A Uniform Mechanism Correlating Dangling-End Stabilization and Stacking Geometry†

J. Isaksson and J. Chattopadhyaya*

Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden

Received December 9, 2004; Revised Manuscript Received February 4, 2005

ABSTRACT: The geometry of the dangling base in 105 published structures (from X-ray/NMR) containing single-stranded overhangs has been analyzed and correlated to the thermodynamic stabilization found (UV) for the corresponding dangling base/closing basepair combination in short oligonucleotides. The study considers most combinations of closing basepairs, sequence and dangling base residue type, attached in both the 3′- and 5′-ends of both DNA and RNA. Linear regression analysis showed a straightforward correlation ($R = 0.873$) between the degree of screening for the hydrogen bonds of the closing basepair provided by the dangling base and the resulting thermodynamic stabilization in both DNA and RNA series with dangling ends either at the 3′- or at the 5′-terminus. Regression analysis of only the datasets from RNA gives an improved correlation, $R = 0.934$, showing that dangling ends on RNA are more ordered than the dangling ends on DNA, $R = 0.376$. This study highlights the gain in the free energy of stabilization owing to the favorable stacking between the dangling nucleobase and the neighboring basepair and the resulting strengthening of the hydrogen bond of the closing basepair. By acting as a hydrophobic cap on the terminal of the DNA or RNA duplex, the dangling-end residue restricts the bulk water access to the terminal basepair, thereby providing it with a microenvironment devoid of water, which consequently enhances its thermodynamic stability, making it energetically comparable to the corresponding internal basepair. Thus, one single structural model consisting of the interplay of the above electrostatic interactions can be used to explain the molecular basis of the observed thermodynamic effects for dangling-end attachment to the 3′- and 5′-ends of both DNA and RNA duplexes, which is a key step toward accurate dangling-end effect prediction.

Unpaired terminal nucleotides naturally occur in various biologically important RNA structures: The acceptor stem in tRNA (*1*−*3*) has a characteristic 3′-NCCA nucleobase overhang that determines its structure, stability, and function. The dangling ends neighboring the codon−anticodon pair have also been found to stabilize the interaction between tRNA and mRNA (*4*, *5*) for protein synthesis in the ribosomal machinery. The efficiency of down regulation of gene expression by RNA interference (RNAi) has been shown to be dependent on a two to three nucleotide overhang (*6*−*9*). Similarly, the structures of pseudoknots (*10*) are stabilized by a complicated interplay between single-stranded and double-stranded motifs. A multitude of literature on the stabilizing interactions of native single-stranded oligonucleotides is available (*9*, *11*−*14*).

Accurate prediction of nucleic acid secondary structure and stability from its primary sequence is a key interest for probe binding optimization in biotechnological applications such as microarrays (*15*) and molecular beacons (*16*). To develop an accurate algorithm (*14*, *17*−*24*) for predicting the thermodynamics of hybridization and folding of long nucleic acid sequences, the free energy contributions from all natural Watson−Crick basepairs including the nearest-neighbor effects as well as all possible alternative basepairing modes, mismatches, and higher ordered structural motifs such as bulges, loops, and dangling ends must be known.

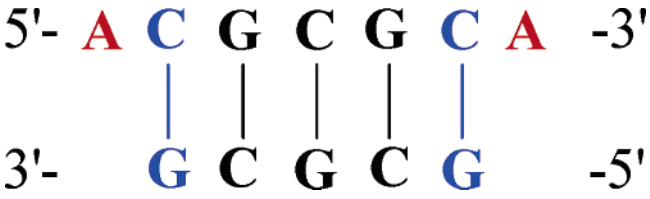

FIGURE 1: Schematic representation of a sequence with 5′- and 3′-*dangling-end* adenosines (red) over C−G *closing basepairs* (blue) attached to the G−C−G core duplex (black).

The dangling-end phenomenon was originally discovered in partially self-complimentary RNA by Martin et al. (*25*, *26*). Since then, the dangling-end effects have been thoroughly and systematically mapped by studying the melting properties of short core duplexes, varying the sequence context by NMR (*27*, *28*) and UV spectroscopy (*29*−*34*). These works have resulted in a full set of thermodynamic parameters for all possible combinations of dangling bases and closing basepairs (Figure 1) attached to short core RNA−RNA duplexes. Bommarito et al. (*35*) have since then extended the database to also include a complete set of thermodynamic parameters for DNA−DNA duplexes.

† Generous financial support from the Swedish Natural Science Research Council (Vetenskapsrådet), the Swedish Strategic Research Foundation (Stiftelsen för Strategisk Forskning), European Union Biotechnology Program (LSHB-CT-2004-005276, Contract No 005276), and Philip Morris USA Inc. is gratefully acknowledged.

* To whom correspondence should be addressed. Phone: +46-18 471 45 77. Fax: +46-18 55 44 95. E-mail: jyoti@boc.uu.se.

10.1021/bi047414f CCC: $30.25
Published on Web 03/17/2005

Attempts have been made to clarify the molecular nature of the dangling-end-promoted stabilization by thermodynamic studies on chemically modified (*36*−*42*) terminal residues with different physicochemical properties. The thermodynamic stabilization by dangling ends in the presence of varying ionic strength (*43*) and different dangling-end lengths (*44*, *45*) has also been explored. Positive trends in dangling-end stabilization have been identified as a result of increasing the size (*39*, *42*), hydrophobicity (*42*), and polarizability (*37*) of the dangling-end part of the molecule. It has also been found that the stability of the dangling-end-containing duplex further depends on the methylation site on the modified nucleobase as well as on the size and electronegativity of the ion substituent (*39*) at the dangling end. However, these studies have so far been unable to provide any general rationale that explains all the observed trends in the native and synthetic dangling-end systems. Moreover, none of the existing studies (*36*−*42*) have so far addressed the key conformational issue if the stabilization of DNA and RNA at the 3′- versus 5′-dangling ends culminate in distinct directional differences in stacking.

To the best of our knowledge, only one structural data mining study (*36*) has been performed so far on the molecules with dangling ends. It shows that the free energies of unpaired nucleotides at the termini of RNA Watson−Crick duplexes correlate with the *probability* that they stack on the duplex or flip out. Stacking in that work (*36*) is however only treated in a two-state on/off model in which stacking is defined as rise ≤ 4 Å, roll/tilt ≤ 30°, and presence of base−base or base−basepair overlap. The stacking geometries are not analyzed in any geometrical detail to show how different bases actually stack with different offsets relative to the closing basepair of different sequences and at different ends of the duplexes. Furthermore, this study only addresses the 3′-effect in RNA sequences and does not address the directional differences in the dangling-end effects between RNA (*29*−*31*, *33*, *34*) and DNA (*35*).

Recent structural studies (*46*) on adenine-rich single-stranded DNA and RNA from this laboratory have demonstrated that they are significantly preorganized in stacked helical forms that are reminiscent of their respective double-stranded helical geometries. The analysis of NOESY footprints and back-calculation of shielding factors show that the single-stranded RNA has a helical stacking pattern that is associated with A-type conformations, while the single-stranded DNA exhibits a pattern that is associated with B-type conformations, leading to considerable directional differences in the nearest-neighbor stacking geometries between the isosequential ssDNA and ssRNA. Partially NMR constrained molecular dynamics (MD) simulations reveal that the electron-rich imidazole of adenine prefers to stack above the electron-deficient pyrimidine of adenine in 5′ → 3′ direction in ssDNA, while, in contrast, the pyrimidine prefers to stack above the imidazole in the 5′ → 3′ direction in ssRNA.

We here report how the interplay of electrostatic and steric interactions between the dangling base and the nearest-neighbor basepair can provide a molecular basis for understanding the differences in the stabilization between the 3′- and 5′-ends in DNA and RNA (*29*−*35*) by correlating the *geometries* of the dangling base stacking (from X-ray and NMR) relative to the closing basepair with the experimental free energies of the corresponding dangling-end sequence from the literature (*29*−*35*).

Table 1: A Total of 105 Structures with Dangling Ends (from X-ray and NMR) Were Used in the Analysis[a]

	3′-end	5′-end
DNA	8	35
RNA	44	18

[a] See databases 1−4 in Supporting Information. The natures of the dangling ends are distributed as follows: 3′-RNA: 44, 5′-RNA: 18, 3′-DNA: 8 and 5′-DNA: 35.

EXPERIMENTAL PROCEDURES

(A) Database Structure Collection. The protein databank (PDB) (*47*) and nucleic acid database (NDB) (*48*) have been surveyed for structures including dangling-end-like motifs, and full details can be found including the relevant identification and citation tags in Supporting Information (databases S1−S4). The NDB search engine was used to find any DNA or RNA duplex structures that extend also to a single-stranded region. Any structure where the dangling base is oriented parallel to the closing basepair was downloaded for analysis together with the attached information from both PDB and NDB. A total of 105 structures were used in the analysis, and the distribution is given in Table 1.

(B) Structural Analysis and Calculation of Screening Factor. The dangling end of each structure together with the two nearest basepairs was extracted from the full structure. A basepairing partner was then built for the dangling base and connected to the phosphate of the closing basepair, without altering any coordinates or minimizing any conformations. The X3DNA package (*49*) was then used to extract all dihedral, helical, and stacking parameters and produce side and top views of the dangling-end conformation.

A procedure similar to the one used to estimate the chemical shielding in single-stranded oligonucleotides (*46*) was used to estimate the influence of the dangling base geometry on the closing basepair. The X−Y displacement of the ring atoms of the dangling base relative to each of the hydrogen bonds of the closing basepair was measured in the program SYBYL (*50*) by moving the dangling base along the *Z*-axis to lie in the same plane as the hydrogen bonds of the closing basepair. The distance from each atom participating in hydrogen bonding to the closest ring atom of the overlapping dangling base was then measured (Figure 2).

The displacement in the X−Y plane (Å) of the atoms in each hydrogen bond (H21−O2, H1−N3, and O6−H41 for G−C and H61−O4 and N1−H3 for A−T/U) from the nearest edge of the dangling nucleobase was then averaged and normalized to a sigmoid function, $y = 1-1/(1+10^{[1-x]/2})$ that has the characteristic on/off behavior centered around an *x*-value of 1 Å, using the scaling factor 0.5. The screening factor of each hydrogen bond (between 0 and 1) was then summed to give a value (between 0 and 3 for G−C closing basepair and between 0 and 2 for A−T/U closing basepair). Thus, the screening value is *proportional* to the sum of how well each individual hydrogen bond of the closing basepair is protected from the bulk solvent by the dangling moiety.

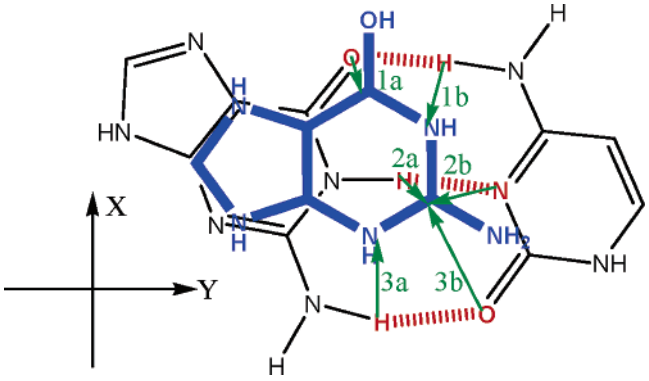

FIGURE 2: Definition of the distances used to calculate the hydrogen bond screening value. The schematic view of the X−Y plane of a G−C closing basepair (black) and a G dangling end (blue) shows the distances 1a, 1b, 2a, 2b, 3a, and 3c used to define the extent of screening. The screening value is defined as the sigmoidal $y_n = 1 - \{1/(1 + 10^{[1-xn]/2})\}$ sum of the average distances of the atoms participating in each hydrogen bond [i.e., $x_n = (1a + 1b)/2$, $(2a + 2b)/2$, and $(3a + 3b)/2$, respectively, for the hydrogen bonds of the GC basepair in the figure and hence the screening $= \sum y_n$].

Thus, the screening factor, *s*, can be written:

$$s = \sum y_n$$

$$y_n = 1 - \frac{1}{1 + 10^{(1 - x_n)/2}}$$

$$x_n = \frac{x_{n_a} + x_{n_b}}{2}$$

where *x* is the distance from the hydrogen bond atom to the closest atom of the dangling base ring (inside the ring produces a negative number, outside produces a positive number), *n* is the index of the hydrogen bond (1, 2, and 3 for G−C and 1 and 2 for A−T/U closing basepairs), and *a* and *b* are indexes for the two atoms forming the hydrogen bond as exemplified in Figure 2.

(C) Thermodynamic Data Collection and Averaging. The thermodynamic data on dangling ends was collected from the systematic studies that were previously made by both salt- and concentration-dependent UV melting studies to be able to predict the stabilization from different overhangs attached to different closing basepairs in DNA (*35*) and RNA (*29*−*34*). The average value was used where multiple values for the same overhang−closing basepair combination but with different core duplexes were available.

RESULTS

(A) Thermodynamic Stabilization of the Duplex by the Dangling End. Thorough systematic probing of the thermodynamic stability added to a DNA or RNA duplex upon attachment of dangling-end residue to the 3′- or 5′-end (*29*−*45*) has provided a complete set of parameters for all possible combinations of single nucleobase overhangs as well as for some mismatches and modified nucleobases. This survey clearly shows that there is a significant difference between the stabilization gained the 3′- and 5′-ends in RNA. Native 3′-dangling ends have a much higher potential ($-\Delta\Delta G^\circ$ = 0.1−1.8 kcal/mol, average: 0.82) for stabilizing duplex formation compared to 5′ RNA dangling ends ($-\Delta\Delta G^\circ$ = 0.0−0.5 kcal/mol, average: 0.22). The directional differences are less pronounced in DNA, but there is a small preference for 5′-DNA dangling ends ($-\Delta\Delta G^\circ = -0.48-0.96$ kcal/mol, average: 0.42) over 3′-dangling ends ($-\Delta\Delta G^\circ = -0.28-0.92$ kcal/mol, average: 0.30).

The recent publication of directional differences in single-strand stacking in ssDNA and ssRNA from our lab (*46*) shows through derived geometries of the adenine−adenine overlaps at each dinucleotide step that the relatively electron-rich imidazole stacks above the relatively electron-deficient pyrimidine in 5′ → 3′ direction in ssDNA, whereas, in contrast, the pyrimidine stacks above the imidazole in the 5′ → 3′ direction in ssRNA (Figure 3A,B).

It is known that the sugar−phosphate backbone torsions in A-RNA and B-DNA families are considerably different and that the essential difference between A- and B-type helices originate from their different puckering modes. Thus, the nucleobase in ssDNA (B-type DNA) with 2′-endo-3′-exo conformation takes up a pseudoequatorial orientation, whereas the nucleobase in ssRNA (A-type RNA) with 2′-exo-3′-endo conformation takes up a pseudoaxial orientation. These subtle torsional differences are finally translated into the distinctive helical and stacking characteristics of the A-type and B-type families. This provides a possible explanation for the previously observed (*29*−*31*, *33*) directional differences in dangling-end stabilization in DNA and RNA. In a canonical RNA duplex where the overhanging base is allowed to stack according to its preferred nonbonded orientation described above (*46*), the nucleobase tends to stack very well over the hydrogen bonds of the closing basepair at the 3′-end, while the same base attached as a dangling end to the 5′-end is steered away from the hydrogen bonds (exemplified by the dangling base geometries in Figure 3D,F). The situation is however reversed for DNA duplex, where a dangling nucleobase in the 5′-end is known (*35*) to stack better than the same nucleobase attached to the 3′-end considering the single-strand stacking alone (exemplified by the dangling base geometries in Figure 3C,E). This has led us to reassess the reported dangling-end effects and qualitatively relate them to the structural observations on the stacking geometries of nonbonded nucleobases in DNA and RNA.

(B) Stabilization of the RNA−RNA Duplex by the Dangling End. It has been established (*29*−*31*, *33*) that 3′-dangling ends on RNA−RNA duplexes hold the highest potential for stabilization ($-\Delta\Delta G^\circ = 0.1-1.8$ kcal/mol). Consequently, that is where it is possible to observe the clearest trends in the stacking geometries:

(1) *It is clear that the size and shape of the dangling-end residue is of primary interest.* When stacking on the same closing basepair, purines (with electron-deficient pyrimidine ring fused with the electron-rich imidazole ring) as a dangling end always provide more stability than pyrimidines. For instance, the $\begin{matrix}5'\text{-CG-}3'\\3'\text{-G−-}5'\end{matrix}$ overhang stabilizes the duplex by 1.7 kcal/mol, while the $\begin{matrix}5'\text{-CU-}3'\\3'\text{-G−-}5'\end{matrix}$ overhang only stabilizes the same duplex by 1.2 kcal/mol (*31*). It has also been shown that by appropriately attaching multiring fused systems [4-desmethylwyosine (*39*), quinolones (*38*), phenazinium (*51*, *52*), pyrene and phenanthrene (*42*), phenylurea and naphthylurea (*41*), cholic acid (*53*, *54*) and a 5′-acylamido

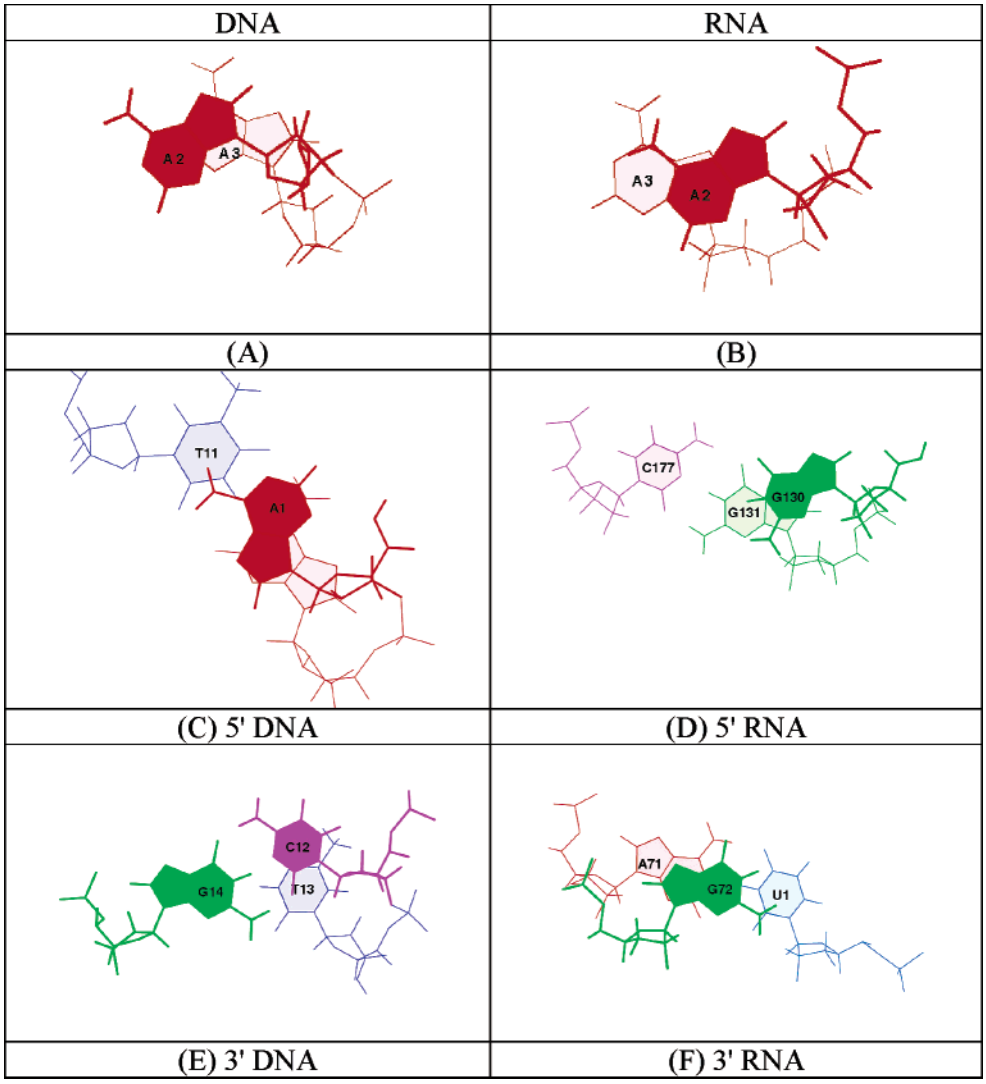

FIGURE 3: Panels A and B show the previously published directionally different stacking geometries of ssDNA (A) and ssRNA (B) [viewed from the 5′-end]. Panels C and D show typical geometries of a 5′-DNA and a 5′-RNA overhang, respectively [viewed from the 5′-end], while panels E and F show typical geometries of a 3′-DNA [viewed from the 5′-end] and a 3′-RNA overhang, respectively [viewed from the 3′-end]. The dangling residues as overhanging bases (panels C−F) also maintain their preferred double-stranded conformation, and hence the dangling-end stacking geometries are consistent with the recently reported (*46*) directional differences between the 3′- and 5′-ends in single-stranded DNA (panel A) and RNA (panel B).

oligonucleotide library (*55*)] to both DNA and RNA duplexes can stabilize it by as much as 4.5 kcal/mol.

(2) *The stabilization correlated with the size of the dangling base also depends on the size of the base it stacks upon.* Comparison reveals that $\begin{matrix}5'\text{-YR-}3'\\3'\text{-R}-\text{-}5'\end{matrix} > \begin{matrix}5'\text{-RR-}3'\\3'\text{-Y}-\text{-}5'\end{matrix} = \begin{matrix}5'\text{-YY-}3'\\3'\text{-R}-\text{-}5'\end{matrix} > \begin{matrix}5'\text{-RY-}3'\\3'\text{-Y}-\text{-}5'\end{matrix}$, i.e., the most favorable sequence is when a purine nucleobase stacks on a pyrimidine base followed by the nucleobases of equal size and the least stabilization is gained when a pyrimidine base stacks on a purine base. For instance, the $\begin{matrix}5'\text{-CA-}3'\\3'\text{-G}-\text{-}5'\end{matrix}$ overhang stabilizes the duplex by 1.8 kcal/mol, while the $\begin{matrix}5'\text{-GA-}3'\\3'\text{-C}-\text{-}5'\end{matrix}$ overhang only stabilizes the same duplex by 1.1 kcal/mol (*31*).

(3) *Perhaps the most important observation is that the nature of the closing basepair plays a central role (29−34).* The same dangling nucleobase stacking on a G−C closing basepair with three hydrogen bonds always provides more stabilization than the A−U closing basepair with two hydrogen bonds. An example of this is that the $\begin{matrix}5'\text{-CA-}3'\\3'\text{-G}-\text{-}5'\end{matrix}$ overhang stabilizes the duplex by 1.8 kcal/mol (*31*), while the $\begin{matrix}5'\text{-UA-}3'\\3'\text{-A}-\text{-}5'\end{matrix}$ overhang only stabilizes the same duplex with 0.6 kcal/mol (*33*) even though both overhangs follow the $\begin{matrix}5'\text{-YR-}3'\\3'\text{-R}-\text{-}5'\end{matrix}$ sequence.

(4) *Finally, studies (44, 45) using different lengths of polynucleotides as dangling ends clearly show that there is a dynamic effect on cooperative intrastrand stacking to dangling-end stabilization.* The stabilization increases for every additional dangling base up to four bases of poly-A: $\Delta\Delta G^\circ$ of $\begin{matrix}5'\text{-CA-}3'\\3'\text{-G}-\text{-}5'\end{matrix} = -1.5$ kcal/mol and $\Delta\Delta G^\circ$ of $\begin{matrix}5'\text{-CAAAA-}3'\\3'\text{-G}-\text{-}\text{-}\text{-}5'\end{matrix} = -2.5$ kcal/mol (*45*).

(C) Stabilization of the DNA−DNA Duplex by the Dangling End. The general trend is that 5′-overhangs tend to

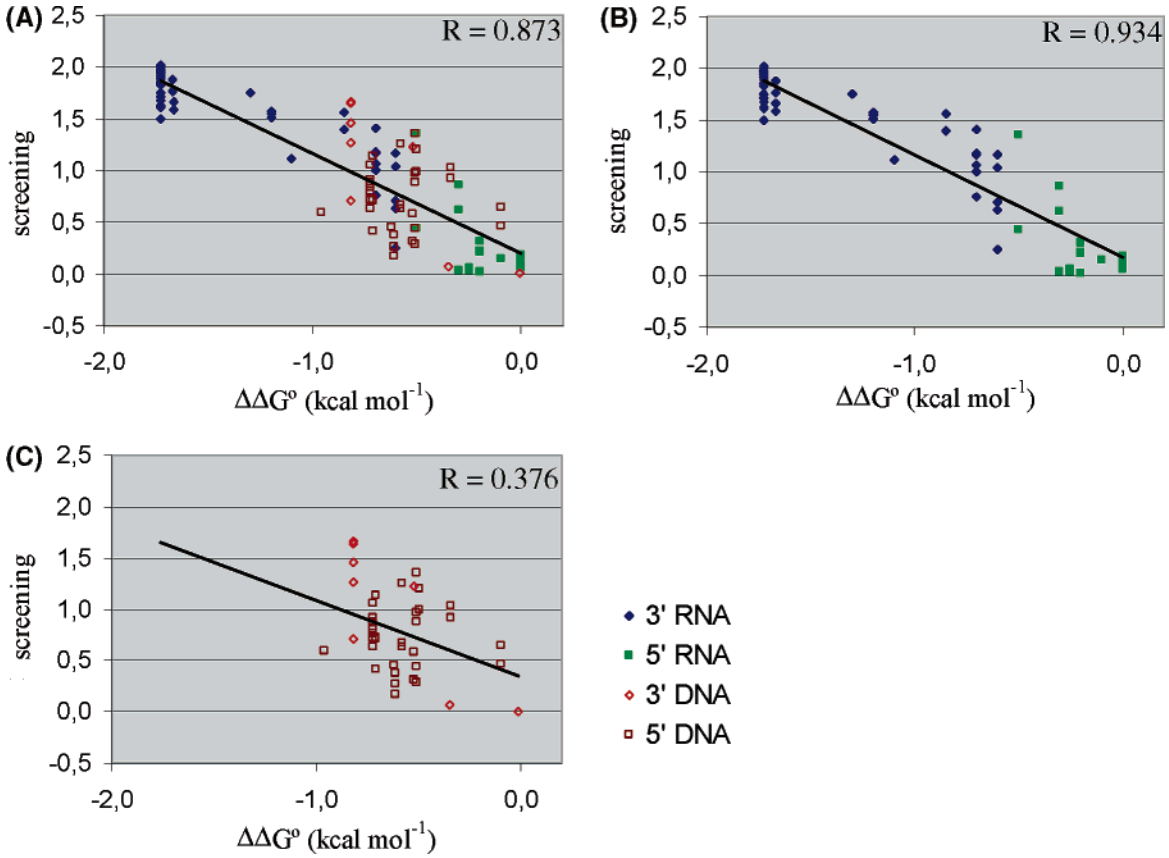

FIGURE 4: Correlation plots of the hydrogen bond screening versus the (thermodynamic stabilization ($\Delta\Delta G°$). (A) The data points from both 3′-RNA (blue solid diamond), 5′-RNA (green solid square), 3′-DNA (red open diamond), and 5′-DNA (brown open square) dangling ends fall on the correlation line and gives a correlation factor, $R = 0.873$. (B) Considering only the points from RNA gives a better correlation, $R = 0.934$, showing that dangling ends on RNA are more ordered than dangling ends on DNA, $R = 0.376$ (C).

stabilize the DNA−DNA duplexes slightly more than the corresponding 3′-overhang. However, the sequence dependence of the stabilizing effects (*35*) are qualitatively less predictable in DNA−DNA duplexes than in RNA−RNA duplexes, which will be further elaborated in Discussion.

(D) Correlation Plots for DNA and RNA. DNA or RNA homo- or heteroduplexes with either dangling-end or tethered aromatic moieties are known to be thermodynamically more stable than the corresponding core duplexes. The NMR studies on these duplexes with a tethered ring system or with the dangling residue (*51*, *52*, *56*−*58*) have shown that the imino protons of the closing basepair in the core duplexes appear as distinct and sharp peaks, as opposed to the corresponding native unprotected terminal basepairs that normally exhibit extensive line-broadening due to rapid exchange of the imino proton with water in the NMR time scale (*59*, *60*). The structural study by NMR on 5′-phenazinium tethered DNA−RNA duplex (*52*) shows that even as the temperature is increased the terminal imino protons significantly resist water exchange compared to the native core duplex. This clearly shows that the tethered ring systems in these modified systems protect the terminal exchangeable hydrogen atoms from chemical exchange with the solvent as they stack over the terminal basepair. This fact, together with the directional differences in base−base stacking between single-stranded DNA and RNA (*46*) and the observed qualitative trends (see above) in previously published thermodynamic data for DNA and RNA (*29*−*40*), has lead us to investigate if there are any correlations between the thermodynamic stabilization and the exact geometry of the dangling base relative to the closing basepair that can uniformly explain the dangling-end effect in both the 3′- and 5′-ends of both DNA and RNA duplexes.

The screening of the hydrogen bonds of the closing basepair in 105 different X-ray and NMR structures (see Databases S1−S4 in Supporting Information for a summary of the full database) of dangling-end structures have been quantified (see Experimental Procedures) and plotted against the thermodynamic stabilization provided by the corresponding dangling-end sequence (*29*−*40*).

The screening value is the inverse sum of the relative displacements of each individual hydrogen bond of the closing basepair with respect to the dangling end. This means that the more efficiently the dangling base can cover all the atoms participating in the hydrogen bonding of the closing basepair, the higher screening value is produced. Thus, there is a higher screening potential for a relatively thermodynamically more stable G−C closing basepair than for an A−T/U closing basepair since there are more atoms participating in hydrogen bonding in the former than in the latter that can potentially be screened by the dangling nucleobase.

The calculated data points for the dangling ends at the 3′-end of RNA, 5′-end of RNA, 3′-end of DNA, and 5′-end of DNA give a correlation factor, $R = 0.873$ (Figure 4A**)** by linear regression analysis. This shows that regardless if the dangling base is attached to the 3′- or 5′-end of the DNA or RNA strand in the duplex, *the thermodynamic stabilization gained is proportional to how well the nucleobase of the dangling nucleotide stacks over the hydrogen bonds of the closing basepair*. The correlation in the regression analysis

is stronger ($R = 0.934$) when only RNA structures are considered alone (Figure 4B). Even though the points calculated from the DNA structures fall reasonably well on the correlation line (Figure 4C), they do not contribute as much to the correlation ($R = 0.376$) because of larger conformational diversity upon crystallization (see below for further discussion). No correlations were found between the free energy gain from the dangling base attachment and the intrastrand, interstrand, and total stacking areas, twist, slide, shift, dx, dy, rise, chi torsion, or phase angle (Figures S1−S11 in Supporting Information).

There are two obvious sources of error in this correlation plot that explain the rather large variation of the screening values on the *Y*-axis.

(1) *The accuracy of the extracted geometries.* Most of the geometries are extracted from X-ray structures and can never fully describe how the dangling end behaves in aqueous solution. In crystal form, the dangling base will always interact in some way with either neighboring identical duplexes between or within the unit cells, metal ions, trapped water, and/or cocrystallized proteins by crystal packing forces that are not present in solution. We have to assume that the dangling end is more likely to crystallize if it can adopt a low-energy conformation that is also populated in solution, with or without bound protein. This is expected to be the main source of error in the correlation plot and stresses the need to analyze a large number of geometries from a large database.

(2) *The dynamics of the cooperative stacking by dangling base with the nearest-neighbor.* If we assume that all dangling ends exist in a two-state equilibrium between a stacked and a destacked geometry and that the stacked geometry perfectly describes the average stacked conformation of the dangling base in solution, then a nucleobase that spends more time in the stacked conformation will provide more protection to the hydrogen bonds of the closing basepair than a nucleobase that spends more time destacked. Therefore, the strength of the stacking interaction between the dangling nucleobase and the core duplex as well as the solvation energy of the dangling-end residue will contribute to the uncertainty along the *Y*-axis.

(E) Mechanism of the Dangling-End Stabilization of the Duplexes. On the basis of the correlations above, *we conclude that the dominant contribution to the thermodynamic gain in stability from all dangling ends comes from how well the dangling base can protect the hydrogen bonds of the closing basepair (Figures 1 and 3) from the solvent, thus giving the terminal basepair a character that more resembles that of an internal basepair (51, 52, 58).* The increased stabilization is most likely to be the result of the combined contributions from two known effects on hydrogen bonds exposed to the bulk solvent: (i) Becker et al (*61*) have elegantly demonstrated that the interior of both DNA and RNA duplexes are devoid of water by showing a significant decrease in acridinium ester hydrolysis inside a complementary double helix compared to mismatched duplexes. Further, Shan et al. have shown that the strength of hydrogen bonding is proportional to the dielectric of the solvent (*62*, *63*). Enzymatic catalysis is proposed to be achieved as the interior of the enzyme has a much lower dielectric than the surrounding solvent, making the hydrogen bonds stabilizing the transition state stronger inside the active site than free in solution, thereby catalyzing the reaction. In a very similar way, the dangling base of a DNA or RNA duplex can be considered a shield against the water surrounding the duplex, thus effectively changing the microenvironment around the terminal bonds in such a way that the bonds experience a much lower dielectric than it would have if it had been fully exposed to the bulk water, thereby increasing the strength of the interstrand hydrogen bond interactions of the terminal basepair (*59*, *60*). (ii) The terminal nucleobases in blunt-ended duplexes normally exhibit increased opening rate and water exchange (*59*, *60*) than the basepairs in the core of the duplex. By restricting the direct accessibility of water to the *exchangeable* hydrogen atoms involved in hydrogen bonding at the terminus by the dangling nucleobase, the exchange rate is made to decrease and thus the lifetime of the Watson−Crick hydrogen bonds is prolonged, thereby further enhancing the overall stability of the complex.

The fact that dangling-end stabilization is correlated to the extent of screening from the bulk water, but completely uncorrelated to the stacking area of the dangling-end base (intra-, interstrand stacking area, as well as total stacking area, Figures S9−S11 in Supporting Information), shows that it is the spatial stereoproximity of the dangling nucleobase and the closing basepair that determines the *potential* of the thermodynamic stabilization and the chemical character of the bases (i.e., the strength of the interaction between the bases) that determines the how *efficiently* that potential is used.

This is the underlying reason the dangling native nucleotides at the 3′-end of RNA has a much higher stabilizing potential than the 5′-end. Elongation of the RNA helix by one nucleobase in the 3′ direction positions it in a way where it is free to interact with the hydrogen bonds of the closing basepair (Figure 3F), while the corresponding elongation in the 5′-direction stereochemically positions the nucleobase away from the hydrogen bonds of the closing basepair (Figure 3D). Thus, even though the interaction area of the dangling nucleobase in the 5′-end (Figure 3D) and the 3′-end (Figure 3F) are very similar (∼4 Å^2), the thermodynamic stabilization potential in the 3′-end is 1.9 kcalmol^{-1} but only 0.5 kcalmol^{-1} in the 5′-end. This is because their potential to interact with the closing basepair is considerably different.

(F) Importance of Dangling-End Dynamics. Once the size, shape, and stereoproximity of the dangling base is such that is has a chance to interact with the hydrogen bonds of the closing basepair, the dynamics of the dangling base also plays an important role as it will determine how much time the dangling base resides in its stacked geometry and thus directly influences how efficiently the terminal basepair of the duplex is screened. Earlier works have pointed out the importance of stacking interaction in oligonucleotide stability and function. Unnatural DNA bases without the possibility to form hydrogen bonds have been shown not to significantly destabilize the duplex formation (*64*−*67*). These unnatural DNA bases (steric mimics) have also been shown to be successfully incorporated in DNA by DNA polymerase with the help of a template strand because of their ability to mimic the shape and stacking interactions of the native base (*68*, *69*). The summation of intrinsic pK_a of the model monomers representing the donor and acceptor components of A−T and C−G basepairs have also been used to estimate the relative contribution of the stacking interaction vis-à-vis

strength of hydrogen bonding in DNA and RNA duplexes (*70*). The stability of the stacked over the destacked conformation is finally determined by the net free-energy of multiple interactions including the electrostatic effects (charges), van der Waals (dipoles and dispersion interactions), and hydrophobic effects between the dangling base and *any* of the nucleobases in the closing basepair. For instance, the difference between $\begin{matrix}5'\text{-CU-}3'\\3'\text{-G}-\text{-}5'\end{matrix}$ ($\Delta\Delta G^\circ = -1.20$ kcal/mol) and $\begin{matrix}5'\text{-CC-}3'\\3'\text{-G}-\text{-}5'\end{matrix}$ ($\Delta\Delta G^\circ = -0.85$ kcal/mol) has been proposed (*36*) to originate from a specific interaction between the O4 of 1-uracilyl and 4-amino-group of the 1-cytosinyl moiety in the former (however, see Discussion below). Another example is the possible role of methyl/π interactions (i.e., between 5-Me of 1-thyminyl with the π cloud of the nearest-neighbor) in dangling-end stabilization. The frequent occurrence of 1-thyminyl-methyl$_{(n)}$/$\pi_{(n-1)}$ interactions in DNA A−T tracts has been statistically identified and evaluated (*71*, *72*). A comparison between DNA-3′-**C** versus DNA-3′-**T** dangling ends show that the **T** overhangs always are more stabilizing than the **C** overhangs by 0.10−0.21 kcal/mol (*35*), showing a possible intrastrand methyl/π interaction stabilizing the stacking interaction. Using the same rationale, the electronic character of the dangling base (or tethered molecule) is expected to affect the dynamics of the dangling end through its ability to stack in a cooperative fashion. The solvation energy of the dangling base will also drive the stacking equilibrium toward the stacked form, thus directly affecting the dynamics of the dangling end. This is in agreement with the recently published results that have shown that the higher the hydrophobicity and polarizability (a_m) of the dangling end, the more stability it provides to the duplex when attached as dangling end (*37*, *42*), even if the stacking areas and geometries are presumably constant. We also propose that the reason that further stabilization is gained when additional dangling bases are added to an overhang sequence is that the additional dangling nucleotides reduce the dynamics of the innermost dangling base by nearest-neighbor intrastrand stacking interactions. This allows the innermost dangling nucleobase to spend more time in a stacked conformation with the closing basepair.

(G) Evaluation. The thermodynamic stabilization is generally characterized by spectroscopic melting studies (*29*−*37*, *39*, *40*, *54*). The melting of oligonucleotides is a bimolecular process, and therefore the strength of the complex can only be increased by either (a) increasing the strength of the interactions between the two molecules, *or* (b) increasing the solvent driven energetic penalty of dissociation.

The *dominating* mechanism of dangling-end stabilization proposed here is a *reduction of the weakening* of one of the already existing interactions between the two strands, the hydrogen bonds of the terminal basepair. In addition to this mechanism, there are possibly other contributions to the stabilization associated with the addition of the dangling end that are not strong enough by themselves to show any correlation with the total free energy gain: (1) Any specific electrostatically attractive interactions between the dangling base (or tethered molecule) and the nucleobase in the opposite strand of the closing basepair can further stabilize the complex formation. (2) Nucleobases are complex psudoaromatic moieties and electrostatic crosstalk (*73*−*75*) between stacking nucleobases could directly modulate the pseudoaromaticity as well as the strength of its hydrogen bonds (pK_a) in addition to the dominant solvent effects discussed above. This has also been suggested by recent theoretical calculations on dimers (*76*). (3) The solvation energy of any part of the dangling base that stacks with *the opposite strand* in duplex form acts as an additional energy penalty of disassociation since this surface will become exposed to the solvent upon duplex melting.

(H) Differences between the Dangling-End Effects in DNA and RNA Duplexes. Prediction of the dangling-end effects in DNA−DNA duplexes holds a much larger uncertainty than the prediction of the corresponding effects in RNA−RNA duplexes. This is most likely be attributed to the fact that the sugar backbone is more flexible in DNA than in RNA (*60*, *77*−*83*). The lower energy barriers associated with DNA deformation (*82*, *83*) makes it more prone to change its structure when being crystallized and/or bound to proteins than RNA, thereby increasing the geometric uncertainty for DNA dangling ends retrieved from X-ray structures. This is confirmed by the fact that the standard deviations of slide, shift, and twist extracted from the RNA dangling-end structures [$\sigma = 0.76$ Å, 1.58 Å, and 11.24°, respectively] are significantly smaller that the ones extracted from DNA [$\sigma = 1.64$ Å, 2.42 Å and 31.65°, respectively] dangling-end structures (Figures S1−S3 in Supporting Information), thus showing the increased polymorphism of the dangling-end base position in DNA compared to RNA.

Another important difference is that the found thermodynamic gain for DNA dangling ends have a much smaller potential (maximum stabilization ∼ 1 kcal/mol) compared to that of RNA (maximum stabilization ∼ 2 kcal/mol). A comparison of the phase angle and chi (χ) torsion of the dangling residue in the database reveals that they indeed maintain their preferred double-stranded conformations (RNA: $P = 15 \pm 14°$, $\chi = -158 \pm 35°$; DNA: $P = 150 \pm 55°$, $\chi = -101 \pm 57°$, Figures S7 and S8 in Supporting Information) and that the DNA dangling bases display a much larger standard deviation than the RNA dangling bases. The reason for this is most likely that the internal motion acts more like a rigid body on the nucleotide level in RNA than in DNA due to the 2′-OH effects. The balance between the steric and anomeric effects and gauche effects of the 2′-oxygen (*84*−*86*) in RNA steers the sugar to a North-type conformation and raises the energy barrier of the North/South equilibrium compared to DNA. The 2′-OH also enables the inter-residual 2′-OH$_{(n)}$-O4′$_{(n+1)}$ hydrogen bond formation (*60*, *83*) in RNA, which further restricts the conformational freedom at room temperature. The absence of these features in DNA is likely the cause of the observed increase in the dynamics of the dangling base step compared to RNA and hence the reduction of the efficiency of the screening.

(I) Relation of the Dangling-End Effects to Established Prediction Algorithms. The nearest-neighbor basepair contribution to the thermodynamic stability of DNA and RNA duplexes has previously been rigorously studied to enable accurate prediction of the complex stability of short duplexes (*17*−*24*). The expanded individual nearest-neighbor model (INN−HB) works on the assumption that each basepair doublet has a contribution to duplex stability and each pair is summed together with an initiation term, an AU/T correction term if the terminal basepair is A−U or A−T

Table 2: A Summary of How the INN−HB Model Predicts the Stability of Internal and Terminal Basepairs Depending in the Sequence Context[a]

		Internal basepair		Terminal basepair	
		G-C	A-U/T	G-C	A-U/T
RNA	Lowest energy (kcal/mol)	-2.2	-1.0	-1.0	-0.5
	Sequence (5'→3')	5'–*ACU* – 3' 3'–*UGA* – 5'	5'–*AAU* – 3' 3'–*UUA* – 5'	5'–*CU* .. – 3' 3'–*GA* .. – 5'	5'–*AA* .. – 3' 3'–*UU* .. – 5'
	Highest energy (kcal/mol)	-2.9	-1.7	-1.7	-1.2
	Sequence (5'→3')	5'–*GCG* – 3' 3'–*CGC* – 5'	5'–*GAC* – 3' 3'–*CUG* – 5'	5'–*GC* .. – 3' 3'–*CG* .. – 5'	5'–*UC* .. – 3' 3'–*AG* .. – 5'
DNA	Lowest energy (kcal/mol)	-1.3	-0.7	-0.6	-0.3
	Sequence (5'→3')	5'–*ACT* – 3' 3'–*TGA* – 5'	5'–*TAT* – 3' 3'–*ATA* – 5'	5'–*CT* .. – 3' 3'–*GA* .. – 5'	5'–*TA* .. – 3' 3'–*AT* .. – 5'
	Highest energy (kcal/mol)	-2.2	-1.4	-1.1	-0.7
	Sequence (5'→3')	5'–*GCG* – 3' 3'–*CGC* – 5'	5'–*GAC* – 3' 3'–*CTG* – 5'	5'–*GC* .. – 3' 3'–*CG* .. – 5'	5'–*TC* .. – 3' 3'–*AG* .. – 5'

[a] Considering a homo-basepair step yields the following values: RNA: G−C: −3.26 kcal/mol, A−U: −0.93 kcal/mol; DNA: G−C: −1.77 kcal/mol, A−T: −1.02 kcal/mol.

instead of G−C and a symmetry-related entropy term for self-complementary duplexes. The initiation term is considered to be a pure entropy term of 1.8−4.1 kcal/mol depending on A−T/U content (*18*, *20*).

The addition of one internal G−C basepair in the RNA−RNA duplex gives a contribution between −2.2 $\binom{5'\text{-ACU-}3'}{3'\text{-UGA-}5'}$ and −2.9 $\binom{5'\text{-GCG-}3'}{3'\text{-CGC-}5'}$ kcal/mol and the addition of one internal A−U basepair gives a contribution between −1.0 $\binom{5'\text{-AAU-}3'}{3'\text{-UUA-}5'}$ and −2.2 $\binom{5'\text{-GAC-}3'}{3'\text{-CUG-}5'}$ kcal/mol (*20*) (Table 2). A terminal G−C basepair only contributes from −1.0 to −1.7 kcal mol^{-1}, which is 1.2−1.9 kcal mol^{-1} less than the most stable internal G−C basepair (*20*) in the INN−HB model. The maximum observed thermodynamic gain from a dangling-end effect on a G−C closing basepair is −1.9 kcal mol^{-1} $\left[\begin{smallmatrix}5'\text{-CA-}3'\\3'\text{-G−-}5'\end{smallmatrix}\right.$ (*31*)]. *This further supports the conclusion that the reduction of the destabilization of the terminal basepairs due to water exchange, by giving the terminal basepair the character of an internal basepair, is potentially dominating the thermodynamic gain.* The corresponding reduction of stability for a terminal A−U basepair is 0.8−1.2 kcal mol^{-1} compared to the most stable internal basepair. The maximum observed thermodynamic gain from a dangling-end effect on a A−U closing basepair is however only −0.7 kcal/mol $\left[\begin{smallmatrix}5'\text{-AA-}3'\\3'\text{-U−-}5'\end{smallmatrix}\right.$ and $\left.\begin{smallmatrix}5'\text{-AG-}3'\\3'\text{-U−-}5'\end{smallmatrix}\right]$ (*33*)], indicating that the stability of a terminal A−U pair may be overestimated in the INN−HB model and thus explain the need for the A−U correction term of +0.45 kcal mol^{-1} being used in this model. This also shows that there is ∼1 kcal mol^{-1} less energy to be gained by screening a terminal A−U pair than a terminal G−C pair from solvent effects.

Using the same reasoning as above for DNA−DNA duplexes shows that a terminal G−C basepair is 1.1−1.6 kcal mol^{-1} less stable than the most stable internal basepair. For A−U, the corresponding value is 0.7−1.1 kcal/mol (Table 2) (*20*). The maximum dangling-end stabilization for a G−C closing basepair is −0.96 and −0.71 kcal mol^{-1} on an A−T closing basepair (*35*). These values are somewhat smaller but in the same region as G−C and A−U basepairs in terms of energy in the RNA−RNA duplex. This clearly suggests that the lack of really strong dangling-end stabilization in the DNA−DNA duplex is a dynamic effect, due to the increased flexibility in DNA compared to RNA.

In light of the correlations presented in this work, we also propose that the underlying reason for significantly different initiation terms in INN−HB algorithms reported based on different sequence libraries of DNA and RNA (*17*−*22*, *24*, *30*) is that the contribution of each base doublet is modulated by sequence-specific differences in internal hydration of the

doublets surroundings due to local deviations from optimal stacking. We also find it likely that the difference in initiation terms between RNA (+3.4 kcal/mol) and DNA (+1.82 kcal/mol) originates from differences in sensitivity to hydration due to differences in geometry and dynamics (owing to the differential sugar conformation, nucleobase orientation as well as the 2′-OH effects) at the terminals of DNA and RNA duplexes. It is also very possible that the increased water exchange destabilization seen at the terminals can reach further into duplexes with A−T-rich ends due to their lower stability compared to G−C basepairs.

DISCUSSION

A full understanding of the underlying mechanism of the dangling-end effects makes it possible to design highly efficient synthetic tethers to be used as stabilizing caps for designing high-affinity probes. In conclusion, there are two components that need to be optimized. First, the geometry of the stacking moiety should be optimized to completely shield the hydrogen bonds of the closing basepair from the bulk water. This means that the dangling moiety must be of appropriate size and shape and have electrostatic compatibility with the closing basepair to be able to occupy the hyperspace over the hydrogen bonds of the closing basepair in an energetically relaxed configuration. Second, the dangling base should be engineered in such a way that its dynamics with respect to the nearest-neighbor is kept to a minimum. The stacked form should be stabilized over the open form by introducing structural elements that favor electrostatic attraction between the dangling modified nucleobase and the closing basepair. The dynamics can also be reduced by adding additional dangling building blocks to the overhang (*44*, *45*) or by reducing the flexibility of the linker by using modified nucleosides that are conformationally restricted.

The proposed mechanism of dangling-end stabilization can also have implications on experiments involving dangling residues. Replacement of guanine with inosine and cytosine with 2-pyrimidinone has been used as a model system to investigate the role of the amines of the closing basepair in dangling-end stabilization (*36*). The removal of one of the amines of the closing basepair leads to a reduced stabilization potential of the dangling-end (fewer available bonds to stabilize of the closing basepair) even though the dangling residue itself is unchanged. Thus, it may be difficult to draw any conclusions from inosine and 2-pyrimidinone substitutions in conjunction with dangling ends.

In an interesting related work (*34*), inosine replacement of guanine is used to estimate the importance of one hydrogen bond at the terminal of short RNA−RNA duplexes. It is found that the more stable a 3′-dangling base overhang is, the less additional stability is gained by completing the corresponding basepair by attaching a basepairing partner to the dangling residue. It is also shown that removing one of the terminal hydrogen bonds by replacing guanine with inosine is less crucial if the affected basepair forms a stable corresponding 3′-dangling base. The authors conclude that hydrogen bonds compete with stacking in a basepair, i.e., strong stacking favors geometries not optimal for hydrogen bonding and vice versa. Another perspective is that when the dangling end provides a lot of stability even without a basepairing partner, the addition of an extra hydrogen bond(s) on that particular base has less impact on the overall stability than when the dangling end is unstable. Consequently, when the corresponding dangling end provides little or no stability to the closing basepair there is a big unused potential for stabilization. Thus, even though the unprotected terminal hydrogen bonds are much weaker than the internal hydrogen bonds (due to the exposure to the bulk water), they can still play a crucial role by stabilizing the closed form over the open form of the terminal basepair and thereby stabilize the core duplex by providing a screen against the bulk water for the neighboring inner basepair.

There are exceptions to the stabilizing trend in DNA. The three overhang sequences: $\begin{matrix}5'\text{-GT-}3' \\ 3'\text{-}-\text{A-}5'\end{matrix}$, $\begin{matrix}5'\text{-AC-}3' \\ 3'\text{-T}-\text{-}5'\end{matrix}$ and $\begin{matrix}5'\text{-AT-}3' \\ 3'\text{-T}-\text{-}5'\end{matrix}$ have been shown to *destabilize* the DNA duplexes with 0.13−0.48 kcal/mol (*35*) instead of stabilizing it. Destabilizing effects of dangling ends are much harder to explain, and we can only speculate that the destabilization of the duplex form comes from the inability of the dangling base to stack over the adenine of the closing basepair in duplex form but not in single-strand form, thereby increasing the solvation energy relative to the core duplex more for the duplex than for the sum of the two single strands and shift the single-strand ⇄ double-strand equilibrium.

IMPLICATIONS

The studies on the mechanism behind dangling-end stabilization highlight the synergy between the stabilization of the hydrogen bond and the favorable stacking interactions from the dangling end in native and modified DNA and RNA duplexes. Upon any mismatch introduction, the optimal stacking in the duplex around the mismatch site becomes perturbed. This causes an increase in the level of hydration around the minor and major grooves, leading to a significant weakening of the interstrand hydrogen bond interactions as evident by the enhanced stability of the matched, mismatched or the hybrid DNA−RNA duplexes by the tethered ligand (*51*, *52*, *58*) as well as by the enhanced rate of hydrolysis of acridinium esters around the mismatched duplex compared to the fully matched duplex (*61*). In high-affinity probe design for hybridization experiments, there is always a balance required between specificity and sensitivity. While the addition of dangling residues to a probe can indeed lead to stronger binding, there is always the risk of getting additional tolerance to mismatches. A thorough understanding of the molecular mechanism of the dangling-end effects could possibly be exploited to design synthetic dangling moieties attached with an engineered rigid linker that can stabilize canonical basepairs but not stack well over the most common mismatches (for example G−U mismatches) at the closing basepair.

Interestingly, since the dangling nucleobases are not normally involved in basepairing, their stacking potential can be more satisfactorily exploited by employing aromatic heterocycles that they both stack well and produce high fluorescence for detection purposes. A similar approach with dangling fluorescent aromatic heterocycles should also be applicable in designing the double-stranded small interfering RNA probes (siRNA) (in which TT dangling nucleotides are

substituted) for inhibiting target mRNA in the RISC, as well as for understanding the mechanism of how the double-stranded siRNA binds to the target RNA with the RISC endonuclease.

SUPPORTING INFORMATION AVAILABLE

Figure S1 plot: thermodynamic stabilization vs shift; Figure S2 plot: thermodynamic stabilization vs slide; Figure S3 plot: thermodynamic stabilization vs twist; Figure S4 plot: thermodynamic stabilization vs *x* displacement; Figure S5 plot: thermodynamic stabilization vs *y* displacement; Figure S6 plot: thermodynamic stabilization vs phase angle; Figure S7 plot: thermodynamic stabilization vs chi; Figure S8 plot: thermodynamic stabilization vs rise; Figure S9 plot: thermodynamic stabilization vs intrastrand stacking area; Figure S10 plot: thermodynamic stabilization vs interstrand stacking area; Figure S11 plot: thermodynamic stabilization vs total stacking area; database 1: double-stranded RNA with 3′-dangling nucleobases; database 2: double-stranded RNA with 5′-dangling nucleobases; database 3: double-stranded DNA with 3′-dangling nucleobases; database 4: double-stranded DNA with 5′-dangling nucleobases. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Limmer, S., Hofmann, H.-P., Ott, G., and Sprinzl, M. (1993) The 3′-terminal end (NCCA) of the tRNA determines the structure and stability of the aminoacul acceptor stem, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6199−6202.
2. Nagan, M. C., Beuning, P., Musier-Forsyth, K., and Cramer, C. J. (2000) Importance of discriminator base stacking interactions: molecular dynamics analysis of A73 microhelix^{Ala} variants, *Nucleic Acids Res. 28*, 2527−2534.
3. Alkema, D., Bell, R. A., Hader, P. A., and Neilson, T. (1981) Invariant Adenosine Residues Stabilize tRNA D Stems, *Fed. Eur. Biochem. Soc. Lett. 136*, 70−71.
4. Yoon, K., Turner, D. H., Tinoco, I. J., Haar, F., and Cramer, F. (1976) The kinetics of binding of U-U-C-A to a dodecanucleotide anticodon fragment from yeast tRNA-Phe, *Nucleic Acids Res. 3*, 2233−2241.
5. Ayer, D., and Yarus, M. (1986) The context effect does not require a fourth base pair location, *Science 231*, 393−395.
6. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells, *Nature 411*, 494−498.
7. Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev. 15*, 188−200.
8. Amarzguioui, M., and Prydz, H. (2004) An algorithm for selection of functional siRNA sequences, *Biochem. Biophys. Res. Commun. 316*, 1050−1058.
9. Ma, J.-B., Ye, K., and Patel, D. J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain, *Nature 429*, 318−322.
10. Du, Z., Giedroc, D. P., and Hoffman, D. W. (1996) Structure of the Autoregulatory Pseudoknot within the Gene *32* Messenger RNA of Bacteriophages T2 and T6: A Model for a Possible Family of Structurally Related RNA Pseudoknots, *Biochemistry 35*, 4187−4198.
11. Kool, E. T. (1997) Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides, *Chem. Rev. 97*, 1473−1487.
12. Kool, E. T. (2001) Hydrogen Bonding, Base Stacking and Steric Effects in DNA Replication, *Annu. Rev. Biophys. Biomol. Struct. 30*, 1−22.
13. Messias, A. C., and Sattler, M. (2004) Structural Basis of Single-stranded RNA Recognition, *Acc. Chem. Res. 37*, 279−287.
14. SantaLucia Jr, J., and Hicks, D. (2004) The Thermodynamics of DNA Structural Motifs, *Annu. Rev. Biophys. Biomol. Struct. 33*, 415−440.
15. Williams, J. C., Case-Green, S. C., Mir, K. U., and Southern, E. M. (1994) Studies of oligonucleotide interactions by hybridization to arrays: the influence of dangling ends on duplex yield, *Nucleic Acids Res. 22*, 1365−1367.
16. Tyagi, S., and Kramer, F. R. (1996) Molecular beacons: probes that fluoresce upon hybridization, *Nature Biotechnol. 14*, 303−308.
17. Breslauer, K. J., Frank, R., Heikens, W., Blöcker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence, *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746−3750.
18. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) Improved Nearest-Neighbor Parameters for Predicting DNA Duplex Stability, *Biochemistry 35*, 3555−3562.
19. Gray, D. M. (1997) Derivation of Nearest-Neighbor Properties from Data on Nucleic Acid Oligomers. I. Simple Sets of Independent Sequences and the Influence of Absent Nearest Neighbors, *Biopolymers 42*, 783−793.
20. Xia, T., SantaLucia Jr, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplex with Watson−Crick Base Pairs, *Biochemistry 37*, 14719−14735.
21. 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.
22. Ma, L., Cheng, C., Miyajima, H., Zhao, Y., and Herdewijn, P. (2002) Predicting melting temperature (Tm) of oligoribonucleotide duplex by neural network, *J. Chemom. 16*, 75−80.
23. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res. 31*, 3406−3415.
24. SantaLucia Jr, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc. Natl. Acad. Sci. U.S.A. 95*, 1460−1465.
25. Martin, F. H., Uhlenbeck, O. C., and Doty, P. (1971) Self-complementary oligoribonucleotides: Adenylic acid-uridylic acid block copolymers, *J. Mol. Biol. 57*, 201−215.
26. Uhlenbeck, O. C., Martin, F. H., and Doty, P. (1971) Self-complementary oligoribonucleotides: Effects of helix defects and guanylic acid-cytidylic acid base pairs, *J. Mol. Biol. 57*, 217−229.
27. Romaniuk, P. J., Hughes, D. W., Gregoire, R. J., Neilson, T., and Bell, R. A. (1978) Stabilizing Effect of Dangling Bases on a Short RNA Double Helix as Determined by Proton Nuclear Magnetic Resonance Spectroscopy, *J. Am. Chem. Soc. 100*, 3971−3972.
28. Neilson, T., Romaniuk, P. J., Alkema, D., Hughes, D. W., Evertt, J. R., and Bell, R. A. (1980) The effects of base sequence and dangling bases on the stability of short ribonucleic acid duplexes, *Nucleic Acids Symp. Ser. 7*, 293−311.
29. Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., and Turner, D. H. (1983) Effects of 3′ Dangling End Stacking on the Stability of GGCC and CCGG Double Helices, *Biochemistry 22*, 6198−6206.
30. Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., and Turner, D. H. (1985) Contributions of Dangling End Stacking and Terminal Base-Pair Formation to the Stabilities of XGGCCp, XCCGGp, XGGCCY, and XCCGGYp Helixes, *Biochemistry 24*, 4533−4539.
31. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) Stability of XGCGCp, GCGCYp and XGCGCYp Helixes: An Emperical Estimate of the Energetics of Hydrogen Bonds in Nucleic Acids, *Biochemistry 25*, 3214−3219.
32. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA Structure Prediction, *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
33. Sugimoto, N., Kierzek, R., and Turner, D. H. (1987) Sequence Dependence for the Energetics of Dangling Ends and Terminal Base Pairs in Ribonucleic Acid, *Biochemistry 26*, 4554−4558.
34. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) Free Energy Increments for Hydrogen Bonds in Nucleic Acid Base Pairs, *J. Am. Chem. Soc. 109*, 3783−3785.
35. Bommarito, S., Peyret, N., and SantaLucia, J., Jr. (2000) Thermodynamic parameters of DNA sequences with dangling ends, *Nucleic Acids Res. 28*, 1929−1934.

36. 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.
37. Rosemeyer, H., and Seela, F. (2002) Modified purine nucleosides as dangling ends of DNA duplexes: the effect of the nucleobase polarizability on stacking interactions, *J. Chem. Soc., Perkin Trans. 2*, 746−750.
38. Tuma, J., Connors, W. H., Stitelman, D. H., and Richert, C. (2002) On the Effect of Covalently Appended Quinolones on Termini of DNA duplexes, *J. Am. Chem. Soc. 124*, 4236−4246.
39. Ziomek, K., Kierzek, E., Biala, E., and Kierzek, R. (2002) The influence of various modified nucleotides placed as 3′-dangling end on thermal stability of RNA duplexes, *Biophys. Chem. 97*, 243−249.
40. Ziomek, K., Kierzek, E., Biala, E., and Kierzek, R. (2002) The termal stability of RNA duplexes containing modified base pairs placed at internal and terminal positions of the oligoribonucleotides, *Biophys. Chem. 97*, 233−241.
41. Nakano, S.-i., Uotani, Y., Nakashima, S., Anno, Y., Fujii, M., and Sugimoto, N. (2003) Large Stabilization of a DNA duplex by the Deoxyadenosine Derivatives Tethering an Aromatic Hydrocarbon Group, *J. Am. Chem. Soc. 125*, 8086−8087.
42. Guckian, K. M., Schweizer, B. A., Ren, R. X.-F., Sheils, C. J., Paris, P. L., Tahmassebi, D. C., and Kool, E. T. (1996) Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA, *J. Am. Chem. Soc. 118*, 8182−8183.
43. Quartin, R. S., and Wetmur, J. G. (1989) Effect of Ionic Strength on the Hybridization of Oligodeoxynucleotides with Reduced Charge Due to Methylphosphonate Linkages to Unmodified Oligodeoxynucleotides Containing the Complementary Sequence, *Biochemistry 28*, 1040−1047.
44. Riccelli, P. V., Mandell, K. E., and Benight, A. S. (2002) Melting studies of dangling-ended DNA hairpins: effects of end length, loop sequence and biotinylation of loop bases, *Nucleic Acids Res. 30*, 4088−4093.
45. Ohmichi, T., Nakano, S.-i., Miyoshi, D., and Sugimoto, N. (2002) Long RNA Dangling End Has Large Energetic Contribution to Duplex Stability, *J. Am. Chem. Soc. 124*, 10367−10372.
46. Isaksson, J., Acharya, S., Barman, J., Cheruku, P., and Chattopadhyaya, J. (2004) Single-Stranded Adenine-Rich DNA and RNA Retain Structural Characteristics of Their Respective Double-Stranded Conformations and Show Directional Differences in Stacking Pattern, *Biochemistry 43*, 15996−16010.
47. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
48. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., and Schneider, B. (1992) The Nucleic Acid Database: A Comprehensive Relational Database of Three-Dimensional Structures of Nucleic Acids, *Biophys. J. 63*, 751−759.
49. Lu, X.-J., and Olson, W. K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structure, *Nucl. Acids Res. 31*, 5108−5121.
50. SYBYL (1995), Tripos Inc., St. Louis, MO.
51. Maltseva, T., Agback, P., Repkova, M. N., Venyaminova, A. G., Ivanova, I. M., Sandström, A., Zarytova, V. F., and Chattopadhyaya, J. (1994) The solution structure of a 3′-phenazinium (Pzn) tethered DNA-RNA duplex with a dangling adenosine: r(5′G-AUUGAA3′): d(5′TCAATC3′-Pzn), *Nucleic Acids Res. 22*, 5590−5599.
52. Maltseva, T., Sandström, A., Ivanova, I. M., Sergeyev, D. S., Zarytova, V. F., and Chattopadhyaya, J. (1993) Structural studies of the 5′-phenazinium-tethered matched and G-A-mismatched DNA duplexes by NMR spectroscopy, *J. Biochem. Biophys. Methods 26*, 173−236.
53. Bleczinski, C. F., and Richert, C. (1999) Steroid-DNA Interactions Increasing Stability, Sequence-Selectivity, DNA/RNA Discrimination and Hypochromicity of Oligonucleotide Duplexes, *J. Am. Chem. Soc. 121*, 10889−10894.
54. Tuma, J., and Richert, C. (2003) Solution Structure of a Steroid-DNA Complex with Cholic Acid Residues Sealing the Termini of a Watson−Crick Duplex, *Biochemistry 42*, 8957−8965.
55. Mokhir, A. A., Tetzlaff, C. N., Herzberger, S., Mosbacher, A., and Richert, C. (2001) Monitored Selection of DNA-hybrids Forming Duplexes with Capped Terminal C:G Pairs, *J. Comb. Chem. 3*, 374−386.
56. Znosko, B. M., Burkard, M. E., Krugh, T. R., and Turner, D. H. (2002) Molecular Recognition in Purine-Rich Internal Loops: Thermodynamic, Structural and Dynamic Consequences of Purine for Adenine Substitutions in 5′(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14978.
57. Lane, A., Martin, S. R., Ebel, S., and Brown, T. (1992) Solution Conformation of a Deoxynucleotide Containing Tandem G−A Mismatched Base Pairs and 3′-Overhanding Ends in d(GT-GAACTT)$_2$, *Biochemistry 31*, 12087−12095.
58. Maltseva, T., Zarytova, V. F., and Chattopadhyaya, J. (1995) Base-pair exchange kinetics of the imino and amino protons of the 3′-phenazinium tethered DNA-RNA duplex, r(5′GAUUGAA3′): d(5′TCAATC3′-Pzn), and their comparison with those of B-DNA duplex, *J. Biochem. Biophys. Methods 30*, 163−177.
59. Denisov, A. Y., Zamaratski, E. V., Maltseva, T. V., Sandström, A., Bekiroglu, S., Altmann, K.-H., Egli, M., and Chattopadhyaya, J. (1998) The Solution Conformation of a Carbocyclic Analog of the Dickerson-Drew Dodecamer: Comparison with its own X-ray Structure and that of the NMR Structure of the Native Counterpart, *J. Biomol. Struct. Dyn. 16*, 547−568.
60. Snoussi, K., and Leroy, J.-L. (2001) Imino Proton Exchange and Base-Pair Kinetics in RNA Duplexes, *Biochemistry 40*, 8898−8904.
61. Becker, M., Lerum, V., Dickson, S., Nelson, N. C., and Matsuda, E. (1999) The Double Helix is Dehydrated: Evidence from the Hydrolysis of Acridinium Ester-Labeled Probes, *Biochemistry 38*, 5603−5611.
62. Shan, S.-o., Loh, S., and Herschlag, D. (1996) The Energetics of Hydrogen Bonds in Model Systems: Implications for Enzymatic Catalysis, *Science 272*, 97−101.
63. Shan, S.-o., and Herschlag, D. (1996) The change in hydrogen bond strength accompanying charge rearrangement: Implications for enzymatic catalysis, *Biochemistry 93*, 14474−14479.
64. Liu, H., Gao, J., and Kool, E. T. (2005) Helix-Forming Properties of Size-Expanded DNA, an Alternative Four-Base Genetic Form, *J. Am. Chem. Soc. 127*, 1396−1402.
65. Matsuda, S., and Romesberg, F. E. (2004) Optimization of Interstrand Hydrophobic Packing Interactions within Unnatural DNA Base Pairs, *J. Am. Chem. Soc. 126*, 14419−14427.
66. Kim, T. W., and Kool, E. T. (2004) A Set of Nonpolar Thymidine Nucleoside Analogues with Gradually Increasing Size, *Org. Lett. 6*, 3949−3952.
67. Alex, H. F., and Kool, E. T. (2005) Novel Benzopyrimidines as Widened Analogues of DNA Bases, *J. Org. Chem. 70*, 132−140.
68. Henry, A. A., Olsen, A. G., Matsuda, S., Yu, C., Geierstanger, B. H., and Romesberg, F. E. (2004) Efforts To Expand the Genetic Alphabet: Identification of a Replicable Unnatural DNA Self-Pair, *J. Am. Chem. Soc. 126*, 6923−6931.
69. Kool, E. T. (2000) Synthetically modified DNAs as substrates for polymerases, *Curr. Opin. Chem. Biol. 4*, 602−608.
70. Acharya, P., Cheruku, P., Chatterjee, S., Babu Karthick, S., Acharya, S., and Chattopadhyaya, J. (2004) The Measurement of Nucleobase pK_a of the Model Mononucleotides shows why RNA-RNA duplex is more stable than DNA-DNA duplex, *J. Am. Chem. Soc. 126*, 2862−2869.
71. Umezawa, Y., and Nishio, M. (2002) Thymine-methyl/π interaction implicated in the sequence-dependent deformability of DNA, *Nucl. Acids Res. 30*, 2183-2192.
72. Wang, S., and Kool, E. T. (1995) Origins of the Large Differences in Stability of DNA and RNA Helices: C-5 Methyl and 2′-Hydroxyl Effects, *Biochemistry 34*, 4125−4132.
73. Acharya, S., Barman, J., Cheruku, P., Chatterjee, S., Acharya, P., Isaksson, J., and Chattopadhyaya, J. (2004) Significant pK_a Perturbation of Nucleobases Is an Intrinsic Property of the Sequence Context in DNA and RNA, *J. Am. Chem. Soc. 126*, 8674−8681.
74. Acharya, P., Acharya, S., Földesi, A., and Chattopadhyaya, J. (2003) Tandem Electrostatic Effect from the First to the Third Aglycon in the Trimeric RNA Owing to the Nearest-Neighbor Interaction, *J. Am. Chem. Soc. 125*, 2094−2100.
75. Acharya, P., Acharya, S., Cheruku, P., Amirkhanov, N. V., Földesi, A., and Chattopadhyaya, J. (2003) Cross-Modulation of the pK_a of Nucleobases in a Single-Stranded Hexameric-RNA Due to Tandem Electrostatic Nearest-Neighbor Interactions, *J. Am. Chem. Soc. 125*, 9948−9961.

76. Guo, D., Sijbesma, R. P., and Zuilhof, H. (2004) π-Stacked Quadruply Hydrogen-Bonded Dimers: π-Stacking Influences H-Bonding, *Org. Lett. 6*, 3667−3670.
77. Shindo, H., Fujiwara, T., Akutsu, H., Matsumoto, U., and Kyogoku, Y. (1985) Phosphorous-31 nuclear magnetic resonance of highly oriented DNA fibers. 1. Static geometry of DNA double helices, *Biochemistry 24*, 887−895.
78. Fujiwara, T., and Shindo, H. (1985) Phosphorous-31 nuclear magnetic resonance of highly oriented DNA fibers. 2. Molecular motions in hydrated DNA, *Biochemistry 24*, 896−902.
79. Thomas, G. J., Benevides, J. M., Overman, S. A., Ueda, T., Ushizawa, K., Saitoh, M., and Tsuboi, M. (1995) Polarized Raman spectra of oriented fibers of A DNA and B DNA: anisotropic and isotropic local Raman tensors of base and backbone vibrations, *Biophys. J. 68*, 1073−1088.
80. Hagerman, P. J. (1997) Flexibility of RNA, *Annu. Rev. Biophys. Biomol. Struct. 26*, 139−156.
81. Hagerman, P. J. (1988) Flexibility of DNA, *Annu. Rev. Biophys. Biophys. Struct. 17*, 265−286.
82. Noy, A., Pérez, A., Lankas, F., Luque, F. J., and Orozco, M. (2004) Relative Flexibility of DNA and RNA: a Molecular Dynamics Study, *J. Mol. Biol. 343*, 627−638.
83. Pan, Y., and MacKerell, A. D., Jr. (2003) Altered structural fluctuations in duplex RNA versus DNA: a conformational switch involving base pair opening, *Nucleic Acids Res. 31*, 7131−7140.
84. Chattopadhyaya, J. (1996) The nature of intramolecular stereoelectronic forces in nucleosides and nucleotides, *Nucleic Acids Symp. Ser. 35*, 111−112.
85. Plavec, J., Thibaudeau, C., and Chattopadhyaya, J. (1996) How do the energetics of the stereoelectronic gauche and anomeric effects modulate the conformation of nucleosides and nucleotides, *Pure Appl. Chem. 68*, 2137−2144.
86. Thibaudeau, C., and Chattopadhyaya, J. (1999) *Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications*, Uppsala University Press, Uppsala, Sweden.

BI047414F