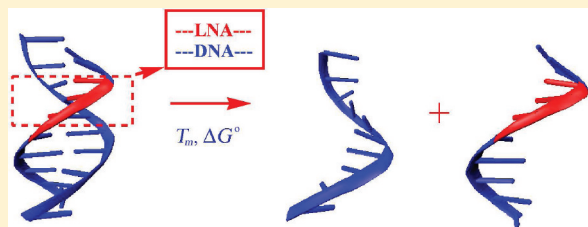# Stability and Mismatch Discrimination of Locked Nucleic Acid−DNA Duplexes

Richard Owczarzy,* Yong You, Christopher L. Groth, and Andrey V. Tataurov

Department of Molecular Genetics and Biophysics, Integrated DNA Technologies, 1710 Commercial Park, Coralville, Iowa 52241, United States

S *Supporting Information*

ABSTRACT: Locked nucleic acids (LNA; symbols of bases, +A, +C, +G, and +T) are introduced into chemically synthesized oligonucleotides to increase duplex stability and specificity. To understand these effects, we have determined thermodynamic parameters of consecutive LNA nucleotides. We present guidelines for the design of LNA oligonucleotides and introduce free online software that predicts the stability of any LNA duplex oligomer. Thermodynamic analysis shows that the single strand−duplex transition is characterized by a favorable enthalpic change and by an unfavorable loss of entropy. A single LNA modification confines the local conformation of nucleotides, causing a smaller, less unfavorable entropic loss when the single strand is restricted to the rigid duplex structure. Additional LNAs adjacent to the initial modification appear to enhance stacking and H-bonding interactions because they increase the enthalpic contributions to duplex stabilization. New nearest-neighbor parameters correctly forecast the positive and negative effects of LNAs on mismatch discrimination. Specificity is enhanced in a majority of sequences and is dependent on mismatch type and adjacent base pairs; the largest discriminatory boost occurs for the central +C·C mismatch within the +T+C+C sequence and the +A·G mismatch within the +T+A+G sequence. LNAs do not affect specificity in some sequences and even impair it for many +G·T and +C·A mismatches. The level of mismatch discrimination decreases the most for the central +G·T mismatch within the +G+G+C sequence and the +C·A mismatch within the +G+C+G sequence. We hypothesize that these discrimination changes are not unique features of LNAs but originate from the shift of the duplex conformation from B-form to A-form.

A locked nucleic acid (LNA) is a useful chemical modification.[1−5] Mixed oligonucleotides consisting of LNA, DNA, and RNA residues have improved polymerase chain reaction (PCR) experiments,[6] single-nucleotide polymorphism assays,[7,8] RNA interference,[1,4] antisense mRNA technology,[2] microRNA profiling and regulation,[9,10] aptamers,[11] LNAzymes,[3] microarrays,[12] and nanomaterials.[13] These applications require that LNA oligonucleotides possess specific melting temperatures (T_m) and free energies of association for complementary sequences ($\Delta G°$).[5]

The thermodynamic stability of nucleic acid duplexes has been described with the nearest-neighbor model, which takes into account energetics of nearest-neighbor base pairs and assumes that interactions beyond neighboring nucleotides can be neglected.[14−17] The total enthalpy and entropy of duplex annealing are calculated by summation of doublet terms

$$\Delta H° = \sum_{i=1}^{N_{bp}-1} \Delta H°_{i,i+1} + \Delta H°_{init} \tag{1}$$

$$\Delta S° = \sum_{i=1}^{N_{bp}-1} \Delta S°_{i,i+1} + \Delta S°_{init} + \Delta S°_{symmetry} \tag{2}$$

where N_{bp} is the number of duplex base pairs. The first term on the right side of eq 1 is the sum over all internal nearest-neighbor doublets ($\Delta H°_{i,i+1}$). The second term ($\Delta H°_{init}$) represents the "initiation" enthalpy, which includes the formation of the duplex first base pair, corrections for the extra hydrogen bond of G·C versus A·T in terminal base pairs,[17] and terminal base−solvent interactions. The initiation parameter varies with the nature of terminal base pairs.[15,16] Equation 2 also includes an entropic symmetry correction ($\Delta S°_{symmetry}$) of −1.4 cal mol^{-1} K^{-1}, which is added when a duplex consists of two identical, self-complementary oligonucleotides.

The nearest-neighbor model accurately predicts thermodynamics and melting temperatures (±1.5 °C) of native oligonucleotides.[15−19] It appears that the nearest-neighbor model also predicts well single-base mismatches[18,20] and some chemical modifications, including single LNAs.[21−24] Because LNAs increase duplex stability and change the specificity of base pairing,[2,5] LNA nearest-neighbor parameters differ significantly from DNA parameters.

(A)

Set 1	TXRD–CGTCA+T**+X**+TCGC
Set 2	TXRD–CGTAC+A**+X**+ATGC
Set 3	TXRD–AGCAC+C**+X**+CTAC
Set 4	TXRD–TGGAC+G**+X**+GACT
Set 5	TXRD–CTATG+A**+X**+GTAAC
Set 6	TXRD–CATC+T**+X**+CAAA
Set 7	TXRD–TAGGA+C**+X**+ACT
Set 8	TXRD–ACTC+G**+X**+TAGG

TXRD ≡ Texas Red dye
IBRQ ≡ Iowa Black RQ quencher
+A,+C,+G,+T ≡ LNA nucleotides

(B)

$$\Delta\Delta H^o$$
$$\Delta\Delta S^o$$
$$\Delta\Delta G^o$$

TXRD — CGTCA**+T+G+T**CGC–3′
IBRQ — GCAGT **A C A**GCG–5′

TXRD — CGTCA**TGT**CGC–3′
IBRQ — GCAGT**ACA**GCG–5′

$$\Delta\Delta H^o\begin{pmatrix} A+T+G+TC \\ T\ A\ C\ AG \end{pmatrix} =$$

$$= \Delta H^o\begin{pmatrix} TXRD-CGTCA+T+G+TCGC \\ IBRQ-GCAGT\ A\ C\ AGCG \end{pmatrix} - \Delta H^o\begin{pmatrix} TXRD-CGTCATGTCGC \\ IBRQ-GCAGTACAGCG \end{pmatrix}$$
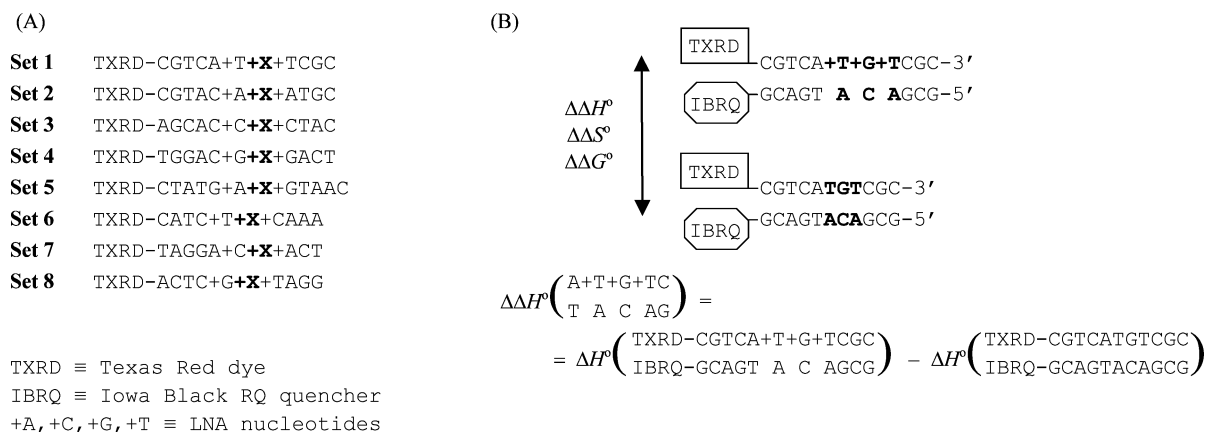
Figure 1. (A) Eight sets of sequences were studied. Various matched and mismatched LNA·DNA base pairs were introduced at the "X·Y" site (X and Y can be A, C, G, T, +A, +C, +G, or +T). (B) Example of analysis. Thermodynamic effects of LNA nucleotides (+N) were determined from the stability differences between LNA-modified duplexes and core DNA sequences.

The LNA parameter set is incomplete and does not cover many useful sequences. Thermodynamic parameters have been published for isolated LNA·RNA base pairs introduced into 2′-O-methyl RNA oligonucleotides[25,26] and for isolated LNA·DNA base pairs.[21] However, many applications benefit from other types of LNA modifications. For example, a triplet of LNA residues appears to maximize mismatch discrimination and improves single-nucleotide polymorphism assays.[5] Fully LNA-modified probes can selectively capture genomic DNA sequences.[27] To determine the parameters for consecutive LNAs, we measured the stability of duplexes using the fluorescence melting method.[28] The energetics of LNA effects was determined from the difference between LNA-modified and native (core) duplexes. Because we used standard experimental conditions (1 M Na$^+$ and pH 7), new parameters are compatible with existing DNA parameters.

■ MATERIALS AND METHODS

Oligonucleotides were synthesized at Integrated DNA Technologies, purified by HPLC,[29] and dialyzed against storage buffer [10 mM Tris-HCl and 0.1 mM Na$_2$EDTA (pH 7.5)].[28] Concentrated oligonucleotide samples were tested by mass spectroscopy (molecular weights were within 2 g/mol) and capillary electrophoresis (>90% pure). DNA concentrations were determined from predicted extinction coefficients (ε) and sample absorbance at 260 nm using the Beer–Lambert law.[29,30] LNA nucleotides were assumed to possess the same extinction coefficients as DNA ones. Coefficients of Texas Red (14400 L mol^{-1} cm^{-1}) or Iowa Black RQ (44510 L mol^{-1} cm^{-1}) were added to the ε of labeled oligonucleotides.

Primary Set of Oligonucleotides. Figure 1A shows the sequences studied. Fluorescent Texas Red dye (TXRD) is attached at the 5′ end of the top strand, and Iowa Black RQ quencher (IBRQ) is attached at 3′ end of the complementary strand. This design efficiently quenches fluorescence when the strands are annealed because the dye and the quencher are in close contact. We use notation of oligonucleotide manufacturers; LNA nucleotides are indicated with + in front of the base symbol (e.g., +A denotes an adenine LNA nucleotide). Cytosine of +C is 5-methylated because oligonucleotide manufacturers usually synthesize the methylated version of LNA cytosine.

The set of DNA duplexes contains a triplet of consecutive LNAs located either in the interior of the strand labeled with

Texas Red or in the interior of the complementary strand labeled with Iowa Black RQ. Eight possible LNA·DNA base pairs (X·Y ≡ +A·T, A·+T, +T·A, T·+A, +C·G, C·+G, +G·C, and G·+C), and 24 mismatches were introduced at the X·Y site. Core duplexes were also measured. They contained DNA·DNA base pairs (X·Y ≡ A·T, T·A, C·G, and G·C), and the same terminal Texas Red–Iowa Black RQ pair was also measured. This design is economical; each oligonucleotide is used in several duplexes. Thirty-six duplexes were melted for each set except for set 3. This set consisted of 27 duplexes because its two sequences, GTAGGGGTGCT-IBRQ and GTA+G+G+GGTGCT-IBRQ, were not obtained with sufficient purity.

For sets 1−4, the same base flanks the X·Y site on the 5′ and 3′ sides. For sets 5−8, the flanking bases are different and each of the four bases (A, T, C, and G) occurs once on the 5′ and 3′ sides of the X·Y base pairs. This design ensures that every possible nearest-neighbor interaction is present several times within the data set.

Figure 1A also shows that duplex lengths range from 10 to 12 bp. Such short sequences are likely to melt in a two-state manner. Nevertheless, non-two-state behavior may occur even for short oligonucleotides if they form stable self-complementary structures, e.g., hairpins or dimers. OligoAnalyzer version 3.1 (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/) confirmed that our sequences do not form such structures.

This paper follows previous conventions to represent duplex sequences.[16] A slash divides the strands in an antiparallel orientation. The sequence is oriented 5′ to 3′ before the slash and 3′ to 5′ after the slash (for example, CA/GT represents the 5′-CA-3′/3′-GT-5′ doublet with two Watson–Crick base pairs). Mismatched nucleotides are underlined or colored red. Ribonucleotides are distinguished from deoxyribonucleotides by the "r" prefix, e.g., rA.

Melting Experiments. We followed our previously described method for fluorescence melting experiments.[28] The melting buffer contained 1 M NaCl, 3.87 mM NaH$_2$PO$_4$, 6.13 mM Na$_2$HPO$_4$, and 1 mM Na$_2$EDTA and was adjusted to pH 7.0 with 1 M NaOH.[30] Buffer reagents of p.a. grade purity were bought from ThermoFisher Scientific (Pittsburgh, PA).

Melting experiments were performed at 13 different total single-strand concentrations (19, 30, 46, 70, 110, 160, 250, 375, 570, and 870 nM and 1.3, 2.0, and 3.0 μM). Duplex samples

were prepared at the highest C_t of 3 μM. Complementary oligonucleotides were mixed in a 1:1 ratio in the melting buffer, heated to 95 °C, and slowly cooled to room temperature. Aliquots of the 3 μM solution were diluted with the melting buffer to make 12 remaining samples. Low-binding Costar microcentrifuge tubes (catalog no. 3207, Corning, Wilkes Barre, PA) were used to reduce the level of binding of oligonucleotides to the tube surface.

We pipetted 25 μL of the melting sample into two wells of a 96-well PCR plate (Extreme Uniform Thin Wall Plates, catalog no. B70501, BIOplastics BV, Landgraaf, The Netherlands). A significant discrepancy between wells alerted us to an erroneous measurement. Using the Bio-Rad iQ5 real-time PCR system, the fluorescence signal in the Texas Red channel was recorded every 0.2 °C while the temperature was increased from 4 to 98 °C and decreased back to 4 °C over two cycles. Subsequent temperature cycles were not used because they were unreliable; T_m sometimes increased, indicating the evaporation of water or degradation of dye. The iQ5 system maintained a temperature rate of 25 °C/h. Analysis was conducted in Microsoft Excel. We programmed VBA software to automate melting profile analysis, including baseline selection using a second-derivative algorithm.[28] The fraction θ was calculated $[\theta = (F - F_L)/(F_U - F_L)]$ from the fluorescence of the DNA sample (F), the fluorescence of the upper linear baseline (F_U), and the fluorescence of the lower linear baseline (F_L). If a duplex melts in a two-state manner, dissociation of the fluorophore from the quencher is likely coupled to the duplex-to-single strand melting transition and θ represents the fraction of melted duplexes.

The melting temperature was defined as the temperature at which $\theta = \frac{1}{2}$. The average standard deviation of T_m values was 0.4 °C. Transition enthalpies, entropies, and free energies were determined from fits to individual melting profiles and from the dependence of melting temperature on DNA concentration.[14,28,31,32] These two analytical methods assume that melting transitions proceed in a two-state manner; that is, intact duplex and unhybridized single strands are dominant, and partially melted duplexes are negligible throughout the melting transition. The methods also assume that transition enthalpies and entropies are temperature-independent. If $\Delta H°$ or $\Delta S°$ values differed more than 15% between these two methods, the duplex did not melt in a two-state manner.[28,32,33] In that case, we excluded $\Delta H°$ or $\Delta S°$ values from further analysis because they were inaccurate.

Stabilizing Effects of LNA Modifications. Locked nucleic acids increase duplex stability and alter the melting transition enthalpy, entropy, and free energy. As shown in Figure 1B, we determined these LNA contributions ($\Delta\Delta H°$, $\Delta\Delta S°$, and $\Delta\Delta G°_{37}$) from the difference between LNA-modified and core duplexes.[28] LNA modifications were located at least five nucleotides from the terminal fluorophore and quencher. In this design, terminal labels do not interact with LNAs and do not influence differential thermodynamic values between modified and core duplexes.

Figure 1B shows an example of the analysis for the Set1−11 duplex. Entering $\Delta H°$ from Table S1 of the Supporting Information, we determined the experimentally measured differential enthalpic change $[\Delta\Delta H°(A+T+G+TC/TACAG)]$ to be $-97.6 - (-86.4) = -11.2$ kcal/mol. In the nearest-neighbor model, this enthalpic contribution is a sum of enthalpies of base pair doublets

$$\Delta\Delta H°(A+T+G+TC/TACAG)$$
$$= \Delta\Delta H°(A+T/TA) + \Delta\Delta H°(+T+G/AC)$$
$$+ \Delta\Delta H°(+G+T/CA) + \Delta\Delta H°(+TC/AG) \quad (3)$$

Rearrangement of eq 3 places unknown LNA parameters on the left side

$$\Delta\Delta H°(+T+G/AC) + \Delta\Delta H°(+G+T/CA)$$
$$= \Delta\Delta H°(A+T+G+TC/TACAG)$$
$$- \Delta\Delta H°(A+T/TA) - \Delta\Delta H°(+TC/AG) \quad (4)$$

The right side of eq 4 contains the experimentally measured enthalpic change and two previously determined nearest-neighbor parameters.[21] McTigue, Peterson, and Kahn investigated the thermodynamics of interactions between LNA·DNA and DNA·DNA base pairs. We used their parameters to account for LNA−DNA interactions that occur in the beginning and at the end of a section of consecutive LNAs. Parameters from their 32NN set (Table 4 of ref 21) were entered into eq 4

$$\Delta\Delta H°(+T+G/AC) + \Delta\Delta H°(+G+T/CA)$$
$$= -11.2 - 1.8 - 1.9$$
$$= -14.9 \text{ kcal/mol} \quad (5)$$

A similar equation was constructed for each LNA duplex. Analogous equations were set up for $\Delta\Delta S°$ and $\Delta\Delta G°_{37}$ contributions.

Determination of LNA Nearest-Neighbor Parameters. Selecting two bases from the set of four (A, T, C, and G) with replacement leads to the creation of 16 nearest-neighbor doublets.[34] Because antiparallel strands of native DNA duplexes exhibit structural symmetry, some doublet sequences are identical, e.g., AC/TG and GT/CA. Therefore, 10 nearest-neighbor parameters are sufficient to represent internal DNA·DNA doublets. No such symmetry exists for LNA·DNA base pairs. The +A+C/TG doublet differs from the +G+T/CA doublet. Sixteen nearest-neighbor parameters are needed for consecutive LNA·DNA base pairs.

We measured 62 perfectly matched LNA duplexes. Sixty of them melted in a two-state manner. Their thermodynamic values were used to determine the parameters. Each of the 16 LNA doublets was well represented in this data set with the following numbers of occurrences: 8 +A+A/TT, 8 +A+C/TG, 8 +A+G/TC, 8 +A+T/TA, 8 +C+A/GT, 6 +C+C/GG, 7 +C+G/GC, 8 +C+T/GA, 8 +G+A/CT, 7 +G+C/CG, 4 +G+G/CC, 8 +G+T/CA, 8 +T+A/AT, 8 +T+C/AG, 8 +T+G/AC, and 8 +T+T/AA.

First, we examined enthalpic effects. Equation 5 was constructed for each LNA duplex. This thermodynamic analysis produced the set of 60 linear equations

$$\mathbf{M}\cdot\mathbf{H}^{n-n} = \mathbf{H}^{exp} \quad (6)$$

where $\mathbf{M}$ is a 60×16 matrix of the number of occurrences for each LNA nearest-neighbor doublet in 60 duplexes, $\mathbf{H}^{n-n}$ is the vector of 16 unknown parameters, and $\mathbf{H}^{exp}$ is the column vector of experimentally measured enthalpic contributions. The parameters reported by McTique et al. were subtracted from the enthalpic contributions as shown in eqs 4 and 5.

Because the number of unknown parameters (16) was less than the number of equations (60), eq 6 was overdetermined.[35] We solved it using singular-value decomposition (SVD)[36] by minimizing χ^2

$$\chi^2 = |(\mathbf{M}\cdot\mathbf{H}^{n-n} - \mathbf{H}^{exp}) \times \boldsymbol{\sigma}_H{}^{-1}|^2 \tag{7}$$

where $\boldsymbol{\sigma}_H$ is the diagonal matrix whose elements are experimental errors of $\Delta\Delta H°$. Because these errors were similar, they were set to a constant value of 3 kcal/mol and the SVD fit was not error weighted. Singular-value decomposition was conducted using Microsoft Excel Add-in, Matrix.xla package, version 2.3.2 (Foxes Team, L. Volpi, http://digilander.libero. it/foxes). Calculations were repeated using the Excel LINEST function, yielding the same values. We also examined matrix $\mathbf{M}$ for degeneracies. The rank of matrix $\mathbf{M}$ was 16. Because the rank was equal to the number of unknown parameters, the matrix had no singular values, and the parameters were unique and linearly independent.[34−36]

Next, we replaced $\Delta\Delta H°$ values with $\Delta\Delta S°$ or $\Delta\Delta G°_{37}$ values in eqs 5−7. Analogous analyses gave us nearest-neighbor parameters for entropies and free energies.

Error Analysis. Error estimates of parameters were obtained from bootstrap simulations.[37] These calculations estimate the dependence of parameter values on the data set. Many bootstrap data sets were created from the original data set. A different value of the parameter was usually determined from each bootstrap data set. The bootstrap estimate of the parameter error is given by the standard deviation of all these parameter values.

In our simulations, the bootstrap data sets were the same size as the original data set; i.e., the sets contained data from 60 duplexes. The duplex data were randomly drawn, with replacement, from the original data set. This means that the entire experimental data set was used in each drawing. This procedure produced bootstrap data sets in which some duplex data from the original data sets were present multiple times and other data were not selected. We generated 5×10^4 bootstrap data sets. Equation 6 was solved for each data set using SVD, and 16 parameters for the consecutive LNAs were determined. If the rank of $\mathbf{M}$ was less than 16, the particular bootstrap data set did not contain all possible nearest-neighbor doublet sequences. Thermodynamic parameters could not be determined in this case; therefore, the bootstrap set was excluded from analysis, and a replacement data set was drawn. Fewer than 3% of the data sets were excluded. Standard deviations and averages were calculated from bootstrap parameter estimates. The average parameters determined from bootstrap analysis agreed with the parameters determined from the original data set.

We have analyzed the error in free energy calculated from entropic and enthalpic contributions ($\Delta\Delta G° = \Delta\Delta H° - T\Delta\Delta S°$). Enthalpies and entropies of DNA melting transitions are correlated.[38] The errors of the enthalpic contribution, $\sigma(\Delta\Delta H°)$, and the entropic contribution, $\sigma(\Delta\Delta S°)$, are also highly correlated; their correlation coefficient is usually above 0.99.[17,21] If the uncertainty in $\Delta\Delta G°$ is estimated by error propagation,[39] the covariance $\mathrm{cov}(\Delta\Delta H°,\Delta\Delta S°)$ significantly decreases the error

$$\sigma(\Delta\Delta G°)^2 = \sigma(\Delta\Delta H°)^2 + T^2\sigma(\Delta\Delta S°)^2$$
$$- 2T\,\mathrm{cov}(\Delta\Delta H°, \Delta\Delta S°)^2 \tag{8}$$

Equation 8 indicates that the free energy is determined more precisely than the enthalpic or entropic contributions alone. The similar error compensation decreases the error in the

melting temperature calculated from $\Delta H°$ and $\Delta S°$.[17] This analysis demonstrates that it is useful to report the $\Delta\Delta H°$ and $\Delta\Delta S°$ parameters in Tables 1−3 beyond their individual errors. If the parameters are rounded to their error estimates, the calculated free energies and melting temperatures may be less precise.

Validation Melting Experiments. Validation sets were measured by ultraviolet spectroscopy as previously described.[30] Absorbance at 268 nm was recorded every 0.1 °C using a Beckman DU 650 spectrophotometer. The temperature was changed at a rate of 25 °C/h in the range from 10 to 98 °C using a high-performance temperature controller (Beckman-Coulter, Brea, CA). Both heating and cooling melting profiles were collected. Sloping baselines were subtracted from the melting profiles,[30] and the melting temperature was defined as the temperature at which the fraction of melted duplexes equaled 0.5.

Nearest-Neighbor Parameters for Single-Base Mismatches. There are 12 possible LNA·DNA mismatches (+A·A, +C·C, +G·G, +T·T, +A·C, +C·A, +A·G, +G·A, +C·T, +T·C, +G·T, and +T·G). In our design, mismatches were located in the center of LNA triplets. Enthalpic, entropic, and free energy effects were determined from the differences between the energetics of LNA mismatch duplexes and core DNA duplexes.

As an example, let us consider the Set1−17 duplex containing the +G·G mismatch. The enthalpic contribution from the A+T +G+TC/TAGAG duplex subsequence is calculated from the difference in the total enthalpy of the Set1−17 (TXRD-CGTCA +T+G+TCGC) and Set1−10 (TXRD-CGTCATGTCGC) duplexes (Table S1 of the Supporting Information)

$$\Delta\Delta H°(\text{A+T}\underline{+G}\text{+TC/TA}\underline{G}\text{AG})$$
$$= -81.2 - (-86.4)$$
$$= 5.2 \text{ kcal/mol} \tag{9}$$

The nearest-neighbor model assumes that this contribution is the sum of four nearest-neighbor doublets. We used the parameters of McTigue et al.[21] for two doublets (A+T/TA and +TC/AG). An equation similar to eq 4 was constructed. The left side contains two unknown parameters

$$\Delta\Delta H°(\text{+T }\underline{+G}\text{/A}\underline{G}) + \Delta\Delta H°(\underline{+G}\text{ +T/}\underline{G}\text{A})$$
$$= \Delta\Delta H°(\text{A+T}\underline{+G}\text{ +TC/TA}\underline{G}\text{AG})$$
$$- \Delta\Delta H°(\text{A+T/TA}) - \Delta\Delta H°(\text{+TC/AG})$$
$$= 5.2 - 1.8 - 1.9$$
$$= 1.5 \text{ kcal/mol} \tag{10}$$

Equation 10 was built for each mismatched duplex. Sequences having the +G·G mismatch were grouped into the subset. The resulting system of linear equations was overdetermined and was solved by SVD analysis. Eight unknown nearest-neighbor parameters (+A+X/TY, +C+X/GY, +G+X/CY, +T+X/AY, +X+A/YT, +X+C/YG, +X+G/YC, and +X+T/YA) were obtained (+X ≡ +G, and Y ≡ G). This procedure was repeated for 12 +X·Y mismatch types, and 96 parameters (8 × 12) were determined. The thermodynamic values of LNA duplexes containing mismatches can be predicted from eqs 1 and 2 using new parameters.

SVD analysis indicated that the number of linearly independent equations and the rank of matrix $\mathbf{M}$ was seven

for mismatches. Eight fitted nearest-neighbor parameters are useful, but they are not a unique solution[34,35,40] because a constraint equation relates the numbers of eight doublets (N)

$$N(+A\underline{+X}/T\underline{Y}) + N(+C\underline{+X}/G\underline{Y})$$
$$+ N(+G\underline{+X}/C\underline{Y}) + N(+T\underline{+X}/A\underline{Y})$$
$$= N(\underline{+X}+A/\underline{Y}T) + N(\underline{+X}+C/\underline{Y}G)$$
$$+ N(\underline{+X}+G/\underline{Y}C) + N(\underline{+X}+T/\underline{Y}A) \quad (11)$$

The constraint decreases the number of unique parameters to seven for each mismatch type. Equation 11 is valid for duplexes containing mismatches within consecutive LNAs.

Unique, linearly independent parameters can be constructed from linear combinations of eight nonunique parameters. The similar constraint limits the number of unique parameters for some DNA mismatches. Allawi and SantaLucia proposed seven linearly independent sequences for DNA mismatches.[41] They added a C·G base pair to nonunique doublets to create linearly independent triplets. Using a similar procedure, we have added the +C·G base pair to LNA doublets and created seven unique triplets (+A$\underline{+X}$+C/T$\underline{Y}$G, +C$\underline{+X}$+C/G$\underline{Y}$G, +G$\underline{+X}$+A/C$\underline{Y}$T, +G$\underline{+X}$+C/C$\underline{Y}$G, +G$\underline{+X}$+G/C$\underline{Y}$C, +G$\underline{+X}$+T/C$\underline{Y}$A, and +T$\underline{+X}$+C/A$\underline{Y}$G). A single LNA mismatch lies in the center. Using SVD analysis, seven parameters for those triplets were determined for each $\underline{+X}\cdot\underline{Y}$ mismatch type (Table S3 of the Supporting Information).

The energetics of any LNA $\underline{+X}\cdot\underline{Y}$ mismatch sequence (+K$\underline{+X}$+M/L$\underline{Y}$N) could also be calculated from unique triplet parameters

$$\Delta\Delta H°(+K\underline{+X}+M/L\underline{Y}N)$$
$$= \sum_{i=1}^{7} [N(\text{triplet}) \times \Delta\Delta H_i°(\text{triplet})] \quad (12)$$

The +K·L and +M·N pairs are Watson−Crick base pairs adjacent to the mismatch. The numbers of seven triplets, N(triplet), are related to the numbers of doublets: N(+A$\underline{+X}$+C/T$\underline{Y}$G) = N(+A$\underline{+X}$/T$\underline{Y}$), N(+C$\underline{+X}$+C/G$\underline{Y}$G) = N(+C$\underline{+X}$/G$\underline{Y}$), N(+G$\underline{+X}$+A/C$\underline{Y}$T) = N($\underline{+X}$+A/$\underline{Y}$T), N(+G$\underline{+X}$+C/C$\underline{Y}$G) = N(+G$\underline{+X}$/C$\underline{Y}$) − N($\underline{+X}$+A/$\underline{Y}$T) − N($\underline{+X}$+G/$\underline{Y}$C) − N($\underline{+X}$+T/$\underline{Y}$A), N(+G$\underline{+X}$+G/C$\underline{Y}$C) = N($\underline{+X}$+G/$\underline{Y}$C), N(+G$\underline{+X}$+T/C$\underline{Y}$A) = N($\underline{+X}$+T/$\underline{Y}$A), and N(+T$\underline{+X}$+C/A$\underline{Y}$G) = N(+T$\underline{+X}$/A$\underline{Y}$). The doublet and triplet parameter sets predict identical thermodynamic values for any LNA mismatch sequence. Both parameter sets are implementations of the same nearest-neighbor model and do not take into account any next-nearest-neighbor or longer interactions.

RESULTS

Nearest-Neighbor Parameters for Consecutive LNAs. Thermodynamic values were measured for the primary oligonucleotide set using fluorescence.[28] The melting process was monitored using Texas Red dye and Iowa Black RQ quencher, which were attached at the termini of duplexes. These labels appear to be optimal for melting experiments, as other fluorophores (FAM, HEX, and TET) do not provide reliable thermodynamic values and may ruin the two-state nature of melting transitions.[28]

Fluorescence versus temperature plots always exhibited single, S-shaped transitions that were reversible. Figure 2 presents examples of averaged melting profiles. Pictured duplexes have a
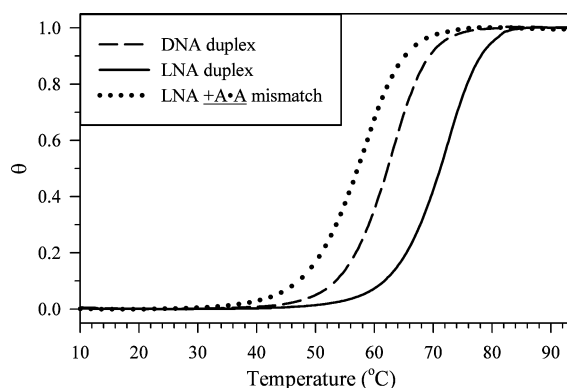

Figure 2. Average fluorescence melting profiles of LNA and core DNA duplexes. The DNA duplex of Set1−01 (TXRD-CGTCATATCGC) is shown (−−−). The first isosequential LNA duplex (Set1−02, TXRD-CGTCA+T+A+TCGC) is perfectly matched (—). The second LNA duplex (Set1−13, TXRD-CGTCA+T$\underline{+A}$+TCGC) contains the $\underline{+A\cdot A}$ mismatch (···). Melting experiments were conducted in 1 M Na$^+$ buffer, and C_t was 375 nM.

TXRD-CGTCA+T$\underline{+A}$+TCGC base sequence. The DNA matched duplex (dashed line) is more stable than the LNA duplex containing the $\underline{+A\cdot A}$ mismatch (dotted line) in Figure 2. This stability order is sequence-dependent and not universally observed. If LNAs cause large duplex stabilization and a single mismatch destabilizes a duplex less, the mismatched LNA duplex will be more stable than the matched DNA duplex of the same base sequence. This occurs often for $\underline{+G\cdot T}$, $\underline{+T\cdot G}$, $\underline{+G\cdot G}$, and $\underline{+G\cdot A}$ mismatches.

Thermodynamic values were extracted from melting profiles. First, the enthalpy, entropy, and free energy were estimated from fits to individual melting profiles.[28] We fitted only data within the transition where fraction θ ranged from 0.15 to 0.85. Second, $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ were determined from graphs of $1/T_m$ versus ln $C_t/4$. These graphs were linear over a 150-fold range of 13 DNA concentrations (Figure S1 of the Supporting Information). When the $1/T_m$ data point deviated from the fitted straight line by a value more than twice the value of the propagated error, it was removed from the fit as an outlier. Fewer than 1% of all graph points were excluded. Melting temperatures and thermodynamic values for the primary data set are presented in Table S1 of the Supporting Information. The enthalpy, entropy, and free energy are negative because they are reported for the annealing reaction, which is customary practice.

Our thermodynamic analysis assumed a two-state nature of melting transitions. When this assumption is valid, both $1/T_m$ versus ln $C_t/4$ plots and fits to melting profiles yield the same results. If thermodynamic values differed more than 15% between these two methods, the specific duplex did not melt in a two-state fashion, and its thermodynamic data were removed from further analysis, averages, and fitting of nearest-neighbor parameters. For the primary data set, average differences between both methods in $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ values were 7.2, 8.3, and 2.5%, respectively.

Duplexes exhibiting deviations from the two-state melting behavior are listed in Table S1 of the Supporting Information. The non-two-state melting transitions may occur when the cooperativity of the melting process is low and the duplex melts in several stages. The oligonucleotides can also fold into alternative stable structures, broadening the melting transition or splitting it into two S-shaped transitions. We did not observe the second transition in any melting profile. Duplexes also have

Table 1. Nearest-Neighbor Parameters for Differences between LNA·DNA and DNA·DNA Base Pairs

sequence[a]	$\Delta\Delta H°$ (kcal/mol)[b]	$\Delta\Delta S°$ (cal mol^{-1} K^{-1})[b]	$\Delta\Delta G°_{37}$ (kcal/mol)	$-310.15\Delta\Delta S°$ (kcal/mol)[c]
+A+A/TT	-2.091 ± 0.8	-4.975 ± 2.3	-0.6 ± 0.1	1.5
+A+C/TG	-2.989 ± 2.0	-6.563 ± 5.3	-1.0 ± 0.4	2.0
+A+G/TC	-4.993 ± 1.5	-10.607 ± 4.2	-1.8 ± 0.3	3.3
+A+T/TA	-7.503 ± 1.7	-20.350 ± 4.5	-1.2 ± 0.3	6.3
+C+A/GT	-5.677 ± 2.0	-12.798 ± 5.4	-1.7 ± 0.4	4.0
+C+C/GG	-7.399 ± 2.6	-16.475 ± 6.5	-2.3 ± 0.5	5.1
+C+G/GC	-3.958 ± 2.7	-8.039 ± 7.4	-1.5 ± 0.4	2.5
+C+T/GA	-7.937 ± 2.3	-20.218 ± 6.0	-1.6 ± 0.5	6.3
+G+A/CT	-5.759 ± 1.6	-12.897 ± 4.4	-1.7 ± 0.3	4.0
+G+C/CG	-6.309 ± 2.5	-16.338 ± 7.0	-1.2 ± 0.4	5.1
+G+G/CC	-5.022 ± 1.5	-9.773 ± 4.2	-2.0 ± 0.3	3.0
+G+T/CA	-8.961 ± 1.7	-23.458 ± 4.7	-1.7 ± 0.3	7.3
+T+A/AT	-3.118 ± 1.8	-4.808 ± 4.9	-1.6 ± 0.4	1.5
+T+C/AG	-0.966 ± 2.2	0.665 ± 5.8	-1.2 ± 0.5	-0.2
+T+G/AC	-1.546 ± 1.6	0.109 ± 4.4	-1.6 ± 0.3	0.0
+T+T/AA	-2.519 ± 1.4	-5.483 ± 3.5	-0.8 ± 0.3	1.7

[a]+N indicates an LNA nucleotide. The sequence orientation is 5′-LNA-3′/3′-DNA-5′. [b]Significant figures are shown beyond error estimates to allow accurate calculations and suppress rounding errors. See Error Analysis in Materials and Methods. [c]Entropic contribution to the Gibbs free energy at 37 °C.

a terminal dye−quencher pair that can interact with neighboring base pairs; this could change duplex melting behavior and local base pair cooperativity. Because fluorescence depends on dye−quencher distance and orientation, the fluorescent signal is more sensitive to non-two-state behavior than the UV absorbance signal. If dissociation of the dye from the quencher does not coincide with duplex melting, discrepancies in thermodynamic analysis are likely to occur and thermodynamic values could be inaccurate.

The majority of duplexes in the data set (>93%) exhibited two-state melting transitions, and the average $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ values of those duplexes were used to determine nearest-neighbor parameters. Table 1 shows the nearest-neighbor parameters for consecutive LNAs. Standard errors were estimated from bootstrap analysis. The free energy values calculated from the Gibbs thermodynamic relation ($\Delta\Delta G°_{37} = \Delta\Delta H° - 310.15\Delta\Delta S°$) agreed within 0.09 kcal/mol with the $\Delta\Delta G°_{37}$ values determined from SVD analysis. This agreement confirms the consistency of our method.

Because $\Delta\Delta G°_{37}$ is negative for all nearest-neighbor doublets in Table 1, consecutive LNAs always stabilize a DNA duplex and the effect is sequence-independent. The most stabilizing doublets are +C+C/GG ($\Delta\Delta G°_{37} = -2.3$ kcal/mol) and +G+G/CC ($\Delta\Delta G°_{37} = -2.0$ kcal/mol). The smallest LNA impact is seen for the +A+A/TT (-0.6 kcal/mol) and +T+T/ AA (-0.8 kcal/mol) sequences. Effects of LNAs on $\Delta\Delta G°_{37}$ are approximately proportional to the duplex fraction of G·C base pairs. Introduction of LNAs stabilizes cytosine-guanine base pairs ∼0.9 kcal/mol more than adenine-thymine base pairs. The $\Delta\Delta S°$ values vary widely from -23.5 to 0.7 cal mol^{-1} K^{-1}.

Software Implementation of New Parameters. Thermodynamic parameters in Table 1 are differential thermodynamic parameters; i.e., they represent deviations from native DNA duplexes. To calculate the total enthalpy for any LNA-modified sequence, one predicts the transition enthalpy for the native DNA duplex ($\Delta H°$) according to eq 1 and adds the differential parameters ($\Delta\Delta H°$) to take into account

LNA effects

$$\Delta H° = \sum_{i=1}^{N_{bp}-1} \Delta H°_{i,i+1} + \Delta H°_{init} + \sum^{LNA} \Delta\Delta H° \tag{13}$$

Both sums of eq 13 contain the same doublet sequences; the difference is in LNA modification (CA/GT vs +C+A/GT). Parameters for the same base sequences could be combined. Addition of differential LNA parameters ($\Delta\Delta H°$) and DNA nearest-neighbor parameters[16] gives full nearest-neighbor LNA parameters ($\Delta H°$)

$$\Delta H°(+K+X/LY)$$
$$= \Delta H°(KX/LY) + \Delta\Delta H°(+K+X/LY) \tag{14}$$

where +K+X/LY is a nearest-neighbor doublet. We present full thermodynamic parameters for consecutive and isolated LNA modifications in Table 2. It is faster and takes fewer computer resources to calculate thermodynamic values from full thermodynamic parameters than from differential ones.

As an example, we present calculations for the perfectly matched 5′-TA+C+AGG-3′ duplex.

$$\Delta H°(TA+C+AGG)$$
$$= \Delta H°(ET/EA) + \Delta H°(TA/AT)$$
$$+ \Delta H°(A+C/TG) + \Delta H°(+C+A/GT)$$
$$+ \Delta H°(+AG/TC) + \Delta H°(GG/CC)$$
$$+ \Delta H°(GE/CE)$$
$$= 2.3 - 7.2 - 5.5 - 14.2 - 7.5 - 8.0 + 0.1$$
$$= -40.0 \text{ kcal/mol} \tag{15}$$

The first and last parameters represent initiation interactions using the concept of a fictitious end base (E).[15,16,34] Transition entropies and free energies can also be casted into full parameters using analogous relationships.

Accuracy of Thermodynamic Parameters for Consecutive LNAs. To verify the analysis and applicability of the

Table 2. Full Thermodynamic Parameters for Perfectly Matched LNA·DNA Base Pairs in 1 M Na$^+$

sequence	$\Delta H°$ (kcal/mol)^a	$\Delta S°$ (cal mol^{-1} K^{-1})^a	$\Delta G°_{37}$ (kcal/mol)
	Consecutive LNA Modifications		
+A+A/TT	−9.991	−27.175	−1.57
+A+C/TG	−11.389	−28.963	−2.44
+A+G/TC	−12.793	−31.607	−3.07
+A+T/TA	−14.703	−40.750	−2.12
+C+A/GT	−14.177	−35.498	−3.11
+C+C/GG	−15.399	−36.375	−4.15
+C+G/GC	−14.558	−35.239	−3.65
+C+T/GA	−15.737	−41.218	−2.92
+G+A/CT	−13.959	−35.097	−3.03
+G+C/CG	−16.109	−40.738	−3.48
+G+G/CC	−13.022	−29.673	−3.82
+G+T/CA	−17.361	−45.858	−3.12
+T+A/AT	−10.318	−26.108	−2.19
+T+C/AG	−9.166	−21.535	−2.48
+T+G/AC	−10.046	−22.591	−3.08
+T+T/AA	−10.419	−27.683	−1.83
	Isolated LNA Modifications^b		
+AA/TT	−7.193	−19.723	−1.09
+AC/TG	−7.269	−18.336	−1.56
+AG/TC	−7.536	−18.387	−1.84
+AT/TA	−4.918	−12.943	−0.89
+CA/GT	−7.451	−18.380	−1.72
+CC/GG	−5.904	−11.904	−2.30
+CG/GC	−9.815	−23.491	−2.50
+CT/GA	−7.092	−16.825	−1.95
+GA/CT	−5.038	−11.656	−1.37
+GC/CG	−10.160	−24.651	−2.65
+GG/CC	−10.844	−26.580	−2.54
+GT/CA	−8.612	−22.327	−1.63
+TA/AT	−7.246	−19.738	−1.14
+TC/AG	−6.307	−15.515	−1.51
+TG/AC	−10.040	−25.744	−2.00
+TT/AA	−6.372	−16.902	−1.13
A+A/TT	−6.908	−18.135	−1.40
A+C/TG	−5.510	−11.824	−1.83
A+G/TC	−9.000	−22.826	−1.88
A+T/TA	−5.384	−13.537	−1.19
C+A/GT	−7.142	−18.333	−1.40
C+C/GG	−5.937	−12.335	−2.24
C+G/GC	−10.876	−27.918	−2.17
C+T/GA	−9.471	−25.070	−1.69
G+A/CT	−7.756	−19.302	−1.74
G+C/CG	−10.725	−25.511	−2.78
G+G/CC	−8.943	−20.833	−2.51
G+T/CA	−9.035	−22.742	−1.96
T+A/AT	−5.609	−16.019	−0.58
T+C/AG	−7.591	−19.031	−1.70
T+G/AC	−6.335	−15.537	−1.56
T+T/AA	−5.574	−14.149	−1.21

^{a}Significant figures are shown beyond error estimates to allow accurate calculations and suppress rounding errors. ^{b}Calculated from the 32NN set of ref 21.

nearest-neighbor model, we used new parameters to predict thermodynamics of the primary data set. New LNA parameters accurately predicted $\Delta H°$, $\Delta S°$, and $\Delta G°$ values for these short duplexes. The average relative errors were 3.3, 3.5, and 2.9%, respectively. This is comparable to the accuracy reported for nearest-neighbor parameters of native nucleic acids where

standard deviations of thermodynamic values ranged from 3 to 8%.[17]

To estimate the robustness of the new parameters, it is important to test their performance with an independent validation set of duplex oligomers that were not used to derive the parameters. We have measured 53 additional LNA-modified duplexes. The oligonucleotides did not have any fluorescent labels or quenchers attached. Their melting transitions were followed using UV spectroscopy.[30] These LNA duplexes ranged from 8 to 10 bp in lengths, from 10 to 88% in G·C content, and from 20 to 60% in LNA content. Figure 3 presents a
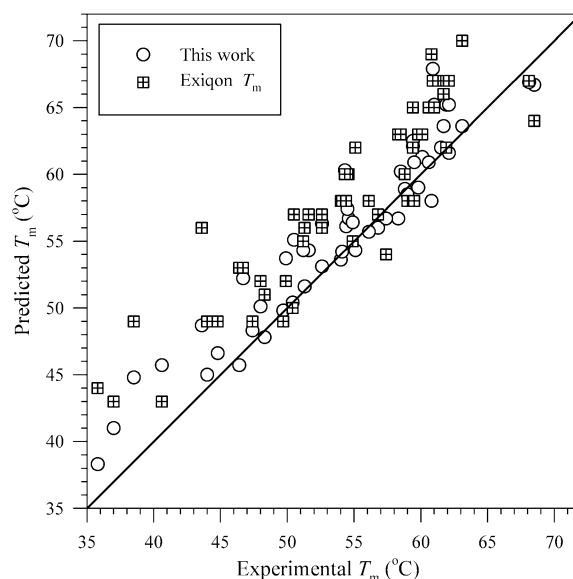

Figure 3. Comparison of predicted and experimentally measured melting temperatures for the validation set of 53 chimeric LNA·DNA duplexes in 1 M Na$^+$. We have predicted melting temperatures from parameters listed in Table 2 (○) and using the Exiqon Tm prediction 1.1 tool (http://lna-tm.com) (⊞).

comparison of experimentally measured melting temperatures with predictions. Good agreement is observed. Additional details are listed in Table S2 of the Supporting Information. The new parameters in Table 2 result in an average T_m prediction error of 2.1 °C (χ^2 = 2549).

Exiqon also developed a thermodynamic model of locked nucleic acids.[42] Because their parameters have not been publicly disclosed and the algorithm has not been described in detail, we relied on T_m predictions that were obtained online using their software. Comparison of experimental melting temperatures reveals that the Exiqon model tends to overestimate melting temperatures for our validation set. The average T_m prediction error is 4.2 °C, and χ^2 is equal to 7981. This level of accuracy agrees with the values reported by the developers where a standard deviation of 5.0 °C was obtained for T_m predictions of chimeric LNA·DNA duplexes.[42] Assuming a normal distribution of measured melting temperatures, probability P of the null hypothesis that this χ^2 difference occurs by random chance is less than 0.01. Thus, a two-tailed F-test for the ratio of χ^2 values[30,35] indicates that the new parameters from Table 2 predict melting temperatures more accurately than the Exiqon software.

Nearest-Neighbor Parameters for Single-Base Mismatches. From the primary data set, we determined nearest-neighbor parameters for single mismatches using SVD analysis. Table 3 shows eight doublet parameters for each of 12 LNA

Table 3. Thermodynamic Parameters for LNA Single Mismatches in 1 M Na⁺

sequence[a]	ΔΔH° (kcal/mol)	ΔΔS° (cal mol⁻¹ K⁻¹)	ΔH° (kcal/mol)	ΔS° (cal mol⁻¹ K⁻¹)
+A·A Mismatch				
+A+A/AT	4.074	9.091	−3.826	−13.109
+A+C/AG	6.033	15.078	−2.367	−7.322
+A+G/AC	2.951	7.993	−4.849	−13.007
+A+T/AA	2.151	2.886	−5.049	−17.514
+A+A/TA	3.671	7.040	−4.229	−15.160
+C+A/GA	2.622	5.037	−5.878	−17.663
+G+A/CA	−0.358	−1.776	−8.558	−23.976
+T+A/AA	9.274	24.746	2.074	3.446
+C·C Mismatch				
+C+A/CT	10.718	27.450	2.218	4.750
+C+C/CG	9.127	21.726	1.127	1.826
+C+G/CC	−0.303	−4.825	−10.903	−32.025
+C+T/CA	5.747	10.483	−2.053	−10.517
+A+C/TC	9.465	20.997	1.065	−1.403
+C+C/GC	−1.522	−7.124	−9.522	−27.024
+G+C/CC	5.033	9.503	−4.767	−14.897
+T+C/AC	12.314	31.458	4.114	9.258
+G·G Mismatch				
+G+A/GT	5.280	12.813	−2.920	−9.387
+G+C/GG	1.661	2.616	−8.139	−21.784
+G+G/GC	2.851	7.392	−5.149	−12.508
+G+T/GA	−0.591	−4.911	−8.991	−27.311
+A+G/TG	2.820	5.574	−4.980	−15.426
+C+G/GG	6.159	15.042	−4.441	−12.158
+G+G/CG	−5.505	−16.121	−13.505	−36.021
+T+G/AG	5.725	13.414	−2.775	−9.286
+T·T Mismatch				
+T+A/TT	3.456	9.151	−3.744	−12.149
+T+C/TG	3.813	8.680	−4.387	−13.520
+T+G/TC	2.154	6.071	−6.346	−16.629
+T+T/TA	0.203	−2.849	−7.697	−25.049
+A+T/TT	2.993	6.093	−4.207	−14.307
+C+T/GT	−0.376	−1.962	−8.176	−22.962
+G+T/CT	1.159	1.778	−7.241	−20.622
+T+T/AT	5.849	15.145	−2.051	−7.055
+A·C Mismatch				
+A+A/CT	6.538	16.649	−1.362	−5.551
+A+C/CG	6.641	15.889	−1.759	−6.511
+A+G/CC	1.251	2.927	−6.549	−18.073
+A+T/CA	3.637	6.295	−3.563	−14.105
+A+A/TC	5.822	12.112	−2.078	−10.088
+C+A/GC	2.632	5.748	−5.868	−16.952
+G+A/CC	−0.277	−2.365	−8.477	−24.565
+T+A/AC	9.890	26.265	2.690	4.965
+C·A Mismatch				
+C+A/AT	−1.344	−6.973	−9.844	−29.673
+C+C/AG	4.239	8.696	−3.761	−11.204
+C+G/AC	0.755	−0.116	−9.845	−27.316
+C+T/AA	4.411	8.483	−3.389	−12.517
+A+C/TA	9.153	21.897	0.753	−0.503
+C+C/GA	−4.714	−15.655	−12.714	−35.555
+G+C/CA	−2.858	−11.329	−12.658	−35.729
+T+C/AA	6.481	15.177	−1.719	−7.023
+A·G Mismatch				
+A+A/GT	10.093	26.574	2.193	4.374
+A+C/GG	−0.053	−0.272	−8.453	−22.672

Table 3. continued

sequence[a]	ΔΔH° (kcal/mol)	ΔΔS° (cal mol⁻¹ K⁻¹)	ΔH° (kcal/mol)	ΔS° (cal mol⁻¹ K⁻¹)
+A·G Mismatch				
+A+G/GC	6.636	18.468	−1.164	−2.532
+A+T/GA	−0.218	−3.666	−7.418	−24.066
+A+A/TG	5.937	13.187	−1.963	−9.013
+C+A/GG	−0.212	−1.079	−8.712	−23.779
+G+A/CG	0.325	0.539	−7.875	−21.661
+T+A/AG	10.407	28.456	3.207	7.156
+G·A Mismatch				
+G+A/AT	5.286	12.798	−2.914	−9.402
+G+C/AG	0.669	−0.947	−9.131	−25.347
+G+G/AC	5.846	16.029	−2.154	−3.871
+G+T/AA	−0.115	−3.913	−8.515	−26.313
+A+G/TA	1.109	−0.148	−6.691	−21.148
+C+G/GA	6.640	16.612	−3.960	−10.588
+G+G/CA	−4.898	−14.756	−12.898	−34.656
+T+G/AA	8.834	22.260	0.334	−0.440
+C·T Mismatch				
+C+A/TT	8.882	22.121	0.382	−0.579
+C+C/TG	5.284	11.900	−2.716	−8.000
+C+G/TC	0.237	−2.115	−10.363	−29.315
+C+T/TA	2.017	0.827	−5.783	−20.173
+A+C/TT	7.708	17.122	−0.692	−5.278
+C+C/GT	−2.299	−8.603	−10.299	−28.503
+G+C/CT	0.738	−1.956	−9.062	−26.356
+T+C/AT	10.273	26.168	2.073	3.968
+T·C Mismatch				
+T+A/CT	1.715	3.953	−5.485	−17.347
+T+C/CG	9.651	23.756	1.451	1.556
+T+G/CC	1.287	2.572	−7.213	−20.128
+T+T/CA	5.503	10.829	−2.397	−11.371
+A+T/TC	6.567	14.599	−0.633	−5.801
+C+T/GC	0.932	0.000	−6.868	−21.000
+G+T/CC	2.547	5.757	−5.853	−16.643
+T+T/AC	8.111	20.754	0.211	−1.446
+G·T Mismatch				
+G+A/TT	2.649	6.802	−5.551	−15.398
+G+C/TG	−5.143	−15.748	−14.943	−40.148
+G+G/TC	−0.110	1.551	−8.110	−18.349
+G+T/TA	−5.813	−17.641	−14.213	−40.041
+A+G/TT	0.670	0.214	−7.130	−20.786
+C+G/GT	−4.262	−12.230	−14.862	−39.430
+G+G/CT	−6.622	−17.610	−14.622	−37.510
+T+G/AT	1.797	4.589	−6.703	−18.111
+T·G Mismatch				
+T+A/GT	2.588	7.261	−4.612	−14.039
+T+C/GG	−1.598	−4.206	−9.798	−26.406
+T+G/GC	3.981	11.635	−4.519	−11.065
+T+T/GA	3.377	6.507	−4.523	−15.693
+A+T/TG	4.836	11.566	−2.364	−8.834
+C+T/GG	−3.596	−9.732	−11.396	−30.732
+G+T/CG	2.167	6.467	−6.233	−15.933
+T+T/AG	4.940	12.895	−2.960	−9.305

The differential parameters[b] header spans ΔΔH° and ΔΔS°; the full parameters header spans ΔH° and ΔS°.

[a]Mismatched bases are underlined. +N indicates an LNA nucleotide. Significant figures are shown beyond error estimates to allow accurate calculations and suppress rounding errors. [b]Parameters for the difference between LNA mismatches and matched DNA·DNA pairs.

mismatch types. Thermodynamic parameters are influenced by flanking base pairs and the type of mismatch. The doublet

format of nearest neighbors simplifies software implementation, but eight parameters for mismatch doublets are not unique, which was demonstrated in Materials and Methods. The constraint equation (eq 11) limits the number of linearly independent parameters to seven for each mismatch type. The unique parameters were constructed in triplet format and are listed in Table S3 of the Supporting Information.

To investigate trends and relationships of mismatch stabilities, we predicted thermodynamic values for all possible LNA triplets with a central mismatch. Matched base pairs flank the mismatch on both 5′ and 3′ sides. There are four possibilities for each flanking base pair (+A·T, +T·A, +C·G, and +G·C). Sixteen triplets, therefore, exist for each mismatch type (+A+X+A/TYT, +A+X+C/TYG, +A+X+G/TYC, +A+X+T/TYA, +C+X+A/GYT, +C+X+C/GYG, +C+X+G/GYC, +C+X+T/GYA, +G+X+A/CYT, +G+X+C/CYG, +G+X+G/CYC, +G+X+T/CYA, +T+X+A/AYT, +T+X+C/AYG, +T+X+G/AYC, and +T+X+T/AYA). There are $4 \times 3 = 12$ mismatch types because three types exist for each LNA nucleotide (for example, +A·A, +A·C, and +A·G for LNA adenine). The total number of unique triplets is therefore $16 \times 12 = 192$. Contributions to the free energy of the duplex transition ($\Delta G°_{37}$) were predicted for these triplets containing LNA mismatches, DNA mismatches, and related perfectly matched sequences using parameters from Tables 2 and 3 and ref 20. Table 1 of ref 20 seems to have typographical errors, so we used parameters from Table 2 of ref 16 instead. To model mismatch effects in the interior of a duplex, initiation free energies were not taken into account.

The LNA triplets were sorted according to free energy contributions. The least stable LNA mismatch is +A+C+T/TCA ($\Delta G°_{37} = 2.7$ kcal/mol). The same C·C mismatch context is also the most destabilizing for DNA·DNA single-base mismatches.[43]

The most stable LNA mismatch is the +G·T mismatch within the context of +G+G+C/CTG (−5.5 kcal/mol). It is interesting that the most stabilizing DNA mismatch occurs in the same sequence context, but it is the G·G mismatch instead, GGC/CGG (−2.2 kcal/mol).

Average $\Delta G°_{37}$ values over 16 triplet contexts produced a trend of decreasing stability for mismatches within consecutive LNA·DNA base pairs: +G·T ≫ +G·G > +T·G ≈ +G·A > +C·A > +T·T > +A·G ≈ +C·T > +A·A > +A·C ≈ +T·C > +C·C. The trend of relative stabilities of RNA·RNA mismatches closely resembles this trend:[44] rG·rU ≫ rG·rG > rU·rU > rA·rC > rC·U > rA·A ≈ rA·rG ≈ rC·rC. The stability trend of DNA·DNA mismatches shows some similarities:[43] G·G > G·T ≈ G·A > T·T ≈ A·A > T·C > A·C > C·C. The main differences between LNAs and DNAs are the higher relative stabilities of +G·T and +C·A mismatches and the lower relative stability of the +A·G mismatch. The order of stability of hybrid RNA·DNA mismatches is between the trends of RNA·RNA and DNA·DNA mismatches.[45,46] The most stable mismatch is the rG·T mismatch, like in RNAs, while the rC·A mismatch has relatively low stability, like in DNAs.

Mismatch Discrimination. To study the dependence of mismatch discrimination on oligonucleotide sequence, the free energy of mismatch discrimination ($\Delta\Delta G°$) was defined as the difference between mismatched and matched duplexes. The $\Delta\Delta G°$ value quantifies the amount of destabilization due to a mismatch. Let us define G·G mismatch discrimination in the +T+G+T/AGA LNA triplet

$$\Delta\Delta G°(\text{LNA}) = \Delta G°_{37}(+\text{T}+\text{G}+\text{T/AGA})$$
$$- \Delta G°_{37}(+\text{T}+\text{G}+\text{T/ACA}) \quad (16)$$

and in the isosequential DNA triplet

$$\Delta\Delta G°(\text{DNA}) = \Delta G°_{37}(\text{TGT/AGA})$$
$$- \Delta G°_{37}(\text{TGT/ACA}) \quad (17)$$

Values of $\Delta\Delta G°$ are positive because the lower stability of the mismatch makes $\Delta G°_{37}$ less negative. The larger the $\Delta\Delta G°$ values, the stronger the destabilization and mismatch discrimination. The positive difference between eqs 16 and 17 $[\Delta\Delta G°(\text{LNA}) - \Delta\Delta G°(\text{DNA})]$ indicates that LNAs increased the level of mismatch discrimination. The negative difference means that LNA modifications decreased the level of mismatch discrimination. We have predicted these free energy differences for the entire set of 192 possible mismatch triplets. Figure 4 shows the range of $\Delta\Delta G°(\text{LNA}) - \Delta\Delta G°(\text{DNA})$
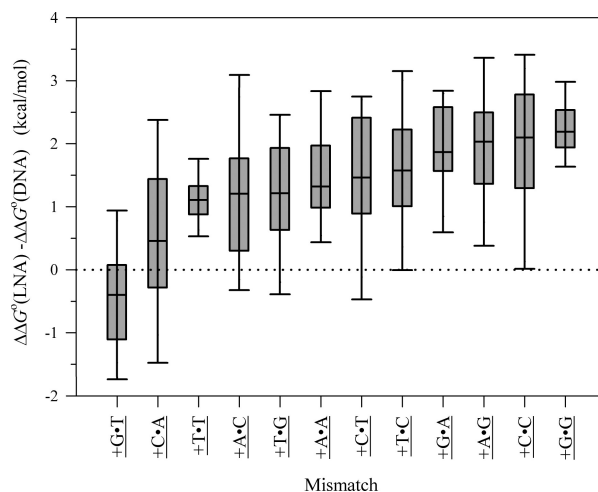

Figure 4. Differences in mismatch discrimination between LNA and DNA sequences. Free energies of duplex destabilization caused by mismatches $[\Delta\Delta G° = \Delta G°_{37}(\text{mismatch}) - \Delta G°_{37}(\text{match})]$ were predicted for triplets containing the central mismatch and all possible neighboring base pair combinations. The difference between these $\Delta\Delta G°$ values between LNA and DNA is plotted. The positive $\Delta\Delta G°(\text{LNA}) - \Delta\Delta G°(\text{DNA})$ value implies that LNA modifications destabilize the mismatch and increase the level of mismatch discrimination relative to DNA. Error bars indicate the range from the minimum to the maximum value. A gray box marks the range from the 25th to 75th percentile, and the horizontal line within the box shows the median value.

values for each mismatch type. LNA modification enhances discrimination for 85% of sequences and weakens it for 8%. Free energy differences are insignificant, that is, between −0.2 and 0.2 kcal/mol, for 7% of mismatches.

Figure 4 shows that LNAs negatively impact discrimination of +G·T mismatches and some +C·A mismatches. It appears that base pairs flanking a mismatch affect discrimination as well. The +G·C base pairs adjacent to a mismatch decrease the level of discrimination, while +A·T or +T·A base pairs increase it. To quantify this effect, we averaged $\Delta\Delta G°(\text{LNA}) - \Delta\Delta G°(\text{DNA})$ differences over possible triplet sequences containing a specific flanking base pair. The order of increasing mismatch discrimination resulting from the flanking base pair is as follows: +G·C < +C·G < +A·T ≈ +T·A [with average

Table 4. Comparison of Predicted and Measured Mismatch Discrimination (ΔT_m) for the Validation Set of LNA Duplexes

Sequence ID	Sequence (5' to 3')[a]	Mismatch	LNA site	Exper. T_m [°C][b]	Predicted T_m [°C]	Exper. ΔT_m [°C][c]	Predicted ΔT_m [°C][c]
VAL-A1	TGACGGAGCGATTCAGC	None	None	70.5	72.2	-	-
VAL-A2	TGACGGAG<u>C</u>GATTCAGC	<u>C</u>•A	None	58.5	60.9	-12.0	-11.3
VAL-A3	TGACGGA+G+C+GATTCAGC	None	+G+C+G/CGC	78.6	79.3	-	-
VAL-A4	TGACGGA+G+<u>C</u>+GATTCAGC	+<u>C</u>•A	+G+<u>C</u>+G/C<u>A</u>C	68.5	70.1	-10.1	-9.2
VAL-B1	CTATCCAGGCATTCGCA	None	None	67.4	69.1	-	-
VAL-B2	CTATCCAG<u>G</u>CATTCGCA	<u>G</u>•T	None	61.1	62.7	-6.3	-6.4
VAL-B3	CTATCCA+G+G+CATTCGCA	None	+G+G+C/CCG	77.4	77.5	-	-
VAL-B4	CTATCCA+G+<u>G</u>+CATTCGCA	+<u>G</u>•T	+G+<u>G</u>+C/C<u>T</u>G	70.7	72.4	-6.7	-5.1
VAL-C1	TTACTGTCAAGGCAACT	None	None	64.3	65.4	-	-
VAL-C2	TTACTGTC<u>A</u>AGGCAACT	<u>A</u>•C	None	52.6	56.9	-11.7	-8.5
VAL-C3	TTACTGT+C+A+AGGCAACT	None	+C+A+A/GTT	73.0	72.8	-	-
VAL-C4	TTACTGT+C+<u>A</u>+AGGCAACT	+<u>A</u>•C	+C+<u>A</u>+A/G<u>C</u>T	61.4	64.0	-11.6	-8.8
VAL-D1	GCGTCAAGCGACATCAT	None	None	69.2	70.3	-	-
VAL-D2	GCGTCAAG<u>C</u>GACATCAT	<u>C</u>•C	None	52.3	58.0	-16.9	-12.3
VAL-D3	GCGTCAA+G+C+GACATCAT	None	+G+C+G/CGC	75.0	77.5	-	-
VAL-D4	GCGTCAA+G+<u>C</u>+GACATCAT	+<u>C</u>•C	+G+<u>C</u>+G/C<u>C</u>C	57.4	64.8	-17.6	-12.7
VAL-E1	CGACTTGTCCATACCTA	None	None	64.7	64.6	-	-
VAL-E2	CGACTTGT<u>C</u>CATACCTA	<u>C</u>•C	None	50.5	54.5	-14.2	-10.1
VAL-E3	CGACTTG+T+C+CATACCTA	None	+T+C+C/AGG	73.7	74.7	-	-
VAL-E4	CGACTTG+T+<u>C</u>+CATACCTA	+<u>C</u>•C	+T+<u>C</u>+C/A<u>C</u>G	53.6	56.6	-20.1	-18.1
VAL-F1	CCATGCGTAGACAAGTG	None	None	65.3	67.4	-	-
VAL-F2	CCATGCGT<u>A</u>GACAAGTG	<u>A</u>•G	None	61.0	63.1	-4.3	-4.3
VAL-F3	CCATGCG+T+A+GACAAGTG	None	+T+A+G/ATC	74.4	77.0	-	-
VAL-F4	CCATGCG+T+<u>A</u>+GACAAGTG	+<u>A</u>•G	+T+<u>A</u>+G/A<u>G</u>C	65.6	67.0	-8.8	-10.0
VAL-G1	CTATCGCATCTAATAAT	None	None	58.1	56.6	-	-
VAL-G2	CTATCGCA<u>T</u>CTAATAAT	<u>T</u>•C	None	48.0	48.6	-10.1	-8.0
VAL-G3	CTATCGC+A+T+CTAATAAT	None	+A+T+C/TAG	63.8	63.6	-	-
VAL-G4	CTATCGC+A+<u>T</u>+CTAATAAT	+<u>T</u>•C	+A+<u>T</u>+C/T<u>C</u>G	50.6	48.1	-13.2	-15.5

[a]Nucleotides of mismatch sites are underlined. [b]C_t was 2 μM. [c]Differences between melting temperatures of mismatched and perfectly matched duplexes.

$\Delta\Delta G°$(LNA) − $\Delta\Delta G°$(DNA) differences of 0.9, 1.2, 1.6, and 1.7 kcal/mol, respectively]. The effect is not dependent on the flanking base pair location because the base pairs on the 5′ side of the mismatch exhibit the same trend as the base pairs on the 3′ side. In agreement with these observations, +G+<u>G</u>+C/C<u>T</u>G and +G+<u>C</u>+G/C<u>A</u>C mismatches exhibit the largest decreases in the level of discrimination from DNA to LNA with $\Delta\Delta G°$(LNA) − $\Delta\Delta G°$(DNA) differences of −1.7 and −1.5 kcal/mol, respectively.

The largest increases in the level of mismatch discrimination, i.e., the most positive $\Delta\Delta G°$(LNA) − $\Delta\Delta G°$(DNA) differences, are seen for the +<u>C</u>·C mismatch in the +T+<u>C</u>+C/A<u>C</u>G triplet (3.4 kcal/mol), the +<u>A</u>·G mismatch in +T+<u>A</u>+G/A<u>G</u>C and +T+<u>A</u>+T/A<u>G</u>A (3.4 kcal/mol), and the +<u>T</u>·C mismatch in +A+<u>T</u>+C/T<u>C</u>G (3.2 kcal/mol). LNAs significantly enhance discrimination of all +<u>G</u>·G, +<u>G</u>·A, +<u>A</u>·A, and +<u>T</u>·T mismatches, as well.

The free energies in Figure 4 were calculated at 37 °C, the temperature of the human body. In some biological applications, for instance, polymerase chain reaction, oligonucleotides are annealed at higher temperatures. Analysis at 60 °C reveals a similar dependence of LNA discriminatory effects on mismatch type (data not shown). However, values of $\Delta\Delta G°$(LNA) − $\Delta\Delta G°$(DNA) increased by ∼0.5 kcal/mol

for the +<u>G</u>·T, +<u>C</u>·A, and +<u>A</u>·C mismatches. This result suggests that the positive effect of LNA on mismatch discrimination increases with temperature. For example, LNAs improve mismatch discrimination, in relative terms with respect to DNA, for half of +<u>G</u>·T mismatches at 60 °C, while such positive effects are rare at 37 °C.

In our analysis, we assumed negligible heat capacity effects ($\Delta C_p \sim 0$). This has also been assumed for previously published thermodynamic parameters, although recent comprehensive studies[47] detected small heat capacity changes, ∼50 cal mol^{-1} K^{-1} bp^{-1}. Because similar mismatch discrimination trends are predicted at different temperatures, the veracity of this assumption does not seem to seriously influence the results of mismatch analysis.

Validation of Nearest-Neighbor Parameters for LNA Mismatches. To test the accuracy of mismatch parameters, we measured the stability of LNA mismatches described in the previous paragraphs. Table 4 lists sequences and their melting temperatures. Neither dye nor quencher was attached to these oligonucleotides. Their melting temperatures were determined using ultraviolet melting experiments. LNA modifications were predicted (1) to decrease the level of mismatch discrimination of VAL-A and VAL-B sequences, (2) not to affect discrimination of VAL-C and VAL-D sequences, and (3) to enhance

mismatch discrimination of VAL-E, VAL-F, and VAL-G sequences. Considering limitations of the nearest-neighbor model,[19,48] the predicted discrimination effects (ΔT_m) agree with experimental measurements for all seven sequence sets.

New LNA mismatch parameters result in an average T_m prediction error of 2.9 °C for the sequences in Table 4. The accuracy of DNA mismatch parameters [20] is the same. For DNA or LNA matched duplexes, average errors of predicted melting temperatures are less than 1.3 °C. The lower accuracy of mismatch predictions suggests that mismatched duplexes are more likely to deviate from assumptions of the nearest-neighbor model and two-state transitions. A small perturbation, like a single-nucleotide mismatch, does not usually break down assumptions of the nearest-neighbor model, but it may increase the magnitude of interactions propagating beyond nearest-neighbor nucleotides. These long-range interactions are often of electrostatic origin and likely become more significant in buffers with low counterion concentrations (<40 mM Na$^+$). We expect the weaker H-bonding interactions and increased nucleobase flexibility at the mismatch site. This potentially decreases cooperativity and increases deviations from the two-state melting behavior.

■ DISCUSSION

Characteristics of Effects of LNA on Duplex Stability. The LNA modifications placed at every second or third nucleotide position are very effective in increasing the duplex stability and affinity for complementary targets.[42] Mismatch discrimination is improved most if the triplet of consecutive LNAs is centered on the mismatch site.[5] A single LNA modification usually discriminates less. We were therefore motivated to study thermodynamics of consecutive LNAs to expand the published nearest-neighbor model of single LNA modifications and improve our understanding of LNA·DNA duplex stability.

We employed the fluorescence melting method to measure the stability of modified oligonucleotide duplexes.[28] This new technology allows measurements for large sets of duplexes with unprecedented speed, and its accuracy is similar to the accuracy of the ultraviolet optical melting method. Using the fluorescence method, the experimental errors of $\Delta H°$, $\Delta S°$, $\Delta G°$, and T_m were 8%, 9%, 4%, and 0.4 °C, respectively. If the duplex melts in the two-state manner, the thermodynamic values are in agreement between both methods. The transition enthalpies and entropies measured using the fluorescence differed by <4% from the values determined by the UV melting method.[28] The free energy values agreed within 2.5% when the optimal Texas Red−Iowa Black RQ pair was attached to the duplex terminus. These differences are similar or smaller than the errors seen in UV melting experiments where the errors of $\Delta H°$, $\Delta S°$, and $\Delta G°$ are ~8, ~8, and ~4%, respectively.[17,41]

The fluorescence melting method relies on the dye− quencher pair attached to one of the duplex termini as shown in Figure 1B. When the duplex melts, the dye and the quencher dissociate, giving the increase in the magnitude of the fluorescence signal. Although the terminal dye−quencher pair stabilizes the duplex, it is attached to both the LNA-modified duplex and the core duplex. The Texas Red−Iowa Black RQ labels therefore change the $\Delta H°$, $\Delta S°$, and $\Delta G°$ values of both duplexes to the same amount. The thermodynamic impact of LNA modification is determined from the difference between the LNA-modified and core duplexes. We have shown previously that these thermodynamic differences ($\Delta\Delta H°$,

$\Delta\Delta S°$, and $\Delta\Delta G°$) are not affected by terminal labels.[28] The stabilizing effect of labels cancels out in this analysis.

Using SVD, we determined nearest-neighbor parameters for consecutive LNA·DNA base pairs. New parameters accurately predict melting temperatures of chimeric LNA·DNA duplexes. The average error was ~2 °C, which is the best accuracy that can be achieved by the nearest-neighbor model.[19,48] If LNA modifications amount to a moderate perturbation of a DNA duplex, new parameters are most accurate. Analysis of the validation data set (Table S2 of the Supporting Information) suggests that accuracy decreases slightly as the percentage of LNA modifications increases. The duplexes of VAL-01− VAL-33 are predicted more accurately (average error of 1.5 °C) than VAL-34−VAL-53 duplexes (3.0 °C). The LNA content is low for the VAL-01−VAL-33 subset (20−25%) and varies from 30 to 60% for the latter subset.

We also predicted melting temperatures for 11 duplexes from published sources where one strand was LNA-modified from 89 to 100%. Initiation parameters for terminal LNAs were assumed to be identical to DNA initiation parameters.[16] Table S4 of the Supporting Information shows results. The average error of T_m predictions was higher for these duplexes (2.7 °C) than the error seen for the set of VAL-01−VAL-33 duplexes (1.5 °C). If an LNA strand is modified ≥50%, LNAs induce structural changes that could propagate beyond neighboring base pairs. In that case, the nearest-neighbor parameters and the model may be less accurate.

Thermodynamic parameters reveal the nature of stabilizing effects. The single strand to helix transition of nucleic acid is usually driven by favorable enthalpic changes associated with an increased level of stacking and H-bonding interactions. Entropic changes are unfavorable. Because single strands explore more degrees of freedom than the strands in the relatively stiff duplex structure, duplex formation incurs the entropic loss. Locked nucleic acids have been reported to alter both transition enthalpy and entropy,[21,49] so the origin of LNA effects is uncertain.

The free energy change due to LNA residues can be divided into enthalpic ($\Delta\Delta H°$) and entropic ($-T\Delta\Delta S°$) components, which are presented in the second and last columns, respectively, of Table 1. The values suggest that the stabilizing effect is of enthalpic origin. Consecutive LNAs induce favorable changes in the transition enthalpy, making it more negative by −1 to −9 kcal/mol per each nearest-neighbor doublet. Changes in the entropic contribution to the free energy (the last column of Table 1) are either unfavorable or negligible. The values of $-T\Delta\Delta S°$ range from 0 to 7 kcal/mol at 37 °C and are smaller in magnitude than $\Delta\Delta H°$. Thus, we conclude that the higher stability of consecutive LNA·DNA base pairs is mostly the result of favorable contributions to the transition enthalpy. This is the case for all nearest-neighbor doublets, confirming that enthalpy drives stabilization of consecutive LNAs regardless of base sequence.

These thermodynamic observations are related to structural changes. Stabilizing enthalpic effects of LNAs are equated with enhanced stacking interactions, potentially improved H-bonding of base pairs, and weakened hydration of the duplex state.[50] The LNA cytosine C5-methyl group, which is not present in the native DNA, may also increase the stacking energies due to additional van der Waals interactions with neighboring bases.[45] The entropic contributions of LNAs originate from backbone conformational preorganization, which is the result of restrictions of ribose flexibility in the C3′-endo

(N-type) conformation.[3,51] Because the modified ribose is similarly constrained in the single strand and in the duplex conformations, it has been argued that the smaller entropic loss occurs upon formation of LNA·DNA rather than DNA·DNA base pairs.

While we observe that the stabilization of consecutive LNAs is driven by enthalpic changes, McTigue et al. reported that the stabilizing effects of a single LNA modification are mostly entropic in origin.[21] Taken together, these findings suggest that the entropic changes characterized by restriction of nucleotide local conformations are achieved by the introduction of a single LNA nucleotide. Additional adjacent LNAs stabilize the duplex further by favorable enthalpic changes. This mechanism may explain the conflicting reports in the literature regarding the origin of LNA stabilization.

Structural studies have shown that both isolated and consecutive LNA residues restrict ribose conformation space and introduce structural changes in the double helix toward A-form. For example, LNAs widen the minor groove and decrease the value of the rise and the twist.[51−54] The ^{1}H NMR experiment with the C+TGA+TA+TGC sequence that contains only isolated LNA modifications failed to show significant changes in base stacking.[52] In contrast, LNAs in the C+TGC+T+TC+TGC sequence containing consecutive modifications enhanced base stacking.[53] Our fluorescence experiments using 2-aminopurine also detected enhanced stacking interactions in LNA triplets.[5] These apparent discrepancies can be reconciled assuming that the energetic impact of a single LNA in the duplex interior is dominated by entropic changes, and the subsequent addition of consecutive LNAs stabilizes duplexes by favorable enthalpic changes that are associated with enhanced stacking interactions.

Energetics of LNA modifications introduced at the duplex terminus may have a different character. Kaur et al. measured impacts of isolated LNA modifications at various positions.[49] The interior modifications decreased entropic loss in agreement with our rationale, but stabilizing effects of the terminal modification were driven by favorable enthalpic changes. We have not studied consecutive LNAs at the duplex terminus.

The A-form helical conformation that is preferred by LNA·DNA duplexes is also dominant in RNA·RNA and RNA·DNA duplexes. In fact, ribose puckering of the LNA·DNA duplex resembles closely the puckering of the RNA·DNA hybrid.[54] However, the structural similarity does not mean the same thermodynamic parameters. The LNA·DNA nearest-neighbor doublets are on average 1.4 kcal/mol more stable than RNA·DNA doublets.[19] For example, the ΔG°_{37} of +C+C/GG is −4.1 kcal/mol, while rCrC/GG is only half as stabilizing, −2.1 kcal/mol. The sequence dependence of parameters is also different. The least stable LNA doublet is +A+A/TT, while the rArA/TT doublet is more stable than five other RNA·DNA doublets. These significant differences reveal that thermodynamic parameters of RNA·DNA duplexes are not good approximations of LNA·DNA thermodynamics. The different composition of the ribose moiety, different patterns of hydration in the minor groove, the extra methyl group of +C, and subtle variations of the helical structure potentially explain these thermodynamic differences.

Enhanced Mismatch Discrimination Is Not Unique for Locked Nucleic Acids. We show in Results that LNA·DNA and RNA·RNA mismatches exhibit a similar trend of stabilities, which deviates from the stability trend of DNA mismatches. To inquire whether the mismatch discrimination is similarly

enhanced in RNA duplexes, like it is enhanced in LNAs, we predicted the free energy of mismatch discrimination (eq 16) for RNA, RNA·DNA, LNA, and DNA triplets. For each mismatch type, the $\Delta\Delta G^\circ$ values were averaged over 16 possible triplet sequences containing the central mismatch. Predictions were based on the established nearest-neighbor parameters for matched LNA, DNA, and RNA base pairs (Table 2 and refs 16, 17, and 19). For mismatches, the complete set of thermodynamic parameters is available for LNA·DNA and DNA·DNA pairs (Table 3 and ref 20). Because parameters for many RNA·DNA mismatches are unknown, we averaged $\Delta\Delta G^\circ$ for eight rG·T sequence contexts reported by the Sugimoto group[45] and predicted the average $\Delta\Delta G^\circ$ values for rA·A, rG·G, and rC·C mismatches. Their parameters were recently determined.[46] The rA·rA, rG·rG, and rC·rC RNA·RNA mismatches were approximated by the algorithm of Davis and Znosko.[44] Mathews, Sabina, Zuker, and Turner parameters were used for the rG·rU mismatch.[18] The RNA calculations were conducted with MELTING version 5.0.3.[55]

Figure 5 shows the average free energies of duplex destabilization due to a mismatch. The general trend of increasing
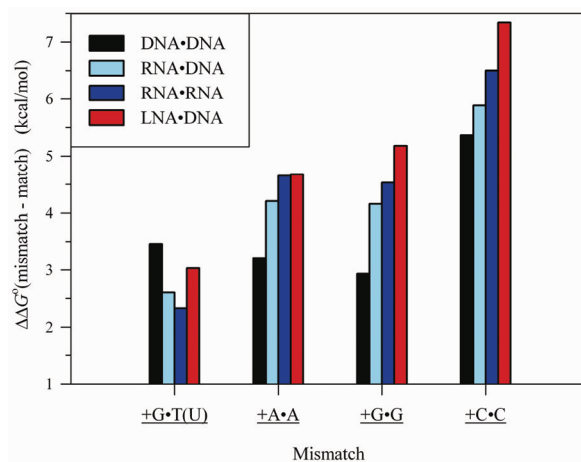

Figure 5. Free energies of duplex destabilization due to a mismatch [$\Delta\Delta G^\circ = \Delta G^\circ_{37}$(mismatch) $- \Delta G^\circ_{37}$(match)] were predicted for triplets containing the central mismatch. The $\Delta\Delta G^\circ$ values are averaged for each mismatch over 16 possible nearest-neighbor base contexts and are plotted as a function of nucleic acid backbone character (DNA, RNA, LNA, and hybrid duplexes). The free energy of rG·T mismatch discrimination was averaged over eight available triplet contexts.[45]

discriminatory power for the A·A, G·G, and C·C mismatches is as follows: DNA·DNA ≪ RNA·DNA < RNA·RNA ≤ LNA·DNA. These mismatches destabilize the LNA·DNA and RNA·RNA duplexes more than the DNA·DNA duplexes. To a lesser degree, the level of mismatch discrimination also increases in RNA·DNA duplexes. The opposite trend is seen for the wobble G·T base pair. The DNA·DNA mismatch shows the strongest discrimination. The +G·T, rG·T, and rG·rU mismatches discriminate less.

This analysis suggests that the enhanced mismatch discrimination is not a unique property of locked nucleic acids but rather the result of structural changes of nucleic acids from B-form to A-form. DNA·DNA duplexes in water solutions are in B-like conformations. The RNA·DNA hybrids fold into structures that are intermediates of A- and B-forms. The RNA·RNA duplexes occur in the A-form conformation, which

is also the structure of LNA·DNA base pairs.[53,54] As the conformational equilibrium is shifted toward the A-form, the level of mismatch discrimination increases. This is likely the result of energetic changes in stacking interactions, H-bonding of base pairs, and hydration envelope when the duplex turns to the A-like conformation.

The one significant structural change from B-form to A-form is the compaction of the rise between base pairs along the helical axis. The rise is significantly smaller in the A-form (0.26 nm) than in the B-form (0.34 nm). Because of the shorter distances, LNA nucleotides in the A-like structure may engage in stronger stacking interactions, which are disrupted by mismatches. If our hypothesis is correct, the enhancements of mismatch discrimination can be expected for any modification that shifts the conformation equilibrium from B-form to A-form, e.g., 2′-O-methyl-RNA, 2′-O-[2-(methoxy)ethyl]-RNA, 2′-deoxy-2′-fluoro-RNA, and N3′→P5′-phosphoramidate-DNA.[56−59] As discussed earlier, +G·T mismatches are the exception; their level of mismatch discrimination decreases when LNA-modified guanine is introduced at the mismatch site. This could be a result of improved stacking interactions of guanine with neighboring bases. These stacking interactions are not significantly weakened by a thymine mismatch because the G·T pair is stabilized by two hydrogen bonds and is well-stacked in the duplex structure. Small pyridine bases are expected to stack less than large purine bases. This may explain the opposite discriminatory effects of LNAs in +T·G versus +G·T mismatches.

Chemical differences among LNA, RNA, and DNA are the composition and conformation of the ribose moiety. Another difference is the C5-methyl group in pyrimidine nucleobases. In RNA, uracil and cytosine are typically unmethylated. In DNA, thymine is C5-methylated and cytosine is not. In LNA nucleotides, both thymine and cytosine are C5-methylated.

Wang and Kool investigated thermodynamic effects of C5-methyl and 2′-OH groups in DNA and RNA duplexes.[60] The methyl group stabilized duplexes on average by 0.25 kcal/mol, and its effects on $\Delta G°$ were largely independent of 2′-hydroxyl effects. The C5-methyl appeared to enhance base stacking. Ziomek et al. studied 5-alkyl and 5-halogen analogues of uracil in $(rArUrCrUrArGrArU)_2$ duplexes.[61] The methyl group stabilized the RNA duplex slightly ($\Delta\Delta G° < 0.1$ kcal/mol). Sugimoto et al. examined thermodynamics of pyridine methyl groups in RNA·DNA mismatches.[45] The rG·dU mismatches were found to be less stable than rG·dT mismatches regardless of neighboring sequence context. The free energy contribution of the thymine C5-methyl was estimated to vary from 0.1 to 0.5 kcal/mol. The methyl moiety likely has a similar thermodynamic impact on the LNA cytosine residue.[62]

The extra methyl group of pyrimidines is not the driver of mismatch discrimination trends. The increase in the level of discrimination occurs in purine mismatches (+A·A and +G·G) and in the sequence contexts that do not contain methylated LNA cytosine. For example, LNAs increase free energies of +A·A mismatch discrimination in the center of the +G+A+G triplet by 1.0 kcal/mol. Further, relative to DNA, the extra C5-methyl group is present in LNA cytosine, but not in RNA cytosine. In both cases, the level of mismatch discrimination increases; i.e., both LNA and RNA duplexes have more discriminatory power than DNA. The presence of the C5-methyl group does not appear to be essential for discriminatory effects.

Oligonucleotide Design and Online Software. Sufficient mismatch discrimination is important for many oligonucleotide applications. Locked nucleic acids enhance discrimination due to two impacts. First, LNAs increase the stability of oligonucleotide probes. This allows the use of shorter sequences with more discriminatory power because the mismatch has a much larger impact on the duplex stability in shorter sequences than in longer ones.[5] This length effect is very significant in duplexes with <30 bp. The $\Delta\Delta G°$ and ΔT_m differences between matched and single-base mismatched duplexes can double when the duplex length is decreased from 25 to 17 bp.

Second, locked nucleic acids can also increase specificity directly if they are located at or next to the mismatch site. We have discovered that the triplet of LNA residues containing the mismatch in the center has the largest discriminatory power; a single LNA modification usually discriminates less.[5] We therefore recommend using the LNA triplet at the mismatch site. This design will increase the discriminatory power for a majority of mismatches (in particular, for A·G, T·C, C·C, G·G, A·A, and T·T). New results also pinpoint several anomalies. LNAs in some +G·T and +C·A mismatches impact discrimination negatively. In these cases, is it not advised to introduce the LNA modifications at the mismatch site, but LNAs could be placed ≥2 bp from the mismatch to increase the stability of the probe−target duplex and make the probe shorter. The short probe will likely exhibit more discriminatory power. Alternatively, the probe could be redesigned to target the complementary strand if it is available in the biological sample. This will change the +G·T mismatch into the +T·G mismatch; the latter one is more likely to show positive effects of LNA on discrimination.

It is also important to optimize the location of mismatches within the probe. The mismatches at the terminus or adjacent to the terminus (penultimate mismatches) show significantly less discrimination than the mismatches in the duplex interior.[5,20] It is preferable to place mismatches at least 3 bp from the termini of the probe−target duplex. Although the mismatch site in the center of the duplex maximizes the discrimination, it is not essential for the mismatch to be located exactly in the center of the oligonucleotide probe. As long as the mismatch is positioned in the interior of the duplex and not next to the termini, its discriminatory power ($\Delta\Delta G°$) will be very close to the maximum.

To help design optimal LNA oligonucleotides, free software is available at the IDT websites http://biophysics.idtdna.com and http://www.idtdna.com. The web tools predict melting temperatures, free energies, and the extent of hybridization using the latest nearest-neighbor parameters, including parameters from Tables 2 and 3. It is important to enter conditions of the experiments (e.g., cation and DNA concentrations) to obtain the relevant predictions. Users can test effects of LNA modifications and mismatches at any location within their sequence. The potential LNA probes can be compared with unmodified probes to estimate benefits of modifications. The probes can be ranked by their mismatch discrimination energetics ($\Delta\Delta G°$ and ΔT_m) and tuned to the hybridization temperature of a specific application. It is often optimal if the probe has a melting temperature 3−5 °C above the annealing temperature. The perfectly matched probe−target duplex will be stable, while the mismatched duplex is likely to be unstable under those conditions and will not give a false positive signal.

Many applications also require that chimeric oligonucleotides bind effectively and exclusively to DNA complements. The design must therefore exclude sequences that can form stable hairpins, dimers, and other self-folding structures. This is important because LNA·LNA base pairs are more stable than isosequential LNA·DNA base pairs.[63] Because thermodynamic parameters for LNA·LNA base pairs, LNA bulges, and hairpin loops are unknown, it is not currently possible to accurately predict the propensity of an LNA oligonucleotide to form self-folding structures. The simple approach is to avoid long stretches of consecutive LNAs. This approach makes stable LNA·LNA duplexes less likely to appear but also unnecessarily impedes probe design. Accurate predictions of LNA·LNA base pair stability would be useful.

The tendency of the base sequence to form hairpins can be estimated by the hairpin function of the IDT OligoAnalyzer tool.[64] The self-dimer function shows the potentially stable structures that can form between two molecules. The hetero-dimer function estimates interactions between the probe and the primers. If the predicted structure contains several consecutive LNA·LNA base pairs, it could be stable enough to compete with the formation of the probe−target duplex and the assay would be negatively impacted. For such sequences, a single LNA modification could be a better choice than the LNA triplet.

The design of real-time PCR hydrolysis probes (e.g., TaqMan probes) calls for additional considerations. This family of assays relies on the 5′ exonuclease activity of the polymerase, which degrades the probe and releases the dye attached to the 5′ terminus of the probe. Locked nucleic acids cannot be introduced at the 5′ terminus of the probe or at the adjacent nucleotide because they would increase nuclease resistance and interfere with the desired probe degradation.

Future Challenges. Although the new parameter set is a significant addition toward a complete thermodynamic model of LNA modifications, parameters for some important LNA structures have yet to be determined (mismatches adjacent to a single LNA modification, LNA·LNA base pairs, bulges, and tandem mismatches). We also do not have parameters for LNAs at duplex termini, although such modifications are employed in PCR primers. The parameters in Table 3 were determined for mismatches located in the interior of a duplex and will not be accurate at the terminus. The mismatch in the terminal or penultimate position often affects duplex stability less than the same mismatch located in the interior, i.e., ≥3 bp from the terminus of the duplex.[20]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information
Thermodynamic values for studied duplexes, figures of $1/T_m$ versus ln C_t fits, unique thermodynamic parameters for triplets containing single-base mismatches, and melting temperatures of validation data sets. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: science@owczarzy.net. Phone: (319) 626-8459. Fax: (319) 626-9621.

■ ABBREVIATIONS:

bp, base pair; C_t, total single-strand concentration; $\Delta G°_{37}$, transition free energy at 37 °C; $\Delta H°$, transition enthalpy; IBRQ, Iowa Black RQ; LNA, locked nucleic acids; $\Delta S°$, transition entropy; SVD, singular-value decomposition; T_m, melting temperature; TXRD, Texas Red.

■ REFERENCES

(1) Braasch, D. A., Jensen, S., Liu, Y., Kaur, K., Arar, K., White, M. A., and Corey, D. R. (2003) RNA interference in mammalian cells by chemically-modified RNA. *Biochemistry 42*, 7967−7975.

(2) Petersen, M., and Wengel, J. (2003) LNA: A versatile tool for therapeutics and genomics. *Trends Biotechnol. 21*, 74−81.

(3) Vester, B., and Wengel, J. (2004) LNA (locked nucleic acid): High-affinity targeting of complementary RNA and DNA. *Biochemistry 43*, 13233−13241.

(4) Fluiter, K., Mook, O. R., and Baas, F. (2009) The therapeutic potential of LNA-modified siRNAs: Reduction of off-target effects by chemical modification of the siRNA sequence. *Methods Mol. Biol. 487*, 189−203.

(5) You, Y., Moreira, B. G., Behlke, M. A., and Owczarzy, R. (2006) Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Res. 34*, e60.

(6) Latorra, D., Campbell, K., Wolter, A., and Hurley, J. M. (2003) Enhanced allele-specific PCR discrimination in SNP genotyping using 3′ locked nucleic acid (LNA) primers. *Hum. Mutat. 22*, 79−85.

(7) Ugozzoli, L. A., Latorra, D., Puckett, R., Arar, K., and Hamby, K. (2004) Real-time genotyping with oligonucleotide probes containing locked nucleic acids. *Anal. Biochem. 324*, 143−152.

(8) Mouritzen, P., Nielsen, A. T., Pfundheller, H. M., Choleva, Y., Kongsbak, L., and Møller, S. (2003) Single nucleotide polymorphism genotyping using locked nucleic acid (LNA). *Expert Rev. Mol. Diagn. 3*, 27−38.

(9) Castoldi, M., Schmidt, S., Benes, V., Hentze, M. W., and Muckenthaler, M. U. (2008) miChip: An array-based method for microRNA expression profiling using locked nucleic acid capture probes. *Nat. Protoc. 3*, 321−329.

(10) Lennox, K. A., and Behlke, M. A. (2010) A direct comparison of anti-microRNA oligonucleotide potency. *Pharm. Res. 27*, 1788−1799.

(11) Darfeuille, F., Reigadas, S., Hansen, J. B., Orum, H., Di Primo, C., and Toulmé, J. J. (2006) Aptamers targeted to an RNA hairpin show improved specificity compared to that of complementary oligonucleotides. *Biochemistry 45*, 12076−12082.

(12) Kierzek, E. (2009) Binding of short oligonucleotides to RNA: Studies of the binding of common RNA structural motifs to isoenergetic microarrays. *Biochemistry 48*, 11344−11356.

(13) Ng, P. S., and Bergstrom, D. E. (2005) Alternative nucleic acid analogues for programmable assembly: Hybridization of LNA to PNA. *Nano Lett. 5*, 107−111.

(14) Borer, P. N., Dengler, B., Tinoco, I. Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices. *J. Mol. Biol. 86*, 843−853.

(15) Owczarzy, R., Vallone, P. M., Gallo, F. J., Paner, T. M., Lane, M. J., and Benight, A. S. (1997) Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers 44*, 217−239.

(16) SantaLucia, J. Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A. 95*, 1460−1465.

(17) Xia, T., SantaLucia, J. Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson−Crick base pairs. *Biochemistry 37*, 14719−14735.

(18) 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.

(19) Sugimoto, N., Nakano, S., 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.

(20) SantaLucia, J. Jr., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct. 33*, 415−440.

(21) 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.

(22) Kierzek, E., Mathews, D. H., Ciesielska, A., Turner, D. H., and Kierzek, R. (2006) Nearest neighbor parameters for Watson-Crick complementary heteroduplexes formed between 2′-O-methyl RNA and RNA oligonucleotides. *Nucleic Acids Res. 34*, 3609−3614.

(23) Sugimoto, N., Satoh, N., Yasuda, K., and Nakano, S. (2001) Stabilization factors affecting duplex formation of peptide nucleic acid with DNA. *Biochemistry 40*, 8444−8451.

(24) Chen, X., Kierzek, R., and Turner, D. H. (2001) Stability and structure of RNA duplexes containing isoguanosine and isocytidine. *J. Am. Chem. Soc. 123*, 1267−1274.

(25) Kierzek, E., Ciesielska, A., Pasternak, K., Mathews, D. H., Turner, D. H., and Kierzek, R. (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2′-O-methyl RNA/RNA heteroduplexes. *Nucleic Acids Res. 33*, 5082−5093.

(26) Pasternak, A., Kierzek, E., Pasternak, K., Fratczak, A., Turner, D. H., and Kierzek, R. (2008) The thermodynamics of 3′-terminal pyrene and guanosine for the design of isoenergetic 2′-O-methyl-RNA-LNA chimeric oligonucleotide probes of RNA structure. *Biochemistry 47*, 1249−1258.

(27) Jacobsen, N., Bentzen, J., Meldgaard, M., Jakobsen, M. H., Fenger, M., Kauppinen, S., and Skouv, J. (2002) LNA-enhanced detection of single nucleotide polymorphisms in the apolipoprotein E. *Nucleic Acids Res. 30*, e100.

(28) You, Y., Tataurov, A. V., and Owczarzy, R. (2011) Measuring thermodynamic details of DNA hybridization using fluorescence. *Biopolymers 95*, 472−486.

(29) Moreira, B. G., You, Y., Behlke, M. A., and Owczarzy, R. (2005) Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability. *Biochem. Biophys. Res. Commun. 327*, 473−484.

(30) Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., and Walder, J. A. (2004) Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry 43*, 3537−3554.

(31) Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry 22*, 256−263.

(32) Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers 26*, 1601−1620.

(33) Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

(34) 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.

(35) Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1996) *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, Cambridge University Press, Cambridge, U.K.

(36) Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., and Benight, A. S. (1992) Studies of DNA dumbbells. I. Melting curves of 17 DNA dumbbells with different duplex stem sequences linked by T4

endloops: Evaluation of the nearest-neighbor stacking interactions in DNA. *Biopolymers 32*, 849−864.

(37) Efron, B., and Tibshirani, R. J. (1993) *An Introduction to the Bootstrap*, Chapman & Hall/CRC, Boca Raton, FL.

(38) Starikov, E. B., and Nordén, B. (2009) Physical rationale behind the nonlinear enthalpy-entropy compensation in DNA duplex stability. *J. Phys. Chem. B 113*, 4698−4707.

(39) Bevington, P. R., and Robinson, D. K. (1992) *Data Reduction and Error Analysis for the Physical Sciences*, WCB/McGraw-Hill, Boston.

(40) Gray, D. M. (1997) Derivation of nearest-neighbor properties from data on nucleic acid oligomers. II. Thermodynamic parameters of DNA·RNA hybrids and DNA duplexes. *Biopolymers 42*, 795−810.

(41) Allawi, H. T., and SantaLucia, J. Jr. (1998) Nearest neighbor thermodynamic parameters for internal G·A mismatches in DNA. *Biochemistry 37*, 2170−2179.

(42) Tolstrup, N., Nielsen, P. S., Kolberg, J. G., Frankel, A. M., Vissing, H., and Kauppinen, S. (2003) OligoDesign: Optimal design of LNA (locked nucleic acid) oligonucleotide capture probes for gene expression profiling. *Nucleic Acids Res. 31*, 3758−3762.

(43) Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J. Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. *Biochemistry 38*, 3468−3477.

(44) Davis, A. R., and Znosko, B. M. (2007) Thermodynamic characterization of single mismatches found in naturally occurring RNA. *Biochemistry 46*, 13425−13436.

(45) Sugimoto, N., Nakano, M., and Nakano, S. (2000) Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. *Biochemistry 39*, 11270−11281.

(46) Watkins, N. E. Jr., Kennelly, W. J., Tsay, M. J., Tuin, A., Swenson, L., Lee, H. R., Morosyuk, S., Hicks, D. A., and Santalucia, J. Jr. (2011) Thermodynamic contributions of single internal rA·dA, rC·dC, rG·dG and rU·dT mismatches in RNA/DNA duplexes. *Nucleic Acids Res. 39*, 1894−1902.

(47) Mikulecky, P. J., and Feig, A. L. (2006) Heat capacity changes associated with nucleic acid folding. *Biopolymers 82*, 38−58.

(48) Sugimoto, N., Honda, K., and Sasaki, M. (1994) Application of the thermodynamic parameters of DNA stability prediction to double-helix formation of deoxyribooligonucleotides. *Nucleosides Nucleotides 13*, 1311−1317.

(49) Kaur, H., Arora, A., Wengel, J., and Maiti, S. (2006) Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes. *Biochemistry 45*, 7347−7355.

(50) Pallan, P. S., Greene, E. M., Jicman, P. A., Pandey, R. K., Manoharan, M., Rozners, E., and Egli, M. (2011) Unexpected origins of the enhanced pairing affinity of 2′-fluoro-modified RNA. *Nucleic Acids Res. 39*, 3482−3495.

(51) Kierzek, E., Pasternak, A., Pasternak, K., Gdaniec, Z., Yildirim, I., Turner, D. H., and Kierzek, R. (2009) Contributions of stacking, preorganization, and hydrogen bonding to the thermodynamic stability of duplexes between RNA and 2′-O-methyl RNA with locked nucleic acids. *Biochemistry 48*, 4377−4387.

(52) Jensen, G. A., Singh, S. K., Kumar, R., Wengel, J., and Jacobsen, J. P. (2001) A comparison of the solution structures of an LNA:DNA duplex and the unmodified DNA:DNA duplex. *J. Chem. Soc., Perkin Trans. 2,*, 1224−1232.

(53) Nielsen, K. E., Singh, S. K., Wengel, J., and Jacobsen, J. P. (2000) Solution structure of an LNA hybridized to DNA: NMR study of the d(CT^{L}GCT^{L}T^{L}CT^{L}GC):d(GCAGAAGCAG) duplex containing four locked nucleotides. *Bioconjugate Chem. 11*, 228−238.

(54) Nielsen, K. E., Rasmussen, J., Kumar, R., Wengel, J., Jacobsen, J. P., and Petersen, M. (2004) NMR studies of fully modified locked nucleic acid (LNA) hybrids: Solution structure of an LNA:RNA hybrid and characterization of an LNA:DNA hybrid. *Bioconjugate Chem. 15*, 449−457.

(55) Le Novère, N. (2001) MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics 17*, 1226−1227.

(56) Lubini, P., Zürcher, W., and Egli, M. (1994) Stabilizing effects of the RNA 2′-substituent: Crystal structure of an oligodeoxynucleotide duplex containing 2′-O-methylated adenosines. *Chem. Biol. 1*, 39−45.

(57) Egli, M., Minasov, G., Tereshko, V., Pallan, P. S., Teplova, M., Inamati, G. B., Lesnik, E. A., Owens, S. R., Ross, B. S., Prakash, T. P., and Manoharan, M. (2005) Probing the influence of stereoelectronic effects on the biophysical properties of oligonucleotides: Comprehensive analysis of the RNA affinity, nuclease resistance, and crystal structure of ten 2′-O-ribonucleic acid modifications. *Biochemistry 44*, 9045−9057.

(58) Ding, D., Gryaznov, S. M., and Wilson, W. D. (1998) NMR solution structure of the N3′ → P5′ phosphoramidate duplex d(CGCGAATTCGCG)$_2$ by the iterative relaxation matrix approach. *Biochemistry 37*, 12082−12093.

(59) Williams, A. A., Darwanto, A., Theruvathu, J. A., Burdzy, A., Neidigh, J. W., and Sowers, L. C. (2009) Impact of sugar pucker on base pair and mispair stability. *Biochemistry 48*, 11994−12004.

(60) 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.

(61) Ziomek, K., Kierzek, E., Biala, E., and Kierzek, R. (2002) The thermal stability of RNA duplexes containing modified base pairs placed at internal and terminal positions of the oligoribonucleotides. *Biophys. Chem. 97*, 233−241.

(62) Sági, J., Szemzö, A., Ébinger, K., Szabolcs, A., Sági, G., Ruff, É., and Ötvös, L. (1993) Base-modified oligodeoxynucleotides. I. Effect of 5-alkyl, 5-(1-alkenyl) and 5-(1-alkynyl) substitution of the pyrimidines on duplex stability and hydrophobicity. *Tetrahedron Lett. 34*, 2191−2194.

(63) Koshkin, A. A., Nielsen, P., Meldgaard, M., Rajwanshi, V. K., Singh, S. K., and Wengel, J. (1998) LNA (Locked Nucleic Acid): An RNA mimic forming exceedingly stable LNA:LNA duplexes. *J. Am. Chem. Soc. 120*, 13252−13253.

(64) Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., Sailor, C. A., Dawson, R. B., and Peek, A. S. (2008) IDT SciTools: A suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res. 36*, W163−W169.