

# Effect of Sodium Ions on RNA Duplex Stability

Zexiang Chen and Brent M. Znosko\*

Department of Chemistry, Saint Louis University, 3501 Laclede Avenue, St. Louis, Missouri 63103, United States

**S** Supporting Information

**ABSTRACT:** The standard sodium concentration for RNA optical melting experiments is 1.021 M. Algorithms that predict  $T_m$ ,  $\Delta G^\circ_{37}$ , and secondary structure from sequence generally rely on parameters derived from optical melting experiments performed in 1.021 M sodium. Physiological monovalent cation concentrations are much lower than 1.021 M. In fact, many molecular biology techniques require buffers containing monovalent cation concentrations other than 1.021 M. Predictions based on the 1.021 M  $\text{Na}^+$  parameters may not be accurate when the monovalent cation concentration is not 1.021 M. Here, we report thermodynamic data from optical melting experiments for a set of 18 RNA duplexes, each melted over a wide range of sodium ion concentrations (71, 121, 221, and 621 mM). Using these data and previously published data for the same sequences melted in 1.021 M  $\text{Na}^+$ , we report  $T_m$  and  $\Delta G^\circ_{37}$  correction factors to scale the standard 1.021 M  $\text{Na}^+$  RNA parameters to other sodium ion concentrations. The recommended  $T_m$  correction factor predicts the melting temperature within 0.7 °C, and the recommended  $\Delta G^\circ_{37}$  correction factor predicts the free energy within 0.14 kcal/mol. These correction factors can be incorporated into prediction algorithms that predict RNA secondary structure from sequence and provide  $T_m$  and  $\Delta G^\circ_{37}$  values for RNA duplexes.

To convert from 1 M  $\text{Na}^+$  to other  $[\text{Na}^+]$ :

$$T_m(2) = T_m(1) + (-1.842 fGC + 2.675) \ln \left[ \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1} \right] - 0.7348 (\ln^2 [\text{Na}^+]_2 - \ln^2 [\text{Na}^+]_1)$$

$$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) + (0.324 fGC - 0.468) \ln \left[ \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1} \right] + 0.133 (\ln^2 [\text{Na}^+]_2 - \ln^2 [\text{Na}^+]_1)$$

RNA is one of the most important biomolecules in all forms of life. RNA, however, needs to fold into appropriate secondary and three-dimensional (3D) structures so that it can function properly.<sup>1,2</sup> Therefore, knowing the secondary and 3D structures of RNA will help scientists better understand its function and mechanism of action.<sup>3,4</sup> However, the number of determined RNA 3D structures is significantly smaller than the soaring number of available RNA sequences.<sup>5</sup> Hence, structure prediction may be the most efficient way to elucidate RNA tertiary structure.

Predicting RNA secondary structure can be an intermediate step in predicting RNA 3D structure.<sup>6</sup> The nearest-neighbor model,<sup>7,8</sup> which is based on sets of adjacent base pairs, is currently the most widely used algorithm for predicting RNA secondary structure from sequence. The nearest-neighbor model can be used to predict the stability of simple Watson–Crick duplexes and duplexes containing more complicated secondary structure motifs such as bulges, internal loops, and hairpins. The parameters used in the nearest-neighbor model were derived from a large series of optical melting experiments for RNA duplexes in salt buffers normally containing 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM  $\text{Na}_2\text{EDTA}$ , which results in a total  $\text{Na}^+$  concentration of 1.021 M.<sup>9</sup>

Cations are crucial for RNA folding and function. The polyanionic backbone of RNA requires cations (specific or non-specific binding) to neutralize the negative charge.<sup>10</sup> Theoretical studies of the relationship between cations and nucleic acids were pioneered by Manning, who proposed the counterion condensation theory.<sup>11</sup> Recently, the Poisson–Boltzmann equation,<sup>12</sup> Monte Carlo simulations,<sup>13</sup> and the tightly bound ion (TBI) theory<sup>14</sup> have also been used to describe the distribution of cations around RNA. A NaCl concentration of 1 M (along

with 20 mM sodium cacodylate and 0.5 mM  $\text{Na}_2\text{EDTA}$ ) was initially chosen by the pioneers of RNA optical melting studies<sup>15</sup> to stabilize short RNA oligonucleotides. Therefore, 1.021 M  $\text{Na}^+$  has become the standard sodium concentration for RNA optical melting experiments, on which secondary structure prediction algorithms are based.

Extracellular and intracellular monovalent cation concentrations, however, are much lower than 1.021 M. In addition, buffer conditions of numerous molecular biology techniques require cation concentrations other than 1.021 M. For instance, polymerase chain reaction experiments usually use buffer conditions that include between 20 and 100 mM monovalent cations.<sup>16</sup> The success of these molecular biology techniques, including antisense RNA and RNAi, is largely dependent on the specific and accurate hybridization between RNA strands.<sup>17</sup> Therefore, it would be beneficial to be able to accurately predict the thermodynamics of RNA, especially the melting temperature ( $T_m$ ) and free energy change ( $\Delta G^\circ_{37}$ ). Many scientists who perform these techniques predict  $T_m$  and  $\Delta G^\circ_{37}$  values of duplexes utilizing the nearest-neighbor model. The major limitation of using the nearest-neighbor model to calculate  $T_m$  and  $\Delta G^\circ_{37}$  is that the parameters in the nearest-neighbor model were derived from RNA duplexes in 1.021 M  $\text{Na}^+$ , which may not be consistent with the thermodynamics under other salt conditions. This difference could lead to unanticipated results or even complete failure of the experiments.

**Received:** June 26, 2013

**Revised:** August 14, 2013

**Published:** October 9, 2013

There have been extensive experimental studies of the relationship between sodium ion concentrations and DNA thermodynamics.<sup>18–24</sup> Recently, a systematic study of the sodium ion dependence of DNA duplex stability was completed by Owczarzy et al.,<sup>17</sup> and correction factors were proposed to adjust the DNA thermodynamic parameters at 1.021 M Na<sup>+</sup> to parameters corresponding to other monovalent cation concentrations. Moreover, Nakano et al.<sup>25</sup> also proposed correction factors for nucleic acids. However, the data in this study were mainly from DNA duplexes, with a few RNA–DNA hybrids and RNA duplexes, so the correction factors may not be accurate for RNA duplexes. Also, the correction factors proposed by Nakano et al.<sup>25</sup> were limited to 100 mM Na<sup>+</sup>. Despite the effort spent on nucleic acids, systematic studies of the relationship between sodium ion concentrations and RNA duplex stability have not been completed.

Here, we report thermodynamic data from optical melting experiments for a set of 18 RNA duplexes, each melted over a wide range of sodium ion concentrations (71, 121, 221, and 621 mM). Using the DNA results of Owczarzy et al.<sup>17</sup> as a guide, we report  $T_m$  and  $\Delta G^\circ_{37}$  correction factors to scale the standard 1.021 M Na<sup>+</sup> RNA parameters to other sodium ion concentrations. These correction factors can be incorporated into prediction algorithms that predict RNA secondary structure from sequence and provide  $T_m$  and  $\Delta G^\circ_{37}$  values for RNA duplexes.

## MATERIALS AND METHODS

### Oligonucleotide Selection, Synthesis, and Purification.

RNA duplexes were selected from the sequences that were used by Xia et al.<sup>9</sup> to derive the RNA nearest-neighbor parameters in 1.021 M Na<sup>+</sup>. The oligomers were ordered from Integrated DNA Technologies, Inc. (Coralville, IA). Purification of oligonucleotides was performed using standard procedures described previously.<sup>26–28</sup>

**Optical Melting Experiments.** All of the strands used here were self-complementary; therefore, mixing of strands was not necessary. After purification, the RNA oligonucleotides were lyophilized and redissolved in melting buffer containing 20 mM sodium cacodylate, 0.5 mM Na<sub>2</sub>EDTA, and 50, 100, 200, and 600 mM NaCl, adjusted to pH 7.0. The resulting total sodium ion concentrations were 71, 121, 221, and 621 mM, respectively. Each duplex was melted at least nine times, using a different concentration each time, to ensure that the total oligonucleotide concentration range was at least 50-fold. Using a heating rate of 1 °C/min on a Beckman-Coulter DU800 spectrophotometer, absorbance versus temperature melting curves were obtained between 0 and 90 °C. For sequences containing at least 50% G-C base pairs, absorbances were measured at 280 nm, while the absorbance of A-U rich oligonucleotides was measured at 260 nm.<sup>15</sup> Meltwin<sup>29</sup> was used to determine the thermodynamic parameters of each duplex. Thermodynamic parameters, which were used in developing correction factors, were derived from the  $1/T_m$  versus  $\ln C_t$  plots, and melting temperatures were calculated at a strand concentration of  $10^{-4}$  M.

**Predicting  $T_m$ .** The accuracy of 10 previously published DNA  $T_m$  correction factors was evaluated with the RNA data collected here. Using the experimental 1.021 M Na<sup>+</sup>  $T_m$  as the starting point, the DNA correction factors were applied to predict  $T_m$  values at 71, 121, 221, and 621 mM Na<sup>+</sup>, which correspond to the RNA data reported here. The accuracy of the 10 models was tested using  $|\Delta T_m|_{ave}$ :

$$|\Delta T_m|_{ave} = \left[ \sum_{j=1}^{j=n} |T_m(j, \text{prediction}) - T_m(j, \text{experiment})| \right] / N \quad (1)$$

For each correction factor, a total of 72 (18 duplexes studied at four different Na<sup>+</sup> concentrations) melting temperatures were predicted and used to calculate  $|\Delta T_m|_{ave}$ .

Although some of the previously published DNA correction factors worked well for the RNA data reported here, the coefficients for some of the previously published DNA correction factors were updated for the RNA data reported here by using the LINEST function of Microsoft Excel. These RNA correction factors were then tested in a similar manner using  $|\Delta T_m|_{ave}$ .

**Predicting  $\Delta G^\circ_{37}$ .** Although there are 10 previously published DNA  $T_m$  correction factors, there is only one previously published DNA  $\Delta G^\circ_{37}$  correction factor.<sup>24</sup> The accuracy of this DNA  $\Delta G^\circ_{37}$  correction factor was evaluated with the RNA data collected here in a similar manner using  $|\Delta \Delta G^\circ_{37}|_{ave}$ :

$$|\Delta \Delta G^\circ_{37}|_{ave} = \left[ \sum_{j=1}^{j=n} \left| \Delta \Delta G^\circ_{37}(j, \text{prediction}) - \Delta \Delta G^\circ_{37}(j, \text{experiment}) \right| \right] / N \quad (2)$$

The coefficient in this previous correction factor was also updated for the RNA data reported here by using the LINEST function of Microsoft Excel. The updated version was then tested using  $|\Delta \Delta G^\circ_{37}|_{ave}$ .

In addition to updating the previous DNA correction factor, we tested several new  $\Delta G^\circ_{37}$  correction factors. The first set of new  $\Delta G^\circ_{37}$  correction factors was derived from the  $T_m$  correction factors. Combining the Gibbs free energy equation and the van't Hoff equation<sup>9</sup> yields the following equation:

$$T_m^{-1} = \frac{\Delta H^\circ - \Delta G^\circ_{37}}{310.15 \Delta H^\circ} + \frac{R \ln C_t}{\Delta H^\circ} \quad (3)$$

For every  $T_m$  correction factor derived, it can be inserted into this equation to yield a  $\Delta G^\circ_{37}$  correction factor. For example, if the  $T_m$  correction factor was

$$T_m^{-1}(2) = T_m^{-1}(1) + 10 \quad (4)$$

Substituting this correction factor (eq 4) into eq 3 would yield

$$\frac{\Delta H^\circ - \Delta G^\circ_{37}(2)}{310.15 \Delta H^\circ} + \frac{R \ln C_t}{\Delta H^\circ} = \frac{\Delta H^\circ - \Delta G^\circ_{37}(1)}{310.15 \Delta H^\circ} + \frac{R \ln C_t}{\Delta H^\circ} + 10 \quad (5)$$

Simplifying this equation results in the corresponding  $\Delta G^\circ_{37}$  correction factor:

$$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) - 3101.5 \Delta H^\circ \quad (6)$$

The accuracy of correction factors derived from this method was evaluated by using  $|\Delta \Delta G^\circ_{37}|_{ave}$ . It is important to note that  $\Delta G^\circ_{37}$  correction factors derived in this way rely on three assumptions.<sup>16</sup> (i) RNA duplexes melted in a two-state process. (ii) Counterion effects were mainly entropic.<sup>24,30,31</sup> (iii) The  $\Delta C_p$  of melting reactions was zero, which means enthalpies and entropies are temperature-independent. All of these assumptions were valid for the oligonucleotides studied here.<sup>16</sup>

**Table 1. Experimental Melting Temperatures of RNA Duplexes at Various Sodium Ion Concentrations<sup>a</sup>**

RNA sequence (5' to 3') <sup>d</sup>	<i>f</i> GC <sup>c</sup>	<i>T</i> <sub>m</sub> <sup>b</sup> (°C)				
		71 mM	121 mM	221 mM	621 mM	1.021 M <sup>e</sup>
CGCGCG	1.00	51.3	53.7	55.2	57.8	57.8
CGGCCG	1.00	55.2	57.4	59.9	61.9	63.2
GCCGGC	1.00	60.1	62.6	65.0	67.6	67.4
GCGCGC	1.00	55.3	57.8	59.9	62.3	62.5
ACCGGU	0.67	43.7	47.2	50.1	52.0	53.9
AGCGCU	0.67	41.3	45.7	47.9	51.4	52.0
CACGUG	0.67	33.6	36.2	38.5	41.0	42.8
CAGCUG	0.67	35.3	37.2	39.7	42.1	43.1
CCAUGG	0.67	35.7	37.0	40.4	43.5	46.4
CCUAGG	0.67	39.7	42.4	44.7	47.4	50.0
CUGCAG	0.67	37.1	39.9	41.9	44.8	45.3
GACGUC	0.67	38.5	41.2	43.1	46.0	46.2
GAGCUC	0.67	40.2	42.7	45.0	47.7	48.7
GCAUGC	0.67	37.8	40.2	42.6	45.8	45.7
AACUAGUU	0.25	34.0	37.0	40.0	43.8	45.7
ACUUAAGU	0.25	32.9	36.2	38.8	42.7	44.0
ACUUAAGU	0.25	29.8	32.4	35.8	39.3	40.3
AGAUAUUCU	0.25	31.0	33.9	37.7	41.7	41.4

<sup>a</sup>*T*<sub>m</sub> values are from the 1/*T*<sub>m</sub> vs ln *C*<sub>t</sub> plots. <sup>b</sup>Calculated for an oligomer concentration of 0.1 mM. <sup>c</sup>All 1.021 M data, except for those of GCCGGC and GCGCGC, are from ref 9. <sup>d</sup>All oligomers are self-complementary and form duplexes in solution. <sup>e</sup>*f*GC is the fraction of G-C base pairs.

**Table 2. Experimental  $\Delta G^{\circ}_{37}$  Values for RNA Duplexes at Various Sodium Ion Concentrations<sup>a</sup>**

RNA sequence (5' to 3') <sup>c</sup>	<i>f</i> GC <sup>d</sup>	$\Delta G^{\circ}_{37}$ (kcal/mol)				
		71 mM	121 mM	221 mM	621 mM	1.021 M <sup>b</sup>
CGCGCG	1.00	−8.12	−8.52	−8.87	−9.23	−9.12
CGGCCG	1.00	−8.76	−9.09	−9.40	−9.70	−9.90
GCCGGC	1.00	−10.23	−10.73	−10.90	−11.41	−11.69
GCGCGC	1.00	−8.85	−9.32	−9.83	−10.29	−10.56
ACCGGU	0.67	−6.85	−7.25	−7.67	−8.07	−8.51
AGCGCU	0.67	−6.32	−6.91	−7.38	−7.82	−7.99
CACGUG	0.67	−5.10	−5.54	−5.94	−6.45	−6.59
CAGCUG	0.67	−5.37	−5.72	−6.18	−6.60	−6.68
CCAUGG	0.67	−5.44	−5.68	−6.34	−7.02	−7.30
CCUAGG	0.67	−6.22	−6.76	−7.18	−7.68	−7.80
CUGCAG	0.67	−5.69	−6.20	−6.53	−7.03	−7.11
GACGUC	0.67	−5.96	−6.44	−6.77	−7.26	−7.35
GAGCUC	0.67	−6.24	−6.70	−7.13	−7.59	−7.98
GCAUGC	0.67	−5.81	−6.26	−6.69	−7.31	−7.38
AACUAGUU	0.25	−5.09	−5.68	−6.24	−6.92	−7.16
ACUUAAGU	0.25	−4.86	−5.53	−6.04	−6.87	−6.98
ACUUAAGU	0.25	−4.37	−4.76	−5.46	−6.11	−6.16
AGAUAUUCU	0.25	−4.45	−5.05	−5.82	−6.59	−6.58

<sup>a</sup> $\Delta G^{\circ}_{37}$  values are from the 1/*T*<sub>m</sub> vs ln *C*<sub>t</sub> plots. <sup>b</sup>All 1.021 M data, except for those of GCCGGC and GCGCGC, are from ref 9. <sup>c</sup>All oligomers are self-complementary and form duplexes in solution. <sup>d</sup>*f*GC is the fraction of G-C base pairs.

A second set of correction factors was derived on the basis of linear or quadratic relationships between  $\Delta G^{\circ}_{37}$  or 1/ $\Delta G^{\circ}_{37}$  and ln[Na<sup>+</sup>]. These resulting  $\Delta G^{\circ}_{37}$  correction factors were similar to the previously published DNA *T*<sub>m</sub> correction factors. The RNA  $\Delta G^{\circ}_{37}$  data reported here and the LINEST function of Microsoft Excel were used to derive the coefficients for these  $\Delta G^{\circ}_{37}$  correction factors. The accuracy of these  $\Delta G^{\circ}_{37}$  correction factors was also evaluated using  $|\Delta\Delta G^{\circ}_{37}|_{\text{ave}}$ .

## RESULTS

**RNA Thermodynamic Parameters.** Eighteen duplexes at five different sodium ion concentrations were melted. Experimental  $\Delta G^{\circ}_{37}$ ,  $\Delta H^{\circ}$ ,  $\Delta S^{\circ}$ , and *T*<sub>m</sub> values are listed in Table S1 of

the Supporting Information. All the oligonucleotides melted in a two-state manner. The experimental *T*<sub>m</sub> and  $\Delta G^{\circ}_{37}$  values for all of the duplexes and all Na<sup>+</sup> concentrations are summarized in Tables 1 and 2, respectively. On average, the *T*<sub>m</sub> values of duplexes in 71, 121, 221, and 621 mM Na<sup>+</sup> were 9.1, 6.4, 3.9, and 1.0 °C lower, respectively, than that of the same duplex in 1.021 M Na<sup>+</sup>. Similarly, duplexes melted in 71, 121, 221, and 621 mM Na<sup>+</sup> were on average 1.62, 1.15, 0.69, and 0.16 kcal/mol less stable, respectively, than the same duplex in 1.021 M Na<sup>+</sup>.

***T*<sub>m</sub> Correction Factors.** RNA duplexes in buffers containing 71–621 mM Na<sup>+</sup> melt at temperatures lower than that of the same duplex in 1.021 M Na<sup>+</sup>. Therefore, *T*<sub>m</sub> correction factors are needed for accurate predictions. Several previously published

Table 3. Previously Published DNA Correction Factors

eq	name	ref	equation	accuracy <sup>c</sup>
DNA $T_m$ Correction Factors				
7	Schildkraut–Lifson $T_m$ equation	13	$T_m(2) = T_m(1) + 16.6 \log \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	7.2 °C
8	Wetmur $T_m$ equation	14	$T_m(2) = T_m(1) + 16.6 \log \frac{[\text{Na}^+]_2(1 + 0.7[\text{Na}^+]_1)}{[\text{Na}^+]_1(1 + 0.7[\text{Na}^+]_2)}$	4.5 °C
9	Frank–Kamenetskii $T_m$ equation	15	$T_m(2) = T_m(1) + (7.95 - 3.057fGC) \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	5.1 °C
10	Marmur–Schildkraut–Doty $T_m$ equation	16, 17	$T_m(2) = T_m(1) + (8.75 - 2.83fGC) \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	6.7 °C
11	$T_m$ and $12.5 \log[\text{Na}^+]$ equation	18	$T_m(2) = T_m(1) + 12.5 \log \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	4.2 °C
12	SantaLucia $T_m$ equation <sup>a,b</sup>	19	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + \frac{0.368N}{\Delta H^\circ} \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	1.4 °C
13	Owczarzy $T_m$ linear equation	12	$T_m(2) = T_m(1) + (-3.22fGC + 6.39) \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	2.3 °C
14	Owczarzy $1/T_m$ linear equation <sup>b</sup>	12	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + (3.85fGC - 6.18) \times 10^{-5} \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	1.5 °C
15	Owczarzy $T_m$ quadratic equation	12	$T_m(2) = T_m(1) + (-4.62fGC + 4.52) \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$ $- 0.985(\ln^2 [\text{Na}^+]_2 - \ln^2 [\text{Na}^+]_1)$	1.5 °C
16	Owczarzy $1/T_m$ quadratic equation <sup>b</sup>	12	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + (4.29fGC - 3.95) \times 10^{-5} \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$ $+ 9.40 \times 10^{-6}(\ln^2 [\text{Na}^+]_2 - \ln^2 [\text{Na}^+]_1)$	1.1 °C
DNA $\Delta G^\circ_{37}$ Correction Factor				
17	SantaLucia $\Delta G^\circ_{37}$ equation <sup>a</sup>	19	$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) - 0.114 \times N \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1}$	0.21 kcal/mol

<sup>a</sup> $N$  is the total number of phosphates in the duplex divided by 2. <sup>b</sup>In this equation,  $T_m$  should be in units of kelvin. <sup>c</sup>As described in Materials and Methods,  $|\Delta T_m|_{\text{ave}}$  is used to evaluate the accuracy of previously published DNA  $T_m$  correction factors in predicting RNA  $T_m$  values, and  $|\Delta \Delta G^\circ_{37}|_{\text{ave}}$  is used to evaluate the accuracy of the previously published DNA  $\Delta G^\circ_{37}$  correction factor in predicting RNA  $\Delta G^\circ_{37}$  values.

DNA  $T_m$  correction factors are listed in Table 3. The SantaLucia<sup>24</sup> and Owczarzy<sup>17</sup> DNA correction factors work particularly well for the RNA data reported here, with  $|\Delta T_m|_{\text{ave}}$  being less than 2.5 °C. Because these worked so well, an attempt was made to further improve these correction factors by deriving updated coefficients based on the RNA data reported here. These newly derived correction factors are listed in Table 4. With the updated coefficients, the accuracy of these models improves, resulting in a  $|\Delta T_m|_{\text{ave}}$  of  $\leq 1.0$  °C. Because of its accuracy ( $|\Delta T_m|_{\text{ave}} = 0.7$  °C) and consistency with the  $\Delta G^\circ_{37}$  correction factor (discussed below), we recommend eq 21 (Table 4) as the  $T_m$  correction factor to be used for  $\text{Na}^+$  concentrations of  $<1.021$  M.

**$\Delta G^\circ_{37}$  Correction Factors.** One previously published DNA  $\Delta G^\circ_{37}$  correction factor is listed in Table 3. This SantaLucia correction factor<sup>24</sup> works particularly well for the RNA data reported here, with a  $|\Delta \Delta G^\circ_{37}|_{\text{ave}}$  value of 0.21 kcal/mol. Because it worked so well, an attempt was made to further improve this correction factor by deriving updated coefficients based on the RNA data reported here. This newly derived correction factor is listed in Table 4. With the updated coefficients, the accuracy of this model improves slightly, resulting in a  $|\Delta \Delta G^\circ_{37}|_{\text{ave}}$  value of 0.18 kcal/mol. Because this was the only DNA  $\Delta G^\circ_{37}$  correction factor available in the literature, several additional  $\Delta G^\circ_{37}$  correction factors were derived and tested. These  $\Delta G^\circ_{37}$  correction

factors are listed in Table 4. Because of its accuracy ( $|\Delta \Delta G^\circ_{37}|_{\text{ave}} = 0.14$  kcal/mol) and relative simplicity, we recommend eq 26 (Table 4) as the  $\Delta G^\circ_{37}$  correction factor to be used for  $\text{Na}^+$  concentrations of  $<1.021$  M.

## DISCUSSION

**Dependence of RNA Duplex Thermal Stability on Sodium Ion Concentration.** As expected, when  $[\text{Na}^+]$  is increased from 71 to 621 mM, RNA  $T_m$  values increase (Table 1). Previous data and theories have suggested that RNA duplexes will become saturated with  $\text{Na}^+$  at high  $\text{Na}^+$  concentrations.<sup>11,14,17,32</sup> As anticipated, increasing the  $\text{Na}^+$  concentration from 621 mM to 1.021 M has very little (increase or decrease) or no effect on the RNA  $T_m$  values. Figure 1 shows the relationship between  $T_m$  and  $\ln[\text{Na}^+]$  for representative RNA oligonucleotides, and it confirms that these RNA duplexes become saturated with  $\text{Na}^+$  at high sodium ion concentrations. Similar observations were made for the relationship between  $\Delta G^\circ_{37}$  and sodium ion concentration, as shown in Table 2 and Figure 2. As expected, when the  $\text{Na}^+$  concentration is increased from 71 to 621 mM, RNA  $\Delta G^\circ_{37}$  values become more negative. Increasing the  $\text{Na}^+$  concentration from 621 mM to 1.021 M has very little (increase or decrease) or no effect on the RNA  $\Delta G^\circ_{37}$  values. Figure 2 shows the relationship between  $\Delta G^\circ_{37}$  and  $\ln[\text{Na}^+]$  for representative RNA



Table 4. Newly Derived RNA Correction Factors

eq	name	equation	accuracy <sup>c</sup>
RNA $T_m$ Correction Factors			
18	modified SantaLucia $T_m$ equation <sup>a,b</sup>	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + \frac{0.3153N}{\Delta H^\circ} \ln \frac{[Na^+]_2}{[Na^+]_1}$	1.0 °C
19	modified Owczarzy $T_m$ linear equation	$T_m(2) = T_m(1) + (-1.842fGC + 4.314) \ln \frac{[Na^+]_2}{[Na^+]_1}$	0.9 °C
20	modified Owczarzy $1/T_m$ linear equation <sup>b</sup>	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + (2.297fGC - 4.575) \times 10^{-5} \ln \frac{[Na^+]_2}{[Na^+]_1}$	0.9 °C
21	modified Owczarzy $T_m$ quadratic equation	$T_m(2) = T_m(1) + (-1.842fGC + 2.675) \ln \frac{[Na^+]_2}{[Na^+]_1} - 0.7348(\ln^2 [Na^+]_2 - \ln^2 [Na^+]_1)$	0.7 °C
22	modified Owczarzy $1/T_m$ quadratic equation <sup>b</sup>	$\frac{1}{T_m(2)} = \frac{1}{T_m(1)} + (2.297fGC - 2.886) \times 10^{-5} \ln \frac{[Na^+]_2}{[Na^+]_1} + 7.575 \times 10^{-6}(\ln^2 [Na^+]_2 - \ln^2 [Na^+]_1)$	0.7 °C
RNA $\Delta G^\circ_{37}$ Correction Factors			
23	$\Delta G^\circ_{37}$ derived equation	$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) - 310.15 \times \Delta H^\circ \times \left[ (2.297fGC - 2.886) \times 10^{-5} \times \ln \frac{[Na^+]_2}{[Na^+]_1} + 7.575 \times 10^{-6} \times (\ln^2 [Na^+]_2 - \ln^2 [Na^+]_1) \right]$	0.14 kcal/mol
24	$\Delta G^\circ_{37}$ linear equation	$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) + (0.324fGC - 0.765) \ln \frac{[Na^+]_2}{[Na^+]_1}$	0.17 kcal/mol
25	$1/\Delta G^\circ_{37}$ linear equation	$\frac{1}{\Delta G^\circ_{37}(2)} = \frac{1}{\Delta G^\circ_{37}(1)} + (-0.0213fGC + 0.0261) \times 10^{-5} \ln \frac{[Na^+]_2}{[Na^+]_1}$	0.19 kcal/mol
26	$\Delta G^\circ_{37}$ quadratic equation	$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) + (0.324fGC - 0.468) \ln \frac{[Na^+]_2}{[Na^+]_1} + 0.133(\ln^2 [Na^+]_2 - \ln^2 [Na^+]_1)$	0.14 kcal/mol
27	$1/\Delta G^\circ_{37}$ quadratic equation	$\frac{1}{\Delta G^\circ_{37}(2)} = \frac{1}{\Delta G^\circ_{37}(1)} + (-0.0213fGC + 0.016) \times 10^{-5} \ln \frac{[Na^+]_2}{[Na^+]_1} - 0.0045(\ln^2 [Na^+]_2 - \ln^2 [Na^+]_1)$	0.17 kcal/mol
28	modified SantaLucia $\Delta G^\circ_{37}$ equation <sup>a</sup>	$\Delta G^\circ_{37}(2) = \Delta G^\circ_{37}(1) - 0.1016 \times N \ln \frac{[Na^+]_2}{[Na^+]_1}$	0.18 kcal/mol

<sup>a</sup> $N$  is the total number of phosphates in the duplex divided by 2. <sup>b</sup>In this equation,  $T_m$  should be in units of kelvin. <sup>c</sup>As mentioned in Materials and Methods,  $|\Delta T_m|_{ave}$  is used here for evaluating the accuracy of the RNA  $T_m$  correction factors, and  $|\Delta \Delta G^\circ_{37}|_{ave}$  is used here for evaluating the accuracy of the RNA  $\Delta G^\circ_{37}$  correction factors.

oligonucleotides. Similar to what was observed for  $T_m$  in Figure 1, saturation of RNA with sodium ions is also observed here at high sodium ion concentrations. In general, RNA duplexes become more thermally stable as sodium ion concentrations increase until they reach a certain saturation point.

**Theoretical Discussion.** Classical counterion condensation theory proposes the effect of sodium ion concentration on melting temperature by the equation<sup>17,33,34</sup>

$$\frac{dT_m}{d(\ln[Na^+])} = \frac{-\alpha RT_m^2}{\Delta H^\circ} \Delta n \quad (29)$$

or its equivalent form

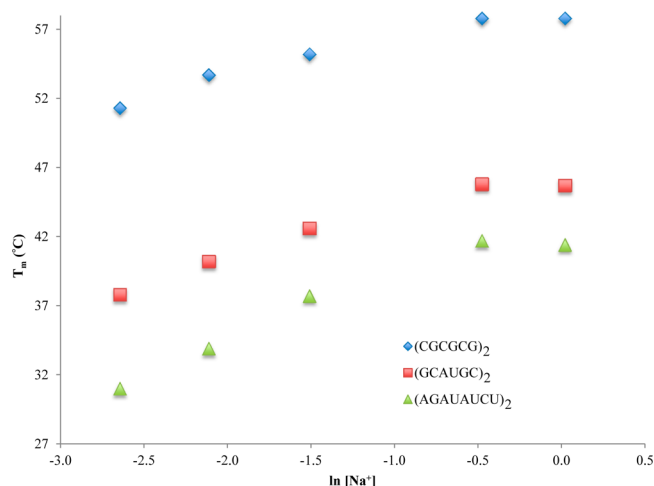
$$\frac{d\left(\frac{1}{T_m}\right)}{d(\ln[Na^+])} = \frac{\alpha R}{\Delta H^\circ} \Delta n \quad (30)$$

where  $\alpha$  is the correction term for the  $Na^+$  activity coefficient,<sup>17,25,33</sup>  $\Delta H^\circ$  is the enthalpy change,  $R$  is the ideal gas constant, and  $\Delta n$  is the

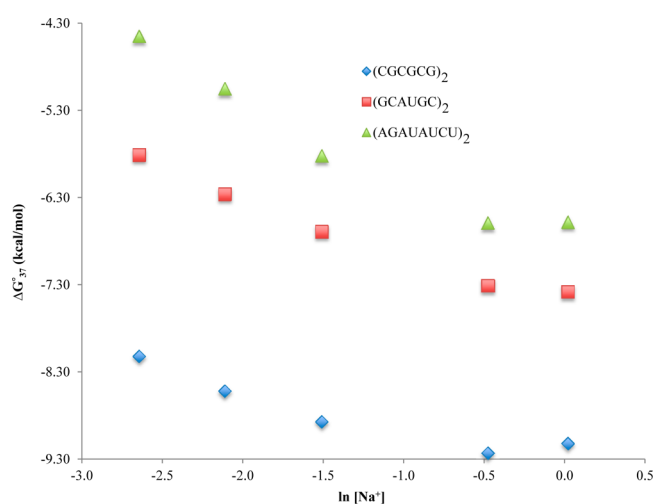
net sodium ion uptake from single strands to a duplex.<sup>17</sup> Duplex RNA has a higher charge density than single strands because the duplex form is more compact than single strands, and this compaction contributes to the uptake of  $Na^+$  during duplex formation.<sup>11</sup> Equations 29 and 30 are the theoretical foundation for deriving the correction factors using the relationship between  $T_m$  or  $1/T_m$  and  $\ln[Na^+]$ .<sup>17</sup>

**$T_m$  Correction Factors.** Previously published DNA  $T_m$  correction factors (Table 3) are mostly based on the functions between melting temperatures (or their reciprocal values) and the common (or natural) logarithm of the sodium ion concentration. The accuracy of 10 previously published correction factors, which range from 1.1 to 7.2 °C, is tested by  $|\Delta T_m|_{ave}$  (Table 3). Under close scrutiny, these correction factors can be sorted into three groups.

The first group consists of three correction factors (eqs 7, 8, and 11) that use only the relationship between  $T_m$  (or  $1/T_m$ ) and  $\ln[Na^+]$  (or  $\log[Na^+]$ ). Equations 7 and 11 are among the most simple correction factors. Although eq 8 is somewhat more



**Figure 1.** Relationship between melting temperature and  $\ln[\text{Na}^+]$  for representative RNA oligomers of different G-C base pair contents: 100% GC, 5'-(CGCGCG)<sub>2</sub>-3'; 66.7% GC, 5'-(GCAUGC)<sub>2</sub>-3'; and 25% GC, 5'-(AGAUAUCU)<sub>2</sub>-3'.



**Figure 2.** Relationship between  $\Delta G^{\circ}_{37}$  and  $\ln[\text{Na}^+]$  for representative RNA oligomers of different G-C base pair contents: 100% GC, 5'-(CGCGCG)<sub>2</sub>-3'; 66.7% GC, 5'-(GCAUGC)<sub>2</sub>-3'; and 25% GC, 5'-(AGAUAUCU)<sub>2</sub>-3'.

complicated, it does not employ any other additional parameters.  $|\Delta T_{\text{m}}|_{\text{ave}}$  values for equations in this group (eqs 7, 8, and 11) are 7.2, 4.5, and 4.2 °C, respectively. It appears as if this type of equation is too simple to accurately describe the relationship between sodium ion concentration and melting temperature. As a result, this type of  $T_{\text{m}}$  correction factor was not pursued further.

The second group is a modification on the first group and consists of only one correction factor. This correction factor (eq 12) introduces two additional parameters into the equation,  $N$  and  $\Delta H^{\circ}$ , to improve the accuracy. Here,  $N$  is half of the total number of phosphates in the duplex, and it is a way to reflect the effect of oligomer length in the correction factor.  $\Delta H^{\circ}$  is the enthalpy change, which could be either an experimental or a predicted value, and is based on the assumption that counterion effects are mainly entropic.<sup>24,30,31</sup> The  $|\Delta T_{\text{m}}|_{\text{ave}}$  value for this equation is 1.4 °C. Because of its accuracy, the coefficients were revised on the basis of the RNA data reported here, resulting in eq 18, with a  $|\Delta T_{\text{m}}|_{\text{ave}}$  value of 1.0 °C (Table 4).

The third group introduces a special parameter into the equations, the fraction of G-C base pairs ( $f_{\text{GC}}$ ). Equations 9, 10, 13, and 14 are linear functions in this group, and eqs 15 and 16 are quadratic functions in this group. The linear functions in this group (eqs 9, 10, 13, and 14) predict RNA  $T_{\text{m}}$  values with  $|\Delta T_{\text{m}}|_{\text{ave}}$  values of 5.1, 6.7, 2.3, and 1.5 °C, respectively. The quadratic functions in this group (eqs 15 and 16) predict RNA  $T_{\text{m}}$  values with  $|\Delta T_{\text{m}}|_{\text{ave}}$  values of 1.5 and 1.1 °C, respectively. These accurate RNA  $T_{\text{m}}$  predictions are not too surprising because these functions that account for  $f_{\text{GC}}$  were previously found to be among the most accurate for predicting DNA  $T_{\text{m}}$  values.<sup>17</sup> Because of their accuracy, the coefficients for the linear and quadratic functions in this group were revised on the basis of the RNA data reported here, resulting in eqs 19–22, with  $|\Delta T_{\text{m}}|_{\text{ave}}$  values of 0.9, 0.9, 0.7, and 0.7 °C, respectively (Table 4). It is important to note that eqs 9 and 10 are the same linear function as eq 13, except for different coefficients. Therefore, when eqs 9, 10, and 13 are revised on the basis of the RNA data reported here, they converge to eq 19.

A previous DNA study<sup>17</sup> took this third type of correction factor even further by expanding it into a more complex form that accounts for sequence dependence by including nearest-neighbor parameters. Because there are 12 unique nearest-neighbor doublets including ends,<sup>35</sup> the expanded version of this correction factor that accounts for nearest neighbors resulted in an increase in the number of fitted parameters from 2 in the linear form of the  $f_{\text{GC}}$  equation to 12 in the linear form of the nearest-neighbor equation and from 3 in the quadratic form of the  $f_{\text{GC}}$  equation to 24 in the quadratic form of the nearest-neighbor equation. Surprisingly, the results of the DNA study show very little improvement in the  $T_{\text{m}}$  prediction. Given its complex form and little improvement in accuracy, the nearest-neighbor version was not investigated here.

In general, eqs 21 and 22 have the best accuracy for predicting RNA  $T_{\text{m}}$  values. They are in quadratic form and could be developed into more complicated forms, such as a cubic or quartic function between  $T_{\text{m}}$  (or  $1/T_{\text{m}}$ ) and  $\ln[\text{Na}^+]$ . However, because  $|\Delta T_{\text{m}}|_{\text{ave}}$  values are already relatively low, and  $T_{\text{m}}$  measurement errors need to be considered at very low  $|\Delta T_{\text{m}}|_{\text{ave}}$  values, we think it is accurate and convenient to utilize eqs 21 and 22 for predictions. Because of the fact that eq 21 (Table 4) is consistent with the  $\Delta G^{\circ}_{37}$  correction factor (discussed below), we recommend it as the  $T_{\text{m}}$  correction factors for RNA in  $\text{Na}^+$  concentrations other than 1.021 M.

Here, we illustrate an example calculation using  $T_{\text{m}}$  correction factor eq 21. We have chosen an independent oligonucleotide, one that was not used in the derivation of the  $T_{\text{m}}$  correction factors proposed here. The example oligonucleotide is 5'-CCAUAUGG-3'/3'-GGUAUACC-5'. Serra et al.<sup>10</sup> measured the  $T_{\text{m}}$  of this oligonucleotide in 0.111 M  $\text{Na}^+$ , and we will use the correction factor to predict this experimental  $T_{\text{m}}$ . Because the experimental  $T_{\text{m}}$  in 1.021 M  $\text{Na}^+$  is not available, we will use the predicted  $T_{\text{m}}$  in 1.021 M  $\text{Na}^+$  based on the standard nearest-neighbor parameters,<sup>9</sup> 51.7 °C. We apply the correction to predict the  $T_{\text{m}}$  at 0.111 M  $\text{Na}^+$ . The correction calculation is shown below:

$$T_{\text{m}}(0.111 \text{ M}) = T_{\text{m}}(1.021 \text{ M}) + (-1.842f_{\text{GC}} + 2.675) \times \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1} - 0.7348(\ln^2 [\text{Na}^+]_2 - \ln^2 [\text{Na}^+]_1) \quad (31)$$

$$T_m(0.111 \text{ M}) = 51.7 + (-1.842 \times 0.5 + 2.675) \ln \frac{0.111}{1.021} - 0.7348(\ln^2 0.111 - \ln^2 1.021) \quad (32)$$

$$T_m(0.111 \text{ M}) = 44.3^\circ\text{C} \quad (33)$$

The  $T_m$  reported in the literature<sup>10</sup> for this oligonucleotide in 0.111 M  $\text{Na}^+$  is 45.7 °C, resulting in a difference of only 1.4 °C between the experimental and predicted temperatures.

**$\Delta G^\circ_{37}$  Correction Factors.** The only DNA  $\Delta G^\circ_{37}$  correction factor available in the literature is eq 17 (Table 3). Similar to some of the  $T_m$  correction factors, it includes  $N$  to account for oligomer length. The  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  of this correction factor is 0.21 kcal/mol. Because of its accuracy, the coefficients were revised on the basis of the RNA data reported here, resulting in eq 28, with a  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  value of 0.18 kcal/mol (Table 4).

Because there was only one DNA  $\Delta G^\circ_{37}$  correction available in the literature, several new  $\Delta G^\circ_{37}$  correction factors were derived and tested. Equation 23 (Table 4) was derived from a  $T_m$  correction factor described above (eq 22). Equation 22 was chosen because it is one of the most accurate  $T_m$  correction factors, and its form is compatible with insertion into eq 3. The  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  of eq 23 is 0.14 kcal/mol.

Other new  $\Delta G^\circ_{37}$  correction factors were derived by simply using the relationship between  $\Delta G^\circ_{37}$  (or  $1/\Delta G^\circ_{37}$ ) and  $\ln[\text{Na}^+]$ . These  $\Delta G^\circ_{37}$  correction factors (eqs 24–27 in Table 4) have similar formats as some of the  $T_m$  correction factors (eqs 19–22 in Table 4). The  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  values for eqs 24–27 are 0.17, 0.19, 0.14, and 0.17 kcal/mol, respectively.

The two  $\Delta G^\circ_{37}$  correction factors having the best accuracy for the RNA data reported here are eqs 23 and 26, both resulting in  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  values of 0.14 kcal/mol. Because both result in the same  $|\Delta\Delta G^\circ_{37}|_{\text{ave}}$  but eq 23 requires an extra parameter ( $\Delta H^\circ$ ), we recommend eq 26 as the  $\Delta G^\circ_{37}$  correction factor for RNA at  $\text{Na}^+$  concentrations other than 1.021 M.

Here, we show an example calculation using  $\Delta G^\circ_{37}$  correction factor eq 26. We have chosen an independent oligonucleotide, one that was not used in the derivation of the  $\Delta G^\circ_{37}$  correction factors proposed here. The example oligonucleotide is 5'-AAGUGAUC-3'/3'-UUCACUAG-5'. Nakano et al.<sup>25</sup> measured the  $\Delta G^\circ_{37}$  of this oligonucleotide in 0.122 M  $\text{Na}^+$ , and we will use the correction factor to predict this experimental  $\Delta G^\circ_{37}$ . Because the experimental  $\Delta G^\circ_{37}$  in 1.021 M  $\text{Na}^+$  is not available, we will use the predicted  $\Delta G^\circ_{37}$  in 1.021 M  $\text{Na}^+$  based on the standard nearest-neighbor parameters,<sup>9</sup> -8.62 kcal/mol. We apply the correction to predict the  $\Delta G^\circ_{37}$  at 0.122 mM  $\text{Na}^+$ . The calculation is shown below:

$$\begin{aligned} \Delta G^\circ_{37}(0.122 \text{ M}) &= \Delta G^\circ_{37}(1.021 \text{ M}) \\ &+ (0.324f\text{GC} - 0.468) \ln \frac{[\text{Na}^+]_2}{[\text{Na}^+]_1} \\ &+ 0.133(\ln^2[\text{Na}^+]_2 - \ln^2[\text{Na}^+]_1) \end{aligned} \quad (34)$$

$$\begin{aligned} \Delta G^\circ_{37}(0.122 \text{ M}) &= -8.62 + (0.324 \times 0.375 - 0.468) \\ &\times \ln \frac{0.122}{1.021} + 0.133(\ln^2 0.122 - \ln^2 1.021) \end{aligned} \quad (35)$$

$$\Delta G^\circ_{37}(0.122 \text{ M}) = -7.30 \text{ kcal/mol} \quad (36)$$

The  $\Delta G^\circ_{37}$  reported in the literature<sup>25</sup> for this oligonucleotide in 0.122 M  $\text{Na}^+$  is -7.26 kcal/mol, resulting in a difference of

only -0.04 kcal/mol between the experimental and predicted free energies.

**Effect of  $\text{Na}^+$  Concentration on  $\Delta H^\circ$  and  $\Delta S^\circ$ .** In the sodium ion concentration range studied here,  $\Delta H^\circ$  is assumed to be independent of  $\text{Na}^+$  concentration.<sup>17,24,30</sup> Figure S1 of the Supporting Information shows the relationship between  $\Delta H^\circ$  and  $\ln[\text{Na}^+]$  for representative oligomers. Considering the proximity of  $\Delta H^\circ$  values in five different sodium ion concentrations and the errors in  $\Delta H^\circ$  in Table S1 and Figure S1 of the Supporting Information, the assumption that  $\Delta H^\circ$  is independent of  $\text{Na}^+$  concentration appears to be valid. Thus, a correction factor for  $\Delta H^\circ$  was not derived.

$\Delta G^\circ_{37}$  and  $T_m$  are typically more accurate than either  $\Delta H^\circ$  or  $\Delta S^\circ$  because of enthalpy–entropy compensation.<sup>9</sup> This is confirmed by both Figure S2 of the Supporting Information, which illustrates the relationship between  $\Delta S^\circ$  and  $\ln[\text{Na}^+]$  for representative oligomers, and the  $\Delta S^\circ$  data in Table S1 of the Supporting Information. Thus, a correction factor for  $\Delta S^\circ$  was not derived.

**Oligomer Length and Sequence.** Previous studies propose different ways to account for the effect of oligomer length and sequence on the stability of DNA at various sodium ion concentrations. SantaLucia et al. incorporate  $N$  into their correction factors for  $T_m$ ,  $\Delta G^\circ_{37}$ , and  $\Delta S^\circ$ .<sup>24</sup> Using  $N$  is a way to account for the effect of oligomer length in the correction factor. However, Owczarzy et al. incorporated only  $f\text{GC}$  into their quadratic correction factor.<sup>17</sup>  $f\text{GC}$  is not a length parameter but rather a sequence-dependent parameter, and the authors state that this correction factor can be used for duplexes ranging from 6 to at least 60 bp in length.<sup>17</sup> Although the correction factor incorporating  $f\text{GC}$  works best for the short RNA duplexes studied here, further studies with longer duplexes are needed to test the accuracy of this correction factor on longer duplexes.

**Range of Sodium Ion Concentrations Appropriate for Correction Factors.** The correction factors derived here were a result of data from RNA melting studies with sodium ion concentrations ranging from 71 mM to 1.021 M. Therefore, it is appropriate to use these correction factors with sodium ion concentrations within this range. Very few experiments are performed in buffers containing >1.021 M  $\text{Na}^+$ , and further studies would need to be done to test the accuracy of the correction factors at these high sodium concentrations. For concentrations below 71 mM  $\text{Na}^+$ , a linear relationship between  $T_m$  and  $\text{Na}^+$  concentration is predicted by counterion condensation theory.<sup>17,33,36</sup> However, the results of a DNA study show that the quadratic form of  $f\text{GC}$  can be used to predict  $T_m$  for <71 mM  $\text{Na}^+$ .<sup>17</sup> Therefore, future work needs to be done to investigate RNA behavior at very low sodium ion concentrations.

**Comparison of Correction Factors to a Generalized Tightly Bound Ion Model.** Tan and Chen<sup>32</sup> previously developed a generalized tightly bound ion (TBI) model to correct RNA  $\Delta G^\circ_{37}$  and  $T_m$  values at 1 M NaCl to other  $\text{Na}^+$  concentrations. In that study, the authors compared their model to a limited data set of experimental data. With the data reported here, a much larger experimental data set is available to compare to their generalized TBI model. Upon comparison of the experimental data reported here to the generalized TBI model, the average difference for  $T_m$  is only 0.97 °C, and the average difference for  $\Delta G^\circ_{37}$  is only 0.16 kcal/mol. Although these differences are slightly larger than the differences resulting from the correction factors derived here, their generalized TBI model works quite well.



## CONCLUSIONS

In summary, the effect of sodium ion concentration on RNA duplex thermal stability was systematically studied. The accuracy of previously published DNA  $T_m$  correction factors and newly derived  $T_m$  correction factors was evaluated using the RNA data obtained here. The newly derived correction factors have higher accuracy than previous correction factors, and eq 21 has the best prediction accuracy, which is 0.7 °C for the RNA data reported here. Similarly, the accuracy of a previously published DNA  $\Delta G^\circ_{37}$  correction factor and newly derived  $\Delta G^\circ_{37}$  correction factors was evaluated using the RNA data obtained here. Equation 26 resulted in an average prediction error of 0.14 kcal/mol for the RNA data reported here and is similar in form to the recommended  $T_m$  correction factor (eq 21). The RNA  $T_m$  (eq 21) and  $\Delta G^\circ_{37}$  (eq 26) correction factors proposed here can be incorporated into RNA secondary structure prediction software to accurately predict  $T_m$  and  $\Delta G^\circ_{37}$  in  $\text{Na}^+$  buffers between 71 mM and 1.021 M.

## ASSOCIATED CONTENT

### Supporting Information

Figures showing the relationship between  $\Delta H^\circ$  and  $\ln[\text{Na}^+]$  and between  $\Delta S^\circ$  and  $\ln[\text{Na}^+]$  for representative RNA oligomers of different G-C base pair contents and a table of experimental RNA thermodynamic parameters for duplex formation. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [znoskob@slu.edu](mailto:znoskob@slu.edu). Phone: (314) 977-8567. Fax: (314) 977-2521.

### Funding

This research was supported by the National Institute of General Medical Sciences of the National Institutes of Health via Grant R15GM085699.

### Notes

The authors declare no competing financial interest.

## REFERENCES

- Thulasi, P., Pandya, L. K., and Znosko, B. M. (2010) Thermodynamic characterization of RNA tri-loops. *Biochemistry* 49, 9058–9062.
- Sheehy, J. P., Davis, A. R., and Znosko, B. M. (2010) Thermodynamic characterization of naturally occurring RNA tetra-loops. *RNA* 16, 417–429.
- Davis, A. R., Kirkpatrick, C. C., and Znosko, B. M. (2011) Structural characterization of naturally occurring RNA single mismatches. *Nucleic Acids Res.* 39, 1081–1094.
- Vanegas, P. L., Hudson, G. A., Davis, A. R., Kelly, S. C., Kirkpatrick, C. C., and Znosko, B. M. (2012) RNA CoSSMos: Characterization of secondary structure motifs—A searchable database of secondary structure motifs in RNA three-dimensional structures. *Nucleic Acids Res.* 40, D439–D444.
- Vanegas, P. L., Horwitz, T. S., and Znosko, B. M. (2012) Effects of non-nearest neighbors on the thermodynamic stability of RNA GNRA hairpin tetraloops. *Biochemistry* 51, 2192–2198.
- Mathews, D. H., and Turner, D. H. (2006) Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* 16, 270–278.
- Mathews, D. H., Sabinampm, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 288, 911–940.
- Badhwar, J., Karri, S., Cass, C. K., Wunderlich, E. L., and Znosko, B. M. (2007) Thermodynamic characterization of RNA duplexes containing naturally occurring 1 × 2 nucleotide internal loops. *Biochemistry* 46, 14715–14724.
- Xia, T., John SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37, 14719–14735.
- Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA* 8, 307–323.
- Manning, G. S. (1978) The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11, 179–246.
- Sharp, K. A., Friedman, R. A., Misra, V., Hecht, J., and Honig, B. (2004) Salt effects on polyelectrolyte–ligand binding: Comparison of Poisson–Boltzmann, and limiting law/counterion binding models. *Biopolymers* 36, 245–262.
- Pack, G. R., Wong, L., and Lamm, G. (1999) Divalent cations and the electrostatic potential around DNA: Monte Carlo and Poisson–Boltzmann calculations. *Biopolymers* 49, 575–590.
- Tan, Z. J., and Chen, S. J. (2006) Nucleic acid helix stability: Effects of salt concentration, cation valence and size, and chain length. *Biophys. J.* 90, 1175–1190.
- Schroeder, S. J., and Turner, D. H. (2009) Optical melting measurements of nucleic acid thermodynamics. *Methods Enzymol.* 468, 371–387.
- Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A., and Walder, J. A. (2008) Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* 47, 5336–5353.
- Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., and Walder, J. A. (2004) Effects of sodiums on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry* 43, 3537–3554.
- Schildkraut, C., and Lifson, S. (1965) Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* 3, 195–208.
- Wetmur, J. G. (1991) DNA probes: Applications of the principles of nucleic acid hybridization. *Crit. Rev. Biochem. Mol. Biol.* 26, 227–259.
- Frank-Kamenetskii, M. D. (1971) Simplification of the empirical relationship between melting temperature of DNA, its GC content and concentration of sodium ions in solution. *Biopolymers* 10, 2623–2624.
- Blake, R. D., and Delcourt, S. G. (1998) Thermal stability of DNA. *Nucleic Acids Res.* 26, 3323–3332.
- Marmur, J., and Doty, P. (1962) Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.* 5, 109–118.
- SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* 35, 3555–3562.
- SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* 95, 1460–1465.
- Nakano, S.-i., Fujimoto, M., Hara, H., and Sugimoto, N. (1999) Nucleic acid duplex stability: Influence of base composition on cation effects. *Nucleic Acids Res.* 27, 2957–2965.
- Wright, D. J., Rice, J. L., Yanker, D. M., and Znosko, B. M. (2007) Nearest neighbor parameters for inosine–uridine pairs in RNA duplexes. *Biochemistry* 46, 4625–4634.
- Christiansen, M. E., and Znosko, B. M. (2009) Thermodynamic characterization of tandem mismatches found in naturally occurring RNA. *Nucleic Acids Res.* 37, 4696–4706.
- Davis, A. R., and Znosko, B. M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry* 46, 13425–13426.
- McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAGGUCUC)<sub>2</sub> by two-dimensional NMR and simulated annealing. *Biochemistry* 35, 14077–14089.



- (30) SantaLucia, J., Jr., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* 33, 415–440.
- (31) Anderson, C. F., Thomas, M., and Record, J. (1995) Salt-nucleic acid interactions. *Annu. Rev. Phys. Chem.* 46, 657–700.
- (32) Tan, Z. J., and Chen, S. J. (2007) RNA helix stability in mixed  $\text{Na}^+/\text{Mg}^{2+}$  solution. *Biophys. J.* 92, 3615–3632.
- (33) Record, M. T., Jr., Anderson, C. F., and Lohman, T. M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association of release, screening and ion effects on water activity. *Q. Rev. Biophys.* 2, 103–178.
- (34) Laing, L. G., Gluick, T. C., and Draper, D. E. (1994) Stabilization of RNA structure by Mg ions specific and non-specific effects. *J. Mol. Biol.* 237, 577–587.
- (35) Gray, D. M. (1997) Derivation of nearest-neighbor properties from data on nucleic acid oligomers. I. Simple sets of independent sequences and the influence of absent nearest neighbors. *Biopolymers* 42, 783–793.
- (36) Record, M. T., Jr., Zhang, W., and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem.* 51, 281–353.