BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 8787–8794 8787
DOI: 10.1021/bi901038s

Portability of the GN(R)A Hairpin Loop Motif between RNA and DNA†

Joshua M. Blose,‡,∥ Kenneth P. Lloyd,§ and Philip C. Bevilacqua*,‡

‡*Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, and*
§*Department of Biochemistry and Molecular Pharmacology, University of Massachusetts, Worcester, Massachusetts 01605.*
∥*Present Address: School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853.*

Received June 18, 2009; Revised Manuscript Received August 7, 2009

ABSTRACT: Hairpins are common nucleic acid secondary structures that serve many structural and functional roles. Recently, we reported that r(UNCG) and r(GNRA) hairpin families use molecular mimicry and electrostatic factors to attain exceptional thermodynamic stability with a CG closing base pair (cbp). Despite having very different overall folds, these tetraloops present the same functionalities and partial charges to the major groove edge of the CG cbp to achieve stability. Herein, we compare the r(GNRA) tetraloop family to the DNA triloop family d(GNA), which is also exceptionally stable with a CG cbp and possesses the same base pairing between the first and last positions of the loop. Nucleobase and functional group modifications were used to investigate interactions of d(GNA) loops with the cbp, which provided for comparison with similar substitutions in r(GNRA) hairpins. Interruption or deletion of loop−cbp interactions in d(GNA) was consistent with electrostatic interactions identified through nonlinear Poisson−Boltzmann (NLPB) calculations, and loop stability changed in a manner consistent with similar loop−cbp interactions for d(GNA) and r(GNRA) loops. We also compared the relationship of ΔG°_{37} and log[Na^+] for d(GNA) and r(GNRA) loops and found a decreased dependence of stability on salt for both loop families when a CG cbp was present. The similarity of the loop−cbp interactions shows portability of this loop−cbp motif across polymer type and loop size and indicates convergence on similar molecular solutions for stability in RNA and DNA.

Hairpins are the most common RNA secondary structure motifs, play roles in cellular processes such as transcription regulation (*1*−*4*) and RNA interference (*5*−*7*), and are often involved in tertiary contacts in complex RNA structures (*8*−*11*). Tetraloops, the most common loops, are often phylogenetically conserved (*12*−*15*) and can be grouped into families on the basis of sequence and structural similarities (*13*, *16*−*22*). For the r(YNMG) and r(GNRA) families, the loops are exceptionally stable in the context of a CG cbp[1] (*20*, *23*−*26*). These loops with the CG cbp may have played an important role in an RNA world, as their occurrence at variable loops in 16S rRNA is inversely correlated with protein content (*12*, *25*).

DNA hairpins, although not as common as RNA hairpins, also play important structural and biological roles. For example, DNA hairpins have been implicated in telomere replication and deletion mutations (*27*−*29*). The GNA triloop family of DNA hairpins is involved in promoter recognition by N4 virion RNA polymerase, and recent crystal structures reveal base-specific contacts with the triloop (*30*, *31*). In addition, mutation of a GNA sequence from GAA to GCA helps promote hairpin formation to regulate centromere folding and condensation (*32*, *33*), and nuclease resistance of the d(GNA) loops has been applied to in vitro protein synthesis (*34*). Other biological instances of d(GNA) loops exist as well. For example, a stable d(cGAAg) loop caps a hairpin that negatively regulates priming for reverse transcription in duck hepatitis B virus (HBV), and the CG cbp was crucial for the full regulatory effect (*35*). In fact, for heron HBV, which contains a related d(GNAB) motif (where B is C, G, or T), where the "B" position is extruded from the loop (*36*), a GC cbp could only partially restore the regulation with a CG cbp (*35*).

Recently, we reported that electrostatic factors contribute to the stability of r(UNMG) (N is any nucleotide, and M is A or C) and r(GNRA) (R is purine) loops with a CG cbp (*26*). Molecular mimicry of functional groups displayed over the major groove edge of the cbp from loop positions 1 and 4 (L1 and L4) plays critical roles in affording significant stability with a CG cbp.

It has also been demonstrated that exceptional stability is afforded to DNA triloops of the d(GNA) family in the context of a CG cbp (*36*−*40*). Given that d(GNA) loops and the r(GNRA) both possess a sheared GA base pair between the first and last positions of the loop (*17*, *37*, *39*, *41*), we hypothesized that the DNA triloops might interact with the CG cbp in a fashion similar to that of their RNA tetraloop counterparts, and that stacking of functional groups at position L1 with those of the CG cbp would be critical for both loops. This idea is chemically reasonable in that our previous studies on the RNA tetraloops indicated no obvious role for the 2′-OH of L1 in this interaction, as it is involved in hydrogen bonding with L3 and L4 away from the cbp (*16*, *17*).

We previously examined the thermodynamic consequences of cbp swaps and nucleobase and functional group substitutions in the r(GNRA) loop via UV melting experiments and compared these to electrostatic potential outputs from calculations using nonlinear Poisson−Boltzmann (NLPB) theory (*26*). In addition, we determined the salt dependence of the free energy for these

†This work was supported by National Science Foundation Grant MCB-0527102 to P.C.B.

*To whom correspondence should be addressed. Telephone: (814) 863-3812. Fax: (814) 865-2927. E-mail: pcb@chem.psu.edu.

[1]Abbreviations: 7dzG, 7-deazaguanine; B, C, G, or T; C3, three-carbon linker; cbp, closing base pair; L, loop position; M, adenine or cytosine; N, any nucleotide; NLPB, nonlinear Poisson−Boltzmann; $P_{10}E_{0.1}$, 10 mM sodium phosphate and 0.1 mM Na_2EDTA (pH 7.0); R, purine; T_M, melting temperature; Y, pyrimidine.

Published on Web 08/14/2009

stable RNA hairpins. In this study, we compare thermodynamic effects of similar substitutions in the loop and cbp of d(GNA) loops as well as the results of salt-dependent UV melting experiments to those results. Across all of these experiments, we find that effects of substitutions and changing ionic strength are similar for d(GNA) and r(GNRA) loops, and that NLPB calculations qualitatively agree with experiments, suggesting that electrostatic factors contribute similarly to the stability of RNA and DNA sheared GA hairpin loop families with a CG cbp. NLPB calculations did reveal differences in surface potential between RNA and DNA, which suggest that the former loops are enhanced with respect to potential for engaging in tertiary interactions.

MATERIALS AND METHODS

DNA Preparation. DNA hairpins have the general sequence 5′-ggaxGCAx′tcc (stem 1) or 5′gaaxGCAx′ttc (stem 2), where x and x′ are complementary nucleotides in the cbp. Two stems were used in order to tune hairpin stability and thus improve upper baselines in UV melts. In the text, loops will be denoted with "xGCAx′". The 7-deazaguanine (7dzG)-substituted sequences were synthesized at the HHMI-Keck Facility at Yale University (New Haven, CT), and all other sequences were synthesized by Integrated DNA Technologies (Coralville, IA). Oligonucleotides were dialyzed into deionized water containing 10 mM NaCl using a microdialysis system, and the purity of all oligonucleotides was confirmed by analytical TLC and estimated to be >95%.

UV Melting Experiments. Prior to melts, DNA was incubated at 90 °C and cooled to room temperature to renature the hairpins. All melts were performed in $P_{10}E_{0.1}$ [10 mM sodium phosphate and 0.1 mM Na_2EDTA (pH 7.0)] with additional NaCl added as appropriate. UV melts were performed using a Gilford Response II spectrophotometer with absorbance at 260 nm acquired every 0.5 °C. UV melts were performed over temperature ranges of 95−5 and 5−95 °C and provided data consistent with the reversibility of the folding transition. Melt data were fit to a two-state model using sloping baselines and analyzed using a Marquadt algorithm for nonlinear curve fitting in KaleidaGraph version 3.5 (Synergy software). Thermodynamic parameters are averages of at least three independently prepared samples, and the melting temperature (T_M) was found to be independent of DNA concentration over a 10-fold range (from 3 to 30 μM), consistent with hairpin formation.

Structural Models and NLPB Calculations. The d(cGAAg) structural model is from coordinates of Protein Data Bank (PDB) entry 1PQT, conformer 7, nucleotides 2−6 (*42*). Conformer 7 was listed by the authors as the most representative. The r(cGAAAg) structural model was drawn from coordinates of PDB entry 2OIU, chain Q, nucleotides 10−15 (*43*), and the r(gGAAAc) structure was drawn from PDB entry 3E5F, chain A, nucleotides 18−23 (*44*). Details of the NLPB model as well as the finite difference procedure used to calculate electrostatic potentials and free energies have been described previously (*26*, *45*−*49*). Calculation of electrostatic potentials and electrostatic free energies was performed using Qnifft version 2.2 (http://crystal.med.upenn.edu/software.html) with partial charges and radii from the cvff91 parameter set from the Discover force field (MSI). The following parameters were used for all calculations: DNA dielectric constant, 2; solvent probe radius, 1.4 Å; ion exclusion radius, 2.0 Å; solvent dielectric constant, 80; and temperature, 310 K. In the calculations, the loop−cbp structures were placed inside a 65^3 lattice, and electrostatic potentials were calculated using a two-step focusing procedure from 20 to 60% lattice fill. The final resolution was 1.6 grids/Å, and potentials converged to 10^{-4} kT/e with charge neutrality maintained within 3%. NLPB output was visualized using the ABPS plug-in (*50*) in PyMol (http://www.pymol.org/).

RESULTS AND DISCUSSION

Similarity between d(GNA) Triloops and r(GNRA) Tetraloops. Although d(GNA) and r(GNRA) loops differ in polymer type, the number of nucleotides in the loop, and the number and position of loop stacking interactions, both possess a sheared GA base pair involving loop positions 1 and 3 for d(GNA) (*42*) and loop positions 1 and 4 for r(GNRA) (*17*, *41*, *43*) (Figure 1A). Moreover, d(GNA) and r(GNRA) use G1 to interact with the CG cbp and present the same functionalities [imino (N7), keto (O6), imino (N1), and amino (N2)] to the CG cbp (Figure 1B).

As tabulated in Figure 1C, partial charges for loop functionalities, as well as distances to cbp functionalities, are similar for d(GNA) and r(GNRA) and are the same on average as distances between stacked bases in the stem of these structures. These properties are consistent with significant loop−cbp stacking for both loops families. Nonetheless, there is a slight shift in the general position of L1 relative to the cbp in DNA, presumably from the inability of the triloop to compact further. As a result, the d(GNA) loop interacts favorably with the H5−H6 portion of C in the cbp in addition to the major groove edge of the cbp. In addition, the favorable interaction of H21 from the amino at G1 to the O6 of G in the cbp is just 2.66 Å in length, although this is compensated by a repulsive interaction between H1 of G1 and the amino group of C in the cbp, with a closest approach distance of 2.95 Å (Figure 1B). The CG cbp contributes similarly for d(GNA) and r(GNRA) loops (see below), perhaps from compensation between a repulsive interaction and an especially stable one.

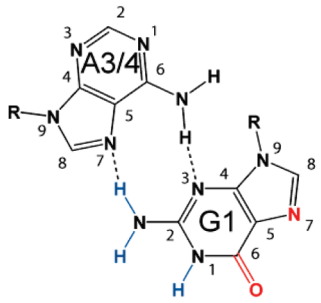
The subsequent sections describe effects of cbp changes, three-carbon (C3) linker insertions, 7dzG substitutions, and increasing salt concentrations on d(GNA) loop stability, with comparison to effects in r(GNRA). Subsequently, these data are interpreted using NLPB models.

Thermodynamic Effects of Changing the Closing Base Pair in d(GNA). Exceptional thermodynamic stability in the context of a CG cbp has been reported previously for both d(GNA) and r(GNRA) loops (*23*, *24*, *26*, *36*−*40*, *51*). To permit direct comparison with substituted loops, we conducted melting experiments with unmodified model d(GCA) loops under the buffer, salt, and stem used herein. These loops are exceptionally stable but slightly less so than their d(GAA) counterparts (*39*), which improves upper baselines from melting experiments. Additionally, to improve the upper baseline in UV melts, we chose a less stable stem (stem 2) for some hairpins (see Materials and Methods, Table 1, and Table S1). It has been shown previously that, outside the cbp, the identity of the stem contributes little to loop stability (*24*). Indeed, we compared the stability of d(cGCAg) loops with two different stems, and after accounting for stem nearest neighbors, obtained loop stabilities within ~0.1 kcal/mol (data not shown), consistent with these expectations.

Sample melts for d(cGCAg) and d(gGCAc) are provided in Figure 2A. For d(GCA), the CG cbp is significantly more stable than its GC cbp counterpart with a $\Delta\Delta G^\circ_{37}$ loop value of 1.7 kcal/mol (Table 1), after considering nearest neighbor effects

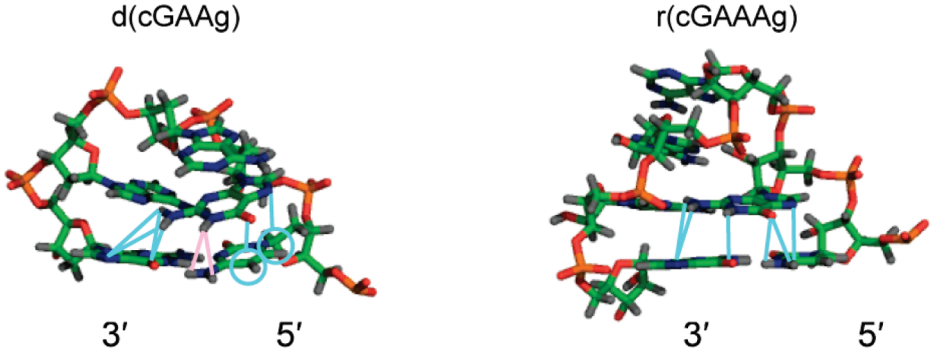
A

Sheared GA Pair



B

Loop Structures



C

Partial Charge Comparison

d(cGAAg)

Loop Nt.	Loop Atom	Loop Atom Charge	CBP Nt.	CBP Atom	CBP Atom Charge	Distance (Å)
G1	N7	-0.5340	C	C6	+0.1850	3.74
				H6	+0.0980	3.73
G1	O6	-0.4580	C	H5	+0.1530	3.27
G1	H1	+0.3360	C	H41	+0.3290	2.95
				H42	+0.3510	3.40
G1	H21	+0.3250	G	O6	-0.4580	2.66
				N7	-0.5340	4.46
G1	H22	+0.3390	G	O6	-0.4580	3.06
				N7	-0.5340	4.16

r(cGAAAg)

Loop Nt.	Loop Atom	Loop Atom Charge	CBP Nt.	CBP Atom	CBP Atom Charge	Distance (Å)
G1	N7	-0.5340	C	H41	+0.3290	3.79
G1	O6	-0.4580	C	H41	+0.3290	3.17
				H42	+0.3510	3.19
G1	H1	+0.3360	G	O6	-0.4580	3.39
G1	H21	+0.3250	G	N7	-0.5340	3.70
G1	H22	+0.3390	G	N7	-0.5340	3.43

FIGURE 1: Similar interactions of d(GNA) and r(GNRA) loops within the loop and with the major groove edge of a CG cbp. (A) Characteristic sheared GA pairing between L1 and L3 or L4 of the loop for the triloop or tetraloop, respectively. Atoms that interact with the CG cbp for both loops are colored according to partial charges: red for negative and blue for positive. Note that all of the interactions come from G of the sheared GA. (B) Three-dimensional structures highlighting interactions between the loop and the cbp with key favorable interactions colored light blue and unfavorable interactions light pink. Light blue circles depict favorable interactions with both the C and its H. The d(cGAAg) structure was from coordinates of PDB entry 1PQT, conformer 7, nucleotides 2−6, and the r(cGAAAg) structure was from coordinates of PDB entry 2OIU, chain Q, nucleotides 10−15. (C) Atoms and partial charges (Discover charge set) of the loop (atoms color coded) and cbp nucleotides. Note that each loop stacks with the major groove edge of the cbp using G at loop position 1. Distances between the loop and cbp functionalities are provided.

in the stem. This value is comparable to the cbp swap in a r(GAAA) loop of 1.4 kcal/mol (Table 1). Thus, the cbp contributes similarly to stability in both d(GNA) and r(GNRA). The slightly more enhanced effect for d(GCA) perhaps could be because in d(GNA) loop−cbp stacking could promote stacking of L2 onto L1 in d(GNA) but not in r(GNRA) (Figure 1B). Notably, the 1.71 kcal/mol cbp stabilization for $\Delta\Delta G^{\circ}_{37}$ loop in d(GCA) is similar to the value of 1.70 kcal/mol for r(UUCG), which has stacking of L3 onto L1 of the loop and is a molecular mimic of r(GNRA) with respect to its loop−cbp interactions (*26*). This observation supports more thermodynamic similarities across RNA and DNA hairpin loops.

Thermodynamic Effects of C3 Linker Interruption of Loop−cbp Interactions in d(GNA). Previously, we used C3 linkers to investigate energetic coupling between positions in nucleic acid hairpins and to gain insight into loop−cbp coupling (*26*, *40*, *51*−*53*). A C3 linker inserted before position 1 of the loop (Table 1, C3 Linker subsection) destabilizes the loop by 1.6 kcal/mol in d(c(C3)GCAg). This large destabilization suggests interruption of strong stacking interactions, consistent with Figure 1B. A similarly large destabilization (1.9 kcal/mol) occurred in r(c(C3)GAAg) (Table 1, right-hand column), supporting similar stacking interactions of L1 with the cbp in both the DNA and RNA motifs.

The 3′-C3 linker insertion is slightly stabilizing in d(cGCA-(C3)g) with a $\Delta\Delta G^{\circ}_{37}$ loop value of −0.37 kcal/mol (Table 1), and the insertion is similarly stabilizing in r(cGAAA(C3)g) with a $\Delta\Delta G^{\circ}_{37}$ loop value of −0.13 kcal/mol. In both the RNA and DNA loops, the A in the sheared GA base pair does not interact appreciably with the CG cbp (Figure 1A,B), again supporting the similarity of the sheared GA-CG cbp motif between DNA and RNA.

Table 1: Thermodynamic Parameters for Hairpin and Loop Formation[a]

sequence	ΔG°_{37} melt[b] (kcal/mol)	stem used	ΔG°_{37} loop[c] (kcal/mol)	$\Delta\Delta G^\circ_{37}$ loop (kcal/mol)	$\Delta\Delta G^\circ_{37}$ loop[d] (rGAAA) (kcal/mol)
			Unmodified		
d(cGCAg)	−2.71 ± 0.05	2	1.03		
d(gGCAc)	−1.68 ± 0.02	1	2.74	1.71[f]	1.38
			C3 Linker		
d(c(C3)GCAg)	−1.12 ± 0.10	2	2.62	1.60[f]	1.87
d(cGCA(C3)g)[e]	−3.92 ± 0.09	1	0.66	−0.37[f]	−0.13
d(g(C3)GCAc)[e]	−0.94 ± 0.06	1	3.48	0.74[g]	1.23
d(gGCA(C3)c)[e]	−2.01 ± 0.09	1	2.41	−0.33[g]	−0.36
			7-dzG		
d(cGCA(7dzg))	−2.45 ± 0.04	2	1.29	0.26[f]	0.84
d((7dzg)GCAc)	−2.29 ± 0.05	1	2.13	−0.61[g]	0.03

[a]All melts were performed in $P_{10}E_{0.1}$, as described in Materials and Methods. [b]ΔG°_{37} values from the UV melt of the hairpin. [c]ΔG°_{37} loop values determined by subtracting the contribution of stem nearest neighbors, which was −4.58 or −4.42 kcal/mol with a CG or GC cbp, respectively, for stem 1 or −3.74 kcal/mol for stem 2 with a CG cbp (*68*). [d]Previously published values of $\Delta\Delta G^\circ_{37}$ loop for similar base pair swaps or substitutions in the r(GAAA) hairpin (*26*). [e]Values have been previously published (*40*, *52*). [f]Values are in comparison to that of d(cGCAg). [g]Values are in comparison to that of d(gGCAc).

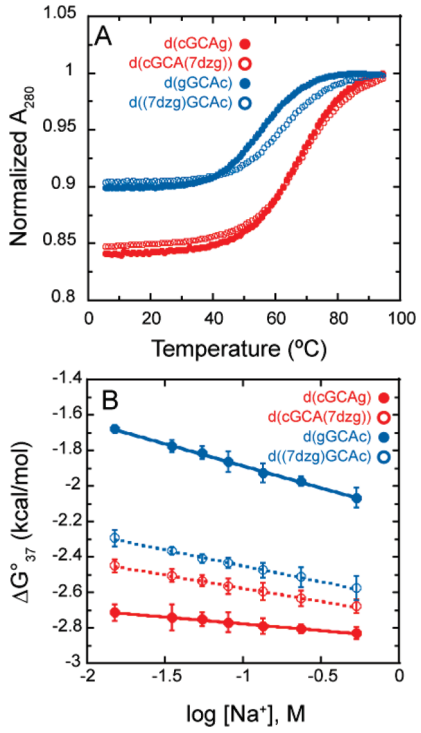


FIGURE 2: Melting curves and salt dependence of the free energy for d(GCA) sequences. (A) Plot of the normalized absorbance at 260 nm vs temperature for representative d(cGCAg), d(gGCAc), and 7dzG-substituted sequences. Only every other data point is shown for the sake of clarity of individual curves. (B) Plot of ΔG°_{37} vs log[Na^+] (M). Slopes and R^2 values for each plot are provided in Table 2. Data for d(cGCAg) are depicted as filled red symbols, data for d(cGCA(7dzg)) as empty red symbols, data for d(gGCAc) as filled blue symbols, and data for d((7dzg)GCAc) as empty blue symbols. Fits to unmodified sequences are solid lines, and fits to 7dzG-modified sequences are dashed lines.

In the context of the GC cbp, the 5′-insertion of the C3 linker in d(g(C3)GCAc) destabilizes the loop by 0.7 kcal/mol, which is less than one-half the loop destabilization for d(c(C3)GCAg). This suggests that the C3 linker interrupts significantly less stable stacking between L1 and the cbp in the context of the GC cbp. In the case of RNA, the 5′-C3 linker insertion is 0.64 kcal/mol less destabilizing for r(g(C3)GAAAc) than for r(c(C3)GAAAg). This observation suggests that stacking of L1 and the cbp is less favorable in the context of the GC cbp in RNA as well. Overall similar effects for 5′-insertion of the C3 linker in the context of CG and GC cbp in DNA and RNA again supports similarity of the loop−cbp motif.

The insertion of the 3′-C3 linker into d(gGCA(C3)g) is stabilizing by −0.33 kcal/mol, which is similar to the stabilization afforded by insertion of the 3′-C3 linker into d(cGCA(C3)g) of −0.37 kcal/mol. This finding supports the model that the A of the sheared GA pair does not significantly interact with the cbp and that the cbp interactions are primarily mediated by L1. Like the insertion of the 3′-C3 linker into DNA, insertion of the C3 linker into RNA is slightly stabilizing, again suggesting that the RNA and DNA loops interact in a similar fashion with the cbp.

Thermodynamic Effects of 7dzG Substitution in the Closing Base Pair in d(GNA). A limited number of functional group substitutions were used previously to examine interactions in d(GNA) and r(GNRA) loops (*26*, *40*, *51*, *52*, *54*). To further probe the energetics of loop−cbp interactions, we substituted 7dzG for G in the cbps. Importantly, this cbp substitution does not change critical hydrogen bonding in the loop and preserves Watson−Crick features in the cbp.

Sample melts for cbp-substituted d(GCA) sequences are provided in Figure 2A and thermodynamic parameters in Table 1 in the 7-dzG subsection. For d(cGCA(7dzg)), there was a slight destabilization of 0.26 kcal/mol as compared to d(cGCAg). We previously melted r(cGAAA(7dzg)) ($\Delta\Delta G = 0.84$ kcal/mol), and although the effect is larger for r(GNRA), the general trend is similar for both loops (Table 1).

In contrast to destabilizing effects of 7dzG in the background of CG cbps, 7dzG substitution in the GC cbp background stabilizes d((7dzg)GCAc) relative to the unmodified sequence by −0.6 kcal/mol. This result suggests that unfavorable interactions between the loop and the less favored GC cbp have been removed. A similar substitution in r((7dzg)GAAAc) appeared to have no thermodynamic effect (0.03 kcal/mol). However, $\Delta\Delta\Delta G^\circ_{37}$ loop for 7dzG substitution in CG versus GC cbp was 0.8 to 0.9 kcal/mol in both DNA and RNA (Table 1).

Table 2: Linear Fit Parameters for ΔG°_{37} vs Log[Na^+] for d(GCA) Triloops[a]

sequence	intercept[b] (kcal/mol)	slope (kcal/mol)	Δslope (kcal/mol)	R^2	slope[c] (rGAAA) (kcal/mol)	Δslope[c] (rGAAA) (kcal/mol)
d(cGCAg) (filled red circles)	−2.85 ± 0.003	−0.08 ± 0.003		0.994	−0.61 ± 0.01	
d(cGCA(7dzg)) (empty red circles)	−2.72 ± 0.004	−0.15 ± 0.003	−0.07 ± 0.004[d]	0.998	−0.60 ± 0.04	0.01 ± 0.04
d(gGCAc) (filled blue circles)	−2.13 ± 0.005	−0.25 ± 0.004	−0.17 ± 0.005[d]	0.998	−0.96 ± 0.03	−0.35 ± 0.03
d((7dzg)GCAc) (empty blue circles)	−2.63 ± 0.005	−0.18 ± 0.005	0.07 ± 0.006[e]	0.997	−0.77 ± 0.04	0.19 ± 0.05

[a]Results are from fits to data plotted in Figure 2B, and the Δslope error is propagated as the square root of the sum of squares from curve fit errors. [b]The *y*-intercept value is defined as ΔG°_{37} extrapolated to 1 M Na^+. [c]The previously published values for a similar substitution in r(GAAA loop) with an identical cbp (*26*). [d]The difference in slopes is relative to d(cGCAg). [e]The difference in slopes is relative to d(gGCAc).

This comparison, which subtracts out stem−stem contributions of 7dzG that might be peculiar to DNA or RNA, further supports similar loop−cbp interaction in DNA and RNA. In the NLPB section below, results of functional group substitutions and C3 linker substitutions will be compared to structural models.

Salt Dependence of ΔG for d(GCA) with a Modified or Unmodified cbp. The salt dependence of hairpin free energy was previously determined for model r(GNRA) loops (*26*). Here, we measure the salt dependencies for d(GNA) loops with CG, GC, and 7dzG-substituted cbps. Hairpin free energies were determined as a function of sodium ion concentration. Results of the UV-monitored, salt-dependent thermal denaturations are shown in Figure 2B, and fitting parameters are provided in Table 2. All loops studied displayed a linear dependence of ΔG°_{37} on log[Na^+] over the measured range of 0.015−0.535 M NaCl (Figure 2B) with reasonable experimental error (Table 2, R^2 values of 0.99). For the d(GNA) loops, the salt dependence was lowest in magnitude for the most stable loop, d(cGCAg). When the cbp was substituted with the destabilizing 7dzG, in d(cGCA(7dzg)), the slope was approximately twice that for the unmodified loop. For d(gGCAc), the least stable loop, its free energy had the greatest salt dependence, which was ~3 times that of d(cGCAg). As mentioned, the effect of the 7dzG substitution is opposite that in the context of a CG cbp which leads the salt dependence of d((7dzg)GCAc) to decrease in magnitude relative to the unsubstituted d(gGCAc) loop.

The observed salt dependencies of the RNA tetraloops and their cbps are generally larger than those observed for DNA triloops (Table 2). For example, slopes ranged from only −0.08 to −0.25 kcal/mol for DNA but from −0.60 to −0.96 kcal/mol for RNA. Also, Δslope has a value of −0.17 kcal/mol for the CG to GC swap in d(GCA) but a value of −0.35 kcal/mol in r(GAAA). The Δslope difference is likely due to the additional phosphate group in the RNA loop (see the next section), while the slope difference has contributions from the stem. Notably, the closest approach of phosphates across the major groove is 5.7 Å for B-form dsDNA, but just 2.7 Å for A-form dsRNA (*55*), leading to greater charge density and dependence on salt for RNA stability. Nonetheless, salt dependencies for hairpin stability are inversely correlated with loop stability for both CG and GC cbps in d(GCA) and r(GAAA) loops, and trends for substituted loops are similar.

Potentials from NLPB Calculations Support Experimental Observations. In the preceding sections, data were presented for effects of changing the cbp, inserting C3 linkers, altering functional groups within the loop and cbp, and increasing the salt concentration on stability of d(GAA) and r(GNRA) loops. In this section, these data are compared to electrostatic potential map outputs from NLPB calculations for structures of d(GNA) and r(GNRA) model loops.

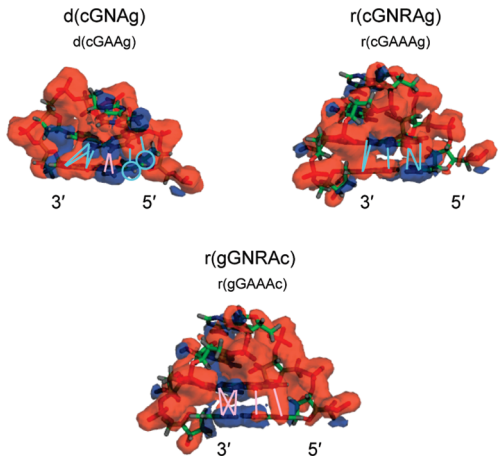


FIGURE 3: Three-dimensional potential contour maps for DNA and RNA loops. Each structure shows the electrostatic potentials calculated at 0.23 M salt and 40% transparency with the underlying structure in stick representation. d(cGAAg) potential contours are −10 kT/e (red) and 10 kT/e (blue), and for r(cGAAAg) and r(gGAAAc), potential contours are −20 kT/e (red) and 10 kT/e (blue). For both loops, most interactions between L1 and the CG cbp contribute to stability, highlighted in light blue, as per Figure 1B. Destabilizing interactions are colored light pink and more prevalent for the r(gGAAAc) than either of the loops with a CG cbp. The d(cGAAg) structure was drawn from coordinates of PDB entry 1PQT, conformer 7, nucleotides 2−6. The r(cGAAAg) structure was drawn from coordinates of PDB entry 2OIU, chain Q, nucleotides 10−15. The r(gGAAAc) structure was drawn from PDB entry 3E5F, chain A, nucleotides 18−23.

Three-dimensional electrostatic potential contours were mapped onto the structures with 40% transparency to reveal the underlying structure in stick representation (Figure 3). The vertical (stacking) interactions between loop and cbp potentials are in concert with the thermodynamic parameters from the UV thermal denaturations of the model DNA and RNA loops with a CG or GC cbp, C3 linker insertions, and functional group substitutions.

For the d(cGAAg) and r(cGAAAg) loops, the favorable loop−cbp electrostatics highlighted in light blue in Figure 1 are also highlighted in light blue in Figure 3 in the context of the three-dimensional potential contour maps. Two potential clashes in the DNA loop are noted in light pink. Notable are the multiple favorable electrostatic interactions of positive (blue) and negative (red) potentials between L1 and the CG cbp for both DNA and RNA loops. For example, O6 of G1 (negative) in both d(cGAAg) and r(cGAAAg) loops interacts with positive potential in the CG cbp. Interaction between L1 and the CG cbp potentials is similar to the favorable electrostatic interaction between the U carbonyl-4 and C amino in the stacking of a 3′-dangling U on a CG base pair in RNA, which contributes −0.5 kcal/mol to stacking (*56*).

The loop−cbp interaction in d(cGAAg) and r(cGAAAg) loops is also similar to that between U1 and the C of the cbp in the r(cUUCGg) loop (*26*). Together, these observations suggest that vertically interacting electrostatic potentials are generally important for stabilizing base stacking in RNA and DNA.

With respect to the GC counterparts, unfortunately a model d(gGNAc) loop is not available in the PDB perhaps because of its inherent flexibility. However, thermodynamic effects of cbp swaps and substitutions suggest more unfavorable interactions with a GC cbp. Indeed, there are considerably more unfavorable loop interactions with like potentials positioned over each other in r(gGAAAc); for example, O6 of G1 in GAAA is repelled by the negative potentials at O6 of G in the GC cbp (Figure 3). On the basis of similarity in the CG−cbp structures, unfavorable stacking of the functional groups of G1 and G of the GC cbp is highly likely in DNA, including interaction of potentials of O6 or N7 of G1 with O6 and N7 of the GC cbp.

Insertions of C3 linker were destabilizing 5′ of the loop in d(c(C3)GNAg) and r(c(C3)GNRAg) loops with penalties of 1.6 and 1.9 kcal/mol, respectively (Table 1). The electrostatic potential contours (Figure 3) reveal that these insertions have the potential to disrupt multiple attractive interactions between L1 and the cbp. The 3′-C3 linker insertions, on the other hand, were slightly stabilizing for both loops, presumably because L4 does not appreciably interact with the cbp. Substitution of 7dzG in the CG cbp is destabilizing for both the DNA triloop and the RNA tetraloop by 0.3 and 0.8 kcal/mol, respectively. Potential maps shown in Figure 3 indicate that 7dzG substitutions would delete attractions between N7 of the cbp and the H2 atoms of G1 of the loop in both DNA and RNA (Figure 3). The difference in the magnitude of this destabilization could be due to the greater distances between interacting groups in DNA (Figure 1C).

For both d(g(C3)GCAc) and r(g(C3)GAAAc), the C3 insertion was destabilizing, with $\Delta\Delta G^{\circ}_{37}$ loop values of 0.7 and 1.2 kcal/mol, respectively, which is at least 0.6 kcal/mol less destabilizing than for the 5′-C3 linker insertion with a CG cbp. Although there is no structure for d(gGNAc), in relation to the r(gGNRAc) structure (Figure 3) there is likely less favorable stacking of the G1 potentials of the cbp in both structures. There are notably many electrostatic clashes (light pink) present in the r(gGAAAc) structure, and the C3 linker insertion probably also relieves these unfavorable interactions in DNA. For example, repulsions between G1 and the cbp, including O6−O6 and N7−N7 interactions, could be interrupted by an insertion of the 5′-C3 linker between the loop and the cbp (Figure 3). Thus, it is likely that for both d(gGNAc) and r(gGNRAc), the C3 linker relieves some unfavorable interactions, giving a net destabilization smaller than that observed with a CG cbp.

In addition to supporting effects of substitutions at the base pair, nucleotide, and functional group levels, the NLPB calculations help explain differences in the salt dependence of hairpin free energy between the DNA and RNA. As shown in Table 2, RNA oligonucleotides have much greater slopes and somewhat greater Δslopes than DNA. The RNA tetraloop has an additional phosphate group compared to the DNA triloop. Examining the surface potentials for both loops on an identical potential scale (Figure 4), one can see that this leads to greater surface potential for the RNA loop, especially near the 5′-side of the loop, which is congruent with divalent metal binding sites (*57*). This negative potential likely also influences tertiary interactions, which are well-known for RNA GAAA loops but not for DNA, even DNAaptamers and DNAzymes.

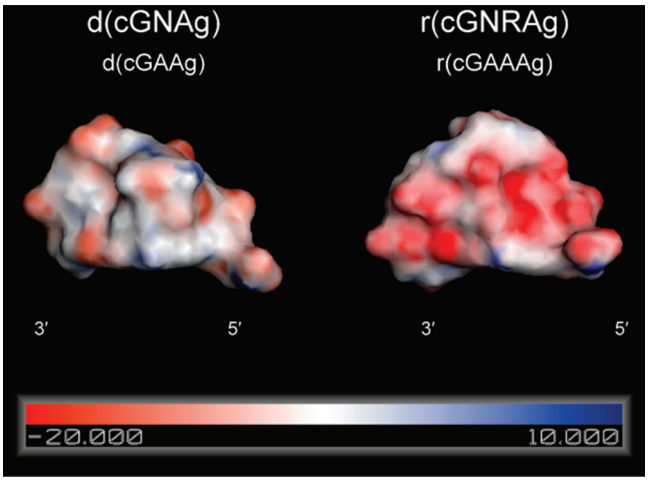


FIGURE 4: Surface potentials of d(cGNAg) and r(cGNRAg) loops. Surface potential maps are drawn with a common scale from NLPB. Maps are calculated using the Discover partial charge parameter set. Note the increase in negative potential at the 5′-side of the loop in the r(cGAAAg) model structure, and the general overall increase in negative surface potential in the RNA tetraloop compared to the DNA triloop.

CONCLUSIONS

Previous studies have established the exceptional stability of some stable RNA tetraloops with a CG cbp (*20*, *23*−*25*), and we previously probed the exceptional stability of the r(UNMG) and r(GNRA) tetraloops using a combination of loop−cbp substitutions and NLPB calculations (*26*). We reported that, in spite of extensive structural differences, these loops are molecular mimics with respect to interaction of functional groups between L1 (and L4) and the major groove edge of the CG cbp (*26*).

The CG cbp has been reported to be critical to the exceptional stability of d(GNA) loops as well (*36*−*40*). It provides stabilization much greater than explained by simple changes to stem nearest neighbors. We sought to compare results between r(GNRA) loops and d(GNA) loops, as both possess a sheared GA base pair whose functional groups interact with the CG cbp. The effects of cbp swaps and substitutions at the nucleobase and functional group levels alter stability in a highly similar fashion in DNA and RNA. Three-dimensional potential contour maps from NLPB calculations support experimental results, showing similar key electrostatic interactions for both loops and revealing how C3 linker insertions and functional group substitutions could interrupt or delete these electrostatic interactions. Moreover, surface potentials for d(GNA) triloops and r(GNRA) tetraloops help explain the difference in magnitude of salt-dependent effects between DNA and RNA. Nonetheless, the loops are not identical. In particular, the DNA loops have much less negative surface potential than their RNA counterparts (Figure 4), consistent with RNA having a greater dependence of stability on salt, as well as potential for tertiary interactions.

Molecular mimicry between r(UNMG) and r(GNRA) tetraloops has been described and likely can be extended to other RNA motifs such as the tandem GA base pair closed by CG base pairs (*58*, *59*). Such recurring motifs form the basis for recent efforts toward modeling RNA (*60*, *61*). With the consideration of d(GNA) loops herein, we can extend stable hairpin loop mimicry from RNA to DNA. Similar motifs in RNA and DNA suggest potential for the latter to form similar structures transiently or permanently in single-stranded DNA in vivo. These types of

interactions could also play significant roles in additional DNA loop motifs such as d(GNNA) and d(GNAB), which are closely related to d(GNA) and prefer a CG cbp to a similar extent (*36*). Molecular mimicry could also stabilize smaller, otherwise unrelated motifs in other nucleic acid-like polymers, including synthetic polymers such as locked nucleic acids (LNA) (*62*−*64*) and possible prebiotic polymers such as threose nucleic acids (TNA) (*65*−*67*).

SUPPORTING INFORMATION AVAILABLE

Full thermodynamic parameters from d(GCA) thermal denaturations. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Zengel, J. M., and Lindahl, L. (1996) A hairpin structure upstream of the terminator hairpin required for ribosomal protein L4-mediated attenuation control of the S10 operon of *Escherichia coli*. *J. Bacteriol. 178*, 2383–2387.
2. Babitzke, P., Schaak, J., Yakhnin, A. V., and Bevilacqua, P. C. (2003) Role of RNA structure in transcription attenuation in *Bacillus subtilis*: The trpEDCFBA operon as a model system. *Methods Enzymol. 371*, 392–404.
3. McGraw, A. P., Bevilacqua, P. C., and Babitzke, P. (2007) TRAP-5′ stem loop interaction increases the efficiency of transcription termination in the *Bacillus subtilis* trpEDCFBA operon leader region. *RNA 13*, 2020–2033.
4. Toulme, F., Mosrin-Huaman, C., Artsimovitch, I., and Rahmouni, A. R. (2005) Transcriptional pausing in vivo: A nascent RNA hairpin restricts lateral movements of RNA polymerase in both forward and reverse directions. *J. Mol. Biol. 351*, 39–51.
5. McManus, M. T., Petersen, C. P., Haines, B. B., Chen, J., and Sharp, P. A. (2002) Gene silencing using micro-RNA designed hairpins. *RNA 8*, 842–850.
6. Boden, D., Pusch, O., Silbermann, R., Lee, F., Tucker, L., and Ramratnam, B. (2004) Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins. *Nucleic Acids Res. 32*, 1154–1158.
7. Ghildiyal, M., and Zamore, P. D. (2009) Small silencing RNAs: An expanding universe. *Nat. Rev. Genet. 10*, 94–108.
8. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science 273*, 1678–1685.
9. Martick, M., and Scott, W. G. (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell 126*, 309–320.
10. Korostelev, A., Trakhanov, S., Laurberg, M., and Noller, H. F. (2006) Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell 126*, 1065–1077.
11. Toor, N., Keating, K. S., Taylor, S. D., and Pyle, A. M. (2008) Crystal structure of a self-spliced group II intron. *Science 320*, 77–82.
12. Wolters, J. (1992) The nature of preferred hairpin structures in 16S-like rRNA variable regions. *Nucleic Acids Res. 20*, 1843–1850.
13. Woese, C. R., Winker, S., and Gutell, R. R. (1990) Architcture of ribosomal RNA: Constraints on the sequence of "tetra-loops". *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467–8471.
14. Gutell, R. R. (1993) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. *Nucleic Acids Res. 21*, 3051–3054.
15. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) A compilation of large subunit (23S and 23S-like) ribosomal RNA structures: 1993. *Nucleic Acids Res. 21*, 3055–3074.
16. Varani, G. (1995) Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct. 24*, 379–404.
17. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol. 264*, 968–980.
18. Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. (2000) The crystal structure of UUCG tetraloop. *J. Mol. Biol. 304*, 35–42.
19. Wu, H., Yang, P. K., Butcher, S. E., Kang, S., Chanfreau, G., and Feigon, J. (2001) A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III. *EMBO J. 20*, 7240–7249.
20. Proctor, D. J., Schaak, J. E., Bevilacqua, J. M., Falzone, C. J., and Bevilacqua, P. C. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry 41*, 12062–12075.
21. Huang, H. C., Nagaswamy, U., and Fox, G. E. (2005) The application of cluster analysis in the intercomparison of loop structures in RNA. *RNA 11*, 412–423.
22. Melchers, W. J., Zoll, J., Tessari, M., Bakhmutov, D. V., Gmyl, A. P., Agol, V. I., and Heus, H. A. (2006) A GCUA tetranucleotide loop found in the poliovirus oriL by in vivo SELEX (un)expectedly forms a YNMG-like structure: Extending the YNMG family with GYYA. *RNA 12*, 1671–1682.
23. Antao, V. P., Lai, S. Y., and Tinoco, I. Jr. (1991) A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res. 19*, 5901–5905.
24. Antao, V. P., and Tinoco, I. Jr. (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res. 20*, 819–824.
25. Williams, D. J., and Hall, K. B. (2000) Experimental and computational studies of the G[UUCG]C RNA tetraloop. *J. Mol. Biol. 297*, 1045–1061.
26. Blose, J. M., Proctor, D. J., Veeraraghavan, N., Misra, V. K., and Bevilacqua, P. C. (2009) Contribution of the closing base pair to exceptional stability in RNA tetraloops: Roles for molecular mimicry and electrostatic factors. *J. Am. Chem. Soc. 131*, 8474–8484.
27. Choi, K. H., and Choi, B. S. (1994) Formation of a hairpin structure by telomere 3′ overhang. *Biochim. Biophys. Acta 1217*, 341–344.
28. Laporte, L., and Thomas, G. J. Jr. (1998) A hairpin conformation for the 3′ overhang of *Oxytricha nova* telomeric DNA. *J. Mol. Biol. 281*, 261–270.
29. Gellert, M. (2002) V(D)J recombination: rag proteins, repair factors, and regulation. *Annu. Rev. Biochem. 71*, 101–132.
30. Gleghorn, M. L., Davydova, E. K., Rothman-Denes, L. B., and Murakami, K. S. (2008) Structural basis for DNA-hairpin promoter recognition by the bacteriophage N4 virion RNA polymerase. *Mol. Cell 32*, 707–717.
31. Murakami, K. S., Davydova, E. K., and Rothman-Denes, L. B. (2008) X-ray crystal structure of the polymerase domain of the bacteriophage N4 virion RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A. 105*, 5046–5051.
32. Chou, S. H., Zhu, L., and Reid, B. R. (1996) On the relative ability of centromeric GNA triplets to form hairpins versus self-paired duplexes. *J. Mol. Biol. 259*, 445–457.
33. Zhu, L., Chou, S. H., and Reid, B. R. (1996) A single G-to-C change causes human centromere TGGAA repeats to fold back into hairpins. *Proc. Natl. Acad. Sci. U.S.A. 93*, 12159–12164.
34. Yoshizawa, S., Ueda, T., Ishido, Y., Miura, K., Watanabe, K., and Hirao, I. (1994) Nuclease resistance of an extraordinarily thermostable mini-hairpin DNA fragment, d(GCGAAGC) and its application to in vitro protein synthesis. *Nucleic Acids Res. 22*, 2217–2221.
35. Habig, J. W., and Loeb, D. D. (2002) Small DNA hairpin negatively regulates in situ priming during duck hepatitis B virus reverse transcription. *J. Virol. 76*, 980–989.
36. Nakano, M., Moody, E. M., Liang, J., and Bevilacqua, P. C. (2002) Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: d(cGNNAg), d(cGNABg),d(cCNNGg), and d(gCNNGc). *Biochemistry 41*, 14281–14292.
37. Hirao, I., Kawai, G., Yoshizawa, S., Nishimura, Y., Ishido, Y., Watanabe, K., and Miura, K. (1994) Most compact hairpin-turn structure exerted by a short DNA fragment, d(GCGAAGC) in solution: An extraordinarily stable structure resistant to nucleases and heat. *Nucleic Acids Res. 22*, 576–582.
38. Sandusky, P., Wooten, E. W., Kurochkin, A. V., Kavanaugh, T., Mandecki, W., and Zuiderweg, E. R. (1995) Occurrence, solution structure and stability of DNA hairpins stabilized by a GA/CG helix unit. *Nucleic Acids Res. 23*, 4717–4725.
39. Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K., and Hirao, I. (1997) GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins. *Biochemistry 36*, 4761–4767.
40. Moody, E. M., and Bevilacqua, P. C. (2003) Thermodynamic coupling of the loop and stem in unusually stable DNA hairpins closed by CG base pairs. *J. Am. Chem. Soc. 125*, 2032–2033.
41. Heus, H. A., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science 253*, 191–194.
42. Padrta, P., Stefl, R., Kralik, L., Zidek, L., and Sklenar, V. (2002) Refinement of d(GCGAAGC) hairpin structure using one- and two-bond residual dipolar couplings. *J. Biomol. NMR 24*, 1–14.

43. Robertson, M. P., and Scott, W. G. (2007) The structural basis of ribozyme-catalyzed RNA assembly. *Science 315*, 1549–1553.
44. Lu, C., Smith, A. M., Fuchs, R. T., Ding, F., Rajashankar, K., Henkin, T. M., and Ke, A. (2008) Crystal structures of the SAM-III/S(MK) riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nat. Struct. Mol. Biol. 15*, 1076–1083.
45. Jayaram, B., Sharp, K. A., and Honig, B. (1989) The electrostatic potential of B-DNA. *Biopolymers 28*, 975–993.
46. Sharp, K. A., and Honig, B. (1990) Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem. 94*, 7684–7692.
47. Honig, B., and Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science 268*, 1144–1149.
48. Misra, V. K., and Honig, B. (1996) The electrostatic contribution to the B to Z transition of DNA. *Biochemistry 35*, 1115–1124.
49. Misra, V. K., and Draper, D. E. (2000) Mg^{2+} binding to tRNA revisited: The nonlinear Poisson-Boltzmann model. *J. Mol. Biol. 299*, 813–825.
50. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037–10041.
51. Moody, E. M., Feerrar, J. C., and Bevilacqua, P. C. (2004) Evidence that folding of an RNA tetraloop hairpin is less cooperative than its DNA counterpart. *Biochemistry 43*, 7992–7998.
52. Moody, E. M., and Bevilacqua, P. C. (2003) Folding of a stable DNA motif involves a highly cooperative network of interactions. *J. Am. Chem. Soc. 125*, 16285–16293.
53. Moody, E. M., and Bevilacqua, P. C. (2004) Structural and energetic consequences of expanding a highly cooperative stable DNA hairpin loop. *J. Am. Chem. Soc. 126*, 9570–9577.
54. SantaLucia, J. Jr., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science 256*, 217–219.
55. Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr. (2000) Nucleic Acids: Structures, Properties, and Functions , University Science Books, Sausalito, CA.
56. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermodynamics of unpaired terminal nucleotides on short RNA helixes correlates with stacking at helix termini in larger RNAs. *J. Mol. Biol. 290*, 967–982.
57. Maderia, M., Horton, T. E., and DeRose, V. J. (2000) Metal interactions with a GAAA RNA tetraloop characterized by ^{31}P NMR and phosphorothioate substitutions. *Biochemistry 39*, 8193–8200.
58. SantaLucia, J. Jr., and Turner, D. H. (1993) Structure of $(rGGCGAGCC)_2$ in Solution from NMR and Restrained Molecular Dynamics. *Biochemistry 32*, 12612–12623.
59. Wu, M., and Turner, D. H. (1996) Solution Structure of $(rGCGGACGC)_2$ by Two-Dimensional NMR and the Iterative Relaxation Matrix Approach. *Biochemistry 35*, 9677–9689.
60. Lemieux, S., and Major, F. (2006) Automated extraction and classification of RNA tertiary structure cyclic motifs. *Nucleic Acids Res. 34*, 2340–2346.
61. Parisien, M., and Major, F. (2008) The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature 452*, 51–55.
62. Braasch, D. A., and Corey, D. R. (2001) Locked nucleic acid (LNA): Fine-tuning the recognition of DNA and RNA. *Chem. Biol. 8*, 1–7.
63. Petersen, M., and Wengel, J. (2003) LNA: A versatile tool for therapeutics and genomics. *Trends Biotechnol. 21*, 74–81.
64. Kaur, H., Babu, B. R., and Maiti, S. (2007) Perspectives on chemistry and therapeutic applications of Locked Nucleic Acid (LNA). *Chem. Rev. 107*, 4672–4697.
65. Schoning, K., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R., and Eschenmoser, A. (2000) Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3′→2′) oligonucleotide system. *Science 290*, 1347–1351.
66. Chaput, J. C., and Szostak, J. W. (2003) TNA synthesis by DNA polymerases. *J. Am. Chem. Soc. 125*, 9274–9275.
67. Ichida, J. K., Zou, K., Horhota, A., Yu, B., McLaughlin, L. W., and Szostak, J. W. (2005) An in vitro selection system for TNA. *J. Am. Chem. Soc. 127*, 2802–2803.
68. 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.