the pK_a values of 5′-CMP and 5′-AMP are not a function of ionic strength (*1*). Thus, salt dependence is not intrinsic to nucleobase protonation.

A slightly more complex system, UUCUU, does exhibit an ionic strength dependence. As ionic strength increases, pK_a values decrease, with a slope of -0.33 ± 0.07, as determined from log plots of data from (*32*). A single strand is a more accurate model of RNA than cytidine or CMP, as it introduces more phosphates to interact favorably with cationic cytosine as well as locally bound metal ions to interact unfavorably. The dependence of the $A^+\cdot C$ mismatch pK_a on ionic strength for the dsDNA oligonucleotides studied here was similar to that for UUCUU, with a slope of -0.34 ± 0.04. This finding is consistent with unfavorable interaction of protonation and locally bound metal ions, as depicted in Scheme 1.

Williams et al. examined folding of d(GCATGC) as a function of Na^+ and found 0.4 ion taken up per strand upon folding (*49*). This value corresponds to uptake of 0.08 Na^+ ion per phosphate. The number of ions released herein upon protonation was determined to be 0.34 per hairpin. Since protonation is a local event for titrations between pH 5 and 9 (see Results), one can expect approximately 4 phosphates to be involved (i.e., two phosphates for the base pair above and two phosphates for the base pair below the mismatch), which gives a number of K^+ ions per phosphate of ∼0.08 (=0.34/4), in good agreement with the d(GCATGC) study.

Possible Biological Relevance of AC pK_a Shifting. The stability of an organism's genetic information is dependent on high-fidelity replication, a process that relies on Watson−Crick base pairing. While this process is generally very reliable (error rates from 10^{-3} to $<10^{-6}$), misincorporation events do occur and can have a large impact on replication (*50*−*52*). It has been reported that misincorporations depend on the spacing between the mismatch and the polymerase active site (*53*); for instance, mismatches up to 4 bp from the polymerase catalytic site lead to misincorporation at the active site. Indeed, a study examining the structures of mismatch errors in a complex with a polymerase found an A·C or C·A mismatch incorporated at a polymerase postinsertion site causes significant disorder in the modeled duplex at the active site (*54*). This observation is consistent with the position-dependent model of pK_a shifting developed here in which the pK_a value depends on the distance between the A·C mismatch and the helix terminus.

When mismatches do occur, multiple DNA proofreading mechanisms exist to repair the damage, each with their own recognition patterns and specificities. Some are specific to $A^+\cdot C$ wobbles. In one specific example, an $A^+\cdot C$ wobble can be repaired by two repair mechanisms, which can result in either correct repair to an AT base pair or transition to a GC pair, thereby introducing more error into the genetic code (*29*).

In addition to roles in replication, protonation extends the functional diversity of DNA and RNA. In previous studies, different classes of protonated RNA nucleobases were discussed, as were potential roles in general acid−base and electrostatic catalysis as an oxyanion hole (*4*, *55*). Our observation of pK_a values near and above neutrality supports such catalytic roles.

The pK_a values for the ionizing $A^+\cdot C$ mismatch are largest when wobble position is internal and the temperature and ionic strength are low. Nature could exploit these characteristics. The mismatches can be located internally and perhaps further stabilized by tertiary structure formation or protein binding. Temperatures are lower in mesophiles and cryophiles, including boreal plants. Lastly, low ionic strength might be afforded by sequestration of the RNA when bound to a protein or internalized in a tertiary structure. All of these strategies make feasible large and biologically relevant pK_a shifts.

ACKNOWLEDGMENT

We thank Dan Herschlag, Scott Showalter, and Jennifer Wilcox for helpful discussions and comments, and David Cacase, Josh Sokoloski, and Jennifer Wilcox for assistance with atomic absorption experiments.

SUPPORTING INFORMATION AVAILABLE

Derivation of global denaturation equations, error analysis for pK_a versus $1/T$ data, imino regions of ^{1}H NMR spectra, atomic absorption data, and supporting references. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Izatt, R. M., Christensen, J. J., and Rytting, J. H. (1971) Sites and thermodynamic quantitities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides. *Chem. Rev. 71*, 439–481.
2. Saenger, W. (1984) Principles of Nucleic Acid Structure, Springer-Verlag, New York.
3. Pasternack, L. B., Lin, S. B., Chin, T. M., Lin, W. C., Huang, D. H., and Kan, L. S. (2002) Proton NMR studies of 5′-d-$(TC)_3(CT)_3(AG)_3$-3′—a paperclip triplex: The structural relevance of turns. *Biophys. J. 82*, 3170–3180.
4. Bevilacqua, P. C., Brown, T. S., Nakano, S., and Yajima, R. (2004) Catalytic roles for proton transfer and protonation in ribozymes. *Biopolymers 73*, 90–109.
5. Tarkoy, M., Phipps, A. K., Schultze, P., and Feigon, J. (1998) Solution structure of an intramolecular DNA triplex linked by hexakis(ethylene glycol) units: d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-TCTCTCTT). *Biochemistry 37*, 5810–5819.
6. Venditti, V., Clos, L., II, Niccolai, N., and Butcher, S. E. (2009) Minimum-energy path for a U6 RNA conformational change involving protonation, base-pair rearrangement and base flipping. *J. Mol. Biol. 391*, 894–905.
7. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry 35*, 6026–6036.
8. Nakano, S., Chadalavada, D. M., and Bevilacqua, P. C. (2000) General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science 287*, 1493–1497.
9. Huppler, A., Nikstad, L. J., Allmann, A. M., Brow, D. A., and Butcher, S. E. (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat. Struct. Biol. 9*, 431–435.
10. Nakano, S., and Bevilacqua, P. C. (2007) Mechanistic characterization of the HDV genomic ribozyme: A mutant of the C41 motif provides insight into the positioning and thermodynamic linkage of metal ions and protons. *Biochemistry 46*, 3001–3012.
11. Gong, B., Chen, J. H., Chase, E., Chadalavada, D. M., Yajima, R., Golden, B. L., Bevilacqua, P. C., and Carey, P. R. (2007) Direct measurement of a pK_a near neutrality for the catalytic cytosine in the genomic HDV ribozyme using Raman crystallography. *J. Am. Chem. Soc. 129*, 13335–13342.
12. Legault, P., and Pardi, A. (1997) Unusual dynamics and pKa shift at the active site of a lead-dependent ribozyme. *J. Am. Chem. Soc. 119*, 6621–6628.
13. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A·C, C·C, G·G, U·C, and U·U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U·U and C·C+ pairs. *Biochemistry 30*, 8242–8251.
14. Pan, B., Mitra, S. N., and Sundaralingam, M. (1999) Crystal structure of an RNA 16-mer duplex $R(GCAGAGUUAAAUCUGC)_2$ with nonadjacent G(syn)·A+(anti) mispairs. *Biochemistry 38*, 2826–2831.
15. Moody, E. M., Lecomte, J. T., and Bevilacqua, P. C. (2005) Linkage between proton binding and folding in RNA: A thermodynamic framework and its experimental application for investigating pKa shifting. *RNA 11*, 157–172.
16. Tang, C. L., Alexov, E., Pyle, A. M., and Honig, B. (2007) Calculation of pK_as in RNA: On the structural origins and functional roles of protonated nucleotides. *J. Mol. Biol. 366*, 1475–1496.
17. Nakano, S., Proctor, D. J., and Bevilacqua, P. C. (2001) Mechanistic characterization of the HDV genomic ribozyme: Assessing the catalytic and structural contributions of divalent metal ions within a multichannel reaction mechanism. *Biochemistry 40*, 12022–12038.
18. Perrotta, A. T., Shih, I., and Been, M. D. (1999) Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme. *Science 286*, 123–126.
19. Ferre-D'Amare, A. R., and Doudna, J. A. (2000) Crystallization and structure determination of a hepatitis delta virus ribozyme: Use of the RNA-binding protein U1A as a crystallization module. *J. Mol. Biol. 295*, 541–556.
20. Nixon, P. L., and Giedroc, D. P. (2000) Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J. Mol. Biol. 296*, 659–671.
21. Kuzmin, Y. I., Da Costa, C. P., Cottrell, J. W., and Fedor, M. J. (2005) Role of an active site adenine in hairpin ribozyme catalysis. *J. Mol. Biol. 349*, 989–1010.
22. Cornish, P. V., and Giedroc, D. P. (2006) Pairwise coupling analysis of helical junction hydrogen bonding interactions in luteoviral RNA pseudoknots. *Biochemistry 45*, 11162–11171.
23. Lipfert, J., Ouellet, J., Norman, D. G., Doniach, S., and Lilley, D. M. (2008) The complete VS ribozyme in solution studied by small-angle X-ray scattering. *Structure 16*, 1357–1367.
24. Spitale, R. C., Volpini, R., Mungillo, M. V., Krucinska, J., Cristalli, G., and Wedekind, J. E. (2009) Single-Atom Imino Substitutions at A9 and A10 Reveal Distinct Effects on the Fold and Function of the Hairpin Ribozyme Catalytic Core. *Biochemistry 48*, 7777–7779.
25. Legault, P., and Pardi, A. (1994) In Situ Probing of Adenine Protonation in RNA by ^{13}C NMR. *J. Am. Chem. Soc. 116*, 8390–8391.
26. Bevilacqua, P. C., and Yajima, R. (2006) Nucleobase catalysis in ribozyme mechanism. *Curr. Opin. Chem. Biol. 10*, 455–464.
27. Muth, G. W., Ortoleva-Donnelly, L., and Strobel, S. A. (2000) A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science 289*, 947–950.
28. Okuda, K., Seila, A. C., and Strobel, S. A. (2005) Uncovering the enzymatic pKa of the ribosomal peptidyl transferase reaction utilizing a fluorinated puromycin derivative. *Biochemistry 44*, 6675–6684.
29. Kim, M., Huang, T., and Miller, J. H. (2003) Competition between MutY and mismatch repair at A × C mispairs in vivo. *J. Bacteriol. 185*, 4626–4629.
30. Giri, I., and Stone, M. P. (2003) Wobble dC·dA pairing 5′ to the cationic guanine N7 8,9-dihydro-8-(N7-guanyl)-9-hydroxyaflatoxin B1 adduct: Implications for nontargeted AFB1 mutagenesis. *Biochemistry 42*, 7023–7034.
31. Moody, E. M., Brown, T. S., and Bevilacqua, P. C. (2004) Simple method for determining nucleobase pKa values by indirect labeling and demonstration of a pKa of neutrality in dsDNA. *J. Am. Chem. Soc. 126*, 10200–10201.
32. Cerrone-Szakal, A. L., Siegfried, N. A., and Bevilacqua, P. C. (2008) Mechanistic characterization of the HDV genomic ribozyme: Solvent isotope effects and proton inventories in the absence of divalent metal ions support C75 as the general acid. *J. Am. Chem. Soc. 130*, 14504–14520.
33. Hwang, T. L., and Shaka, A. J. (1995) Water suppression that works: Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J. Magn. Reson., Ser. A 112*, 275–279.
34. Feig, A. L., and Uhlenbeck, O. (1999) The role of metal ions in RNA biochemistry. In The RNA World (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) 2nd ed., pp 287−319, Cold Spring Harbor Laboratory Press, Plainview, NY.
35. Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S., and Singh, R. M. (1966) Hydrogen ion buffers for biological research. *Biochemistry 5*, 467–477.
36. Siegfried, N. A., Metzger, S. L., and Bevilacqua, P. C. (2007) Folding cooperativity in RNA and DNA is dependent on position in the helix. *Biochemistry 46*, 172–181.
37. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry 38*, 14214–14223.
38. Lu, Y., Weers, B., and Stellwagen, N. C. (2001) DNA persistence length revisited. *Biopolymers 61*, 261–275.
39. Kebbekus, P., Draper, D. E., and Hagerman, P. (1995) Persistence length of RNA. *Biochemistry 34*, 4354–4357.
40. Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley & Sons, New York.
41. Allawi, H. T., and SantaLucia, J., Jr. (1998) Nearest-neighbor thermodynamics of internal A·C mismatches in DNA: Sequence dependence and pH effects. *Biochemistry 37*, 9435–9444.
42. Guo, M., Spitale, R. C., Volpini, R., Krucinska, J., Cristalli, G., Carey, P. R., and Wedekind, J. E. (2009) Direct Raman Measurement of an Elevated Base pK_a in the Active Site of a Small Ribozyme in a Precatalytic Conformation. *J. Am. Chem. Soc. 131*, 12908–12909.
43. Ohmichi, T., Nakano, S., Miyoshi, D., and Sugimoto, N. (2002) Long RNA dangling end has large energetic contribution to duplex stability. *J. Am. Chem. Soc. 124*, 10367–10372.
44. Olmsted, M. C., Anderson, C. F., and Record, M. T., Jr. (1989) Monte Carlo description of oligoelectrolyte properties of DNA oligomers: Range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits. *Proc. Natl. Acad. Sci. U.S.A. 86*, 7766–7770.
45. Zhang, W., Ni, H., Capp, M. W., Anderson, C. F., Lohman, T. M., and Record, M. T., Jr. (1999) The importance of coulombic end effects: Experimental characterization of the effects of oligonucleotide

flanking charges on the strength and salt dependence of oligocation (L8+) binding to single-stranded DNA oligomers. *Biophys. J. 76*, 1008–1017.
46. Lamm, G., and Pack, G. R. (1990) Acidic domains around nucleic acids. *Proc. Natl. Acad. Sci. U.S.A. 87*, 9033–9036.
47. Hanlon, S., Wong, L., and Pack, G. R. (1997) Proton equilibria in the minor groove of DNA. *Biophys. J. 72*, 291–300.
48. Fersht, A. (1985) Enzyme Structure and Mechanism, 2nd ed., Freeman, New York.
49. Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., and Turner, D. H. (1989) Laser Temperature-Jump, Spectroscopic, and Thermodynamic Study of Salt Effects on Duplex Formation by dGCATGC. *Biochemistry 28*, 4283–4291.
50. Echols, H., and Goodman, M. F. (1991) Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem. 60*, 477–511.
51. Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol. 28*, 83–126.
52. Kunkel, T. A., and Bebenek, K. (2000) DNA replication fidelity. *Annu. Rev. Biochem. 69*, 497–529.
53. Carver, T. E., Jr., Hochstrasser, R. A., and Millar, D. P. (1994) Proofreading DNA: Recognition of aberrant DNA termini by the Klenow fragment of DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A. 91*, 10670–10674.
54. Johnson, S. J., and Beese, L. S. (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell 116*, 803–816.
55. Narlikar, G. J., and Herschlag, D. (1997) Mechanistic aspects of enzymatic catalysis: Lessons from comparison of RNA and protein enzymes. *Annu. Rev. Biochem. 66*, 19–59.