Thermostability Trends of TNA:DNA Duplexes Reveal Strong Purine Dependence

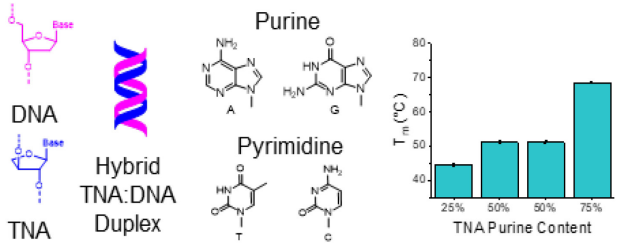
Hershel H. Lackey[†], Eric M. Peterson[†], Zhe Chen[†], Joel M. Harris[†], and Jennifer M. Heemstra[*,†,‡]

[†] Department of Chemistry, University of Utah, Salt Lake City, UT, 84112, USA

[‡] Department of Chemistry, Emory University, Atlanta, GA, 30322, USA

*To whom correspondence should be addressed. Tel: +1 404 727 7766: jen.heemstra@emory.edu

ABSTRACT



The development of high fidelity polymerases and streamlined synthesis of threose nucleic acid (TNA) triphosphates and phosphoramidites has made TNA accessible as a motif for generating nuclease-resistant high-affinity aptamers, antisense oligos, and synthetic genetic biopolymers. Little is known, however, about the thermostability trends of TNA:DNA duplexes. Here we investigate the thermostability of 14 TNA:DNA duplexes with the goal of elucidating the fundamental factors governing TNA:DNA duplex stability. We find that purine content in TNA significantly influences the stability and conformation of TNA:DNA duplexes. Low TNA purine content destabilizes duplexes, with T_m values often 5 °C lower than analogous DNA:DNA and RNA:DNA duplexes. By contrast, TNA:DNA duplexes having high TNA purine content display greater stability than DNA:DNA or RNA:DNA duplexes having the same sequences. High TNA purine content leads TNA:DNA duplexes to adopt conformations similar to TNA:TNA (A-form) configuration, whereas duplexes with low TNA purine content have conformations more similar to DNA:DNA (B-form) configuration. These insights provide a basis for understanding and predicting TNA:DNA duplex stability, which is anticipated to guide the practical use of TNA in biotechnology applications.

Keywords: XNA, TNA, purine, nearest neighbor, nucleic acid dynamics, pre-organization

Nucleic acids provide an information-rich architecture that can be harnessed for a wide variety of applications in biotechnology and medicine beyond the confines of native cellular environments. DNA and RNA, however, are quickly degraded in nuclease-rich environments, significantly hindering their utility in many of these applications. To address this challenge, chemical modifications to the structure of DNA and RNA can be employed to reduce recognition by nucleases, slowing or preventing degradation [1, 2]. Researchers have developed non-natural nucleic acid systems, or XNAs, that are capable of Watson-

Crick base pairing with themselves and with DNA or RNA [3]. This recognition capability and compatibility with polymerase enzymes is necessary for the transfer and replication of their genetic information, which in turn is critical for biological evolution and other biotechnology applications. However, many of the structural motifs that enable an XNA to serve as a polymerase substrate allow the XNA to serve as a nuclease substrate, making them subject to polymerase degradation (albeit at a slower rate than DNA or RNA) [4].

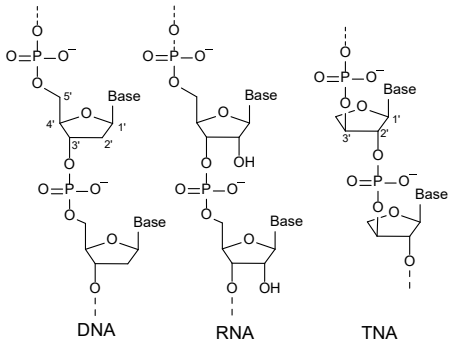


Figure 1. Chemical structures of DNA, RNA, and TNA. DNA and RNA connectivity is 5'-3', while TNA connectivity is 3'-2'.

TNA (α-l-(3'-2')-threofuranosyl nucleic acid) (Figure 1) has emerged as a particularly promising XNA for biological applications, as it is capable of base pairing and enzymatic replication, yet has an exceptionally high resistance to degradation by nucleases. TNA hybridizes through Watson-Crick base-pairing with complementary strands of TNA, DNA, and RNA [5-7], and can adopt secondary structures similar to DNA or RNA, including hairpins and G-quadruplexes [5, 8]. TNA shows greater stability in nuclease-rich environments compared to DNA or RNA, as well as the modified nucleic acids 2'F RNA and FANA [4], surviving more than a week under nuclease-rich conditions[4, 9]. In addition, recent research has provided convenient strategies for the synthesis of TNA phosphoramidites and triphosphates [10-13] and generated engineered polymerases for high-fidelity transcription and reverse transcription between TNA and DNA [14, 15]. These capabilities make TNA an attractive candidate for the development of nuclease-resistant aptamers for use in cellular environments, and such aptamers have been evolved for protein and small-molecule targets [9, 15-17].

While TNA aptamers hold significant promise, their utility is hindered by the lack of comprehensive information regarding the sequence-specific thermodynamics of TNA hybridization. For example, enzymatic replication of TNA requires the melting temperature (T_m) between a TNA primer and a DNA template to be tuned to enable hybridization at the optimal temperature for polymerase extension efficiency [14]. In addition, a better understanding of the conformation and stability of TNA:DNA duplexes and their relationship to substrate binding and extension kinetics could enable the development of improved TNA polymerases [18]. Models for TNA duplex stability are also critical to the post-selection engineering of TNA aptamers, especially in applications such as structure-switching biosensors which

require a precise balancing the thermodynamics of duplex stability and substrate binding [19, 20]. TNA also has utility for applications beyond the development and use of aptamers, for example in antisense therapeutics [21] or potentially as a genetic polymer in artificial biological systems [22]. These applications similarly require detailed understanding and prediction of the thermodynamics of hybridization between TNA and natural nucleic acids.

Significant effort has been devoted to elucidating the conformation of TNA in Watson-Crick base-paired duplexes. NMR-based analysis of a TNA:TNA duplex indicates that the threose sugars adopt a C4'-*exo* conformation with the 2'- and 3'- substituents assuming a quasi-diaxial orientation [23]. X-ray crystallography of a TNA nucleotide (nt) embedded in a self-complementary DNA strand shows a similar conformation for the threose sugar [24, 25]. Circular dichroism (CD) experiments indicate that TNA:DNA duplexes appear to adopt an A-form conformation similar to RNA:RNA duplexes [26]. Given these data, TNA is believed to be more structurally analogous to RNA than DNA, with relatively limited flexibility in its sugar backbone [24]. An investigation of the dynamics of a TNA:DNA heteroduplex using NMR revealed asymmetric fraying at the 3'-3' end (the 3' terminus of TNA is functionally equivalent to the 5' terminus of DNA/RNA in duplex formation), which differs significantly from the more symmetrical fraying seen in natural nucleic acid duplexes [26]. This result suggests that trends in TNA duplex thermodynamic stability may differ from those observed for natural nucleic acids, but there has been relatively little work focused on studying these structure-stability relationships.

From the limited data reported on TNA:DNA duplex stability, a correlation can be seen with TNA purine content. However, most of the sequences used in these studies contain only A and T nucleotides (nt) [5-7] which may not be representative of the overall thermostability trends [27], especially since tracts of A and T/U are known to induce conformational changes in DNA:DNA [28] and stability differences in RNA:DNA duplexes [29]. To date, there have only been a few studies reporting thermodynamic parameters of TNA:DNA duplexes having GC content, and these indicate conflicting results. In one example, two TNA:DNA duplexes, each having eight base pairs (bp), 50% GC content, and 50% purine content, were reported to have T_m values approximately 10 °C lower than the homologous RNA:DNA, DNA:RNA, or DNA:DNA duplexes [5], while in another the T_m of a 10 bp TNA:DNA duplex having 60% GC content and 60% purine content was only 1 °C lower than its homologous DNA:DNA duplex [26]. These differing results indicate that DNA:TNA duplex stability may have a complex relationship with GC content, purine content, and sequence, and that a more systematic study of these parameters is needed to understand trends in duplex stability for use in biotechnology applications.

With that aim, we sought to systematically probe the effect of sequence order, duplex length, GC content, and purine content on the thermostability of 14 TNA:DNA duplexes. We initially surveyed a series of 8-mer TNA:DNA duplexes with constant nearest neighbor (NN) parameters, but varying order. We observed no significant variation in duplex stability, suggesting that the thermodynamic parameters of TNA:DNA duplex formation can be predicted using the NN model with no higher-order sequence dependence.

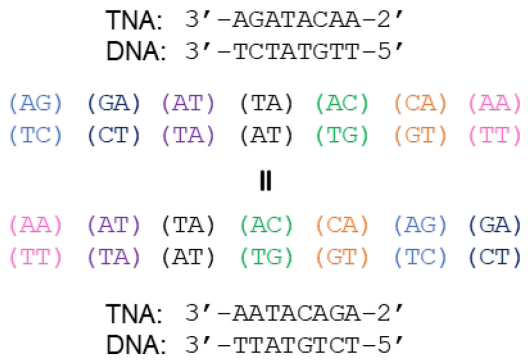
Within this set, TNA:DNA duplexes having higher GC content, but the same purine content, were found to be more stable, indicating that TNA:DNA duplexes rely on Watson-Crick base pairing in a similar paradigm to natural nucleic acid duplexes[30]. We find that TNA purine content increases TNA:DNA duplex stability to a greater degree than RNA purine content in RNA:DNA duplexes. Structural investigations using CD spectroscopy reveal that TNA:DNA duplex conformation is highly dependent on purine content; duplexes with moderate or high TNA purine content adopt an A-form conformation similar to TNA:TNA, and TNA:DNA duplexes having low purine content adopt a unique conformation that appears to be a hybrid of A- and B-form. Purine content also affects the relationship between duplex length and stability. In duplexes with high purine content, increasing length from 8 to 12 bp increased the stability of TNA:DNA duplexes more than RNA:DNA or DNA:DNA duplexes, while duplexes having low TNA purine content show weaker length dependence. Together, these data reveal the key factors influencing TNA:DNA duplex stability and provide a heuristic model that can be applied toward the design of TNA sequences for biotechnology applications.

RESULTS

Sequence Ordering Effect on the Thermostability of TNA:DNA Duplexes

The NN model has historically served as the basis for predicting thermodynamic properties for the formation of a nucleic acid duplex by summing the empirically determined energy contribution of each adjacent NN bp term in a duplex [31, 32]. The simpilicity of these models, as well as their improved accuracy over those accounting only for GC content and duplex length, has led to wide implementation for a variety of nucleic acid duplexes, including DNA:DNA [33], RNA:DNA [34], RNA:RNA [35], and LNA:DNA [36], and multiple online calculators applying these models are available for public use [37, 38]. Despite increasing interest in using TNA for biological and biomedical applications, the amenability of TNA:DNA duplex formation to prediction using an NN model has yet to be explored. Moreover, little is known about the long-range interactions of TNA NN basepairs, and NMR studies indicate that TNA duplexes could denature in a more asymmetric mechanism than either DNA:DNA or RNA:RNA duplexes [26]. These and other factors could preclude the use of NN models to predict the thermodynamic properties of TNA:DNA duplex formation.

To investigate the feasiblity of modeling TNA:DNA duplex stability using NN parameters, we adopted procedures previously used to study RNA:DNA duplexes [34]. In this prior study, the melting point of two duplexes having the same NN parameters, but different sequence order, were compared (Scheme 1). To test the impact of long-range ordering on duplex stability, we synthesized coupled 8 nt TNA sequences, wherein both sequences in each couple had the same NN parameters but varying sequence order (Table 1). If stability depends only on the NN content, and not sequence order over multiple nucleotides, we would expect these coupled sequences to have similar thermodynamic stabilities. For most TNA sequences having identical NN terms, the nucleotides at each terminus were maintained constant, with the exception of the TNA:DNA duplexes in Table 1, Entry 7 and 8. These duplexes were selected to investigate whether terminal nucleotides have a significant impact on the stability of the complex.



Scheme 1. Design of sequences to test applicability of NN model. Coupled TNA:DNA duplexes have the same NN units, but varying sequence order.

Table 1. Comparison of thermodynamic parameters for TNA:DNA duplex formation between coupled sequences having identical NN sets.[a]

Entry	Hybrid Duplex	ΔH° (kcal/mol)	ΔS° (cal/mol·K)	$\Delta G°_{37}$ (kcal/mol)	T_m (°C)
1	t(3'-AGATACAA-2') d(3'-TCTATGTT-5')	-60.7 ± 1.1	-178 ± 4	-5.6 ± 0.1	25.2 ± 0.3
2	t(3'-AATACAGA-2') d(3'-TTATGTCT-5')	-61.2 ± 1.1	-180 ± 4	-5.6 ± 0.1	25.1 ± 0.4
3	t(3'-AAGCGTAG-2') d(3'-TTCGCATC-5')	-62.3 ± 0.3	-177 ± 1	-7.4 ± 0.1	36.1 ± 0.2
4	t(3'-AGCGTAAG-2') d(3'-TCGCATTC-5′)	-60.8 ± 0.1	-173 ± 1	-7.1 ± 0.1	35.0 ± 0.2
5	t(3'-CTACGCTT-2') d(3'-GATGCGAA-5')[b]	-41.8 ± 0.2	-117 ± 1	-5.5 ± 0.1	20.3 ± 0.2
6	t(3'-CTTACGCT-2') d(5'-GAATGCGA-3')[b]	-54.3 ± 0.8	-158 ± 5	-5.4 ± 0.3	22.8 ± 0.1
7	t(3'-AGTCCTGA-2') d(3'-TCAGGACT-5')[b]	-61.2 ± 0.7	-184 ± 3	-4.1 ± 0.2	20.0 ± 0.1
8	t(3'-CTGAGTCC-2') d(3'-GACTCAGG-5')	-57.6 ± 0.1	-169 ± 7	-4.8 ± 0.1	22.9 ± 0.3
9	t(3'-GAGCCGTG-2') d(3'-CTCGGCAC-5')	-63.3 ± 0.7	-177 ± 2	-8.3 ± 0.1	40.8 ± 0.2
10	t(3'-GCCGTGAG-2') d(3'-CGGCACTC-5')	-67.2 ± 1.2	-190 ± 4	-8.2 ± 0.1	39.9 ± 0.1

[a]Experiments were performed in 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA, pH = 7.0, with an total oligo concentration of 10 μM (5 μM of each oligo) unless otherwise noted.
[b]Complexes displayed non-two state melting, while their ΔH and ΔS values are unreliable, their ΔG and T_m remain pertinent for comparison. See SI for details.
[c] Errors represent one standard deviation.

Before embarking on our study of TNA:DNA duplexes, we first validated our choice of the NN duplex couples in Table 1 by performing thermodynamic analysis of DNA:DNA duplexes having equivalent sequences. UV melting analysis was performed for each duplex at 10 μM total oligonucleotide

concentration (5 μM of each oligonucleotide) in 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA, pH 7.0 aqueous buffer. The absorbance of the solution at 260 nm was monitored as the temperature was lowered slowly from 80 °C to 4 °C and then returned to 80 °C. The thermodynamic parameters for duplex formation were determined from each melting curve using van't Hoff analysis, in which the melting curve was fit to a non-linear equation as previously described [36, 39]. The DNA:DNA duplexes having the same NN units displayed similar thermodynamic parameters to one another, with average differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$ and T_m between the two duplexes of only 1.6%, 1.7%, 2.7% and 0.9 °C, respectively (SI Table 1). In addition, the thermodynamic parameters were close to those predicted using empirical models developed by Sugimoto and Weber [33, 40], with an average difference in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$ and T_m of 5.5 %, 6.9 %, 3.2 %, and 0.4 °C, respectively (SI Table 1). These differences are within the level of error typically reported for UV melting curve analysis, indicating that for these DNA:DNA duplex sequences, the thermodynamics of duplex formation can be obtained using UV melting curve analysis and the NN model can accurately predict the thermodynamic parameters of duplex formation. These results also demonstrate that the single-stranded DNA oligonucleotides do not form secondary structures having sufficient stability to interfere with duplex formation.

Having validated our sequence set, we next investigated the impact of NN ordering on the thermodynamics of TNA:DNA hybridization . The duplexes in Table 1 were analyzed using UV melting analysis under the same conditions described above for the DNA:DNA duplexes. The thermodynamic parameters were similar between the duplexes having the same NN terms, with an average difference in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$ and T_m of 7.7 %, 9.0%, 4.4%, and 1.5 °C, respectively. These differences are similar to those observed for the RNA:DNA duplexes, suggesting that the thermodynamic properties of TNA:DNA duplex formation can be predicted using an NN model. In addition, the data suggest that the precise location of GC content or purine content within the strand does not significantly impact the stability of the duplex.

Purine Content Effect on Duplex Stability and Conformation

Previous work evaluating the influence of purine content on TNA:DNA duplex stability focused on a series of four 16 bp TNA:DNA duplexes having 0%, 50%, and 100% adenine and no GC content [5]. The two duplexes composed of 50% purine content exhibited T_m values of 41°C and 47 °C, whereas the stability of the A_{16}:T_{16} duplexes varied significantly depending on backbone assignment. Duplexes having the A_{16} TNA and T_{16} DNA had a T_m of 68 °C, compared to 32 °C when the backbone identities were reversed. Comparison of these duplexes with homologous RNA and DNA duplexes showed that the T_m for the t(A_{16})/d(T_{16}) duplex is higher than for r(A_{16})/r(U_{16}), while the T_m for the t(T_{16})/d(A_{16}) duplex is lower than than for either d(T_{16})/d(A_{16}) or r(U_{16})/d(A_{16}). However, it is unclear whether these represent universal trends in purine content, or if they are restricted to duplexes having only A and T nucleotides.

To explore this question, we investigated TNA:DNA duplexes having all four nucleotides to determine whether the same trends would be observed, and if so, the magnitude of the impact compared to RNA:DNA duplexes. Working from our 8'mer set in Table 1, we utilized sequences in which GC content was held constant at 50% while purine content was varied between 25-75% (Figure 2). T_m values were found to be significantly higher (36.1 °C and 35.0 °C) for duplexes having 75% purine content compared to duplexes having 50% (20.0 °C) or 25% purine content (20.3 and 22.8 °C). Furthermore, in comparing these TNA:DNA duplexes to their RNA:DNA analogues, sequences having 75% purine content showed similar stability, while sequences having 50% or 25% purine content showed significantly lower stability compared to their RNA:DNA counterparts (Figure 2). Comparison of the remaining 8 bp TNA:DNA duplexes having variable GC content to their RNA:DNA analogues supported this trend, with higher TNA purine content resulting in more stable duplexes (SI Figure S5). Specifically, duplexes having 75% and 62.5 % TNA purine content exhibited stabilities similar to their analogous RNA:DNA duplexes (ΔT_m of *ca* +2 °C and -3 °C respectively), whereas duplexes having 37.5% TNA purine content exhibited significant destabilization compared to the RNA:DNA equivalent (ΔT_m of *ca* -10 °C). These results indicate that while GC content and purine content both strongly influence overall stability, there is no apparent cooperativity between these two factors.

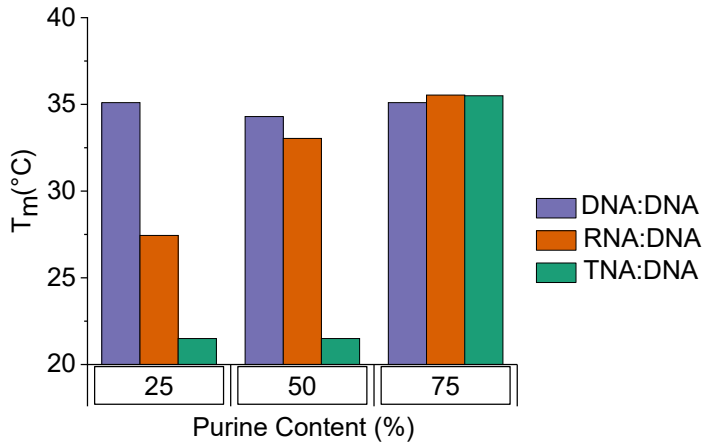


Figure 2. Average T_m values for duplexes having 50% GC content and varying purine content. T_m values for RNA:DNA duplexes are estimated from NN parameters developed by Sugimoto and coworkers [34].

Although TNA:DNA duplexes having low TNA purine content showed significantly reduced stability compared to DNA:DNA or RNA:DNA duplexes, they were still capable of strong hybridization to complementary TNA sequences. Using the 8'mer duplex sequences above having 25% TNA purine content, we performed UV melting analysis on the analogous TNA:TNA duplexes, with thermodynamic parameters summarized in Table S4. The TNA duplexes showed T_m values of 34.2 °C and 35.5 °C,

similar to homologous DNA duplexes, having T_m values of 34.9 °C and 34.8 °C. The similarity of T_m values between DNA and TNA homoduplexes with low purine content demonstrates that isolating purine content onto one strand does not destabilize all TNA duplexes.

To determine whether purine content plays a similar role in duplexes of longer lengths, four 12'mer TNA sequences were synthesized (Table 2). The TNA sequences were designed to each have 50% GC content. To allow for comparison of thermodynamic parameters across TNA:DNA, TNA:TNA, TNA:RNA, RNA:DNA, and DNA:DNA duplexes, the TNA entries in Table 2 were designed to be complementary to one another (Entry 2/3 with 50% TNA purine content and Entry 1/4 with 25% and 75% purine content). Similar to the shorter strands, the thermostability of the 12 bp TNA:DNA duplexes is highly dependent on the TNA purine content. The 12 bp TNA:DNA duplexes with 50% purine content showed similar T_m values (51.0 and 51.1 °C). In contrast, the TNA:DNA duplex having 25% TNA purine content displayed a reduced T_m value of 44.6 °C, while the TNA:DNA duplex having 75% TNA purine content showed the highest stability with a T_m of 68.5 °C.

Table 2. Thermodynamic parameters for TNA:DNA duplex formation of 12'mers having 50% GC content and varied TNA purine content.[a,b]

Entry	Hybrid Duplex	TNA Purine (%)	ΔH° (kcal/mol)	ΔS° (cal/mol·K)	$\Delta G°_{37}$ (kcal/mol)	T_m (°C)
1	t(3'-ACGTCATTCCTC-2') d(3'-TGCAGTAAGGAG-5')	25	-76.2 ± 0.5	-216 ± 1	-9.4 ± 0.1	44.6 ± 0.1
2	t(3'-GCAATGTTCAGC-2') d(3'-CGTTACAAGTCG-5')	50	-84.3 ± 1.5	-236 ± 5	-11.2 ± 0.1	51.1 ± 0.1
3	t(3'-GCTGAACATTGC-2') d(3'-CGACTTGTAACG-5')	50	-78.3 ± 0.9	-217 ± 3	-10.9 ± 0.1	51.0 ± 0.1
4	t(3'-GAGGAATGACGT-2') d(3'-CTCCTTACTGCA-5')	75	-123.8 ± 3.5	-339 ± 10	-19.6 ± 0.5	68.5 ± 0.2

[a]UV melting analysis conditions are provided in Table 1.

[b]Errors represent one standard deviation.

As highlighted in the data above, TNA purine content significantly influences TNA:DNA stability. To illustrate this effect relative to other nucleic acid duplexes, melting temperatures for various homo- and heteroduplexes having the same sequence were measured and are compared in Figure 3. For 12'mer duplexes having 50% purine content, TNA:DNA duplexes showed similar results to RNA:DNA duplexes, and were less stable than any homoduplex or the TNA:RNA duplexes (Figure 3A). The 12'mer heteroduplexes having 25% or 75% TNA purine content showed significantly more variation, as was expected (Figure 3B). Swapping the purine content from TNA to DNA decreases the T_m of a TNA:DNA heteroduplex by ~25 °C, and the TNA:DNA duplex is significantly more sensitive to purine content than the RNA:DNA duplex, which shows a shift in T_m of only ~8°C when swapping purine content. Interestingly, the effect of swapping purine content is similar for TNA:RNA and DNA:RNA duplexes. We

were also surprised to observe that the TNA:DNA and TNA:RNA duplexes having high TNA purine content showed greater stability than TNA or DNA homoduplexes.

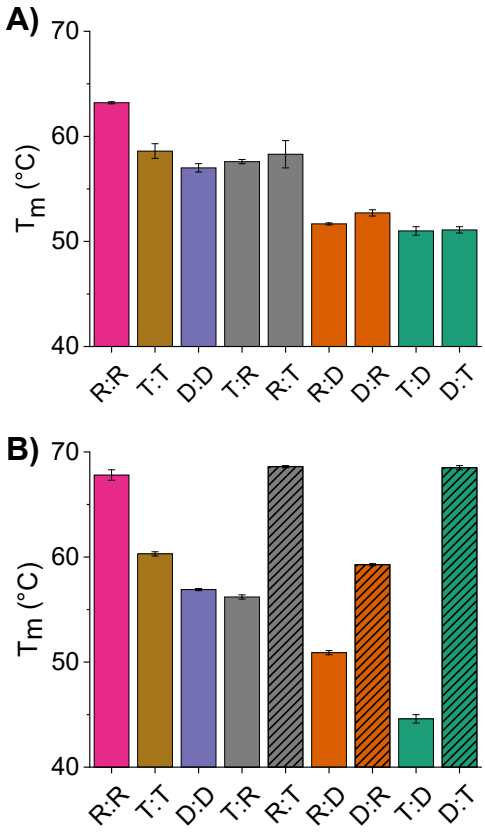


Figure 3. (A) T_m values for 12'mer duplexes having 50% GC content and 50% purine content (Table 2, Entry 2 & 3). (B) T_m values for 12'mer duplexes having 50% GC content and split of 25% and 75% purine content (Table 2, Entry 1 & 4). For TNA:DNA and TNA:RNA hetero-duplexes, hashed bar represents duplex having 75% TNA purine content; for DNA:RNA heteroduplex, hashed bar represents duplex having 75% RNA purine content.

We hypothesized that the impact of TNA purine content could be related to a change in duplex conformation, and thus CD spectroscopy was used to explore the conformational differences between the 12'mer duplexes (Figure 4 A & B). Spectra for duplexes having 50% purine content in each sequence (Table 2, Entries 2 & 3) are plotted in Figure 4A according to nucleic acid backbone type. DNA:DNA duplexes displayed the expected spectral characteristics of a B-form conformation, with a positive long wavelength peak between 260-280 nm and a negative peak at approximately 245 nm [41-43] and RNA:RNA duplexes displayed the expected spectral characteristics of an A-form conformation, with a prominent positive peak at 260 nm and a strong negative peak at 210 nm [41]. The TNA:DNA and TNA:TNA duplexes show similar spectra to one another, but with a hybrid of the characteristics of A- and B-form conformations [26]. The similarity in CD spectra for the two TNA:DNA duplexes each having 50% TNA purine content suggests that duplex conformation is not significantly impacted by the order of nucleotides

in the sequence. The RNA:TNA duplexes have CD spectra that fall between those of the homologous RNA:RNA and TNA:TNA duplexes and suggest an A-form conformation (Figure S6A). Compared to the duplexes having 50% purine content in each strand, duplexes having mixed purine content show more varied structures (Figure 4B). As expected, the DNA:DNA duplex adopts a B-form conformation, while the RNA:RNA duplex adopts an A-form conformation. For TNA:DNA duplexes, high TNA purine content leads to a conformation similar to TNA:TNA, whereas the low TNA purine content duplex shows a large positive peak at approximately 275 nm and a shallow negative peak at approximately 210 nm, potentially indicating a unique conformation with some aspects of both A and B form configurations. Overall, the CD spectra demonstrate that TNA purine content does significantly impact duplex conformation, and we hypothesize that this could be at least in part responsible for the influence on duplex stability.

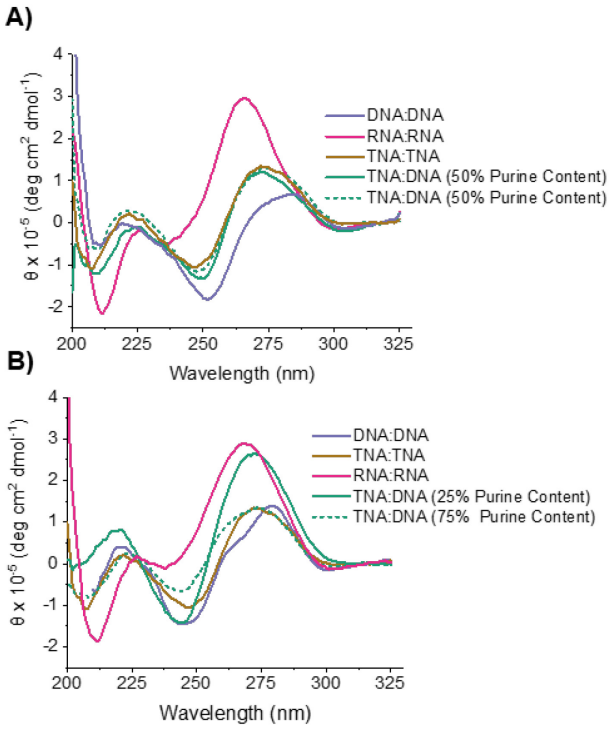


Figure 4. (A) CD spectra of sequences analogous to Table 2 entries 2 and 3, having 50% purine content. RNA:RNA (pink), TNA:TNA (brown), TNA:DNA (smooth green), DNA:TNA (dashed green), DNA:DNA (blue). (B) CD spectra of sequences analogous to Table 2 entries 1 and 4, having 25% and 75% purine content. Spectral color scheme is identical color scheme as part A. TNA:DNA duplex containing 25% TNA purine content (green-smoothed), TNA:DNA containing 75% TNA purine content (green-dashed).

GC Content Effect on Duplex Stability

GC content is known to increase the stability of nucleic acid duplexes [27, 34]; however, a meta-analysis of multiple studies suggests that GC content has slightly less influence on duplex stability for RNA:DNA than for DNA:DNA [44]. Given our observations that TNA:DNA duplexes behave more like RNA:DNA with

regard to structure and the impact of purine content, we were curious whether TNA:DNA duplexes also shared similar trends for the effect of GC content. To approach this question, we returned to the 8'mer duplexes, holding purine content constant at 75%, while varying GC content at 25% or 50% (Table 1, Entry 1-4). As expected, the data in Figure 5 indicate that increasing the GC content of TNA:DNA duplexes increases stability. Backbone identity may have a small impact on the magnitude of this effect, as the T_m difference between the 25% and 50% GC content duplexes is slightly smaller for TNA:DNA duplexes compared to RNA:DNA or DNA:DNA duplexes. However, additional investigation using a larger sequence set would be needed to better understand whether these small differences in stability changes are significant.

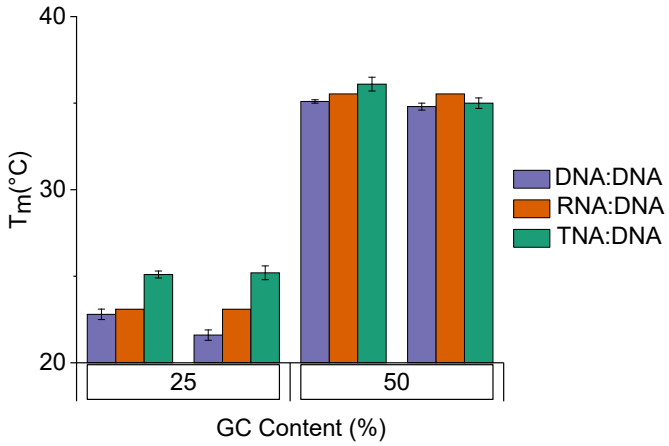


Figure 5. Effect of GC content on duplex stability as a function of backbone identity. RNA:DNA T_m values are estimated from NN parameters determined by Sugimoto Lab [34].

Interplay of Duplex Length with Purine Content in Determining Stability

The impact of duplex length on stability is well-understood for duplexes containing DNA and RNA, but in the case of TNA-containing duplexes, relatively little is known regarding these trends. The length dependence of TNA duplexes has potential to be especially interesting given that the repeating backbone unit in TNA is one atom shorter than DNA or RNA, and in long heterodoplex sequences the cumulative effect of this difference could result in pronounced instability. To investigate the impact of duplex length on stability, the GC content was held constant at 50%, and the duplex length was varied. We also systematically varied the TNA purine content to investigate the interplay of this factor with duplex length. As shown in Figure 6A, increasing the length of a TNA:DNA duplex from 8 to 12 bp unsurprisingly increases the stability of the duplex, and this trend is observed for duplexes having 25%, 50%, and 75% TNA purine content.

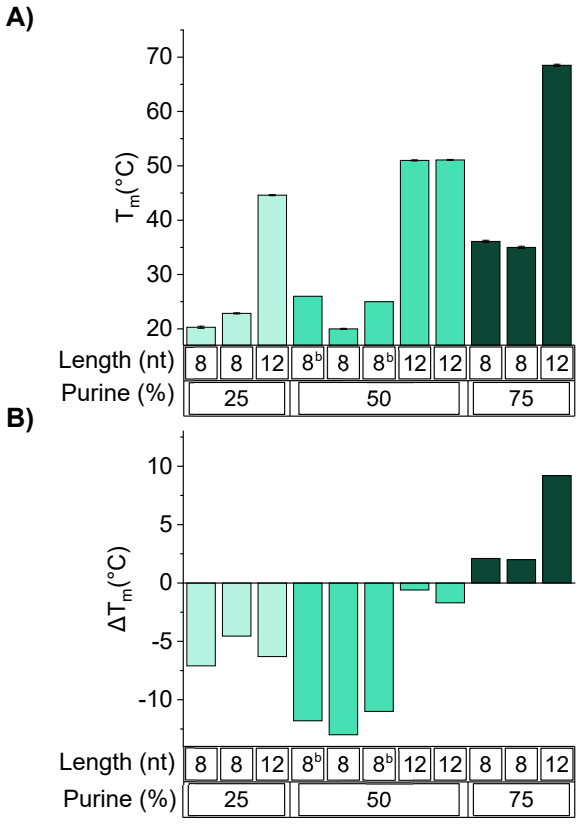


Figure 6. Stability of TNA:DNA duplexes having 50% GC content as a function of length and TNA purine content. A) T_m values of TNA:DNA duplexes grouped by purine content. B) ΔT_m values for TNA:DNA duplexes and analogous RNA:DNA duplexes at varying purine contents and lengths.

[a]ΔT_m is the calculated difference between the T_m of the TNA:DNA duplex and homologous RNA:DNA duplex. 8 bp RNA:DNA duplexes T_m values are estimated from NN reported by Sugimoto unless otherwise noted [34].
[b]Previously reported T_m values for TNA:DNA and RNA:DNA duplexes [5]. ΔT_m is calculated directly from analogous RNA:DNA duplex.

Interestingly, TNA purine content played a large role in how the length of the duplex affected its stability compared to RNA:DNA duplexes (Figure 6B). In TNA:DNA duplexes having 25% TNA purine content, duplex stability remained significantly lower than that of homologous RNA:DNA duplexes at lengths of both 8 and 12 bp. At 50% TNA purine content, 8 bp TNA:DNA duplexes were still significantly destabilized compared to homologous RNA:DNA duplexes, but 12 bp TNA:DNA duplexes showed only minor destabilization compared to homologous RNA:DNA duplexes. For TNA:DNA duplexes having 75% purine content, the stabilities of 8 bp duplexes were slightly higher than those of homologous RNA:DNA duplexes, while 12 bp TNA:DNA duplexes displayed significantly higher stability than RNA:DNA duplexes (Figure 6B). These data suggest that purine content impacts the relationship between duplex length and thermostability for TNA:DNA duplexes differently than for analogous RNA:DNA duplexes. At low (25%) TNA purine content, the effect of duplex length is similar for TNA:DNA and RNA:DNA, but at 50% and

75% TNA purine content, increasing duplex length is disproportionally beneficial for stability in TNA:DNA relative to RNA:DNA.

DISCUSSION

Threose purine content is a dominant factor in the stability and conformation of TNA:DNA duplexes, which parallels the trends observed for naturally occurring RNA:DNA duplexes. Increasing RNA purine content increases the stability of RNA:DNA duplexes,[29, 45-47] and also like TNA, increasing purine content in RNA causes the duplex to adopt a more A-form like conformation.[29, 46, 48-49] NMR studies show that the deoxyribose sugars adjust conformation in the RNA:DNA duplex in relation to the RNA oligonucleotide. For example, the ribose sugars in an RNA:DNA duplex adopt a C3'-*endo* configuration, the same configuration adopted in RNA:RNA duplexes,[50] while the 2'-deoxyribose sugars adopt a mix of C3'-*endo*/ C2'-*endo* configurations,[51-52] differing from the *C2'-endo* configuration the 2'-deoxyribose sugar typically adopts in a DNA:DNA duplex. Thus, formation of an RNA:DNA duplex is dependent on the DNA oligonucleotide to adjust to the more static RNA structure. Recent molecular dynamics simulations have shown that deoxyribose purine content greatly effects the flexibility of the DNA to conform to the RNA sequence. Specifically, a DNA oligonucleotide with high pyrimidine content (smaller nucleobases) can more easily conform to a complementary RNA oligonucleotide than a DNA sequence with high purine content, in which conformational flexibility is presumably more limited by the larger purine bases. Thus, high pyrimidine content in DNA sequences facilitates the conformation change needed to hybridize with RNA in an A-form duplex, leading to more favorable thermodynamics of hybridization.

Similar to RNA, the TNA sugar conformation is constrained, with the threose sugar adopting a C4'-*exo* configuration.[23-25] As a result, flexibility is required in the 2'-deoxyribose sugar to form a TNA:DNA duplex.[24] The data presented here show that increased purine content in TNA (and thus increased pyrimidine content in DNA) leads to stronger duplex formation with DNA, and that this increased purine content drives the duplex to adopt conformations more similar to a RNA:RNA duplex (A-form). This suggests a related mechanism and conformational requirements for duplex formation of TNA and RNA with DNA. However, while TNA does follow the same trends as RNA, it displays a much greater range of duplex stabilities with DNA, which may be attributable to the more rigid nature of the TNA backbone. Similar to other artificial nucleic acids having rigid backbones, TNA may pre-organize into a helical conformation, and thus enjoy increased duplex stability with complementary sequences that provide a good fit to this structure or have high conformational flexibility to adapt to it. Overall, the theories discussed above support our observed trends for TNA hydbridization, and future biophysical studies of TNA will continue to enhance our understanding of the similarities and differences between TNA, natural nucleic acids, and other XNAs.

CONCLUSION

Overall, TNA and RNA share general thermodynamic trends when hybridizing to complementary DNA sequences, experiencing similar impact from variations in GC content, length, and purine content. However, the impact of TNA purine content is magnified relative to this effect in RNA duplexes. The current study leaves some questions unanswered and highlights the need for further investigation of TNA:DNA duplex structure and stability. However, the data and trends reported here shed light on a significant new portion of the TNA duplex stability landscape. We anticipate that this will prove to be enabling for the rational design of TNA sequences to be used in biotechnology applications.

MATERIALS AND METHODS

Synthesis of Oligonucleotides

DNA and RNA phosphoramidites were purchased from Glen Research (Sterling, VA). TNA phosphoramidites were synthesized according to previously published procedures [13]. Oligonucleotides were synthesized by the DNA/Peptide Core Facility, part of the Health Sciences Center Cores at the University of Utah. The oligos were desalted and lyophilized after synthesis and used without additional purification. Oligos were resuspended in degassed, Millipore water to give approximately 100 μM concentration stock solutions. The concentration of each oligo was determined by absorbance at 260 nm using molar extinction coefficient determined through the IDT oligo calculator (https://www.idtdna.com/calc/analyzer). TNA molar extinction coefficients were approximated using equivalent DNA sequences. Oligo solutions were aliquoted and stored at 4 °C for short-term use or -20 °C for long-term use.

Melting Curve Experiments and Determination of Thermodynamic Parameters

Parameters for melting curve experiments were adapted from previously published procedures [54]. All melting curve experiments were performed in 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA, pH = 7.0. Oligo concentrations for each experiment were 5 μM of each strand for a total concentration of 10 μM. Melting experiments were performed on a Shimadzu UV-1800 combined with the TMSPC-8 Thermal Melt Analysis System using an eight-well quartz cuvette. Samples of 125 μL were placed in each sample cell of the cuvette and capped with mineral oil. Samples were brought to their max temperature (typically 80 °C) and allowed to equilibrate for 15 minutes to allow for full denaturation. The temperature was then decreased to 4 °C at a rate of 1.0 °C/min or 0.5 °C/min. The temperature was then increased back to the starting temperature at the same rate. Absorbance was monitored at 260 and 315 nm with readings taken every ~0.5 °C for both the forward and reverse runs. Absorbance at 315 nm was monitored to control for the formation of bubbles within the samples. Melting curves where the reverse curve was not within 5% of the forward curve were discarded. All samples were run in triplicate unless otherwise noted. Melting curves were fit to a non-linear equation using MATLAB, based on previously described fitting procedures [36, 39]; Matlab scripts are included in Supplementary Data.

CD Spectroscopy

CD spectra for experiments were obtained on AVIV Model 410 CD Spectrometer. Oligos were suspended in a 1.0 M NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA, pH = 7.0 buffer solution. Samples were measured at 25 °C. Oligo concentrations for each experiment were 5 μM of each sequence for a total concentration of 10 μM. Using a 0.1 cm pathlength cuvette, three scans of each solution, from 325 to 200 nm wavelength, were obtained, and these scans were averaged. CD spectra were baseline corrected and smoothed using the adjacent averaging function in Origin.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: (insert link)

Supplementary tables, supplementary figures, UV thermal denaturation curves, and Matlab scripts used to fit van't Hoff models to data (PDF link).

Melting point prediction software used in this paper is available online. Umelt software is available, courtesy of the Wittwer Lab, at (https://www.dna.utah.edu/umelt/umelt.html). IDT Oligo Calculator is available at (http://www.idtdna.com/calc/analyzer).

AUTHOR INFORMATION

Coresponding Author

*E-mail: jen.heemstra@emory.edu

ORCID

Jennifer Heemstra: 0000-0002-7691-8526

Present Address

Department of Chemistry, Emory University, Atlanta, GA, 30322, USA

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

HHL was supported by the Air Force Institute of Technology Civilian Institute Program. This work was supported by the DARPA Folded Non-Natural Polymers with Biological Function (Fold F(x)) Program [N66001-14-2-4054]. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of DARPA. This work was also supported by the National Science Foundation (DMR 1822262 and CHE 1818781 to J. Heemstra and CHE 1608949 to J. Harris). The authors would like to thank Dr. Peter Flynn for his review of the manuscript.

REFERENCES

[1] Meek, K. N., Rangel, A. E., and Heemstra, J. M. (2016) Enhancing aptamer function and stability via in vitro selection using modified nucleic acids, *Methods 106*, 29-36.
[2] Deleavey, Glen F., and Damha, Masad J. (2012) Designing Chemically Modified Oligonucleotides for Targeted Gene Silencing, *Chemistry & Biology 19*, 937-954.
[3] Anosova, I., Kowal, E. A., Dunn, M. R., Chaput, J. C., Van Horn, W. D., and Egli, M. (2016) The structural diversity of artificial genetic polymers, *Nucleic Acids Research 44*, 1007-1021.
[4] Culbertson, M. C., Temburnikar, K. W., Sau, S. P., Liao, J.-Y., Bala, S., and Chaput, J. C. (2016) Evaluating TNA stability under simulated physiological conditions, *Bioorganic & Medicinal Chemistry Letters 26*, 2418-2421.
[5] Schöning, K.-U., 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.
[6] Schöning, K.-U., Scholz, P., Wu, X., Guntha, S., Delgado, G., Krishnamurthy, R., and Eschenmoser, A. (2002) The α-L-Threofuranosyl-(3’→2’)-oligonucleotide System (‘TNA'): Synthesis and Pairing Properties, *Helvetica Chimica Acta 85*, 4111-4153.
[7] Yang, Y.-W., Zhang, S., McCullum, E. O., and Chaput, J. C. (2007) Experimental Evidence That GNA and TNA Were Not Sequential Polymers in the Prebiotic Evolution of RNA, *Journal of Molecular Evolution 65*, 289-295.
[8] Liao, J.-y., Anosova, I., Bala, S., Van Horn, W. D., and Chaput, J. C. (2017) A parallel stranded G-quadruplex composed of threose nucleic acid (TNA), *Biopolymers 107*, e22999-n/a.
[9] Rangel, A. E., Chen, Z., Ayele, T. M., and Heemstra, J. M. (2018) In vitro selection of an XNA aptamer capable of small-molecule recognition, *Nucleic Acids Research 46*, 8057-8068.
[10] Sau, S. P., and Chaput, J. C. (2017) A Gram-Scale HPLC-Free Synthesis of TNA Triphosphates Using an Iterative Phosphorylation Strategy, *Organic Letters 19*, 4379-4382.
[11] Bala, S., Liao, J.-Y., Mei, H., and Chaput, J. C. (2017) Synthesis of α-l-Threofuranosyl Nucleoside 3’-Monophosphates, 3’-Phosphoro(2-Methyl)imidazolides, and 3’-Triphosphates, *The Journal of Organic Chemistry 82*, 5910-5916.
[12] Sau, S. P., and Chaput, J. C. (2016) A one-pot synthesis of α-l-threofuranosyl nucleoside triphosphates (tNTPs), *Bioorganic & Medicinal Chemistry Letters 26*, 3271-3273.
[13] Sau, S. P., Fahmi, N. E., Liao, J.-Y., Bala, S., and Chaput, J. C. (2016) A Scalable Synthesis of α-l-Threose Nucleic Acid Monomers, *The Journal of Organic Chemistry 81*, 2302-2307.
[14] Yu, H., Zhang, S., Dunn, M. R., and Chaput, J. C. (2013) An Efficient and Faithful in Vitro Replication System for Threose Nucleic Acid, *Journal of the American Chemical Society 135*, 3583-3591.
[15] Pinheiro, V. B., Taylor, A. I., Cozens, C., Abramov, M., Renders, M., Zhang, S., Chaput, J. C., Wengel, J., Peak-Chew, S.-Y., McLaughlin, S. H., Herdewijn, P., and Holliger, P. (2012) Synthetic Genetic Polymers Capable of Heredity and Evolution.
[16] Yu, H., Zhang, S., and Chaput, J. C. (2012) Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor, *Nature Chemistry 4*, 183.
[17] Mei, H., Liao, J.-Y., Jimenez, R. M., Wang, Y., Bala, S., McCloskey, C., Switzer, C., and Chaput, J. C. (2018) Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues, *Journal of the American Chemical Society*.
[18] Chim, N., Shi, C., Sau, S. P., Nikoomanzar, A., and Chaput, J. C. (2017) Structural basis for TNA synthesis by an engineered TNA polymerase, *Nature Communications 8*, 1810.
[19] Cho, E. J., Lee, J.-W., and Ellington, A. D. (2009) Applications of Aptamers as Sensors, *Annual Review of Analytical Chemistry 2*, 241-264.
[20] Pfeiffer, F., and Mayer, G. (2016) Selection and Biosensor Application of Aptamers for Small Molecules, *Frontiers in Chemistry 4*, 25.

[21] Liu, L. S., Leung, H. M., Tam, D. Y., Lo, T. W., Wong, S. W., and Lo, P. K. (2018) α-l-Threose Nucleic Acids as Biocompatible Antisense Oligonucleotides for Suppressing Gene Expression in Living Cells, *ACS Applied Materials & Interfaces 10*, 9736-9743.
[22] Appella, D. H. (2009) Non-natural nucleic acids for synthetic biology, *Current Opinion in Chemical Biology 13*, 687-696.
[23] Ebert, M.-O., Mang, C., Krishnamurthy, R., Eschenmoser, A., and Jaun, B. (2008) The Structure of a TNA−TNA Complex in Solution: NMR Study of the Octamer Duplex Derived from α-(l)-Threofuranosyl-(3′-2′)-CGAATTCG, *Journal of the American Chemical Society 130*, 15105-15115.
[24] Pallan, P. S., Wilds, C. J., Wawrzak, Z., Krishnamurthy, R., Eschenmoser, A., and Egli, M. (2003) Why Does TNA Cross-Pair More Strongly with RNA Than with DNA? An Answer From X-ray Analysis, *Angewandte Chemie International Edition 42*, 5893-5895.
[25] Wilds, C. J., Wawrzak, Z., Krishnamurthy, R., Eschenmoser, A., and Egli, M. (2002) Crystal Structure of a B-Form DNA Duplex Containing (l)-α-Threofuranosyl (3′→2′) Nucleosides: A Four-Carbon Sugar Is Easily Accommodated into the Backbone of DNA, *Journal of the American Chemical Society 124*, 13716-13721.
[26] Anosova, I., Kowal, E. A., Sisco, N. J., Sau, S., Liao, J.-y., Bala, S., Rozners, E., Egli, M., Chaput, J. C., and Van Horn, W. D. (2016) Structural Insights into Conformation Differences between DNA/TNA and RNA/TNA Chimeric Duplexes, *ChemBioChem 17*, 1705-1708.
[27] SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proceedings of the National Academy of Sciences 95*, 1460-1465.
[28] Hagerman, P. J. (1990) Sequence-Directed Curvature of DNA, *Annual Review of Biochemistry 59*, 755-781.
[29] Lesnik, E. A., and Freier, S. M. (1995) Relative Thermodynamic Stability of DNA, RNA, and DNA:RNA Hybrid Duplexes: Relationship with Base Composition and Structure, *Biochemistry 34*, 10807-10815.
[30] Watson, J. D., and Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid, *Nature 171*, 737.
[31] DeVoe, H., and Tinoco, I. (1962) The stability of helical polynucleotides: Base contributions, *Journal of Molecular Biology 4*, 500-517.
[32] Gray, D. M., and Tinoco, I. (1970) A new approach to the study of sequence-dependent properties of polynucleotides, *Biopolymers 9*, 223-244.
[33] Sugimoto, N., Nakano, S.-i., Yoneyama, M., and Honda, K.-i. (1996) Improved Thermodynamic Parameters and Helix Initiation Factor to Predict Stability of DNA Duplexes, *Nucleic Acids Research 24*, 4501-4505.
[34] Sugimoto, N., Nakano, S.-i., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995) Thermodynamic Parameters To Predict Stability of RNA/DNA Hybrid Duplexes, *Biochemistry 34*, 11211-11216.
[35] Xia, T., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson−Crick Base Pairs, *Biochemistry 37*, 14719-14735.
[36] McTigue, P. M., Peterson, R. J., and Kahn, J. D. (2004) Sequence-Dependent Thermodynamic Parameters for Locked Nucleic Acid (LNA)−DNA Duplex Formation, *Biochemistry 43*, 5388-5405.
[37] Tm Tool v1.5b; Wittwer Lab, Salt Lake City, Utah, USA. Accessed 15 March 2018 https://www.dna.utah.edu/tm/tool.html.
[38] OligoAnalyzer 3.1.; IDT, Coralville, Iowa, USA. Accessed 15 March 2018 https://www.idtdna.com/calc/analyzer.
[39] Wang, X., Hoshika, S., Peterson, R. J., Kim, M.-J., Benner, S. A., and Kahn, J. D. (2017) Biophysics of Artificially Expanded Genetic Information Systems. Thermodynamics of DNA Duplexes

Containing Matches and Mismatches Involving 2-Amino-3-nitropyridin-6-one (Z) and Imidazo[1,2-a]-1,3,5-triazin-4(8H)one (P), *ACS Synthetic Biology 6*, 782-792.
[40] Weber, G. (2015) Optimization method for obtaining nearest-neighbour DNA entropies and enthalpies directly from melting temperatures, *Bioinformatics 31*, 871-877.
[41] Kypr, J., Kejnovská, I., Renčiuk, D., and Vorlíčková, M. (2009) Circular dichroism and conformational polymorphism of DNA, *Nucleic Acids Research 37*, 1713-1725.
[42] Bloomfield, V. A. (2000) *Nucleic acids : structures, properties, and functions*, Sausalito, Calif. : University Science Books.
[43] Berova, N., Nakanishi, K. j., and Woody, R. (2000) *Circular dichroism : principles and applications*, 2nd ed.. ed., New York : Wiley-VCH.
[44] Vologodskiĭ, A. V. (2015) *Biophysics of DNA*, Cambridge ; New York : Cambridge University Press.
[45] Gyi, J. I., Conn, G. L., Lane, A. N., and Brown, T. (1996) Comparison of the Thermodynamic Stabilities and Solution Conformations of DNA·RNA Hybrids Containing Purine-Rich and Pyrimidine-Rich Strands with DNA and RNA Duplexes, *Biochemistry 35*, 12538-12548.
[46] Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) Sequence Specific Thermodynamic and Structural Properties for DNA.cntdot.RNA Duplexes, *Biochemistry 33*, 5298-5304.
[47] Shaw, N. N., and Arya, D. P. (2008) Recognition of the unique structure of DNA:RNA hybrids, *Biochimie 90*, 1026-1039.
[48] Pramanik, S., Nagatoishi, S., Saxena, S., Bhattacharyya, J., and Sugimoto, N. (2011) Conformational Flexibility Influences Degree of Hydration of Nucleic Acid Hybrids, *The Journal of Physical Chemistry B 115*, 13862-13872.
[49] Suresh, G., and Priyakumar, U. D. (2014) DNA-RNA hybrid duplexes with decreasing pyrimidine content in the DNA strand provide structural snapshots for the A- to B-form conformational transition of nucleic acids, *Physical Chemistry Chemical Physics 16*, 18148-18155.
[50] Chou, S. H., Flynn, P., and Reid, B. (1989) Solid-phase synthesis and high-resolution NMR studies of two synthetic double-helical RNA dodecamers: r(CGCGAAUUCGCG) and r(CGCGUAUACGCG), *Biochemistry 28*, 2422-2435.
[51] Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., and Reid, B. R. (1993) The DNA strand in DNA.cntdot.RNA hybrid duplexes is neither B-form nor A-form in solution, *Biochemistry 32*, 4207-4215.
[52] Fedoroff, O. Y., Salazar, M., and Reid, B. R. (1993) Structure of a DNA : RNA Hybrid Duplex: Why RNase H Does Not Cleave Pure RNA, *Journal of Molecular Biology 233*, 509-523.
[53] Kool, E. T. (1997) Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides, *Chemical Reviews 97*, 1473-1488.
[54] Mergny, J.-L., and Lacroix, L. (2004) Analysis of Thermal Melting Curves, *Oligonucleotides 13*, 22.