

Molecular Recognition of 6'-N-5-Hexynoate Kanamycin A and RNA 1x1 Internal Loops Containing CA Mismatches[†]

Tuan Tran^{‡,§} and Matthew D. Disney^{*,§}

[‡]Department of Chemistry and The Center of Excellence in Bioinformatics and Life Sciences, University at Buffalo, The State University of New York, Natural Sciences Complex, Buffalo, New York 14260, United States, and

[§]Department of Chemistry, The Scripps Research Institute, Scripps Florida, 130 Scripps Way #3A1, Jupiter, Florida 33458, United States

Received October 27, 2010; Revised Manuscript Received December 14, 2010

ABSTRACT: In our previous study to identify the RNA internal loops that bind an aminoglycoside derivative, we determined that 6'-N-5-hexynoate kanamycin A prefers to bind 1x1 nucleotide internal loops containing C·A mismatches. In this present study, the molecular recognition between a variety of RNAs that are mutated around the C·A loop and the ligand was investigated. Studies show that both loop nucleotides and loop closing pairs affect binding affinity. Most interestingly, it was shown that there is a correlation between the thermodynamic stability of the C·A internal loops and ligand affinity. Specifically, C·A loops that had relatively high or low stability bound the ligand most weakly whereas loops with intermediate stability bound the ligand most tightly. In contrast, there is no correlation between the likelihood that a loop forms a C-A⁺ pair at lower pH and ligand affinity. It was also found that a 1x1 nucleotide C·A loop that bound to the ligand with the highest affinity is identical to the consensus site in RNAs that are edited by adenosine deaminases acting on RNA type 2 (ADAR2). These studies provide a detailed investigation of factors affecting small molecule recognition of internal loops containing C·A mismatches, which are present in a variety of RNAs that cause disease.

RNA has diverse structures that are important for biological function (1–7). Unlike DNA, RNA is single stranded and has secondary and tertiary structures derived from noncanonical pairing interactions. One example of a noncanonical pair is C·A, which can form a variety of structures such as a non-Watson–Crick pair, a wobble C-A⁺ pair, a reverse Hoogsteen pair, and a reverse wobble pair (Figure 1D–G). These structures are energetically different (8). The least stable loops are the non-Watson–Crick (Figure 1D) and the reverse Hoogsteen (Figure 1F) pairs in which no hydrogen bonds form at the Watson–Crick face. The reverse wobble pair (Figure 1G) is found in parallel duplexes (9). C-A⁺ pairs form when A is protonated, leading to a wobble geometry (Figure 1E) (8, 10). The structure is similar to the geometry of wobble GU pairs in an antiparallel duplex (11). Thus, C-A⁺ pairs fit into a helix similarly to wobble GU pairs or Watson–Crick pairs (9, 11, 12).

The function of RNA is often dictated by the presence of C·A pairs. For example, a C·A loop is the preferred substrate of adenosine deaminase acting on RNA type 2 (ADAR2)¹ (13).

ADARs modify the A in the C·A loop to an inosine (13–15). In hepatitis delta virus (HDV), ADAR is utilized to diversify the genome, for example, changing a stop codon to a tryptophan codon (16). Global ADAR substrate analysis reveals that editing is context dependent: the A in 1x1 nucleotide C·A loops is edited more efficiently in 5'-WAK-3' (where W is A or U and K is G or U) than A's in other contexts (17). A variety of inherited diseases are also due to defects in ADAR editing of RNA transcripts. These include editing of the 2C subtype of the serotonin receptor, which has been implicated in contributing to Prader–Willi syndrome (18, 19). A variety of other central nervous system disorders are also caused by malfunctions in RNA editing (20). Since CA loops are often the preferred editing site(s) in RNA, small molecules that target editing sites could provide lead ligands to treat diseases caused by misediting by ADAR. Additionally, mutations in mitochondrial tRNAs that cause neuromuscular disorders are often due to insertion of CA 1x1 nucleotide internal loops (21–24). Thus, many diseases are due to insertion or altered processing of RNAs with 1x1 nucleotide CA internal loops.

Information on the RNA secondary structures that bind small molecules has been previously used to target RNAs that cause disease. For example, information on the preferred binding sites of kanamycin and Hoechst 33258 derivatives have been used to rationally design cell-permeable, modularly assembled ligands targeting the RNAs that cause myotonic dystrophy types 1 and 2, spinocerebellar ataxia type 3, and Huntington's disease (25–28). It is likely that a similar approach could be used to target diseases that are caused by CA loops.

Previously, it was identified that RNAs with C·A loops were one of the preferred targets of 6'-N-5-hexynoate kanamycin A (KanHex) by using a selection-based approach (Figure 2) (29, 30).

[†]This work was funded by the National Institutes of Health (R01-GM079235).

^{*}To whom correspondence should be addressed. E-mail: Disney@scripps.edu. Phone: 561-228-2203. Fax: 561-228-2147.

¹Abbreviations: ADAR, adenosine deaminase acting on RNA; A-site, aminoacyl-tRNA site; BSA, bovine serum albumin; DMSO, dimethyl sulfoxide; dNTP, deoxyribonucleoside triphosphate; HDV, hepatitis delta virus; HEPES, N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid; IL, internal loop; KanHex, 6'-N-5-hexynoate kanamycin A; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; RNA-PSP, RNA privileged space predictor; rRNA, ribosomal RNA; RT-PCR, reverse transcriptase-polymerase chain reaction; SDS, sodium dodecyl sulfate; TBTA, tris(benzyltriazolylmethyl)amine; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride

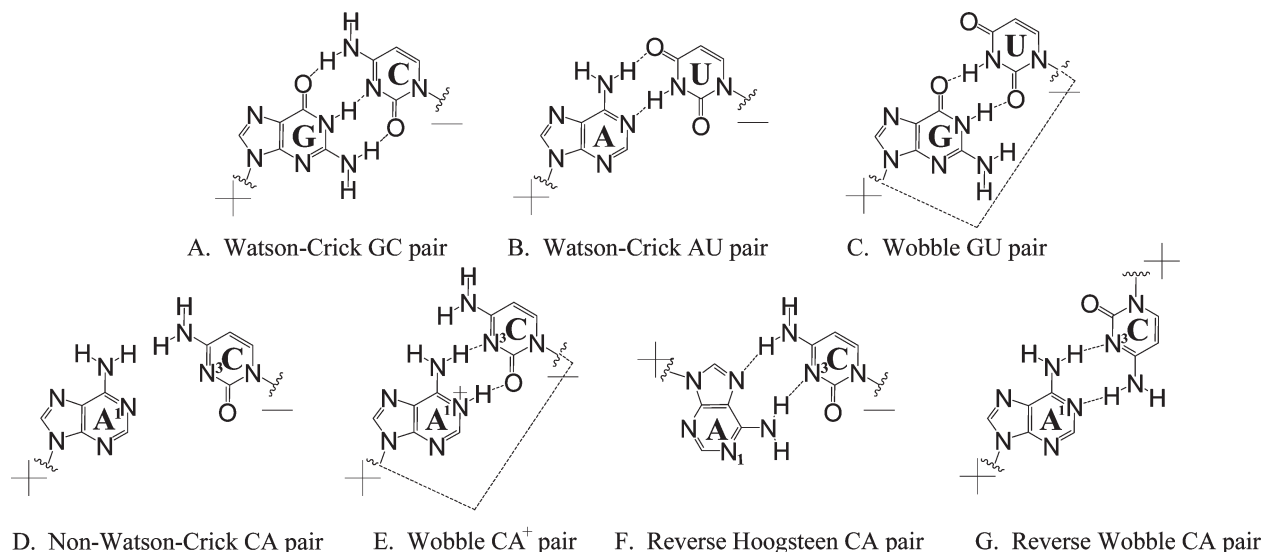


FIGURE 1: Canonical base pairs (A–C) and four possible CA pairs (D–G) in RNA. Plus and minus signs represent the direction of strands (+/- represent antiparallel strands, and ++ represent parallel strands).

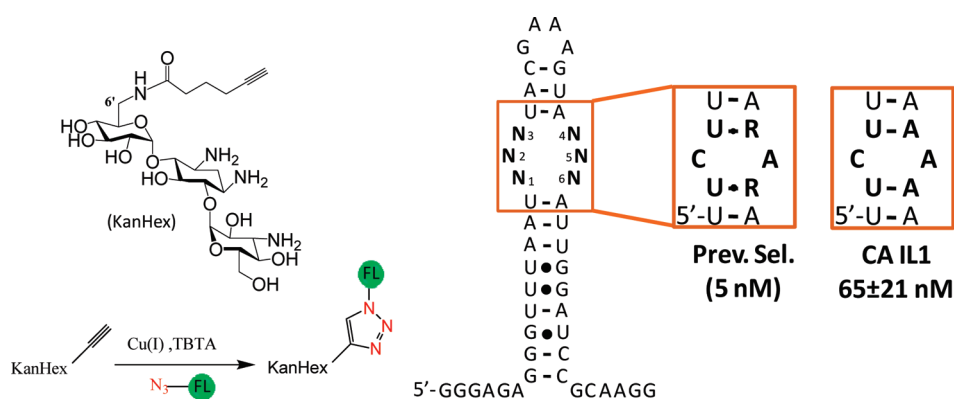


FIGURE 2: The 3x3 nucleotide internal loop library (ILL) selection with 6'-N-5-hexynoate kanamycin A (KanHex) using a microarray platform (30). Prev. Sel. indicates the consensus sequence and the corresponding dissociation constant from a previous selection. CA ILX indicates an internal loop containing a C across from an A; the associated value is the measured K_d from this study.

Selected RNA–KanHex interactions occurred with dissociation constants ranging from 5 to 12 nM at pH 7.0. No loops were selected that were closed by two GC pairs, indicating that the thermodynamic stability of the loop could be an important determinant in high-affinity binding. Various studies have shown that the identity and orientation of loop closing pairs affect loop stability (31–34). We are unaware of any other studies that investigate how ligand affinity is correlated to loop stability.

Herein, the molecular recognition of KanHex by 1x1 nucleotide C·A loops is reported. Specifically, the features in the C·A loops that provide the optimal KanHex binding site were investigated by mutating and deleting loop nucleotides and mutating loop closing pairs. The stabilities of some loops at pH 7.0 and 5.5 were also determined using optical melting. Importantly, there is a correlation between loop stability at pH 7.0 and ligand affinity. That is, loops with intermediate stabilities bind the ligand with the highest affinity. Loops that are relatively more or less stable bind more weakly. In contrast, there is no correlation between the formation of C-A⁺ pairs at lower pH and affinity. It was also found that the consensus RNA loop binder for this ligand is similar to the consensus ADAR2 editing site (13). Thus, there could be similarities in molecular recognition of CA

loops by KanHex and the RNA-binding domain of the ADAR2 protein. These studies could enable the design of compounds that bind specifically to RNAs edited by ADAR2 and modulate editing. Such compounds have not been reported in the literature.

MATERIALS AND METHODS

Aminoglycoside Derivatives. KanHex and the fluorescein-labeled derivatives of all aminoglycosides were synthesized from the parent aminoglycoside as previously described (29, 30).

DNA Templates and PCR Amplification of DNA Templates Encoding Selected RNAs. All DNA templates were purchased from Integrated DNA Technologies, Inc. (IDT) and used without further purification unless noted otherwise. The DNA templates encoding the mutated RNAs were PCR-amplified in 50 μ L of 1 \times PCR buffer (10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.1% Triton X), 4.25 mM MgCl₂, 0.33 mM dNTPs, 2 μ M each primer (5'-CCTTGCGGATCCAAT and 5'-GGCCGGA-TCCTAATACGACTCACTATAGGGAGAGGGTTTAAT), and 0.2 μ L of Taq DNA polymerase. The DNA was amplified by 25 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 1 min. All PCR reactions were analyzed by gel electrophoresis on a 5% agarose gel stained with ethidium bromide.

RNA Transcription and Purification. RNA oligonucleotides were transcribed using an RNAMaxx transcription kit (Stratagene) according to the manufacturer's protocol with 12.5 μL of the amplified DNA from the PCR reaction described above. After transcription, 1 unit of DNase I (Invitrogen) was added, and the sample was incubated for an additional 30 min at 37 °C. The transcribed RNAs were then purified by gel electrophoresis on a denaturing 15% polyacrylamide gel. The RNAs were visualized by UV shadowing and extracted into 300 mM NaCl by tumbling overnight at 4 °C. The resulting solution was concentrated with 2-butanol and ethanol precipitated. The RNAs were dissolved in 150 μL of nanopure water, and the concentrations were determined by their absorbances at 260 nm and the corresponding extinction coefficients. Extinction coefficients were determined using HyTher version 1.0 (35, 36). The parameters used by HyTher are based on information about the extinction coefficients of nearest neighbors in RNA (37).

Fluorescence Binding Assays. Dissociation constants were determined using an in solution, fluorescence-based assay. RNA was folded in HBI (8 mM Na_2PO_4 , pH 7.0, 185 mM NaCl, 0.1 mM EDTA) and 40 $\mu\text{g}/\text{mL}$ BSA at 60 °C for 5 min. The solution was allowed to cool slowly to room temperature. Then, fluorescently labeled KanHex was added to a final concentration of 50 nM. Serial dilutions (1:2) were completed in $1\times$ HBI (HBI supplemented with 40 $\mu\text{g}/\text{mL}$ BSA and 50 nM fluorescently labeled KanHex). The solutions were incubated for 30 min at room temperature and then transferred to wells of a black 96-well plate. Fluorescence intensity was measured on a Bio-Tek FLX-800 plate reader. The change in fluorescence intensity as a function of RNA concentration was fit to eq 1 (38):

$$I = I_0 + 0.5\Delta\epsilon\{([\text{FL}]_0 + [\text{RNA}]_0 + K_t) - (([\text{FL}]_0 + [\text{RNA}]_0 + K_t)^2 - 4[\text{FL}]_0[\text{RNA}]_0)^{0.5}\} \quad (1)$$

where I is the observed fluorescence intensity, I_0 is the fluorescence intensity in the absence of RNA, $\Delta\epsilon$ is the difference between the fluorescence intensity in the absence of RNA and the fluorescence intensity in the presence of infinite RNA concentration, $[\text{FL}]_0$ is the concentration of the fluorescently labeled KanHex, $[\text{RNA}]_0$ is the concentration of the selected internal loop or control RNA, and K_t is the dissociation constant.

Optical Melting Experiments and Determination of Thermodynamic Parameters. Single-stranded oligonucleotides were purchased from IDT and purified by preparative thin-layer chromatography using 55:35:10 (v/v/v) 1-propanol: ammonium hydroxide:water as the mobile phase. The product was identified by UV shadowing and was extracted from the silica by tumbling in NANOpure water for 3 h. All nonself-complementary duplexes were formed by mixing the corresponding oligonucleotides in an equimolar ratio.

Melting experiments were completed in 8 mM Na_2HPO_4 (pH 5.5 or 7.0), 180 mM NaCl, and 0.1 mM Na_2EDTA using a Beckman Coulter DU800 UV-vis spectrometer with an attached Peltier heater. Melting curves of absorbance versus temperature were acquired at 260 nm with a heating rate of 1 °C/min from 12 to 89 °C. Melting curves were fit to a two-state model using the MeltWin program (<http://www.meltwin.com>) (39). The program fits the shape of each curve with sloping base lines and temperature independent ΔH° and ΔS° using a nonlinear least-squares algorithm (39–41). Thermodynamic parameters for the nonself-complementary duplexes were also obtained by plotting

the inverse of the melting temperature of the oligonucleotides duplex (T_M^{-1}) versus $\ln(C_T/4)$. The curves were then fit to eq 2 (42):

$$T_M^{-1} = \frac{R}{\Delta H^\circ} \ln\left(\frac{C_T}{4}\right) + \frac{\Delta S^\circ}{\Delta H^\circ} \quad (2)$$

where R is the gas constant, 1.987 cal $\text{K}^{-1} \text{mol}^{-1}$.

In order to determine if a loop of interest likely forms a C-A⁺ pair at lower pH's, the stabilities of the duplexes were compared at pH 7.0 and 5.5. The difference in free energies, $\Delta\Delta G^\circ_{37, \text{pH}}$, was calculated according to eq 3 (based on the ΔG°_{37} values calculated from eq 2):

$$\Delta\Delta G^\circ_{37, \text{pH}} = \Delta G^\circ_{37, \text{pH } 5.5} - \Delta G^\circ_{37, \text{pH } 7.0} \quad (3)$$

Determination of Loop Stability. The stabilities of the CA internal loops, $\Delta G^\circ_{37 \text{ loop}}$, were calculated using to eq 4 (43):

$$\Delta G^\circ_{37 \text{ loop}} = \Delta G^\circ_{37 \text{ duplex with loop}} - \Delta G^\circ_{37 \text{ duplex without loop}} + \Delta G^\circ_{37 \text{ interrupted base stack}} \quad (4)$$

where $\Delta G^\circ_{37 \text{ duplex with loop}}$ is the free energy for formation of a duplex containing a CA internal loop as calculated from eq 2, $\Delta G^\circ_{37 \text{ duplex without loop}}$ is the free energy for formation of the duplex without the CA internal loop as calculated from eq 2, and $\Delta G^\circ_{37 \text{ interrupted base stack}}$ is the free energy for the nearest neighbor that was disrupted by insertion of the internal loops (40, 44).

RESULTS AND DISCUSSION

We previously reported a two-dimensional combinatorial screening (2DCS) selection in which the RNAs from a 3x3 nucleotide internal loop library that bind 6'-N-5-hexynoate kanamycin A (KanHex) were identified (Figure 2) (30). It was determined that the RNA-binding ligand prefers symmetric internal loops that display a cytosine across from an adenine (C·A) closed by UA or UG base pairs (Figure 2) and binds with a stoichiometry of 1:1 (30). Furthermore, 62.5% of selected loops contain C across from A whereas only 33% of the 4096-membered library contains this motif (30). This represents a significant bias in the selection (p -value = 0.0124, or there is less than a 1.24% chance the loops were selected randomly). The consensus loop that has the highest affinity was a 1x1 nucleotide loop containing a CA mismatch closed by two UA base pairs (5 nM) (30). In addition, no selected internal loop was closed by two CG/GC base pairs. These results suggest that both CA loops and their closing pairs are important for high affinity binding to KanHex. Thus, we studied these factors by determining the affinities of loops with mutated or deleted nucleotides and mutated closing pairs.

The Identity and Position of Loop Nucleotides Are Important for Binding KanHex. The various internal loops and their closing base pairs (Figures 2–4) were inserted into the same RNA cassette used in the 2DCS selection experiment (Figure 2). The starting construct contains a CA mismatch closed by two UA base pairs (CA IL1, Figure 2). CA IL1 binds KanHex with a dissociation constant of 65 nM. The loop was then systematically mutated in order to determine the effect of loop nucleotides and loop closing base pairs on molecular recognition (Figure 3). The first set of mutations generated fully paired RNAs, CA IL2, IL3, and IL4 (Figure 3). There was no change in fluorescence intensity when up to 1 μM CA IL2, IL3, or IL4 was added to 50 nM KanHex, indicating that the fully paired RNAs do not bind the ligand or bind very weakly.

In order to determine if the C and/or the A in the loop are important for the molecular recognition of KanHex, one of the nucleotides or the other was deleted. Deletion of A₅ generates a C₂ bulge (CA IL5, Figure 3), and binding was abolished ($K_d > 1 \mu\text{M}$). Interestingly, deletion of the C₂ residue, which creates an A₅ bulge (CA IL6, Figure 3), resulted in a dissociation constant of 83 nM. Mutating C₂ to A₂ results in a one nucleotide bulge in which the A is on the other side of the helix, or CA IL7 (Figure 3). CA IL7 binds weakly to KanHex with a $K_d > 1 \mu\text{M}$. Switching the orientation of both closing base pairs in CA IL7 to 5'-AU-3' (generating CA IL8) does not rescue binding affinity.

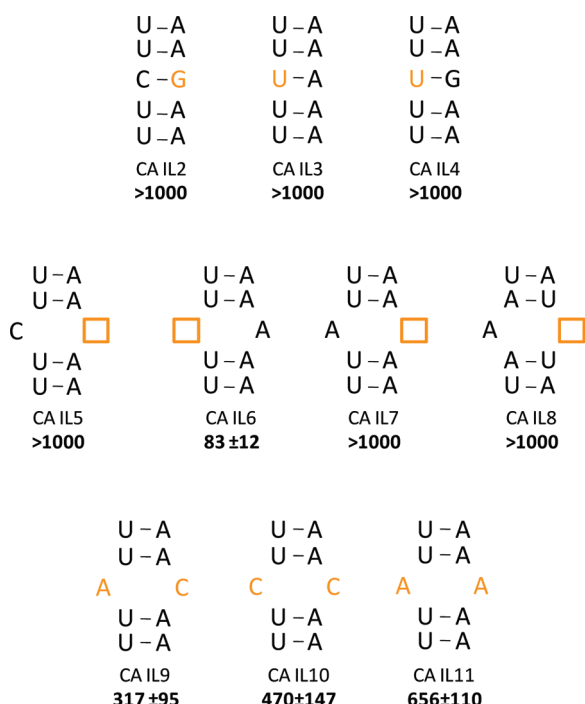


FIGURE 3: The RNA constructs used to determine how the identity and position of loop nucleotides affect binding to KanHex. Nucleotides are derived from the boxed region in Figure 2. Dissociation constants are reported in nanomolar.

CC (CA IL10) and AA (CA IL11) mismatches were also studied. They bind KanHex about 8- and 10-fold more weakly, respectively, than CA IL1. Interestingly, a previous study showed that kanamycin A bound to 1x1 nucleotide CC loops present in the 5' untranslated region (UTR) of thymidylate synthase mRNA with weak dissociation constants ($> 2 \mu\text{M}$) (45). In summary, the identity of the loop nucleotides, particularly A₅, is important for high-affinity binding to KanHex.

Loop Closing Pairs Affect KanHex Binding Affinity. To investigate the importance of C·A loop closing base pairs, they were systematically altered. Of the 36 possible CA mismatches with different closing base pairs, 16 were studied (Figure 4). Series A contains RNAs in which the closing base pairs are mutated to afford U·G or G·U wobble base pairs, series B contains RNAs in which the stability of closing base pair is increased by incorporating one CG or GC base pair, and series C RNAs have two CG or GC closing base pairs.

In series A, the most deleterious mutation switches the orientation of both 5'-UA-3' closing base pairs to 5'-AU-3' (CA IL17) as no binding of KanHex is observed ($K_d > 1000 \text{ nM}$). Similarly, when both closing base pairs were mutated to 5'-GU-3' (CA IL15), a K_d of 261 nM was obtained. Interestingly, if one of the 5'-AU-3' pairs in CA IL17 is replaced with a 5'-UG-3' pair, affording CA IL16, binding is rescued ($K_d = 42 \text{ nM}$). Likewise, if one of the 5'-GU-3' closing pairs in CA IL15 is replaced with 5'-UA-3', then binding affinity is restored (CA IL13). Replacement of one or both 5'-UA-3' pairs with 5'-UG-3' pairs does not significantly affect affinity (CA IL12 and IL14). This suggests that the identity and orientation of closing base pairs are important for binding to KanHex. This may be due to the ligand forming direct contacts to the closing pair or other factors such as electrostatic interactions.

Next, the loop closing pairs were systematically altered to contain one or two CG/GC closing base pairs (series B and C, Figure 4). In series B, only one GC or CG base pair was introduced to the RNA construct, and the highest affinity C·A internal loop was CA IL19 (15 nM). Other mutations were also well tolerated, either not affecting affinity or increasing affinity slightly as compared to the starting construct RNA. The exception is CA

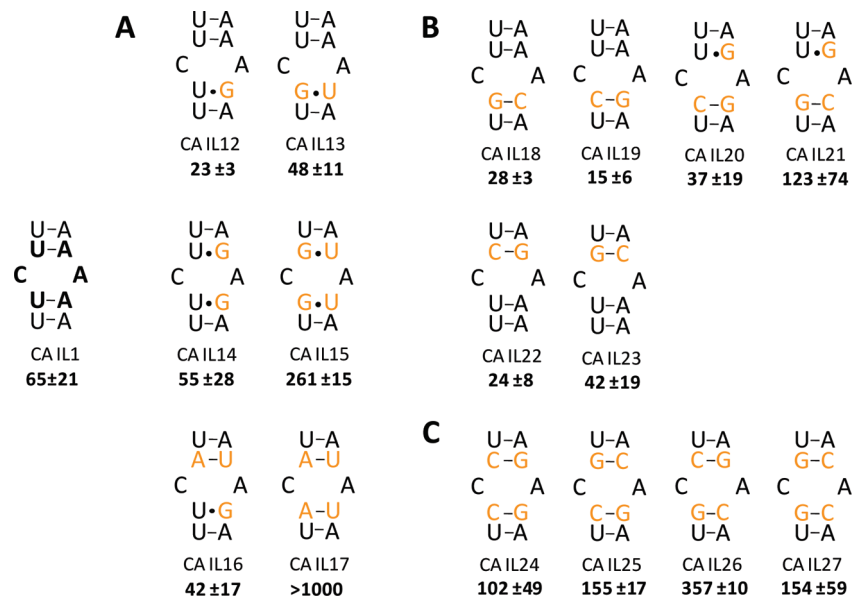
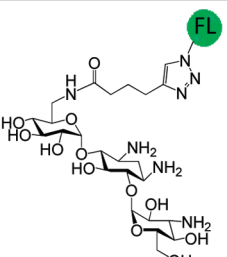
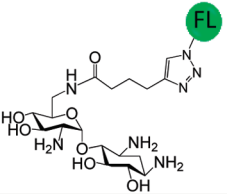
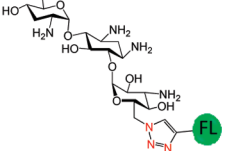
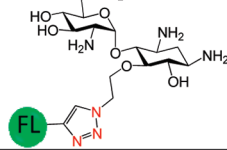
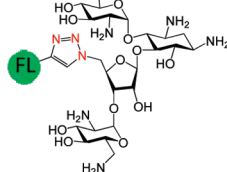


FIGURE 4: The RNA constructs used to determine how loop closing pairs affect binding affinity to KanHex. Nucleotides are derived from the boxed region in Figure 2. Dissociation constants are reported in nanomolar.

IL21, which contains a 5'GC and a 3'UG closing pair and binds with a dissociation constant of ~ 125 nM. In contrast, when both closing pairs are mutated to GC pairs, a 2–5-fold decrease in affinity as compared to the starting construct RNA is observed (series C, Figure 4).

Table 1: Selectivity of CA IL19 for Binding to Fluorescently Labeled Aminoglycosides

Aminoglycoside	K_d (nM)	Selectivity
	15 ± 6	--
	747 ± 23	50
	793 ± 116	53
	>5000	>333
	181 ± 21	12

In summary, the important features in the RNA for binding to KanHex are (1) a purine at position 4, (2) the presence of at least one 5'-UA-3', 5'-UG-3', or 5'-CG-3' closing base pair, and (3) the absence of two 5'-AU-3', 5'-GU-3', or 5'-GC-3' closing base pairs. Evidently, the orientation and identity of the loop closing pairs are important for tight binding of KanHex to internal loops containing CA mismatches.

Aminoglycoside Selectivity for Binding CA IL19. A variety of aminoglycosides are known to interact with RNA, and most often the RNA binds with highest affinity to the aminoglycoside with the greatest number of amino groups (typically neomycin B) (46, 47). Thus, there is often little aminoglycoside specificity in binding RNA targets when the aminoglycosides have a similar number of amino groups. In order to determine if the consensus C·A most specifically binds KanHex over other aminoglycosides, a series of binding assays were completed for CA IL19. The results of these studies are summarized in Table 1.

KanHex is the highest affinity ligand that binds CA IL19; specificities for other aminoglycosides range from 12- to >330 -fold. Not surprisingly, the next highest affinity binder is the neomycin B derivative (with the largest number of amino groups (six)). The largest difference in affinity is observed between KanHex and the neamine derivative which has the smallest number of amino groups. Interestingly, KanHex and the neamine derivative have the same number of amino groups. Thus, the interaction of the KanHex to CA IL19 is due to specific interactions with the structure of the aminoglycosides and is not due to simple charge–charge interactions. Similar specificities have been observed with aminoglycosides binding to RNA internal and hairpin loops when the RNAs have been selected to bind to aminoglycosides via 2DCS (30, 48–51). Other studies have suggested that aminoglycoside recognition is shape-dependent (47). The results presented here suggest that high-affinity binding is also sequence dependent.

Thermodynamic Stability of the C·A Loop Is Correlated with Binding Affinity. The fact that C·A internal loops containing two CG/GC closing base pairs bind more weakly to KanHex suggested that loop stability might play a role in binding affinity. Previous studies have shown that closing pairs affect the stability of RNA 1x1 nucleotide internal loops, with loops closed by two GC pairs being the most stable (32). Additional reports have elucidated other factors that affect stability (33, 34). It is also possible that some C·A loops form wobble C-A⁺ pairs, which

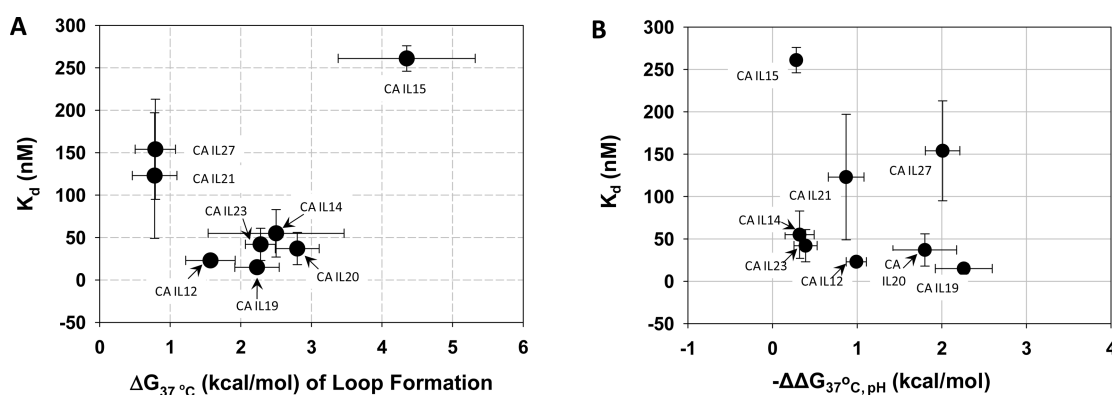


FIGURE 5: (A) Correlation of affinity of C·A internal loops for KanHex and the thermodynamic stabilities of the loops. The sequences for the loops are shown in Table 2. (B) The affinity of C·A internal loops with KanHex as a function of thermodynamic stability at different pHs. $\Delta\Delta G_{37}^{\circ, \text{pH}}$ is calculated by taking the difference between $\Delta G_{37}^{\circ, \text{pH } 5.5}$ and $\Delta G_{37}^{\circ, \text{pH } 7.0}$. There is no correlation between affinity and the likelihood that a loop forms a C-A⁺ pair at lower pH.

Table 2: Thermodynamic Stability of Duplex Formation of RNAs with and without AC Mismatches at pH 7.0 or 5.5

	duplexes with loops									
	average of two-state melt curve fits					linear fit to $1/T_M$ vs $\ln(C_T/4)$ plots				
	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (cal/(K·mol))	$-\Delta G^\circ_{37}$ (kcal/mol)	$-T_M$ (°C)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (cal/(K·mol))	$-\Delta G^\circ_{37}$ (kcal/mol)	$-T_M$ (°C)	$\Delta G^\circ_{37, \text{loop}}$ (kcal/mol)	$\Delta\Delta G_{\text{pH}}$ (kcal/mol)
CA IL12	86.47 ± 10.95	251.39 ± 35.08	8.50 ± 0.08	44.2	99.48 ± 5.69	292.90 ± 18.09	8.63 ± 0.09	43.7	1.57 ± 0.35	0.86
3'-GGCUGGC-5'	90.75 ± 9.27	262.23 ± 29.29	9.42 ± 0.20	47.2	95.23 ± 3.61	276.44 ± 11.35	9.49 ± 0.09	47.0		
CA IL14	79.23 ± 8.40	227.69 ± 26.24	8.61 ± 0.128	45.4	79.22 ± 0.94	227.70 ± 2.98	8.60 ± 0.02	45.3	2.50 ± 0.96	0.14
3'-GGCUGGC-5'	67.62 ± 11.75	191.30 ± 37.51	8.29 ± 0.17	45.3	75.87 ± 7.40	217.35 ± 23.41	8.46 ± 0.22	45.1		
CA IL15	74.33 ± 29.18	217.14 ± 94.63	6.99 ± 0.33	38.9	109.47 ± 3.35	331.05 ± 10.86	6.79 ± 0.02	37.7	4.35 ± 0.97	0.28
3'-GGCUGGC-5'	57.84 ± 14.51	164.18 ± 47.52	6.92 ± 0.40	39.1	119.90 ± 4.21	365.57 ± 13.66	6.51 ± 0.035	37.0		
CA IL19	83.08 ± 13.72	231.99 ± 41.95	11.13 ± 0.72	55.2	79.85 ± 5.27	222.11 ± 16.16	10.97 ± 0.27	55.2	2.23 ± 0.31	2.26
3'-GGCUGGC-5'	90.24 ± 12.15	250.78 ± 36.94	12.46 ± 0.70	58.8	103.27 ± 3.45	290.32 ± 10.47	13.23 ± 0.20	58.5		
CA IL20	82.50 ± 21.80	231.10 ± 66.92	10.82 ± 1.01	54.0	65.69 ± 1.55	179.24 ± 4.77	10.09 ± 0.07	54.8	2.80 ± 0.31	1.80
3'-GGCUGGC-5'	82.38 ± 17.92	228.84 ± 54.69	11.40 ± 0.958	56.5	91.52 ± 6.72	256.73 ± 20.53	11.89 ± 0.37	56.3		
CA IL21	87.21 ± 14.16	142.12 ± 43.51	12.12 ± 0.87	58.2	89.14 ± 3.15	248.01 ± 9.58	12.22 ± 0.18	58.1	0.78 ± 0.32	0.87
3'-GGCUGGC-5'	93.05 ± 11.44	264.20 ± 34.55	13.11 ± 0.72	60.1	94.92 ± 1.76	263.84 ± 5.31	13.09 ± 0.11	60.0		
CA IL23	99.35 ± 12.45	283.05 ± 38.39	11.56 ± 0.55	53.6	119.34 ± 2.47	344.72 ± 7.93	12.43 ± 0.11	53.1	2.28 ± 0.21	0.39
3'-GGCUGGC-5'	81.95 ± 11.80	227.58 ± 35.95	11.36 ± 0.65	56.4	95.40 ± 1.65	268.76 ± 5.05	12.04 ± 0.08	53.6		
CA IL27	75.71 ± 7.75	203.58 ± 23.23	12.57 ± 0.53	63.9	79.10 ± 1.69	213.67 ± 5.08	12.83 ± 0.12	63.8	0.79 ± 0.29	2.01
3'-GGCUGGC-5'	116.26 ± 18.0	320.74 ± 53.22	16.78 ± 1.47	67.0	93.10 ± 2.05	252.34 ± 6.06	14.84 ± 0.16	67.4		

	average of two-state melt curve fits					linear fit to $1/T_M$ vs $\ln(C_T/4)$ plots				
	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (cal/(K·mol))	$-\Delta G^\circ_{37}$ (kcal/mol)	$-T_M$ (°C)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (cal/(K·mol))	$-\Delta G^\circ_{37}$ (kcal/mol)	$-T_M$ (°C)	$-\Delta G^\circ_{37}$ (kcal/mol)	$-T_M$ (°C)
duplex 12	57.52 ± 8.88	152.79 ± 27.01	10.13 ± 0.51	57.7	56.84 ± 1.13	150.53 ± 4.46	10.15 ± 0.06	58.1	10.15 ± 0.06	58.1
3'-GGCUGGC-5'	58.44 ± 10.67	155.43 ± 32.62	10.23 ± 0.57	58.0	53.38 ± 1.18	139.75 ± 3.62	10.04 ± 0.06	58.8	10.04 ± 0.06	58.8
duplex 14	80.12 ± 5.31	223.72 ± 16.37	10.73 ± 0.23	54.2	80.42 ± 2.42	224.64 ± 7.44	10.75 ± 0.11	54.2	10.75 ± 0.11	54.2
3'-GGCUGGC-5'	71.76 ± 1.81	199.3 ± 5.73	9.84 ± 0.18	52.0	92.42 ± 2.20	263.70 ± 6.83	10.63 ± 0.08	51.4	10.63 ± 0.08	51.4
duplex 15	110.48 ± 9.29	320.39 ± 29.02	11.11 ± 0.31	50.4	98.18 ± 3.89	282.15 ± 12.07	10.67 ± 0.14	50.6	10.67 ± 0.14	50.6
3'-GGCUGGC-5'	91.02 ± 6.25	247.13 ± 18.41	14.38 ± 0.54	66.3	102.31 ± 1.86	280.58 ± 5.50	15.28 ± 0.15	66.0	15.28 ± 0.15	66.0
duplex 19	98.78 ± 3.88	271.44 ± 11.64	14.59 ± 0.27	64.6	104.45 ± 2.29	288.38 ± 6.84	15.00 ± 0.17	64.4	15.00 ± 0.17	64.4
3'-GGCUGGC-5'	95.85 ± 8.87	262.87 ± 26.38	14.32 ± 0.67	64.4	111.31 ± 0.89	308.88 ± 2.65	15.51 ± 0.07	64.2	15.51 ± 0.07	64.2
duplex 20	114.69 ± 2.51	316 ± 7.42	16.42 ± 0.21	66.3	119.47 ± 2.07	330.96 ± 6.14	16.82 ± 0.17	66.2	16.82 ± 0.17	66.2
3'-GGCUGGC-5'	91.56 ± 6.41	241.02 ± 18.78	16.81 ± 0.58	76.2	92.39 ± 2.84	243.45 ± 8.33	16.88 ± 0.25	76.1	16.88 ± 0.25	76.1
duplex 21										
duplex 23										
3'-GGCUGGC-5'										
duplex 27										
3'-GGCUGGC-5'										

duplexes without loops

would also contribute to loop stability. To address thermodynamic stability and structural differences of internal loops containing CA mismatches and their effects on binding affinity, optical melting experiments were completed. All 9-mer oligonucleotides duplexes (nonself-complementary) are composed of 5'- and 3'-constant flanking regions, 5'-CCG-3'/3'-GGC-5', and the C·A internal loops with different closing base pairs were inserted between them. The stabilities of the duplexes with internal loops were then compared to the corresponding fully paired duplex to calculate the free energy of loop formation (thermodynamic stability of the loop, eq 3).

The thermodynamic stabilities of eight C·A internal loops were determined. As shown in Table 2, the range of free energy values for loop formation is in good agreement with previous reports (8, 31). C·A internal loops containing CG/GC closing pairs are the most stable (CA IL27, $\Delta G^\circ_{37} = 12.83 \pm 0.12$ kcal/mol). In contrast, C·A internal loops containing two 5'-GU-3' closing base pairs are the least stable (CA IL15, $\Delta G^\circ_{37} = 6.79 \pm 0.02$ kcal/mol). Interestingly, CA IL19 is less stable than CA IL23 by 1.46 kcal/mol; the two RNAs have the same closing pairs but in different orientations. Similarly, CA IL20 and CA IL21 have the same closing pairs but in different orientations, and their stabilities differ by about 1.4 kcal/mol. Hence, the stability of C·A internal loops depends on both the orientation and the identity of the loop closing base pairs, in good agreement with previous reports (31–34).

In general, the internal loops with intermediate affinity bind to KanHex with the highest affinity (Figure 5A). Less stable and more stable loops bind more weakly. The two most stable loops are CA IL21 and CA IL27. These loops bind weakly to KanHex with dissociation constants of 123 and 154 nM, respectively. CA IL15 is the least stable loop studied (destabilizing by 4.35 kcal/mol) and binds to the ligand with a K_d of 261 nM. The stabilities of the other loops studied (CA IL12, IL14, IL19, IL20, and IL23) are intermediate, ranging from 1.57 to 2.80 kcal/mol (Table 2). These loops bind with the highest affinity, having dissociation constants ranging from 15 to 55 nM (Figure 4).

One of the structures that a CA mismatch can form is a wobble C-A⁺ pair (Figure 1E). The pK_a for adenine N1 is 3.5; thus it is usually deprotonated at physiological pH. However, large pK_a shifts have been observed for adenine residues in the leadzyme (52), the hairpin ribozyme (53), the hepatitis delta virus ribozyme (54), and the Varkud satellite ribozyme (55). Previous studies have also shown that the protonation state of adenine in an C·A mismatch is context- and sequence-dependent (8, 10). Since the affinity of the RNA internal loop–ligand complex is correlated to loop stability, we also determined if affinity was correlated with formation of C-A⁺ pairs at lower pHs.

The thermodynamic parameters for the eight CA internal loops that were studied at pH 7.0 were determined at pH 5.5 (Table 2, red). Based on these data, CA IL19 has the largest change in stability, $\Delta\Delta G^\circ_{37, \text{pH}}$, while CA IL15 has the smallest $\Delta\Delta G^\circ_{37, \text{pH}}$. This suggests that C·A mismatch in CA IL19 but not in CA IL15 is likely a C-A⁺ wobble pair at pH 5.5. However, there is no correlation between the affinity of C·A internal loops for KanHex and $\Delta\Delta G^\circ_{37, \text{pH}}$ (Figure 5B) and therefore the likelihood that the loop forms a C-A⁺ pair at lower pHs. For example, CA IL27 ($K_d = 154$ nM) has a similar $\Delta\Delta G^\circ_{37, \text{pH}}$ as CA IL19 ($K_d = 15$ nM) but binds weakly to KanHex. CA IL14 and CA IL23 have similar $\Delta\Delta G^\circ_{37, \text{pH}}$'s as CA IL15 ($K_d = 261$ nM) but bind to KanHex with K_d 's of 55 ± 28 and 42 ± 19 nM, respectively.

Without high resolution structures, it is difficult to determine if the presence of KanHex affects the protonation state of the CA mismatches or vice versa (56). Binding assays completed at lower pHs cannot give insight into this matter since lowering the pH also changes the protonation state of the ligand; both affect binding affinity. Likewise, optical melting experiments completed in the presence of the ligand cannot deconvolute thermodynamic contributions due to binding of ligand and a change in protonation state of the C-A⁺ mismatch if one occurs. Since the highest affinity CA mismatches are those with intermediate stability, it is likely that formation of a C-A⁺ pair would decrease affinity.

SUMMARY

The molecular recognition of KanHex by C·A internal loops was investigated by systematically mutating the loop nucleotides and closing base pairs and studying their affinities. Results revealed that the identity and orientation of the closing pairs affect binding of KanHex as do the identity and position of loop nucleotides. Interestingly, loops with intermediate thermodynamic stability bind most tightly to KanHex while those that are relatively more or less stable bind more weakly. Affinity is not correlated with the likelihood that a loop forms a C-A⁺ wobble pair at lower pHs.

ACKNOWLEDGMENT

We thank Jessica Childs-Disney for assistance with manuscript preparation. M.D.D. dedicates this work to Kaitlyn Disney.

REFERENCES

1. Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862.
2. Zaug, A. J., and Cech, T. R. (1986) The intervening sequence RNA of Tetrahymena is an enzyme. *Science* 231, 470–475.
3. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849–857.
4. Winkler, W. C., Cohen-Chalamish, S., and Breaker, R. R. (2002) An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15908–15913.
5. Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419, 952–956.
6. Gallego, J., and Varani, G. (2001) Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. *Acc. Chem. Res.* 34, 836–843.
7. von Ahlsen, U., and Noller, H. F. (1993) Footprinting the sites of interaction of antibiotics with catalytic group I intron RNA. *Science* 260, 1500–1503.
8. Chen, G., Kennedy, S. D., and Turner, D. H. (2009) A CA(+) pair adjacent to a sheared GA or AA pair stabilizes size-symmetric RNA internal loops. *Biochemistry* 48, 5738–5752.
9. Leontis, N. B., Stombaugh, J., and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isosteric matrices. *Nucleic Acids Res.* 30, 3497–3531.
10. Jang, S. B., Hung, L. W., Chi, Y. I., Holbrook, E. L., Carter, R. J., and Holbrook, S. R. (1998) Structure of an RNA internal loop consisting of tandem C-A⁺ base pairs. *Biochemistry* 37, 11726–11731.
11. Pan, B., Mitra, S. N., and Sundaralingam, M. (1998) Structure of a 16-mer RNA duplex r(GCAGACUAAAUCUGC)₂ with wobble C·A⁺ mismatches. *J. Mol. Biol.* 283, 977–984.
12. Varani, G., and McClain, W. H. (2000) The G × U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.* 1, 18–23.
13. Wong, S. K., Sato, S., and Lazinski, D. W. (2001) Substrate recognition by ADAR1 and ADAR2. *RNA* 7, 846–858.
14. Bass, B. L. (1997) RNA editing and hypermutation by adenosine deamination. *Trends Biochem. Sci.* 22, 157–162.

15. Das, A. K., and Carmichael, G. G. (2007) ADAR editing wobbles the microRNA world. *ACS Chem. Biol.* 2, 217–220.
16. Casey, J. L. (2006) RNA editing in hepatitis delta virus, in *Hepatitis Delta Virus*, pp 67–89, Springer-Verlag, Berlin.
17. Lehmann, K. A., and Bass, B. L. (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. *Biochemistry* 39, 12875–12884.
18. Burns, C. M., Chu, H., Rueter, S. M., Hutchinson, L. K., Canton, H., Sanders-Bush, E., Emeson, R. B., Niswender, C. M., Copeland, S. C., and Herrick-Davis, K. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* 387, 303–308.
19. Morabito, M. V., Abbas, A. I., Hood, J. L., Kesterson, R. A., Jacobs, M. M., Kump, D. S., Hachey, D. L., Roth, B. L., and Emeson, R. B. (2010) Mice with altered serotonin 2C receptor RNA editing display characteristics of Prader-Willi syndrome. *Neurobiol. Dis.* 39, 169–180.
20. Morabito, M. V., and Emeson, R. B. (2009) RNA editing as a therapeutic target for CNS disorders. *Neuropsychopharmacology* 34, 246.
21. Wittenhagen, L. M., and Kelley, S. O. (2002) Dimerization of a pathogenic human mitochondrial tRNA. *Nat. Struct. Biol.* 9, 586–590.
22. Roy, M. D., Wittenhagen, L. M., and Kelley, S. O. (2005) Structural probing of a pathogenic tRNA dimer. *RNA* 11, 254–260.
23. Wittenhagen, L. M., and Kelley, S. O. (2003) Impact of disease-related mitochondrial mutations on tRNA structure and function. *Trends Biochem. Sci.* 28, 605–611.
24. Zifa, E., Giannouli, S., Theotokis, P., Stamatis, C., Mamuris, Z., Stathopoulos, C., Ling, J., Roy, H., Qin, D., Rubio, M. A., Alfonso, J. D., Fredrick, K., Ibba, M., Wani, A. A., Ahanger, S. H., Bapat, S. A., Rangrez, A. Y., Hingankar, N., Suresh, C. G., Barnabas, S., Patole, M. S., and Shouche, Y. S. (2007) Mitochondrial tRNA mutations: clinical and functional perturbations. *RNA Biol.* 4, 38–66.
25. Disney, M. D., Lee, M. M., Pushechnikov, A., and Childs-Disney, J. L. (2010) The role of flexibility in the rational design of modularly assembled ligands targeting the RNAs that cause the myotonic dystrophies. *ChemBioChem* 11, 375–382.
26. Pushechnikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., and Disney, M. D. (2009) Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3. *J. Am. Chem. Soc.* 131, 9767–9779.
27. Lee, M. M., Childs-Disney, J. L., Pushechnikov, A., French, J. M., Sobczak, K., Thornton, C. A., and Disney, M. D. (2009) Controlling the specificity of modularly assembled small molecules for RNA via ligand module spacing: targeting the RNAs that cause myotonic muscular dystrophy. *J. Am. Chem. Soc.* 131, 17464–17472.
28. Lee, M. M., Pushechnikov, A., and Disney, M. D. (2009) Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS Chem. Biol.* 4, 345–355.
29. Disney, M. D., and Childs-Disney, J. L. (2007) Using selection to identify and chemical microarray to study the RNA internal loops recognized by 6'-N-acylated kanamycin A. *ChemBioChem* 8, 649–656.
30. Childs-Disney, J. L., Wu, M., Pushechnikov, A., Aminova, O., and Disney, M. D. (2007) A small molecule microarray platform to select RNA internal loop-ligand interactions. *ACS Chem. Biol.* 2, 745–754.
31. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry* 38, 14214–14223.
32. Davis, A. R., and Znosko, B. M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry* 46, 13425–13436.
33. Davis, A. R., and Znosko, B. M. (2010) Structural characterization of naturally occurring RNA single mismatches. *Nucleic Acids Res.* (in press, doi:10.1093/NAR/gpk793).
34. Davis, A. R., and Znosko, B. M. (2010) Positional and neighboring base pair effects on the thermodynamic stability of RNA single mismatches. *Biochemistry* 49, 8669–8679.
35. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. *Biochemistry* 38, 3468–3477.
36. 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.
37. Puglisi, J. D., and Tinoco, I., Jr. (1989) Absorbance melting curves of RNA. *Methods Enzymol.* 180, 304–325.
38. Wang, Y., and Rando, R. R. (1995) Specific binding of aminoglycoside antibiotics to RNA. *Chem. Biol.* 2, 281–290.
39. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)₂ by two-dimensional NMR and simulated annealing. *Biochemistry* 35, 14077–14089.
40. Xia, T., SantaLucia, J., Jr., 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.
41. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* 22, 256–263.
42. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices. *J. Mol. Biol.* 86, 843–853.
43. Gralla, J., and Crothers, D. M. (1973) Free energy of imperfect nucleic acid helices. 3. Small internal loops resulting from mismatches. *J. Mol. Biol.* 78, 301–319.
44. Mathews, D. H., Sabina, 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.
45. Tok, J. B., Cho, J., and Rando, R. R. (1999) Aminoglycoside antibiotics are able to specifically bind the 5'-untranslated region of thymidylate synthase messenger RNA. *Biochemistry* 38, 199–206.
46. Wong, C. H., Hendrix, M., Priestley, E. S., and Greenberg, W. A. (1998) Specificity of aminoglycoside antibiotics for the A-site of the decoding region of ribosomal RNA. *Chem. Biol.* 5, 397–406.
47. Tor, Y., Hermann, T., and Westhof, E. (1998) Deciphering RNA recognition: aminoglycoside binding to the hammerhead ribozyme. *Chem. Biol.* 5, R277–283.
48. Aminova, O., Paul, D. J., Childs-Disney, J. L., and Disney, M. D. (2008) Two-dimensional combinatorial screening identifies specific 6'-acylated kanamycin A- and 6'-acylated neamine-RNA hairpin interactions. *Biochemistry* 47, 12670–12679.
49. Disney, M. D., Labuda, L. P., Paul, D. J., Poplawski, S. G., Pushechnikov, A., Tran, T., Velagapudi, S. P., Wu, M., and Childs-Disney, J. L. (2008) Two-dimensional combinatorial screening identifies specific aminoglycoside-RNA internal loop partners. *J. Am. Chem. Soc.* 130, 11185–11194.
50. Paul, D. J., Seedhouse, S. J., and Disney, M. D. (2009) Two-dimensional combinatorial screening and the RNA Privileged Space Predictor program efficiently identify aminoglycoside-RNA hairpin loop interactions. *Nucleic Acids Res.* 37, 5894–5907.
51. Tran, T., and Disney, M. D. (2010) Two-dimensional combinatorial screening of a bacterial rRNA A-site-like motif library: defining privileged asymmetric internal loops that bind aminoglycosides. *Biochemistry* 49, 1833–1842.
52. Legault, P., and Pardi, A. (1997) Unusual dynamics and pK_a shift at the active site of a lead-dependent ribozyme. *J. Am. Chem. Soc.* 119, 6621–6628.
53. Guo, M., Spitale, R. C., Volpini, R., Krucinska, J., Cristalli, G., Carey, P. R., and Wedekind, J. E. (2009) Direct Raman measurement of an elevated base pK_a in the active site of a small ribozyme in a precatalytic conformation. *J. Am. Chem. Soc.* 131, 12908–12909.
54. Nakano, S., Chadalavada, D. M., and Bevilacqua, P. C. (2000) General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* 287, 1493–1497.
55. Wilson, T. J., Li, N. S., Lu, J., Frederiksen, J. K., Piccirilli, J. A., and Lilley, D. M. (2010) Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11751–11756.
56. Kaul, M., Barbieri, C. M., Kerrigan, J. E., and Pilch, D. S. (2003) Coupling of drug protonation to the specific binding of aminoglycosides to the A site of 16 S rRNA: elucidation of the number of drug amino groups involved and their identities. *J. Mol. Biol.* 326, 1373–1387.