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

pubs.acs.org/JPCA

Article

Calculations of pK_a Values for a Series of Naturally Occurring Modified Nucleobases

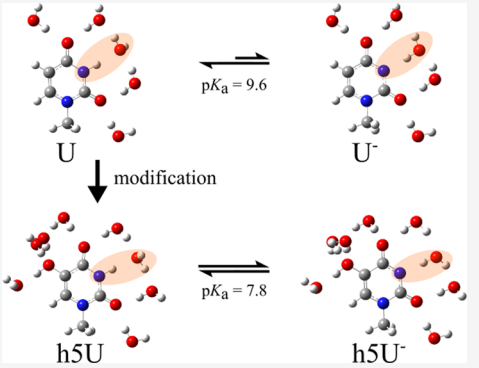
Evan L. Jones, Alan J. Mlotkowski, Sebastien P. Hebert, H. Bernhard Schlegel, and Christine S. Chow*

Cite This: *J. Phys. Chem. A* 2022, 126, 1518−1529

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Modified nucleobases are found in functionally important regions of RNA and are often responsible for essential structural roles. Many of these nucleobase modifications are dynamically regulated in nature, with each modification having a different biological role in RNA. Despite the high abundance of modifications, many of their characteristics are still poorly understood. One important property of a nucleobase is its pK_a value, which has been widely studied for unmodified nucleobases, but not for the modified versions. In this study, the pK_a values of modified nucleobases were determined by performing ab initio quantum mechanical calculations using a B3LYP density functional with the 6-31+G(d,p) basis set and a combination of implicit−explicit solvation systems. This method, which was previously employed to determine the pK_a values of unmodified nucleobases, is applicable to a variety of modified nucleobases. Comparisons of the pK_a values of modified nucleobases give insight into their structural and energetic impacts within nucleic acids.



INTRODUCTION

RNA serves a myriad of cellular roles, which rely on complex higher-order structures and biomolecular interactions.[1−3] RNA secondary and tertiary structures are mainly governed by electrostatic forces, specifically hydrogen bonding.[4] The determination of RNA structure by experimental means, such as X-ray crystallography and cryogenic electron microscopy, is challenging and has not kept pace with emerging sequence information, particularly in cases of RNAs with varying modification states. Obtaining accurate values for the thermodynamic stabilities of individual hydrogen bonds would aid in the ability to predict secondary and tertiary structures of a given RNA by computational methods. Such approaches would enhance our understanding of how various modified RNAs function.[5−7]

There are currently over 150 known RNA modifications.[8] There is a wide range of nucleoside modifications that alter the sugar or base moieties, from simple methylations to more complex additions or isomerizations. Some modifications to the nucleobases are illustrated in Figure 1. While most studies over the past 70 years have focused on characterizing the chemical compositions and locations of these RNA modifications, less effort has been made to understand the physical contributions of protonation state and pK_a values within RNA systems and structures. By determining the effects of modification on the physical characteristics of the nucleobase, one can infer the impact on downstream folding and three-dimensional structures of RNAs.

Because of the challenges of site-specific incorporation of modifications into RNA,[9] determining their impacts within

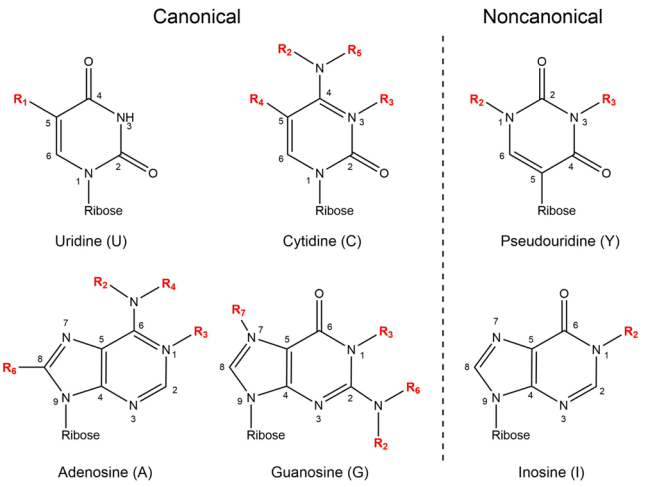
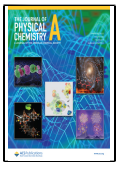


Figure 1. Structures of four canonical and two noncanonical nucleosides. The structures of pyrimidines (uridine, cytidine, and pseudouridine) and purines (adenosine, guanosine, and inosine) with nucleobase numbering schemes are shown. The modification positions are indicated with the R1 to R7 functional groups.

Received: December 28, 2021
Revised: February 12, 2022
Published: February 24, 2022



varying sequence and structural contexts is even more daunting. In the absence of experimentally determined thermodynamic parameters, it has been shown that by using the pK_a value of an acceptor and donor pair, it is possible to estimate the hydrogen bond strength.[4,10−12] This principle has been applied to nucleosides, which base pair through hydrogen bonding in nucleic acids. In certain cases, greater differences in pK_a values correlated with weaker base pairing between complementary nucleosides.[11]

Modifications to the nucleobase have been observed to cause shifts in the pK_a values (i.e., protonation/deprotonation states of the nucleobases),[12] which then affects base-pairing preferences. For example, deamination of adenosine to form inosine results in a base-pair preference for cytidine rather than the canonical uridine.[13] Any modifications to RNA bases that result in changes in pK_a values will impact RNA−RNA interactions. Therefore, our goal was to determine how various modifications alter the pK_a values of the nucleobase portions of the nucleosides, which can then be applied to studies with higher-order RNA structures.

Protonated nucleobases occur at important locations in RNA tertiary structures, including regions containing non-canonical base pairs. Protonated nucleobases have roles in ribozyme catalysis,[14,15] peptide-bond formation in ribosomes,[16] and tRNA structure formation.[17] Adenine is one of the most well-studied protonated nucleobases and has been observed in the protonated form when mismatched with cytosine in both DNA and RNA,[18−21] as well as in A-G base pairs.[22] Cytosines are protonated in the four-stranded i-motif structure, which involves two C−C$^+$ base pairs.[23]

Modifications in RNA (i.e., tRNA, rRNA, and mRNA) are highly prevalent.[24,25] Such chemical alterations to the nucleosides have been shown to induce pK_a shifts, either directly or indirectly. For example, purine methylation at N1 or N7 leads to a direct change in the nucleoside pK_a values to favor the protonated forms of both 1-methyladenosine (m1A) and 7-methylguanosine (m7G; please note that all abbreviations used follow the short name convention from MODOMICS[8]).[26,27] tRNA contains many examples of nucleoside modifications with altered pK_a values such as protonated m7G46, which participates in a triple-base interaction, and m1A58, which plays a role in T-loop interactions.[17,28] In mRNA, m1A modifications have been suggested to play roles in stabilizing alternative translation start sites.[29−31]

Indirect effects may occur when nucleobase modifications enhance protonation of an unmodified base-pairing partner. Examples include methylation of U2552 in rRNA, which promotes protonation of C2556 and formation of a more stable base pair compared to the unmodified, neutral counterpart.[32] Methylation of O6 in guanosine enhances base-pair formation with protonated cytidine in duplex DNA.[33,34] Beyond these few examples, there is a multitude of modifications in nucleic acids with enigmatic functions.

Experimental methods to determine the pK_a values for organic compounds,[35] such as nucleobases and nucleosides,[36] include potentiometry and nuclear magnetic resonance (NMR), fluorescence, and UV−visible spectroscopies.[35] Contemporary pK_a measurements of modifications in varying contexts (nucleobases, nucleosides, nucleotides, oligomers, etc.) have employed NMR spectroscopy, in which peaks for the HA and A$^-$ species are used to determine the pK_a values.[35,37] The experimental detection methods rely on precise pH titrations to determine the ratio of protonated and deprotonated species, which can be challenging.[35] UV−visible spectroscopy has also been used for pK_a determinations of nucleobases, with differing spectra for the protonated and deprotonated forms.[35] The four canonical nucleobases and nucleosides of RNA have unique spectra, and the extent of change in their UV spectra as a function of pH (i.e., changes in the ratio of HA and A$^-$ species) is also unique.[37] The pK_a values of the four canonical nucleosides have been studied extensively under various experimental conditions with good agreement of the values.[38] In contrast, less than 25% of the known modified nucleosides have experimentally determined pK_a values, mainly because of the limitations in isolation methods or challenges in synthesizing authentic modified nucleoside samples. Many of the modifications require specialized isolation or synthetic procedures because of their varying functional groups and physical properties.[36]

Computational chemistry allows for investigation of pK_a values at the molecular level, with good accuracy.[39−43] Previous theoretical studies to determine pK_a values used continuum solvation to account for the effects of the aqueous environment. More recent work demonstrated that adding explicit water molecules to a continuum solvation model yielded better approximations to experimental pK_a calculations.[44−48] Thapa et al. applied a combination of explicit water molecules with continuum solvation to model the pK_a values of methyl-substituted nucleobases and found that by using four water molecules located in hydrogen-bonding positions around the methyl-substituted nucleobase, one could obtain good agreement with experimental pK_a values.[49] Thapa's work also showed that using a B3LYP density functional with a 6-31+G(d,p) basis set gave good agreement with experimentally determined pK_a values of nucleosides.[49] In this current study, we used a B3LYP density functional with a 6-31+G(d,p) basis set to determine pK_a values for the modified nucleobases. In addition, we used a set of explicit waters at all of the hydrogen-bonding positions in the first solvation shell along with continuum solvation to avoid bias when selecting the locations of water molecules for differing nucleobases such as guanine and uracil. By comparing our calculated theoretical pK_a values to experimentally determined pK_a values, we were able to generate a linear, two-parameter adjustment to the data and obtain previously undetermined pK_a values for a range of modified nucleobases.

■ COMPUTATIONAL METHODS

The free energies of nucleobases were calculated using the Gaussian suite of programs.[50] The method used was based on a previous study by Thapa et al.[49] in which the B3LYP density functional with the 6-31+G(d,p) basis set was used for pK_a calculations. Our primary goal was to determine the pK_a values of modified nucleosides using theoretical approaches; however, it is difficult to model nucleosides or nucleotides with high accuracy in complex biological environments. Previous work has shown that reliable pK_a values can be obtained by substituting the sugar moiety with a methyl group. This approach was used in the current study to minimize the computational cost.

Modified nucleobases exist in a solvated environment, therefore a physiological environment must be replicated as best as possible to obtain biologically relevant results. To balance accuracy with computational costs, an implicit−explicit solvation system was used. An implicit solvation method uses a polarizable continuum surrounding the molecule to model the

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

effects of solvation. In the implicit−explicit approach, individual waters were placed at specified positions around the molecular structure, and the entire system was surrounded by an implicit solvation model. Previous work showed that increasing the number of explicit waters in the first solvation shell improved the agreement with experimentally determined pK_a values.[47] The implicit−explicit solvation approach was used for methyl-substituted nucleosides, and better agreement was found as the number of explicit waters increased from one to four.[49]

In this study, we recognize the importance of using an appropriate number of water molecules in the first solvation shell of the various unmodified and modified nucleobases. Explicit water molecules were placed at every hydrogen-bonding position,[45−49] and the SMD implicit solvation method of Marenich, Cramer, and Truhlar[51] was used to model the bulk water in the environment. Because the number of hydrogen-bonding positions depended on the nature of the modified nucleobases, the number of explicit water molecules varied. For example, 1,*N*4-dimethyl cytidine (*N*4-methyl cytidine for the corresponding nucleoside) allowed for four explicit water molecules (Figure 2). The specific locations and orientations of the waters used for each nucleobase structure are given in the Supporting Information (SI).

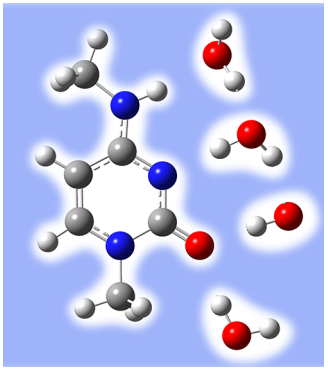


Figure 2. Example of the implicit−explicit solvation system that was used in the computational study of 1,*N*4-dimethyl cytidine. The blue field represents the implicit solvation field with explicit waters placed where hydrogen bonds can occur with the structure.

To find the lowest energy structures for the neutral form of each modified nucleobase, various combinations of hydrogen-bonding patterns were optimized and the lowest energy structures were selected for protonation/deprotonation. For instances in which modifications could impede the number of interacting water molecules, multiple combinations of molecular arrangements were tested. The maximum number of waters that could fit into the hydrogen-bonding positions in the first solvation shell were used. If multiple conformers were present, the structure containing the lowest total energy was selected (i.e., the syn over the anti conformation of m6A). Once the lowest energy neutral species was determined, protonated or deprotonated species were fully optimized with the same locations and hydrogen-bonding patterns of the explicit water molecules.

Theoretical methods for determining pK_a values have been extensively reviewed and studied.[40−43] The free energy difference of the neutral and ionic species can be used to determine the pK_a for protonation or deprotonation. For deprotonation of a neutral, AH → A^- + H^+, the free energy for the acid dissociation is determined by using eq 1.

$$\Delta G_{\text{dissoc(aq)}} = G_{\text{anion(aq)}} + G_{\text{proton(aq)}} - G_{\text{neutral(aq)}} \quad (1)$$

For protonation of a neutral, the corresponding free energy for the acid dissociation, AH^+ → A + H^+, is determined by using eq 2.

$$\Delta G_{\text{dissoc(aq)}} = G_{\text{neutral(aq)}} + G_{\text{proton(aq)}} - G_{\text{cation(aq)}} \quad (2)$$

The experimental estimate of the standard Gibbs free energy of a proton in water is −270.297 kcal/mol ($G_{\text{proton(aq)}}$), which is the Gibbs free energy of a gas phase proton, plus the change in standard state from 1 atm to 1 M, and the aqueous solvation free energy of a proton (−265.9 kcal/mol).[52,53] The pK_a value is obtained from the free energy for acid dissociation, $\Delta G_{\text{dissoc(aq)}}$, as shown in eq 3.

$$pK_a = \frac{\Delta G_{\text{dissoc(aq)}}}{2.303RT} \quad (3)$$

After the theoretical pK_a values were determined, a linear adjustment factor based on the line of best fit was derived for pyrimidines and purines by comparing the calculated values to those that had experimentally determined pK_a values. The adjusted p$K_a$values for pyrimidines and purines were determined by factoring calculated pK_a values into eqs 4 and 5, respectively.

$$\text{adjusted p}K_a = (\text{calculated p}K_a + 0.05)/1.15 \quad (4)$$

$$\text{adjusted p}K_a = (\text{calculated p}K_a - 0.25)/1.16 \quad (5)$$

RESULTS AND DISCUSSION

Determination of Nucleobase pK_a Values. Previous work demonstrated that applying the B3LYP/6-31+G(d,p) level of theory in conjunction with an implicit−explicit solvation method provided pK_a values of methyl-substituted nucleobases[49] that corresponded well with the experimental values for complete nucleosides. Although the calculations in the present study were done on methylated nucleobases, the names of the nucleosides are used in our comparisons to the experimental values. Because of structural differences of modifications, differing numbers of water molecules were used for each nucleobase, and each structure had a unique number of potential hydrogen-bonding interactions. The maximum number of water molecules for each structure was used, placing each water at potential hydrogen-bonding locations. By applying these methods to the methyl-substituted versions of the modified nucleosides, the pK_a values were calculated for 33 modified methyl-substituted nucleobases, 22 of which have known pK_a values for the corresponding nucleosides.

By comparing the experimental and theoretical pK_a values of the known modified pyrimidines and purines, a linear adjustment for the calculated results was applied to account for possible biases that arise from using the combination of theoretical parameters. All experimental pK_a values were obtained from the literature for nucleosides, except for three deoxynucleosides (m5C, hm5C, and f5C). Because of their different chemical and electronic structures (e.g., one vs two rings), different adjustment parameters were applied for pyrimidines and purines. In our pyrimidine and purine data sets, the mean absolute errors (MAEs) obtained were 0.4 and

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

0.6 pK_a units, respectively. The typical observed standard deviation in the theoretical pK_a calculations was 1.2 pK_a units before adjustment. We note that there are a combination of challenges in obtaining accurate experimental pK_a values, such as obtaining pure samples, and low accuracy of instrumentation,[36] which can negatively impact the correlation between experimental and theoretical pK_a values. In addition, there are small, intrinsic errors in the calculation because of the use of a limited number of stationary, explicit water molecules, the particular basis set, the density functional theory, or the polarizable continuum model (PCM) employed.

Comparisons of our adjusted theoretical data to the experimental values of the modified nucleosides from the literature show similar trends in changes to the pK_a values as a result of modification. For a few of the studied purines, we did not observe the expected number of explicit waters after protonation/deprotonation. This issue became evident during geometry optimization when water molecules were observed to drift from the expected positions. For example, if conventional hydrogen-bonding principles are followed, we would expect m2G to coordinate six water molecules; however, during optimization one water would often drift away. In order to maintain consistency in our focus of only the first solvation sphere, one water molecule was removed and the structures were reoptimized (optimized structure of m2G is shown in the SI).

Pyrimidines. A total of 16 different methyl-substituted pyrimidine nucleobases were examined, of which four lack experimentally determined pK_a values. The N3 position is the primary physiological site of deprotonation in uridine modifications and the protonation site in cytidine modifications. One uridine modification, pseudouridine (Y), contains two sites of deprotonation at the N1 and N3 positions (Figure 1). By plotting the experimental against the calculated values (Figure 3), we found a small systematic 0.05 pK_a unit shift for

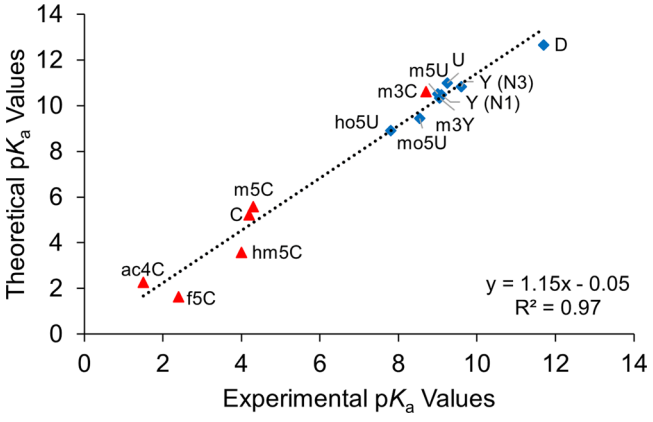


Figure 3. Plot and linear regression fit for the theoretical and experimental pK_a values of the pyrimidines are shown. Uridine and its modifications are indicated with blue diamonds, and cytidine and its modifications are marked by red triangles.

the theoretical data. A two-parameter adjustment for pyrimidine modifications was used and a linear fit of $1.15x - 0.05$ with an R^2 value of 0.97 was obtained with a MAE of 0.4 pK_a units, indicating good agreement between the experimental and adjusted values (Figure 3). As an example, the calculated pK_a value for 1-methyluracil (the methylated nucleobase representing uridine) is 10.98, but a value of 9.6 was obtained after applying the adjustment factor, which is comparable to the literature value of 9.25 for the parent nucleoside uridine.[36]

After the adjustment is applied, there is good agreement between the experimental and calculated pK_a values (Table 1) for modified uridines (structures are shown in Figure 1 and SI). A MAE of 0.2 pK_a units is observed for the uridines. Due to our model constraints, the structures used for the pK_a calculations for m5U and m1Y are the same. Methylation at C5 (e.g., m5U) has the smallest effect on the pK_a value for uridine deprotonation at the N3 position, whereas oxidation or hydrogenation results in larger shifts in the pK_a values with the following trend: ho5U < mo5U < m5U (m1Y) < Y ≈ U ≪ D < m5D (range 7.8−11.6). The greatest difference observed between the adjusted and experimental pK_a values is 0.6 pK_a units for the N3 position of dihydrouridine (D).

Of the studied modified pyrimidine nucleobases, the addition of a hydroxyl group at position C5 (ho5U) leads to the largest decrease in pK_a value (ΔpK_a 1.8 compared to U; all ΔpK_a values are reported as absolute values) for the uridine series. The mo5U nucleobase, which has been studied primarily in the context of anticodon loops in tRNA, forms from enzymatic conversion of the hydroxyl at the C5 position (ho5U) to a methoxy group (mo5U).[59] This methylation of the OH group slightly increases the pK_a value (8.3). Consistent with experimental observations, methylation at the C5 position (m5U) results in only a small decrease in the N3 pK_a value compared to the parent uridine (ΔpK_a 0.4). In contrast, saturation of the C5−C6 double bond to give D leads to a significantly higher pK_a value for the N3 position when compared to uridine (ΔpK_a 1.5). Absence of the double bond and loss of aromaticity may contribute to the increased pK_a value, and good agreement is observed between the experimental and adjusted values.[60] Methylation of D (m5D) further increases the pK_a value (ΔpK_a 0.5 compared to D).

Pseudouridine (Y) is unique compared to the other uridine modifications because it contains two deprotonatable sites (N1 and N3) as a result of the isomerization reaction of U, replacing the N1−C1′ linkage with a C5−C1′ glycosidic bond to generate the C-glycoside.[61] This modification allows for formation of an additional hydrogen bond at the N1 position. The pK_a value of the N3 position in Y is similar to that of U. Experimental observations showed that the N1 position of Y has a lower pK_a value (9.1) than N3 (9.6). The adjusted pK_a values from our computations follow the same trend with the N3 value being higher (9.5) than the N1 (9.2). Methylation of Y at the N1 or N3 position produces 1-methylpseudouridine (m1Y) or 3-methylpseudouridine (m3Y), respectively. Both methylation events have minimal effects on the adjusted pK_a values (N3 of m1Y and N1 of m3Y).

Modifications on U were shown to cause both increases and decreases in the pK_a values. These altered pK_a values due to modifications could alter the pairing interactions with adenosine.[12] Two oxygen-containing modified nucleobases (mo5U and ho5U) have lower pK_a values for N3 relative to U. These altered values may impact the nucleotide roles as wobble-pairing partners in tRNA.[62]

A modest change in the N3 pK_a value of Y compared to U is unlikely to impact the stability of Y-A pairs. However, the additional N1H of Y has been observed to form a water bridge with the phosphate backbone in several RNAs.[63−66] The presence of these water bridges correlates with increased stability of certain pseudouridylated RNAs when the

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

Table 1. Experimental and Theoretical pK_a Values for Modified Uridines and 1-Methyluracils

name	abbreviation	measured position	experimental pK_a[a]	adjusted theoretical pK_a
5-hydroxyuridine	ho5U	N3	7.8[54]	7.8
5-methoxyuridine	mo5U	N3	8.54[55]	8.3
3-methylpseudouridine	m3Y	N1	9.05[56]	9.0
5-methyluridine	m5U	N3	9[57]	9.2
1-methylpseudouridine	m1Y	N3		9.2
pseudouridine	Y	N1	9.1[58]	9.2
		N3	9.6[58]	9.5
uridine	U	N3	9.25[36]	9.6
dihydrouridine	D	N3	11.7[36]	11.1
5-methyldihydrouridine	m5D	N3		11.6

[a]The pK_a values were determined by NMR spectroscopy[36,58] and UV titration.[36,54,57]

Table 2. Experimental and Theoretical pK_a Values for Modified Cytidines and 1-Methylcytosines

name	abbreviation	measured position	experimental pK_a[a]	adjusted theoretical pK_a
5-formylcytidine	f5C	N3	2.4[71]	1.5
*N*4-acetylcytidine	ac4C	N3	1.5[72]	2.0
5-hydroxymethylcytidine	hm5C	N3	4[71]	3.2
*N*4,*N*4-dimethylcytidine	m4,4C	N3		4.1
*N*4-methylcytidine	m4C	N3		4.4
cytidine	C	N3	4.2[73]	4.6
5-methylcytidine	m5C	N3	4.3[74]	4.9
3-methylcytidine	m3C	N4	8.7[75]	9.3

[a]The pK_a values were determined by NMR spectroscopy[71] and UV titration.[73−75]

modification is in duplex regions.[67,68] Methylation of the N3 position on pseudouridine has minimal impact on the N1 pK_a value, but reduces the number of protonatable sites on Y. In one known example of Y methylation, the N3 modification has been shown to stabilize RNA structures when present in hairpin loops, but the added stability is not correlated with a change in the pK_a value.[69] Stabilization from the methyl group could be important for long-range interactions in the ribosome.

In contrast to uridine, cytidines generally undergo protonation at N3, and the modifications, except m5C, display decreases in the calculated pK_a values for this position (Table 2). A MAE of 0.6 pK_a units is observed for the cytidines. Methylation on the C5 position (m5C) is observed in both DNA and RNA, and both hm5C and f5C are intermediates in the demethylation of m5C.[70] Experimental pK_a values for the RNA nucleosides are lacking for hm5C and f5C, thus the deoxynucleoside pK_a values for these two modifications and m5C were used for the linear regression and general comparison. The adjusted pK_a values for the cytidine modifications examined range from 1.5 to 4.9 with the following trend: f5C < ac4C < hm5C < m4,4C < m4C < C < m5C. The largest error is observed for f5C, which required more explicit water molecules in the first solvation shell leading to more water orientations.

For cytidine-based modifications, most contribute to a decrease in the pK_a value for N3. The largest decrease in pK_a value (ΔpK_a 3.1 compared to C) is attributed to the addition of a formyl group at C5 to generate f5C. Nucleobases m4,4C and ac4C containing an N4 dimethyl or acetyl, respectively, also display increased acidity at the N3 position. The hydroxymethyl group at C5 (hm5C) has a similar effect of decreasing the pK_a value (ΔpK_a 1.4 compared to C) at N3. In contrast, single methylation at the N4 position has a negligible impact on the pK_a value, with a decrease of 0.2 pK_a units compared to C. Consistent with a similar modification in U (m5U), there is no significant change in the pK_a value at the N3 of C upon methylation at C5 (m5C). Methylation at the N3 position (m3C), which has been shown to favor the imino tautomer, is protonated at the N4 position rather than N3 (Figure 4).[76]

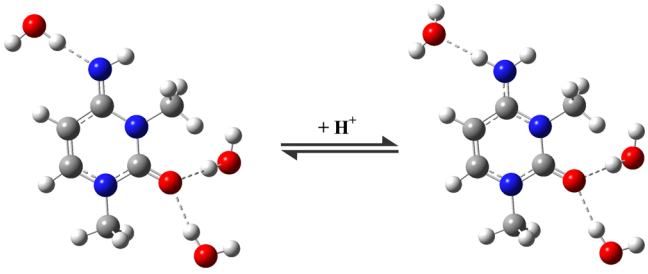


Figure 4. Structure of m3C. Protonation at the N4 position of m3C results in a delocalized positive charge.

Acetylation of the N4 of cytidine (ac4C) has been observed to increase translation efficiency when present in the wobble position of mRNA codons,[77] as well as increase the thermal stability of G-ac4C base pairs in DNA.[78] Methylation at the N3 position (m3C) results in an electronic rearrangement of the nucleobase upon protonation. This modification leads to a delocalized positive charge, following protonation at N4 (Figure 4). As a result, m3C can no longer participate in canonical Watson−Crick (WC) interactions. Similarly, dimethylation at the N4 position (m4,4C) would disrupt conventional WC base pairs. Despite a loss of WC pairing ability, these modifications still have the potential to impact noncanonical pairing within RNA tertiary structures.

Demethylation of m5C to C occurs through two oxidized intermediates, hm5C and f5C, in which the N3 pK_a values are lower compared to C. The decrease in pK_a value for f5C is

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

likely due to its ability to hydrogen bond with the N4 amino proton.[79] We would predict that these changes in pK_a values would result in altered base-pairing interactions with G. In fact, this has been observed with both hm5C and f5C, in which these modifications increase the melting temperatures of RNA sequences when participating in base pairing.[80,81]

Purines. Of the 17 purines studied (with 35 protonation/deprotonation sites), only nine (specifically 12 protonation/deprotonation sites) had previously reported pK_a values. For purine nucleobases, the primary sites of protonation and deprotonation are the N1 and N7 positions. Adenosines are protonated at N1, whereas guanosines are deprotonated at N1. Both purines can be protonated at the N7 position. Similar to our pyrimidine data set, application of the adjustment factor results in good agreement with the experimental pK_a values. The known experimental pK_a values were plotted against the calculated values of purines to give a linear fit of $1.16x + 0.25$, with an R^2 value of 0.94 (Figure 5). The high correlation value

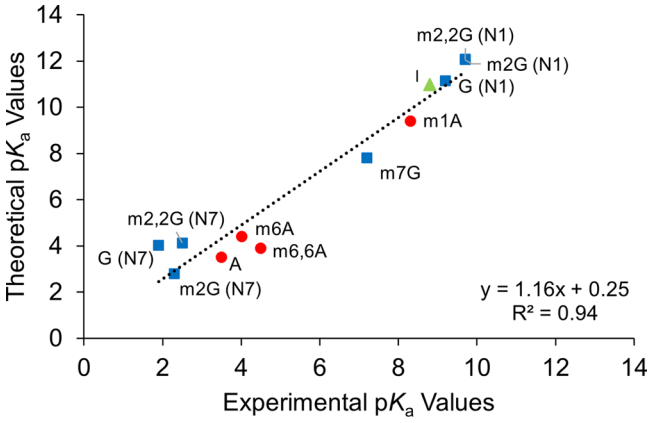


Figure 5. Plot and linear regression fit for the theoretical and experimental values of purines are shown. Values for adenosine and its modifications are shown using red circles; guanosine and its modifications are shown using blue squares. A green triangle is used for inosine.

indicates that our purine data set is in good agreement with available experimental values (MAE of 0.6). There are two ranges of adjusted pK_a values at the N1 position observed in the purine set, −1.7−5.3 for adenosines (Table 3) and 3.0−10.2 for guanosines (Table 4), with the latter having a higher degree of agreement with the experimental values.

Of all the nucleobases studied here, adenosine modifications have the poorest agreement between the adjusted theoretical and experimental pK_a values; however, only four experimental pK_a values for adenosines were found in the literature. Adenosine has a diversity of modifications with a wide range of pK_a values (Table 3). Adenosine and its modifications are primarily protonated at the N1 position, with the following increasing order of pK_a values: f6A ≪ m8A ≈ hm6A < A < m6,6A < m6A < m2,8A < m2A (range of −1.7 to 5.3). The pK_a values for the N7 position of adenosine follow the trend of: hm6A < A < f6A < m1A < m6,6A < m2A ≈ m6A < m8A ≪ m2,8A (range of 0.3−4.1).

The modification with the lowest pK_a value, f6A, forms from the oxidation of m6A by fat mass and obesity-associated protein (FTO) using hm6A as an intermediate.[84] (Figure 6) The f6A nucleobase shows the greatest decrease in pK_a at the N1 position (ΔpK_a of 4.5 compared to A). Additionally, the N6 position, which contains the carboxamide group, has a considerably lower pK_a value (8.7) compared to that of A (17.0). In contrast, the N7 position of f6A displays a modest increase in pK_a value (ΔpK_a of 0.2). The intermediate, hm6A, has a lower pK_a value at N1 (2.1), without a large change in pK_a value at N6 (17.5).

Methylation of adenosine at the C8 position (m8A) leads to changes in the pK_a values at the N1 and N7 positions (decreased by 0.8 and increased by 1.7 units, respectively). In contrast, methylation at the N6 (m6A) results in a pK_a increase at both N1 (ΔpK_a 0.8) and N7 (ΔpK_a 1.4). Dimethylation at the N6 position (m6,6A) leads to a modest increase in the pK_a value (ΔpK_a 0.3 compared to A) at N1 and a larger increase at N7 (ΔpK_a 0.8). Methylation at the C2 position (m2A) has the largest increase in pK_a for the N1 position (ΔpK_a 2.5) and an increase at the N7 position (ΔpK_a 1.3). Additional methylation of either m2A or m8A to form dimethylated m2,8A, decreases the acidity at N7 to give a relatively high pK_a value (4.1), while the N1 position showed a similar increase as was observed in m2A (ΔpK_a 2.0 compared to A). Methylation at N1 in adenosine (m1A) results in an electronic configuration that has been shown to favor protonation at N6 of the nucleobase under physiological conditions (pK_a 7.9).[26] The electronic rearrangement in m1A also slightly impacts the pK_a value of the neighboring N7 position, which is 0.5 units higher than that of A.

As discussed above, methylation on adenosine generally results in an increase in the pK_a values, with one exception, the N1 position of m8A. In contrast, the N1 position becomes more acidic when converted from hm6A to f6A and is therefore unlikely to be protonated. This conversion process of hm6A to f6A is important for demethylation at N6. The *N*6-methyl group is removed through oxidation by FTO to form adenosine (Figure 6). In the m6A binding pocket of FTO, both the N1 and N6 of m6A are within hydrogen-bonding distance of active site amino acids.[85] Since the pK_a values distinctly change at both the N1 and N6 positions between the different modifications in the demethylation pathway they could potentially serve a specific mechanistic role in the demethylation of m6A.[85,86]

Adenosine methylated at the N6 position (m6A) exists in both anti and syn conformations, in which only the anti conformer can form WC base pairs. We found that the syn conformer has a lower total energy and an N1 pK_a value of 3.6 (Table 3). A unique situation arises for this modification in which the N1 pK_a value of 4.1 for the anti conformer has better agreement with the experimental value of 4.01 (SI). For the anti conformer of m6A, the increased N1 pK_a value could lead to altered m6A-U base-pair stability (compared to an A-U pair). It has been shown experimentally that there are variations in stability for sequences containing m6A,[87,88] perhaps a reflection of this conformational flexibility.

The m1A nucleobase is unique in that it required a modification in the placement of explicit waters for the calculation. The m1A is one of two systems studied here that is positively charged at physiological pH. Attempts to place a water molecule at the N6 position in the neutral structure of m1A failed because of poor hydrogen bonding. Therefore, a water bridge was introduced to properly coordinate the positions of the remaining water molecules to the nucleobase under neutral conditions (Figure 7). Protonation occurs at the N6 position as shown in Figure 7. The error on the adjusted pK_a value for m1A is in line with the MAE for purines.

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

Table 3. Experimental and Theoretical pK_a Values for Modified Adenosine and 9-Methyladenines

name	abbreviation	measured position	experimental pK_a[a]	adjusted theoretical pK_a
*N*6-formyladenosine	f6A	N1		−1.7
		N6		8.7
		N7		0.8
8-methyladenosine	m8A	N1		2.0
		N7		2.3
*N*6-hydroxymethyladenosine	hm6A	N1		2.1
		N6		17.5
		N7		0.3
adenosine	A	N1	3.5[82]	2.8
		N6		17.0
		N7		0.6
*N*6,*N*6-dimethyladenosine	m6,6A	N1	4.50[83]	3.1
		N7		1.4
*N*6-methyladenosine	m6A	N1	4.01[83]	3.6
		N6		18.3
		N7		2.0
2,8-dimethyladenosine	m2,8A	N1		4.8
		N7		4.1
2-methyladenosine	m2A	N1		5.3
		N7		1.9
1-methyladenosine	m1A	N6	8.3[26]	7.9
		N7		1.1

[a]The pK_a values of were determined by UV titration[82] and potentiometric titrations.[83]

Table 4. Experimental and Theoretical pK_a Values for Modified Guanosine and 9-Methylguanosines

name	abbreviation	measured position	experimental pK_a[a]	adjusted theoretical pK_a
*O*6-methylguanosine	m6G	N1		3.0
		N7		2.9
*N*7-methylguanosine	m7G	N1	7.2[90]	6.5
inosine	I	N1	8.8[91]	9.2
		N7		0.5
guanosine	G	N1	9.2[92]	9.4
		N7	1.90[92]	3.3
*N*2,*N*2-dimethylguanosine	m2,2G	N1	9.7[93]	10.2
		N7	2.5[93]	3.3
*N*2-methylguanosine	m2G	N1	9.7[93]	10.2
		N7	2.3[93]	2.2
1-methylinosine	m1I	N7		0.8
1-methylguanosine	m1G	N7		3.0

[a]The pK_a values were determined by electrophoretic mobility[93] and UV titration.[90,91]

Although the N1 methylation on A prevents canonical WC base pairing, its prevalence in mRNA and tRNA may enhance tertiary interactions (e.g., Hoogsteen pairing with the N6 and N7 positions) within the folded RNA.[89]

The calculated guanosine pK_a values are in good agreement with the experimental values. Fewer guanosine modifications were studied here because of the higher nature of complexity of guanosine modifications (e.g., wyosine, which contains a tricyclic core structure). We limited our study to various sites of methylation on guanosine (Table 4). Inosine, the deaminated form of guanosine, was also studied. The observed trend is that the position of methylation has a bigger effect on the N1 pK_a than the N7 value. The trends in the pK_a values of the N1 position (deprotonation) are m6G ≪ m7G ≪ I < G < m2,2G = m2G (range of 3.0 to 10.2). Modifications result in changes to the N7 pK_a values with the following trend: I < m1I ≪ m2G < m6G ≈ m1G < G = m2,2G (range of 0.5 to 3.3).

The biggest change in the N1 pK_a value (ΔpK_a 6.4) is observed for the nucleobase methylated at the O6 position (m6G). The modified base exists in the enol form, which has a ketimine functionality through N1, similar to A (the N1 pK_a values are 3.0 and 2.8 for m6G and A, respectively). However, the N7 pK_a value of m6G is similar to that of G (2.9 vs 3.3, respectively).

Methylation at O6 (m6G) has been observed as DNA damage.[94] Under physiological conditions, this modified nucleobase is unable to form canonical WC base pairs with C due to the presence of the ketimine. At low pH, protonation of the ketimine promotes m6G−C interactions.[33] This base pair is similar to previously observed A−C mismatches in RNA in which the pair is stabilized by protonation at low pH.[95] In DNA systems, m6G has been observed to pair with thymine (T) at pH 7.[33] This dual ability of m6G to pair with C or T could have varying biological consequences such as errors in

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

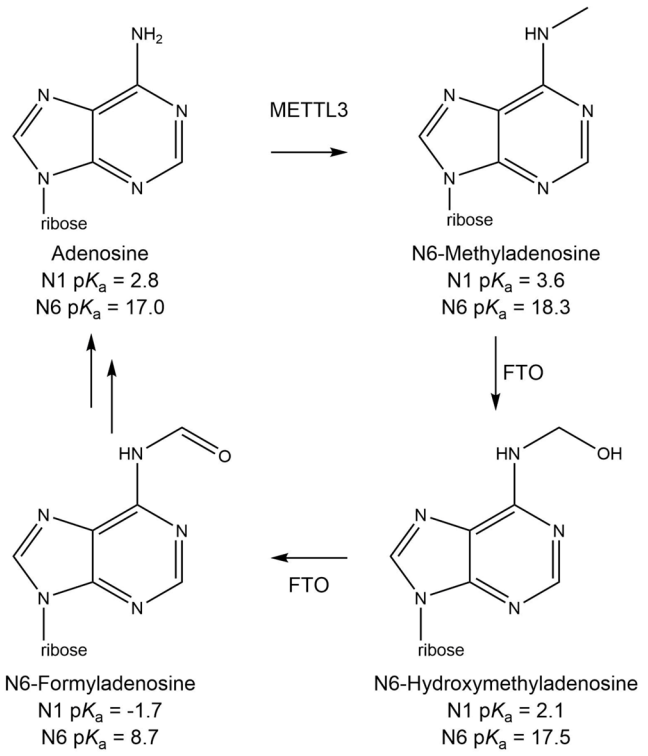


Figure 6. Pathway for the demethylation cycle of m6A. Once hm6A is converted to f6A the formyl group is released as formic acid following further oxidation. Each modification's relevant pK_a value is listed below the structure.

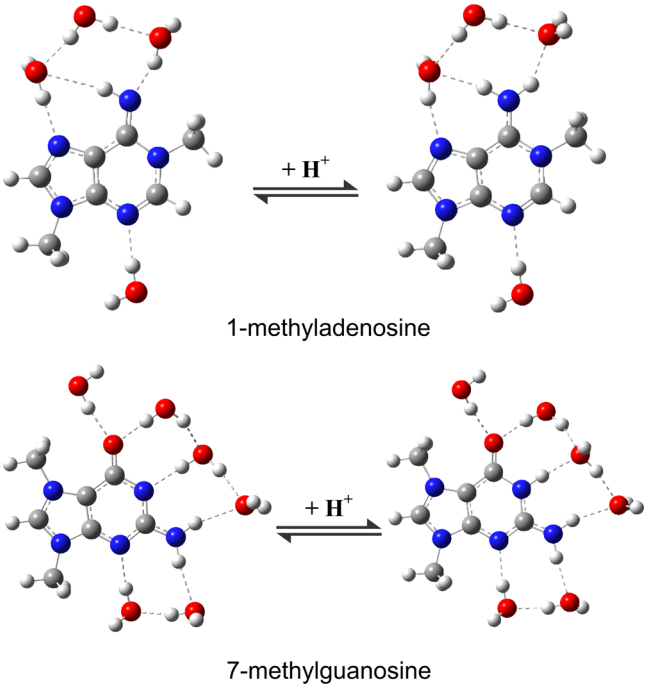


Figure 7. Two nucleobases have been reported to be protonated under physiological conditions: 1-methyladenosine (top) and 7-methylguanosine (bottom). The neutral structures are shown on the left, while the protonated structures are shown on the right.

replication or, in the case of RNA, alternative structure formation.

Methylation at the N1 position to give m1G has minimal impact on the N7 pK_a value (pK_a value of 3.0) compared to the parent G. In contrast, methylation at the N2 position forms m2G, which displays an increased pK_a value at N1 (ΔpK_a 0.8) and a decreased value at N7 (ΔpK_a 1.1). Further methylation of m2G (m2,2G) leads to an increase in the pK_a value at the N7 position (ΔpK_a 1.1 for singly vs doubly methylated G). If the N7 position is methylated (m7G), even more substantial changes are observed for the N1 pK_a value (ΔpK_a 2.9 compared to G). As a consequence, m7G is one of a few nucleobases that is protonated at physiological pH (structures of neutral and protonated m7G are shown in Figure 7).[96]

The *N*7-methylguanosine (m7G) has several interesting roles such as participating in a triple-base interaction (e.g., position 46 in tRNA) with protonation at N1 under physiological conditions.[28] The protonation of the m7G nucleobase also enhances stacking with aromatic residues in proteins.[97] For other methylated nucleobases such as m2G, the modification can lead to a conformational change as mentioned earlier with m6A (here the rotamers are referred to as cis/trans; see structures in SI). Minimum energy calculations were performed on both rotamers, with the trans having lower energy.

Consistent with experimental values, an increase in the pK_a of m2G is observed. Despite the addition of a methyl group and increase in the N1 pK_a value, m2G has been reported to have a negligible effect on the stability of RNA duplexes with noncanonical secondary structures (adjacent mismatches).[98] Unlike *N*-methylation of A (m6A) in RNA duplexes, the cis versus trans positioning of the m2G methyl group within mismatch sites has been shown to be iso-energetic.[98,99] Further methylation at the N2 position to form m2,2G was shown to have a destabilizing effect on RNA duplexes when studied in the context of sheared m2,2G-A base pairs; however, this trend was not observed in larger, more complex RNA structures.[100] Dimethylation of N2 has been shown in tRNA to prevent alternative folding by restricting WC pairing with C.[101]

Inosine (I) is derived from adenosine in biological systems, but more closely resembles guanosine in structure. Inosine is formed from the deamination of adenosine. The adjusted pK_a for deprotonation of the N1 position of I (9.2) closely aligns with that of G (9.4). In contrast, the pK_a value of the N7 position of I (0.5) aligns more closely with the parent nucleoside A (0.6) than with G (3.3). Methylation of I at the N1 position results in 1-methylinosine (m1I). Because of the keto group at the C6 position, m1I does not undergo tautomerization and is, therefore, unable to rearrange electronically for protonation at the O6 position. The pK_a value of the N7 position (0.8) of m1I is slightly increased compared to I.

Inosine plays an important role when present at the wobble position in tRNA anticodons by allowing interactions with more than one codon.[102] The modified nucleobase's ability to pair with U, C, or A expands the possible number of anticodon−codon interactions. In contrast, the canonical nucleobase A only pairs with U at the same position.[102] The differences in pK_a values of the three I pairs are similar to the nonmodified counterparts, that is, I−A (6.4) vs A−U (6.8), I−C (4.6) vs G−C (4.8), and I−U (0.4) vs G−U (0.2). Consistent with this observation, the presence of I in DNA can cause U to C mutations in the complementary strand.[103] The I−C base-pair interaction has been shown in RNA systems to have a similar stability as G−C pairs.[104]

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

Similar to m2,2G, m1G is unable to participate in canonical WC pairing. Methylation at the N1 position exists as a post-transcriptional modification in RNA or as damage to DNA at specific sites.[105] At G37 in tRNA, *N*1-methylation was found to cause base flipping, therefore preventing a frameshifting event.[106,107] The formation of the tertiary amine in m1G disfavors protonation at the N1 position. In contrast to other examples presented in which methylation impacted pK_a values and base-pair stabilities, methylation at the N1 position has little impact on the nucleobase pK_a value, providing an example in which modification predominantly has a structural effect on the nucleic acid.

Previous reports have shown that the difference in pK_a values between an acceptor and donor pair can be used to estimate the strength of a hydrogen bond, referred to as the ΔpK_a rule.[10] This rule has been applied to predict stabilities in model nucleic acid systems,[11] in which a decrease in the ΔpK_a indicated stronger base pairing. We used the ΔpK_a rule with adjusted theoretical pK_a values and compared to published stability data. In one report of duplex RNAs containing m4C, there was a slight decrease in duplex stability when the modification was present.[108] This decrease in stability for G-m4C correlated with an observed increased ΔpK_a compared to G−C. In contrast, other studies showed an opposite correlation between oligonucleotide stabilities and ΔpK_a between donor and acceptor pairs. More specifically, duplexes containing hm5C and f5C modifications resulted in higher melting temperatures compared to C despite having a larger ΔpK_a.[80,81] Because of these different observed correlations, more work needs to be done to understand the relationship between ΔpK_a and nucleic acid stabilities, which are affected by the presence of multiple hydrogen bonds, base stacking, and nearest-neighbor effects.

CONCLUSIONS

Previous theoretical calculations established a reliable methodology using a combined implicit−explicit solvation system to determine the pK_a values for nucleobases. Here, we built on that method by using an expanded first solvation shell to calculate the pK_a values of 33 modified nucleobases. A total of 52 pK_a values were calculated, half of which have never been reported, with R^2 values of 0.97 and 0.94 for pyrimidines and purines, respectively. This marks the largest study of pK_a values of modified nucleobases, with a compilation of available experimental and theoretical pK_a data for the modifications. By applying an adjustment to our purine and pyrimidine pK_a data sets based on comparison of theoretical and experimental values of nucleosides, closer approximations to the experimental data were generated. The MAEs were 0.4 and 0.6 for pyrimidines and purines, respectively.

The model system using the first explicit solvation shell in combination with the B3LYP/6-31+G(d,p) level of theory provides reliable pK_a values, but with larger deviations between experimental and theoretical values at acidic compared to basic sites. Certain modifications were more challenging than others. For example, the waters weakly associated with the N6H of m1A in the neutral form, but formed stronger interactions after protonation. Additionally, larger and more complex structures required more computation time and optimizations would fail more frequently. Despite these challenges, the overall trends for relative pK_a values of the modified nucleobases are consistent with those from experiment, except in the cases of m6A/m6,6A and ac4C/f5C. Methylation of m6A to m6,6A was reported previously to increase the experimental pK_a value, but this is contradicted by our adjusted theoretical data. Reevaluation of the pK_a values of these nucleobases using identical experimental methods may lead to better correlations. Future work to determine theoretical pK_a values of more complex nucleobases using a combined explicit-implicit solvation system may have to account for more water molecules at strategic placements. Many of these complex nucleobases are lacking experimental values, therefore computation done in conjunction with experiment would provide more reliable adjusted pK_a values.

Modifications perturb the nucleobase pK_a values. Both the type and position of the modification influence the pK_a values compared to those of the unmodified counterparts. For example, monomethylations of G result in a large range in pK_a values depending on the modification sites. Similarly, our calculated values for the intermediates of the m6A demethylation process (i.e., hm6A and f6A) show a wide range of pK_a values. These results reveal that the N6 position of f6A (pK_a 8.7) can be deprotonated under physiological conditions, as observed experimentally for mo5U (pK_a 8.54).[55] The ΔpK_a value has been used to correlate helix stabilities and base-pairing preferences, but caution is needed when applying this rule. For example, we observe poor correlation with ΔpK_a and stability in several cases that we examined (e.g., N4-acetylation of C in G−C pairs;[78] N6-methylation of A in A−U pairs[88]).

To fully understand how modifications affect the stability of complex structures, other factors need to be considered in addition to the pK_a values. Some modifications, such as m2,2G, block WC base-pairing ability. As such, those modified nucleobases may participate in noncanonical interactions within larger structural contexts. Because of our decision to use methyl substitution in place of (deoxy)ribose on our nucleobase models, some methylated systems could not be studied such as *N*1-methylpseudouridine (m1Y) (i.e., the simplified structure used to model m1Y would be the same as m5U). Future work may therefore involve expanding the methyl substituent to differentiate the (deoxy)ribose attachment site from a modification. Subsequent studies on the effects of pK_a changes on structure will need to account for the nearest-neighbor effects in order to obtain a deeper understanding of how the pK_a values impact nucleic acid structure.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.1c10905.

Optimized structures and Cartesian coordinates of methyl-substituted nucleobases (PDF)

Details on calculations of pK_a values (XLSX)

AUTHOR INFORMATION

Corresponding Author

Christine S. Chow − *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0002-5538-4553; Email: cchow@wayne.edu

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529

Authors

Evan L. Jones − *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0001-7640-8780

Alan J. Mlotkowski − *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0002-7384-6432

Sebastien P. Hebert − *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States*

H. Bernhard Schlegel − *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0001-7114-2821

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.1c10905

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank John SantaLucia, Jr., and Bishnu Thapa for helpful discussions. This work was supported in part by grants from the National Science Foundation (CHE1856437 and CHE1904754). The Wayne State University High Performance Computing Services Department provided computational support. A.J.M. thanks Wayne State University for a Thomas C. Rumble Fellowship.

■ REFERENCES

(1) Thieffry, D.; Sarkar, S. Forty Years Under the Central Dogma. *Trends Biochem. Sci.* **1998**, *23*, 312−316.

(2) Woese, C. R. Translation: in Retrospect and Prospect. *RNA* **2001**, *7*, 1055−1067.

(3) Guil, S.; Esteller, M. RNA−RNA Interactions in Gene Regulation: the Coding and Noncoding Players. *Trends Biochem. Sci.* **2015**, *40*, 248−256.

(4) Šponer, J.; Leszczynski, J.; Hobza, P. Electronic Properties, Hydrogen Bonding, Stacking, and Cation Binding of DNA and RNA Bases. *Biopolymers* **2001**, *61*, 3−31.

(5) SantaLucia, J. Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-Neighbor Thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460−1465.

(6) Roost, C.; Lynch, S. R.; Batista, P. J.; Qu, K.; Chang, H. Y.; Kool, E. T. Structure and Thermodynamics of N6-Methyladenosine in RNA: a Spring-Loaded Base Modification. *J. Am. Chem. Soc.* **2015**, *137*, 2107−2115.

(7) Vaidyanathan, P. P.; AlSadhan, I.; Merriman, D. K.; Al-Hashimi, H. M.; Herschlag, D. Pseudouridine and N^6-Methyladenosine Modifications Weaken PUF Protein/RNA Interactions. *RNA* **2017**, *23*, 611−618.

(8) Boccaletto, P.; Machnicka, M. A.; Purta, E.; Piatkowski, P.; Baginski, B.; Wirecki, T. K.; de Crécy-Lagard, V.; Ross, R.; Limbach, P. A.; Kotter, A.; et al. MODOMICS: a Database of RNA Modification Pathways. 2017 Update. *Nucleic Acids Res.* **2018**, *46*, D303−D307.

(9) Chow, C. S.; Mahto, S. K.; Lamichhane, T. N. Combined Approaches to Site-Specific Modification of RNA. *ACS Chem. Biol.* **2008**, *3*, 30−37.

(10) Shan, S.-o.; Loh, S.; Herschlag, D. The Energetics of Hydrogen Bonds in Model Systems: Implications for Enzymatic Catalysis. *Science* **1996**, *272*, 97−101.

(11) Acharya, P.; Cheruku, P.; Chatterjee, S.; Acharya, S.; Chattopadhyaya, J. Measurement of Nucleobase pK_a Values in Model Mononucleotides Shows RNA-RNA Duplexes to be More Stable than DNA-DNA Duplexes. *J. Am. Chem. Soc.* **2004**, *126*, 2862−2869.

(12) Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. Calculation of pK_as in RNA: on the Structural Origins and Functional Roles of Protonated Nucleotides. *J. Mol. Biol.* **2007**, *366*, 1475−1496.

(13) Bass, B. L.; Weintraub, H. An Unwinding Activity that Covalently Modifies its Double-Stranded RNA Substrate. *Cell* **1988**, *55*, 1089−1098.

(14) Nakano, S.-i.; Chadalavada, D. M.; Bevilacqua, P. C. General Acid-Base Catalysis in the Mechanism of a Hepatitis Delta Virus Ribozyme. *Science* **2000**, *287*, 1493−1497.

(15) Shih, I.-h.; Been, M. D. Involvement of a Cytosine Side Chain in Proton Transfer in the Rate-Determining Step of Ribozyme Self-Cleavage. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1489−1494.

(16) Muth, G. W.; Ortoleva-Donnelly, L.; Strobel, S. A. A Single Adenosine with a Neutral pK_a in the Ribosomal Peptidyl Transferase Center. *Science* **2000**, *289*, 947−950.

(17) Holbrook, S. R.; Sussman, J. L.; Warrant, R. W.; Kim, S.-H. Crystal Structure of Yeast Phenylalanine Transfer RNA: II. Structural Features and Functional Implications. *J. Mol. Biol.* **1978**, *123*, 631−660.

(18) Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. Nitrogen-15-labeled Oligodeoxynucleotides. 3. Protonation of the Adenine N1 in the A·C and A·G Mispairs of the Duplexes {d[CG(^{15}N^1)-AGAATTCCCG]}$_2$ and {d[CGGGAATTC(^{15}N^1)ACG]}$_2$. *J. Am. Chem. Soc.* **1991**, *113*, 5486−5488.

(19) Legault, P.; Pardi, A. Unusual Dynamics and pK_a Shift at the Active Site of a Lead-Dependent Ribozyme. *J. Am. Chem. Soc.* **1997**, *119*, 6621−6628.

(20) Huppler, A.; Nikstad, L. J.; Allmann, A. M.; Brow, D. A.; Butcher, S. E. Metal Binding and Base Ionization in the U6 RNA Intramolecular Stem-Loop Structure. *Nat. Struct. Biol.* **2002**, *9*, 431−435.

(21) Siegfried, N. A.; O'Hare, B.; Bevilacqua, P. C. Driving Forces for Nucleic Acid pK_a Shifting in an A+·C Wobble: Effects of Helix Position, Temperature, and Ionic Strength. *Biochemistry* **2010**, *49*, 3225−3236.

(22) Morse, S. E.; Draper, D. E. Purine-Purine Mismatches in RNA Helices: Evidence for Protonated G·A Pairs and Next-Nearest Neighbor Effects. *Nucleic Acids Res.* **1995**, *23*, 302−306.

(23) Gehring, K.; Leroy, J.-L.; Guéron, M. A Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs. *Nature* **1993**, *363*, 561−565.

(24) Agris, P. F. In *Progress in Nucleic Acid Research and Molecular Biology*; Cohn, W. E., Moldave, K., Eds.; Elsevier, 1996; Vol. *53*; pp 79−129. DOI: DOI: 10.1016/s0079-6603(08)60143-9.

(25) Nachtergaele, S.; He, C. Chemical Modifications in the Life of an mRNA Transcript. *Annu. Rev. Genet.* **2018**, *52*, 349−372.

(26) Macon, J. B.; Wolfenden, R. 1-Methyladenosine. Dimroth Rearrangement and Reversible Reduction. *Biochemistry* **1968**, *7*, 3453−3458.

(27) Pfleiderer, W. Purine, IV. über die Methylierung des 9-Methylguanins und die Struktur des Herbipolins. *Justus Liebigs Ann. Chem.* **1961**, *647*, 167−173.

(28) Agris, P. F.; Sierzputowska-Gracz, H.; Smith, C. Transfer RNA Contains Sites of Localized Positive Charge: Carbon NMR Studies of [13C] Methyl-Enriched *Escherichia coli* and Yeast tRNAPhe. *Biochemistry* **1986**, *25*, 5126−5131.

(29) Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M.; et al. Topology of the Human and Mouse m^6A RNA Methylomes Revealed by m^6A-seq. *Nature* **2012**, *485*, 201−206.

(30) Meyer, K. D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C. E.; Jaffrey, S. R. Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3 UTRs and Near Stop Codons. *Cell* **2012**, *149*, 1635−1646.

(31) Dominissini, D.; Nachtergaele, S.; Moshitch-Moshkovitz, S.; Peer, E.; Kol, N.; Ben-Haim, M. S.; Dai, Q.; Di Segni, A.; Salmon-Divon, M.; Clark, W. C.; et al. The Dynamic N(1)-Methyladenosine

Methylome in Eukaryotic Messenger RNA. *Nature* **2016**, *530*, 441−446.
(32) Blanchard, S. C.; Puglisi, J. D. Solution Structure of the A Loop of 23S Ribosomal RNA. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3720−3725.
(33) Williams, L. D.; Shaw, B. R. Protonated Base Pairs Explain the Ambiguous Pairing Properties of O6-Methylguanine. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 1779−1783.
(34) Gaffney, B. L.; Goswami, B.; Jones, R. A. Nitrogen-15-Labeled Oligodeoxynucleotides. 7. Use of ^{15}N NMR to Probe H-Bonding in an O^{6}MeG·C Base Pair. *J. Am. Chem. Soc.* **1993**, *115*, 12607−12608.
(35) Reijenga, J.; van Hoof, A.; van Loon, A.; Teunissen, B. Development of Methods for the Determination of pK_a Values. *Anal. Chem. Insights* **2013**, *8*, 53−71.
(36) Hall, R. *The Modified Nucleosides in Nucleic Acids*; Columbia University Press: New York, 1971.
(37) Thaplyal, P.; Bevilacqua, P. C. Experimental Approaches for Measuring pK_a's in RNA and DNA. *Methods Enzymol.* **2014**, *549*, 189−219.
(38) Saenger, W. *Principles of Nucleic Acid Structure*; Springer Science & Business Media, 2013. DOI: DOI: 10.1007/978-1-4612-5190-3.
(39) Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. Prediction of the pK_a of Carboxylic Acids Using the ab initio Continuum-Solvation Model PCM-UAHF. *J. Phys. Chem. A* **1998**, *102*, 6706−6712.
(40) Shields, G. C.; Seybold, P. G. *Computational Approaches for the Prediction of pKa Values*; CRC Press, 2013. DOI: DOI: 10.1201/b16128.
(41) Ho, J.; Coote, M. L. A Universal Approach for Continuum Solvent pK_a Calculations: Are we There Yet? *Theor. Chem. Acc.* **2010**, *125*, 3−21.
(42) Ho, J. Predicting pK_a in Implicit Solvents: Current Status and Future Directions. *Aust. J. Chem.* **2014**, *67*, 1441−1460.
(43) Seybold, P. G.; Shields, G. C. Computational Estimation of pK_a Values. *WIREs Comput. Mol. Sci.* **2015**, *5*, 290−297.
(44) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Adding Explicit Solvent Molecules to Continuum Solvent Calculations for the Calculation of Aqueous Acid Dissociation Constants. *J. Phys. Chem. A* **2006**, *110*, 2493−2499.
(45) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A., III Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models. *J. Phys. Chem. B* **2008**, *112*, 9709−9719.
(46) Zhang, S. A Reliable and Efficient First Principles-Based Method for Predicting pK_a Values. III. Adding Explicit Water Molecules: Can the Theoretical Slope be Reproduced and pK_a Values Predicted More Accurately? *J. Comput. Chem.* **2012**, *33*, 517−526.
(47) Thapa, B.; Schlegel, H. B. Improved pK_a Prediction of Substituted Alcohols, Phenols, and Hydroperoxides in Aqueous Medium Using Density Functional Theory and a Cluster-Continuum Solvation Model. *J. Phys. Chem. A* **2017**, *121*, 4698−4706.
(48) da Silva, E. F.; Svendsen, H. F.; Merz, K. M. Explicitly Representing the Solvation Shell in Continuum Solvent Calculations. *J. Phys. Chem. A* **2009**, *113*, 6404−6409.
(49) Thapa, B.; Schlegel, H. B. Calculations of pK_a's and Redox Potentials of Nucleobases with Explicit Waters and Polarizable Continuum Solvation. *J. Phys. Chem. A* **2015**, *119*, 5134−5144.
(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A. et al. *Gaussian 09*, revision D.01; 2014.
(51) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(52) Camaioni, D. M.; Schwerdtfeger, C. A. Comment on "Accurate Experimental Values for the Free Energies of Hydration of H^+, OH^-, and H_3O^+. *J. Phys. Chem. A* **2005**, *109*, 10795−10797.
(53) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006**, *110*, 16066−16081.
(54) La Francois, C. J.; Jang, Y. H.; Cagin, T.; Goddard, W. A. I.; Sowers, L. C. Conformation and Proton Configuration of Pyrimidine Deoxynucleoside Oxidation Damage Products in Water. *Chem. Res. Toxicol.* **2000**, *13*, 462−470.
(55) Sochacka, E.; Lodyga-Chruscinska, E.; Pawlak, J.; Cypryk, M.; Bartos, P.; Ebenryter-Olbinska, K.; Leszczynska, G.; Nawrot, B. C5-Substituents of Uridines and 2-Thiouridines Present at the Wobble Position of tRNA Determine the Formation of Their Keto-Enol or Zwitterionic Forms - a Factor Important for Accuracy of Reading of Guanosine at the 3′-End of the mRNA Codons. *Nucleic Acids Res.* **2017**, *45*, 4825−4836.
(56) Matsuda, A.; Pankiewicz, K.; Marcus, B. K.; Watanabe, K. A.; Fox, J. J. Synthesis of 3-Methylpseudouridine and 2-Deoxy-3-Methylpseudouridine. *Carbohydr. Res.* **1982**, *100*, 297−302.
(57) Fox, J. J.; Yung, N.; Davoll, J.; Brown, G. B. Pyrimidine Nucleosides. I. A New Route for the Synthesis of Thymine Nucleosides. *J. Am. Chem. Soc.* **1956**, *78*, 2117−2122.
(58) Luyten, I.; Pankiewicz, K.; Watanabe, K.; Chattopadhyaya, J. Determination of the Tautomeric Equlibrium of Ψ-Uridine in the Basic Solution. *J. Org. Chem.* **1998**, *63*, 1033−1040.
(59) Murao, K.; Hasegawa, T.; Ishikura, H. 5-Methoxyuridine: A New Minor Constituent Located in the First Position of the Anticodon of tRNAAla, tRNAThr, and tRNAVal From *Bacillus subtilis*. *Nucleic Acids Res.* **1976**, *3*, 2851−2860.
(60) Nawrot, B.; Malkiewicz, A.; Smith, W. S.; Sierzputowska-Gracz, H.; Agris, P. F. RNA Modified Uridines VII: Chemical Synthesis and Initial Analysis of tRNA D-Loop Oligomers with Tandem Modified Uridines. *Nucleos. Nucleot. Nucl.* **1995**, *14*, 143−165.
(61) Raychaudhuri, S.; Conrad, J.; Hall, B. G.; Ofengand, J. A Pseudouridine Synthase Required for the Formation of two Universally Conserved Pseudouridines in Ribosomal RNA is Essential for Normal Growth of Escherichia coli. *RNA* **1998**, *4*, 1407−1417.
(62) Näsvall, S. J.; Chen, P.; Björk, G. R. The Wobble Hypothesis Revisited: Uridine-5-Oxyacetic Acid is Critical for Reading of G-Ending Codons. *RNA* **2007**, *13*, 2151−2164.
(63) Arnez, J. G.; Steitz, T. A. Crystal Structure of Unmodified tRNA(Gln) Complexed with Glutaminyl-tRNA Synthetase and ATP Suggests a Possible Role for Pseudo-Uridines in Stabilization of RNA Structure. *Biochemistry* **1994**, *33*, 7560−7567.
(64) Davis, D. R. Stabilization of RNA Stacking by Pseudouridine. *Nucleic Acids Res.* **1995**, *23*, 5020−5026.
(65) Auffinger, P.; Westhof, E. RNA Hydration: Three Nanoseconds of Multiple Molecular Dynamics Simulations of the Solvated tRNA^{Asp} Anticodon Hairpin. *J. Mol. Biol.* **1997**, *269*, 326−341.
(66) Newby, M. I.; Greenbaum, N. L. Investigation of Overhauser Effects Between Pseudouridine and Water Protons in RNA Helices. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12697−12702.
(67) Newby, M. I.; Greenbaum, N. L. A Conserved Pseudouridine Modification in Eukaryotic U2 snRNA Induces a Change in Branch-Site Architecture. *RNA* **2001**, *7*, 833−845.
(68) Sumita, M.; Desaulniers, J.-P.; Chang, Y.-C.; Chui, H. M.-P.; Clos, L.; Chow, C. S. Effects of Nucleotide Substitution and Modification on the Stability and Structure of Helix 69 From 28S rRNA. *RNA* **2005**, *11*, 1420−1429.
(69) Chui, H. M.-P.; Desaulniers, J.-P.; Scaringe, S. A.; Chow, C. S. Synthesis of Helix 69 of Escherichia coli 23S rRNA Containing Its Natural Modified Nucleosides, m3Ψ and Ψ. *J. Org. Chem.* **2002**, *67*, 8847−8854.
(70) Huang, W.; Lan, M.-D.; Qi, C.-B.; Zheng, S.-J.; Wei, S.-Z.; Yuan, B.-F.; Feng, Y.-Q. Formation and Determination of the Oxidation Products of 5-Methylcytosine in RNA. *Chem. Sci.* **2016**, *7*, 5495−5502.
(71) Dai, Q.; Sanstead, P. J.; Peng, C. S.; Han, D.; He, C.; Tokmakoff, A. Weakened N3 Hydrogen Bonding by 5-Formylcytosine and 5-Carboxylcytosine Reduces Their Base-Pairing Stability. *ACS Chem. Biol.* **2016**, *11*, 470−477.

(72) Van Montagu, M.; Molemans, F.; Stockx, J. Preparation of Cytidine, Cytidylic Acids and Ribonucleic Acid Specifically Acetylated in the Exocyclic Amino Group of Cytosine. *Bull. Soc. Chim. Belg.* **1968**, *77*, 171−179.
(73) Shugar, D.; Fox, J. J. Spectrophotometric Studies of Nucleic Acid Derivatives and Related Compounds as a Function of pH. I. Pyrimidines. *Biochim. Biophys. Acta* **1952**, *9*, 199−218.
(74) Fox, J. J.; Van Praag, D.; Wempen, I.; Doerr, I. L.; Cheong, L.; Knoll, J. E.; Eidinoff, M. L.; Bendich, A.; Brown, G. B. Thiation of Nucleosides. II. Synthesis of 5-Methyl-2'-Deoxycytidine and Related Pyrimidine Nucleosides[1]. *J. Am. Chem. Soc.* **1959**, *81*, 178−187.
(75) Brookes, P.; Lawley, P. D. 254. The Methylation of Cytosine and Cytidine. *J. Chem. Soc.* **1962**, 1348−1351.
(76) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. Tautomerism in Cytosine and 3-Methylcytosine. A Thermodynamic and Kinetic Study. *J. Am. Chem. Soc.* **1976**, *98*, 6338−6349.
(77) Arango, D.; Sturgill, D.; Alhusaini, N.; Dillman, A. A.; Sweet, T. J.; Hanson, G.; Hosogane, M.; Sinclair, W. R.; Nanan, K. K.; Mandler, M. D.; et al. Acetylation of Cytidine in mRNA Promotes Translation Efficiency. *Cell* **2018**, *175*, 1872−1886.
(78) Wada, T.; Kobori, A.; Kawahara, S.-i.; Sekine, M. Synthesis and Properties of Oligodeoxyribonucleotides Containing 4-N-Acetylcytosine Bases. *Tetrahedron Lett.* **1998**, *39*, 6907−6910.
(79) Szulik, M. W.; Pallan, P. S.; Nocek, B.; Voehler, M.; Banerjee, S.; Brooks, S.; Joachimiak, A.; Egli, M.; Eichman, B. F.; Stone, M. P. Differential Stabilities and Sequence-Dependent Base Pair Opening Dynamics of Watson−Crick Base Pairs with 5-Hydroxymethylcytosine, 5-Formylcytosine, or 5-Carboxylcytosine. *Biochemistry* **2015**, *54*, 1294−1305.
(80) Riml, C.; Lusser, A.; Ennifar, E.; Micura, R. Synthesis, Thermodynamic Properties, and Crystal Structure of RNA Oligonucleotides Containing 5-Hydroxymethylcytosine. *J. Org. Chem.* **2017**, *82*, 7939−7945.
(81) Wang, R.; Luo, Z.; He, K.; Delaney, M. O.; Chen, D.; Sheng, J. Base Pairing and Structural Insights into the 5-Formylcytosine in RNA Duplex. *Nucleic Acids Res.* **2016**, *44*, 4968−4977.
(82) Harkins, T. R.; Freiser, H. Adenine-Metal Complexes[1,2]. *J. Am. Chem. Soc.* **1958**, *80*, 1132−1135.
(83) Martin, D.; Reese, C. Some Aspects of the Chemistry of N(1)- and N(6)-Dimethylallyl Derivatives of Adenosine and Adenine. *J. Chem. Soc. C* **1968**, *14*, 1731−1738.
(84) Fu, Y.; Jia, G.; Pang, X.; Wang, R. N.; Wang, X.; Li, C. J.; Smemo, S.; Dai, Q.; Bailey, K. A.; Nobrega, M. A.; et al. FTO-Mediated Formation of N^6-Hydroxymethyladenosine and N^6-Formyladenosine in Mammalian RNA. *Nat. Commun.* **2013**, *4*, 1−8.
(85) Zhang, X.; Wei, L.-H.; Wang, Y.; Xiao, Y.; Liu, J.; Zhang, W.; Yan, N.; Amu, G.; Tang, X.; Zhang, L.; et al. Structural Insights into FTO's Catalytic Mechanism for the Demethylation of Multiple RNA Substrates. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 2919−2924.
(86) Toh, J. D. W.; Crossley, S. W. M.; Bruemmer, K. J.; Ge, E. J.; He, D.; Iovan, D. A.; Chang, C. J. Distinct RNA *N*-Demethylation Pathways Catalyzed by Nonheme Iron ALKBH5 and FTO Enzymes Enable Regulation of Formaldehyde Release Rates. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 25284−25292.
(87) Liu, B.; Shi, H.; Rangadurai, A.; Nussbaumer, F.; Chu, C.-C.; Erharter, K. A.; Case, D. A.; Kreutz, C.; Al-Hashimi, H. M. A Quantitative Model Predicts How m^6A Reshapes the Kinetic Landscape of Nucleic Acid Hybridization and Conformational Transitions. *Nat. Commun.* **2021**, *12*, 1−17.
(88) Zou, S.; Toh, J. D.; Wong, K. H.; Gao, Y.-G.; Hong, W.; Woon, E. C. N^6-Methyladenosine: a Conformational Marker that Regulates the Substrate Specificity of Human Demethylases FTO and ALKBH5. *Sci. Rep.* **2016**, *6*, 1−12.
(89) Helm, M.; Giegé, R.; Florentz, C. A Watson-Crick Base-Pair-Disrupting Methyl Group (m1A9) is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNALys. *Biochemistry* **1999**, *38*, 13338−13346.
(90) Lawley, P. D.; Brookes, P. Further Studies on the Alkylation of Nucleic Acids and Their Constituent Nucleotides. *Biochem. J.* **1963**, *89*, 127−138.
(91) Beaven, G. H.; Holiday, E. R.; Johnson, E. A. Optical Properties of Nucleic Acids and Their Components. In *The Nucleic Acids*; Chargaff, E., Davidson, J., Eds.; Academic Press Inc., 1955; Vol. *1*, pp 493−553. DOI: DOI: 10.1002/jps.3030440720.
(92) Levene, P. A.; Bass, L. W. *Nucleic Acids*; Chemical Catalog Company: New York, 1931.
(93) Smith, J. D.; Dunn, D. B. The Occurrence of Methylated Guanines in Ribonucleic Acids From Several Sources. *Biochem. J.* **1959**, *72*, 294−301.
(94) Loveless, A. Possible Relevance of O−6 Alkylation of Deoxyguanosine to the Mutagenicity and Carcinogenicity of Nitrosamines and Nitrosamides. *Nature* **1969**, *223*, 206−207.
(95) Gao, X. L.; Patel, D. J. NMR Studies of A·C Mismatches in DNA Dodecanucleotides at Acidic pH. Wobble A(*Anti*)·C(*Anti*) Pair Formation. *J. Biol. Chem.* **1987**, *262*, 16973−16984.
(96) Marcotrigiano, J.; Gingras, A.-C.; Sonenberg, N.; Burley, S. K. Cocrystal Structure of the Messenger RNA 5′ Cap-Binding Protein (eIF4E) Bound to 7-Methyl-GDP. *Cell* **1997**, *89*, 951−961.
(97) Hu, G.; Gershon, P. D.; Hodel, A. E.; Quiocho, F. A. mRNA Cap Recognition: Dominant Role of Enhanced Stacking Interactions Between Methylated Bases and Protein Aromatic Side Chains. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7149−7154.
(98) Rife, J. P.; Cheng, C. S.; Moore, P. B.; Strobel, S. A. N2-Methylguanosine is Iso-Energetic With Guanosine in RNA Duplexes and GNRA Tetraloops. *Nucleic Acids Res.* **1998**, *26*, 3640−3644.
(99) Engel, J. D.; Von Hippel, P. H. Effects of Methylation on the Stability of Nucleic Acid Conformations: Studies at the Monomer Level. *Biochemistry* **1974**, *13*, 4143−4158.
(100) Pallan, P. S.; Kreutz, C.; Bosio, S.; Micura, R.; Egli, M. Effects of N^2,N^2-Dimethylguanosine on RNA Structure and Stability: Crystal Structure of an RNA Duplex with Tandem m^2_2G:A pairs. *RNA* **2008**, *14*, 2125−2135.
(101) Steinberg, S.; Cedergren, R. A Correlation Between N2-Dimethylguanosine Presence and Alternate tRNA Conformers. *RNA* **1995**, 886.
(102) Crick, F. H. Codon−Anticodon Pairing: The Wobble Hypothesis. *J. Mol. Biol.* **1966**, *19*, 548−555.
(103) Alseth, I.; Dalhus, B.; Bjørås, M. Inosine in DNA and RNA. *Curr. Opin. Genet. Dev.* **2014**, *26*, 116−123.
(104) Strobel, S. A.; Cech, T. R.; Usman, N.; Beigelman, L. The 2,6-Diaminopurine Riboside·5-Methylisocytidine Wobble Base Pair: An Isoenergetic Substitution for the Study of G·U Pairs in RNA. *Biochemistry* **1994**, *33*, 13824−13835.
(105) Zhou, H.; Kimsey, I. J.; Nikolova, E. N.; Sathyamoorthy, B.; Grazioli, G.; McSally, J.; Bai, T.; Wunderlich, C. H.; Kreutz, C.; Andricioaei, I.; et al. m^1A and m^1G Disrupt A-RNA Structure Through the Intrinsic Instability of Hoogsteen Base Pairs. *Nat. Struct. Mol. Biol.* **2016**, *23*, 803−810.
(106) Bjork, G. R.; Wikstrom, P. M.; Bystrom, A. S. Prevention of Translational Frameshifting by the Modified Nucleoside 1-Methylguanosine. *Science* **1989**, *244*, 986−989.
(107) Hong, S.; Sunita, S.; Maehigashi, T.; Hoffer, E. D.; Dunkle, J. A.; Dunham, C. M. Mechanism of tRNA-Mediated +1 Ribosomal Frameshifting. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 11226−11231.
(108) Mao, S.; Sekula, B.; Ruszkowski, M.; Ranganathan, S. V.; Haruehanroengra, P.; Wu, Y.; Shen, F.; Sheng, J. Base Pairing, Structural and Functional Insights into N4-Methylcytidine (m4C) and N4, N4-Dimethylcytidine (m42C) Modified RNA. *Nucleic Acids Res.* **2020**, *48*, 10087−10100.

https://doi.org/10.1021/acs.jpca.1c10905
J. Phys. Chem. A 2022, 126, 1518−1529