BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

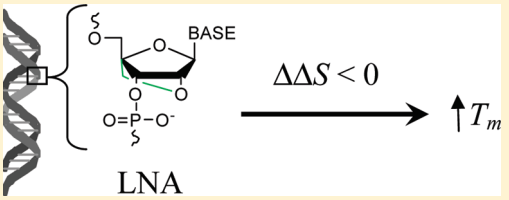
pubs.acs.org/biochemistry

Role of the Heat Capacity Change in Understanding and Modeling Melting Thermodynamics of Complementary Duplexes Containing Standard and Nucleobase-Modified LNA

Curtis B. Hughesman, Robin F. B. Turner, and Charles A. Haynes*

Michael Smith Laboratories and Department of Chemical and Biological Engineering or Department of Chemistry, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

ABSTRACT: Melting thermodynamic data obtained by differential scanning calorimetry (DSC) are reported for 43 duplexed oligonucleotides containing one or more locked nucleic acid (LNA) substitutions. The measured heat capacity change (ΔC_p) for the helix-to-coil transition is used to compute the changes in enthalpy and entropy for melting of an LNA-bearing duplex at the T_m of its corresponding isosequential unmodified DNA duplex to allow rigorous thermodynamic analysis of the stability enhancements provided by LNA substitutions. Contrary to previous studies, our analysis shows that the origin of the improved stability is almost exclusively a net reduction ($\Delta\Delta S° < 0$) in the entropy gain accompanying the helix-to-coil transition, with the magnitude of the reduction dependent on the type of nucleobase and its base pairing properties. This knowledge and our average measured value for ΔC_p of 42 ± 11 cal mol^{-1} K^{-1} bp^{-1} are then used to derive a new model that accurately predicts melting thermodynamics and the increased melting temperature (ΔT_m) of heteroduplexes formed between an unmodified DNA strand and a complementary strand containing any number and configuration of standard LNA nucleotides A, T, C, and G. This single-base thermodynamic (SBT) model requires only four entropy-related parameters in addition to ΔC_p. Finally, DSC data for 20 duplexes containing the nucleobase-modified LNAs 2-aminoadenine (D) and 2-thiothymine (H) are reported and used to determine SBT model parameters for D and H. The data and model suggest that along with the greater stability enhancement provided by D and H bases relative to their corresponding A and T analogues, the unique pseudocomplementary properties of D-H base pairs may make their use appealing for in vitro and in vivo applications.



Though they provide the elegantly simple yet extremely powerful language of life, the nucleotides adenine (a), thymine (t), guanine (g), and cytosine (c), through their specific backbone and base pairing chemistry, limit the thermodynamic behavior and other functional properties that can be engineered into oligonucleotides created for research, diagnostic, and therapeutic applications. This has prompted extensive research into chemical modifications that allow facile tuning of the stability and functional properties of oligonucleotides to enhance their performance within specific applications and technologies.[1] A number of unnatural nucleotide chemistries have been used to increase duplex stability while maintaining sequence specific recognition of cDNA or RNA strands. These studies show that stability enhancements can be achieved with chemistries that increase the level of preorganization of the modified single strand through a reduction in the number of conformational degrees of freedom, including rotations around internal bonds.[2] The locked nucleic acid (LNA) is one example of a nucleotide that is conformationally restricted through chemical modification.[3−5] In LNA nucleotides, a 2′-O,4′-C-methylene link is introduced into the ribose ring to constrain or "lock" the sugar moiety into an N-type (3′-endo) conformation.[5−8] The incorporation of an LNA into a complementary duplex is also thought to cause local changes in the helix structure to an A-type form.[9−12] Together, these novel properties of LNA have been shown to improve duplex stability,[4,6−8,13] mismatch discrimination,[14] and resistance to degradation by nucleases.[15,16]

The 2′-O,4′-C-methylene bridge can be introduced into any of the DNA nucleosides to create LNA-adenine (A), LNA-thymine (T), LNA-guanine (G), or LNA-cytosine (C). Oligonucleotides containing A, T, G, and C substitutions can be synthesized using standard phosphoramidite chemistry employed in DNA synthesis.[17] Moreover, the incorporation of LNA bases does not alter oligonucleotide solubility, allowing LNAs to be strategically substituted into DNA or RNA for focused functional design of reagents for in vitro and in vivo applications.[18,19] Such rational design, however, requires methods for predicting the effect of LNA substitutions on duplex stability and function.

The design and application of single-stranded (ss) DNA reagents (e.g., primers, probes, etc.) have benefited significantly from the development of thermodynamic models, most notably nearest-neighbor thermodynamic (NNT)-type models that accurately predict the duplex stability and melting temperature (T_m) from sequence knowledge. The development of an equivalent model for oligonucleotides bearing LNA (and/or other modified nucleotide) substitutions would obviously offer similar

Received: February 11, 2011
Revised: May 2, 2011
Published: May 06, 2011

benefits. To date, two models have been developed to predict the T_m of a complementary duplex formed with an oligonucleotide containing one or more LNA substitutions. The first model, termed "OligoDesign", was developed by Exiqon Inc.[20] It is restricted to calculation of T_m, and the details regarding the structure and parameters of the model have not been published. However, it is stated to have a standard deviation of ±5.0 °C for prediction of T_m values for duplexes with LNA substitutions, and a convenient interface for using the model is available at http:\\lna-tm.com.

The second model is specifically designed to predict the T_m of LNA-DNA "mixmers", which are short complementary duplexes containing individual LNA-DNA base pairs flanked on both sides by DNA-DNA base pairs.[21] It is an NNT-type model that builds on the widely used "unified NNT model" for unmodified duplex DNA[22] through the addition of 64 regressed LNA nearest-neighbor (NN) parameters, 32 of which are used to compute $\Delta\Delta H^\circ$ ($=\Delta H^\circ_{LNA} - \Delta H^\circ_{DNA}$), where ΔH° is the standard enthalpy change for the helix-to-coil transition, and the remaining 32 are used to compute $\Delta\Delta S^\circ$. The model therefore assumes that LNA substitutions enhance duplex stability by altering both ΔH° and ΔS°. Given the importance of the hydrogen bonding and base stacking forces that promote duplex DNA formation, as well as the chain organization processes that oppose it, this assumption seems reasonable. Indeed, upon first inspection, the limited ΔH° and ΔS° data available for melting of duplexes containing LNA substitutions appear to be consistent with this assumption,[6,7,21,23,24] though they do not provide a definitive mechanism for stability enhancement.

However, the heat capacity change (ΔC_p) accompanying the helix-to-coil transition was typically not measured or included in these previous studies. More importantly, ΔC_p is not included in the LNA NNT model described in ref 21. By assuming $\Delta C_p = 0$, that model treats ΔH° as temperature invariant and computes $\Delta\Delta H^\circ$ from melting thermodynamics data as $\Delta H^\circ_{LNA}(T_{m,LNA}) - \Delta H^\circ_{DNA}(T_{m,DNA})$, where $T_{m,LNA} \neq T_{m,DNA}$ because of the stability enhancement provided by the LNA substitution(s). If $\Delta C_p \neq 0$ and is positive, this treatment of the data is incorrect and will create an analysis error that biases $\Delta\Delta H^\circ$ to values greater than 0 and thus an improper accounting of the mechanism of action of an LNA substitution. There are now numerous data available showing that ΔC_p for the melting of unmodified duplex DNA is non-zero and significant,[25−34] and one recent study has reported a non-zero (positive) ΔC_p for an LNA-bearing duplex.[35] These data show that both ΔH°_{LNA} and ΔH°_{DNA} are temperature-dependent state functions, and $\Delta\Delta H^\circ$ must then be computed at a common temperature, with $T_{m,DNA}$ being the most appropriate when one is trying to understand the mechanism of action of an LNA. $\Delta\Delta H^\circ$ is therefore properly given by $\Delta H^\circ_{LNA}(T_{m,DNA}) - \Delta H^\circ_{DNA}(T_{m,DNA})$, where $\Delta H^\circ_{LNA}(T_{m,DNA})$ is computed using the measured enthalpy change $\Delta H^\circ_{LNA}(T_{m,LNA})$ and the measured value of ΔC_p. We sought to determine if this more rigorous thermodynamic treatment would yield an improved understanding of the thermodynamic origin of the stability enhancement provided by LNA nucleotides.

Here, differential scanning calorimetry (DSC) is used to measure helix-to-coil transition thermodynamics for a library of unmodified and LNA-substituted duplexes. We find that ΔC_p per base pair (ΔC_p^{bp}) is non-zero, positive, and, within the accuracy of the experiment, unchanged by LNA substitution(s). By accounting for ΔC_p^{bp}, we show that ΔT_m ($=T_{m,LNA} - T_{m,DNA}$) can be predicted using a simple four-parameter model, which we have named the "single-base thermodynamic" (SBT) model, with an accuracy and precision similar to those of the previously described LNA nearest-neighbor thermodynamic model[21] requiring a total of 64 parameters. The regressed single-base (SB) parameters are entropy-related and offer insight into the mechanism by which the four standard LNA nucleobases provide their stability enhancements. Moreover, we show that the SBT model, which works at the individual base pair level, can be used to predict ΔT_m values for both mixmers and "gapmers", which are duplex sequences containing one or more neighboring (adjacent) LNA-DNA base pairs.

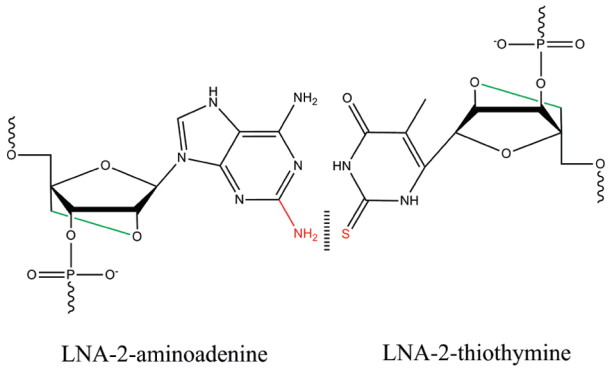


Figure 1. Structure of LNA-2-aminoadenine and LNA-2-thiothymine shown in base pairing configuration to demonstrate steric hindrance between the 2-amino (red) and 2-thioketo (red) groups of the respective bases. The 2′-O,4′-C-methylene bridge of LNA is also identified (green).

We also report DSC-derived melting thermodynamics data and SBT model parameters for the modified nucleotides LNA-2-aminoadenine (D) and LNA-2-thiothymine (H)[36−38] and show that duplexes substituted with D and/or H bases are more stable than their isosequential duplexes bearing A and/or T, respectively. Finally, we demonstrate that the D-H base pair (Figure 1) is pseudocomplementary, a property that can be exploited in the functional design of oligonucleotides to minimize the formation of undesired stable secondary structures.[39−41]

MATERIALS AND METHODS

Sequence Design. Short complementary duplexes containing A and/or T substitutions were designed from 12 duplex DNA reference sequences known to possess a two-state melting transition from repeated DSC scans. From these 12 reference sequences, a total of 20 duplexes with one, two, or three A and/or T substitutions were designed such that all 32 possible trinucleotide sequences xMy (where x and y are a, t, g, or c and M is A or T) were represented (Table 1). The modified ss within each duplex was designed as a mixmer with each LNA separated from any other LNA, as well as from the 5′ and 3′ termini, by a minimum of two DNA nucleotides. Each modified ss was also designed to preclude formation of stable ss structures (e.g., hairpins and homodimers), as well as undesired bimolecular products such as slipped duplexes with the cDNA strand. A second set of sequences used to analyze stability enhancements provided by the two base-modified LNAs (D and H) was then created via substitution of A with D and T with H within the same set of duplexes (Table 2).

This basic design strategy was also used to study duplexes with G and/or C substitutions, where the set of LNA substituted duplexes (Table 1) were derived from 14 DNA reference duplexes.

Table 1. DSC-Determined Thermodynamic Parameters for Duplexes between DNA Oligonucleotides with and without A, T, G, and/or C Substitutions[a]

name	sequence	C_T (μM)	ΔC_p (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta\Delta G°_{37}$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$\Delta\Delta S°$ (cal mol^{-1} K^{-1})	ΔT_m (°C)
UD1	gcccagcg	100	557	10.0	57.6	154	56.9				
UL1a	gcccAgcg	100	556	10.5	58.9	156	59.3	0.4	−0.1	−1.6	2.4
UL1b	cgcTgggc	100	439	10.9	59.4	157	61.4	0.7	−0.2	−2.9	4.5
UL1c	gccCagcg	100	466	11.2	62.0	164	62.4	1.0	1.8	2.5	5.5
UL1d	cgCtgGgc	100	394	13.3	65.8	169	72.8	2.8	1.9	−2.8	15.9
UD2	agacctagt	100	440	8.1	56.0	154	46.2				
UL2	agAcctAgt	100	445	10.2	61.1	164	56.7	1.8	0.4	−4.7	10.5
UD3	ccatgtccc	100	509	9.8	65.9	181	53.4				
UL3	ccatgTccc	100	351	11.1	67.6	182	59.6	1.1	−0.5	−5.1	6.1
UD4*	tgcacgcta	100	538	10.0	55.0	145	58.2				
UL4*	tgcAcgcta	100	393	10.3	55.2	145	60.0	0.3	−0.5	−2.3	1.9
UD5	atgctcatgc	100	472	10.5	65.0	176	57.1				
UL5a	atgctcAtgc	100	445	11.1	67.8	183	59.3	0.5	1.8	4.0	2.2
UL5b	atgcTcatgc	100	472	11.2	65.8	176	60.7	0.6	−1.0	−5.1	3.7
UD6	ggcacgcttcg	75	469	14.8	88.4	237	68.2				
UL6a	cgAagcgTgcc	75	499	16.8	92.7	245	75.0	1.7	1.0	−2.4	6.8
UL6b	ggcaCgcttcg	75	343	16.1	90.2	239	73.2	1.2	0.1	−3.5	5.0
UL6c	cgaagCgtGcc	75	302	17.1	90.0	235	77.6	2.0	−1.2	−10.6	9.4
UD7	aagttctcttat	75	531	10.1	74.2	207	52.0				
UL7a	aagTtctcTtat	75	372	11.8	76.8	209	59.2	1.5	−0.1	−5.3	7.1
UL7b	aagtTctctTat	75	563	12.2	80.0	218	60.1	1.8	1.2	−1.9	8.0
UD8	aactatgaaact	75	516	10.7	81.9	229	52.9				
UL8a	aactAtgaaAct	75	498	12.1	83.6	231	58.4	1.2	−1.0	−7.2	5.4
UL8b	agtTtcaTagtt	75	510	12.7	84.9	233	60.4	1.7	−0.8	−8.1	7.4
UD9	aaatagagaattc	75	610	11.0	88.9	251	52.8				
UL9	aaAtagAgaaTtc	75	850	13.3	93.5	259	60.4	1.9	−1.9	−12.2	7.6
UD10	gaaacagttaaag	75	474	12.3	96.6	272	56.1				
UL10	gaAacagttAaag	75	387	12.8	96.0	268	58.0	0.4	−1.3	−5.6	1.8
UD11	aacatagattacat	50	480	12.9	98.4	276	56.8				
UL11a	aacatagattAcat	50	558	13.6	100.2	279	58.9	0.6	0.6	0.0	2.1
UL11b	atgTaatctaTgtt	50	543	15.4	104.7	288	64.1	2.3	2.3	0.2	7.4
UL11c	aacatagAttacat	50	412	13.5	98.2	273	59.0	0.6	−1.1	−5.4	2.2
UL11d	atgtaaTctatgtt	50	663	13.3	95.3	264	59.2	0.4	−4.7	−16.5	2.5
UD12	ttcttatagatacaag	50	504	14.1	113.6	321	57.8				
UL12a	ttcttatagatacAag	50	622	14.5	114.5	322	59.0	0.4	0.1	−0.8	1.2
UL12b	ctTgtatcTataAgaa	50	647	17.4	119.3	329	66.8	2.9	−0.1	−9.4	8.9
UL12c	ttcttataGataCaag	50	416	16.1	113.9	315	64.2	1.8	−2.4	−13.6	6.3
UL12d	cttGtatCtataagaa	50	604	16.4	117.4	326	64.1	2.0	0.0	−6.5	6.3
UD13	ctggagc	100	418	7.1	46.2	126	40.6				
UL13	ctGgagc	100	375	7.7	48.1	130	44.9	0.6	0.3	−1.0	4.3
UD14*	ggtgccaa	100	382	8.6	52.0	140	49.8				
UL14*	ggtgCcaa	100	267	9.7	51.6	135	57.2	1.0	−2.4	−10.9	7.4
UD15*	acgtcttcg	100	583	9.1	57.0	154	51.4				
UL15*	acgtCttcg	100	380	10.5	61.3	164	58.3	1.3	1.7	1.4	6.9
UD16	tagagggcagac	75	399	14.6	91.0	247	66.4				
UL16a	tagaggGcagac	75	523	15.5	92.9	249	69.5	0.8	0.2	−2.0	3.2
UL16b	tagaGggcagac	75	542	16.1	93.9	251	71.4	1.2	0.2	−3.4	5.0
UD17	ctcgggaacgcc	75	519	16.6	100.1	269	71.2				
UL17a	ctcGggaacgcc	75	317	17.3	100.8	269	73.4	0.6	0.0	−1.7	2.2
UL17b	ggcgttccCgag	75	552	17.1	98.4	262	73.5	0.3	−3.0	−10.5	2.3
UL17c	ggcgttCccgag	75	307	17.5	97.8	259	75.4	0.7	−3.5	−13.7	4.2
UL17d	ctcggGaacgcc	75	333	17.1	98.1	261	73.9	0.4	−2.9	−10.7	2.7

dx.doi.org/10.1021/bi200223s |*Biochemistry* 2011, 50, 5354–5368

Table 1. Continued

name	sequence	C_T (μM)	ΔC_p (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta\Delta G°_{37}$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$\Delta\Delta S°$ (cal mol^{-1} K^{-1})	ΔT_m (°C)
UD18	gatctgaggtact	75	523	14.2	98.4	271	62.5				
UL18a	gatctgagGtact	75	500	15.3	99.9	273	66.4	1.0	−0.4	−4.6	3.8
UL18b	agtaCctcagatc	75	400	15.9	101.6	276	67.9	1.6	1.0	−1.8	5.4
UL18c	agtacCtcagatc	75	486	15.5	101.4	277	66.5	1.2	1.1	−0.3	4.0
UD19	gccctcgcacgtc	75	674	18.8	109.9	294	75.2				
UL19	gccctcGcacgtc	75	476	19.3	112.0	299	76.1	0.4	1.7	3.9	0.9
UD20	cgaaagcctcggc	75	749	17.1	102.7	276	71.8				
UL20	cgaaaGcctCggc	75	701	19.4	107.0	282	78.8	1.8	−0.6	−7.7	7.0
UD21	gcatgcccgtgacac	50	650	21.4	129.0	347	76.2				
UL21	gcatgcCcgtgacac	50	476	22.4	127.6	339	79.7	0.8	−3.1	−12.5	3.5
UD22	agtagtaatcacacc	50	645	16.6	117.1	324	64.9				
UL22	agtaGtaatCacacc	50	737	18.4	118.8	324	70.1	1.5	−2.2	−11.7	5.2
UD23	caacttgatattaata	50	609	14.5	114.9	324	58.9				
UL23	caaCttGatattaata	50	554	17.3	119.2	329	66.4	2.4	0.2	−7.3	7.5

[a] Measured thermodynamic values ($\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m) are for the helix-to-coil transition and are the average of two-state and calorimetric analysis. For the mean experimental values reported, the estimated errors (standard deviations) for ΔC_p, $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m were determined to be 25, 2, 3, and 3% and 0.5 °C, respectively. $\Delta H°$ and $\Delta S°$ are reported at the T_m of the sample, while $\Delta\Delta H°$ and $\Delta\Delta S°$ are at the T_m of the reference isosequential DNA duplex and computed from the measured thermodynamic values using the measured ΔC_p for the given duplex.

Table 2. DSC-Derived Thermodynamic Data for Complementary DNA Duplexes in Which One Strand Contains D and/or H Substitutions[a]

name	sequence	C_T (μM)	ΔC_p (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta\Delta G°_{37}$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$\Delta\Delta S°$ (cal mol^{-1} K^{-1})	ΔT_m (°C)
USL1a	gcccDgcg	100	463	11.4	61.5	162	63.8	1.2	0.7	−1.6	6.9
USL1b	cgcHgggc	100	501	11.5	60.2	157	65.2	1.2	−1.6	−9.3	8.3
USL2	agDcctDgt	100	414	11.0	62.7	167	61.1	2.6	0.5	−6.6	14.9
USL3	ccatgHccc	100	330	11.5	66.9	179	61.9	1.5	−1.7	−10.4	8.5
USL4	tgcDcgcta	100	384	11.3	60.2	158	63.9	1.1	3.0	6.1	5.8
USL5a	atgctcDtgc	100	436	12.0	71.4	192	62.8	1.4	3.9	8.2	5.8
USL5b	atgcHcatgc	100	311	12.1	69.5	185	63.8	1.4	2.4	3.3	6.7
USL6a	cgDagcgHgcc	75	449	17.5	92.1	241	78.2	2.2	−0.7	−9.6	10.0
USL7a	aagHtctcHtat	75	425	12.3	75.8	205	61.9	1.9	−2.6	−14.6	9.8
USL7b	aagtHctctHat	75	333	12.3	73.4	197	62.7	2.0	−4.4	−20.6	10.7
USL8a	aactDtgaaDct	75	418	13.2	87.0	238	62.2	2.3	1.2	−3.5	9.3
USL8b	agtHtcaHagtt	75	474	13.8	85.7	232	64.8	2.7	−1.8	−14.6	11.9
USL9	aaDtagDgaaHtc	75	600	14.4	92.7	252	65.1	2.9	−3.6	−20.9	12.3
USL10	gaDacagttDaag	75	343	13.6	96.8	268	60.7	1.2	−1.4	−8.2	4.6
USL11a	aacatagattDcat	50	276	14.3	100.4	278	61.4	1.4	0.7	−2.1	4.7
USL11b	atgHaatctaHgtt	50	591	16.7	105.1	285	68.6	3.3	−0.3	−11.7	11.8
USL11c	aacatagDttacat	50	464	14.0	98.3	272	60.7	1.0	−2.0	−9.5	4.0
USL11d	atgtaaHctatgtt	50	526	14.0	94.2	259	62.0	0.9	−7.0	−25.5	5.2
USL12a	ttcttatagatacDag	50	854	14.8	114.8	323	59.7	0.6	−0.3	−2.9	1.8
USL12b	ctHgtatcHataDgaa	50	610	18.5	117.0	318	71.0	3.8	−4.6	−27.0	13.1

[a] Measured thermodynamic values ($\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m) are for the helix-to-coil transition and are the average of two-state and calorimetric analysis. For the mean experimental values reported, the estimated errors (standard deviations) for ΔC_p, $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m were determined to be 25, 2, 3, and 3% and 0.5 °C, respectively. $\Delta H°$ and $\Delta S°$ are reported at the T_m of the sample, while $\Delta\Delta H°$ and $\Delta\Delta S°$ are at the T_m of the reference isosequential DNA duplex and computed from the measured thermodynamic values using the measured ΔC_p for the given duplex.

Thus, results for each LNA-containing duplex could be directly compared to results for its corresponding isosequential DNA duplex. As before, all 32 possible trinucleotide xMy sequences containing a G or C central substitution were represented. However, two of these duplexes failed to show the required two-state melting behavior, resulting in three trinucleotide sequences (cGa, cGt, and gCa) being excluded from our analysis.

Finally, the reference duplex sequences that were modified to study terminal and tandem (gapmer-type) LNA substitutions were 11-mers that contain an equal representation of all 10 possible

dx.doi.org/10.1021/bi200223s |*Biochemistry* 2011, 50, 5354–5368

DNA nearest neighbors. Single terminal LNA substitutions at either the 5′ or 3′ end were introduced into a total of three reference duplexes. All four standard LNA bases were tested at both positions. Likewise, 16 tandem LNA sequences were created from three reference duplexes. To avoid any end effects, all tandem LNA substitutions were placed a minimum of two nucleotides from the 5′ or 3′ end.

Oligonucleotide Synthesis and Purification. Oligonucleotides containing D and/or H substitutions were synthesized and then purified via high-performance liquid chromatography by Exiqon Inc. (Vedbæk, Denmark). All other DNA- and LNA-containing oligonucleotides used in this study were obtained from Integrated DNA Technologies (Coralville, IA). All oligonucleotides were resuspended in buffer containing 1 M NaCl, 10 mM Na_2HPO_4, and 1 mM Na_2EDTA (pH 7.0) and quantified with a UV spectrophotometer (Cary 1E) at 80 °C using extinction coefficients provided by the supplier.

Differential scanning calorimetry was used to collect excess heat capacity (ΔC_p^{ex}) versus temperature (T) data for the helix-to-coil transition. All experiments were performed on a Microcal (Northampton, MA) VP-DSC instrument, with a nominal cell volume of 0.52811 mL. Data were collected by scanning from 1 to 100 °C at a rate of 1 °C/min. Both the sample and reference cells were initially filled with buffer solution, and DSC data were collected for a minimum of two cycles to establish baseline buffer scans. Samples with total strand concentrations (C_T) of 50, 75, or 100 μM containing equimolar amounts of two complementary strands were then dynamically loaded during the down scan part of the cycle, and data were collected for a minimum of two cycles. Final ΔC_p^{ex} versus T profiles used for data analysis were obtained by subtracting the average baseline scans from those obtained for duplex melting after normalizing versus the concentration, corrected for strand purities. These corrected thermograms were then used to determine ΔC_p as well as two-state ($\Delta H^\circ_{2\text{-st}}$, $\Delta S^\circ_{2\text{-st}}$, and $T_m^{2\text{-st}}$) and calorimetric (ΔH°_{cal}, ΔS°_{cal}, and T_m^{cal}) data as described previously.[34] Thermodynamic data reported throughout this paper are the average of two-state (model-dependent) and calorimetric (model-independent) values.

UV spectroscopy was also used to collect helix-to-coil transition data using a Varian (Santa Clara, CA) Cary 1E spectrophotometer equipped with a 12-cell Peltier temperature controller and sample temperature probes. All UV-monitored melt (UVM) experiments were conducted at 260 nm by scanning from 10 to 98 °C at a scan rate of 0.5 °C/min to obtain absorbance (A_{260}) versus T profiles. Cuvettes with path lengths of 10 mm or 1 mm were loaded with buffer or duplex samples with a C_T of 7.5 or 75 μM, respectively, and caps were fitted to prevent evaporation during heating. UVM data were analyzed using a two-state thermodynamic model that includes contributions from a non-zero ΔC_p as previously described.[32] Raw UVM curves were transformed to compute the fraction of strands in the random-coil state (θ) by fitting the pretransition and post-transition baselines to

$$A_{260}(T) = (m_{pre}^{UVM}T + b_{pre}^{UVM})(1-\theta) + (m_{post}^{UVM}T + b_{post}^{UVM})\theta \quad (1)$$

where m_{pre}^{UVM} and b_{pre}^{UVM}, and m_{post}^{UVM} and b_{post}^{UVM}, are the slopes and intercepts of the pre- and post-transition baselines, respectively. The enthalpy ΔH°_{UVM} and entropy ΔS°_{UVM} changes at the measured T_m for the helix-to-coil transition were then determined by fitting the fractional curve to eqs 2 and 3

$$\theta(T) = \frac{-K_d(T) + \sqrt{K_d^2(T) + 2K_d(T)C_T}}{C_T} \quad (2)$$

$$K_d(T) = \exp\left[-\frac{\Delta H^\circ_{UVM} + \Delta C_p(T - T_m)}{RT} + \frac{\Delta S^\circ_{UVM} + \Delta C_p \ln(T/T_m)}{R}\right] \quad (3)$$

where $K_d(T)$ is the temperature-dependent equilibrium constant for the dissociation (helix-to-coil) reaction. The ΔC_p required in eqs 2 and 3 was computed from the strand length and the average ΔC_p^{bp} value of 42 cal mol^{-1} K^{-1} bp^{-1}. T_{max} was used as an estimate of T_m in eq 3 during the first solution iteration. After $\Delta H^\circ_{UVM}(T_m)$ and $\Delta S^\circ_{UVM}(T_m)$ had been determined, a new T_m value was then determined from eq 4 written for the helix-to-coil transition

$$T_m = \frac{\Delta H^\circ_{UVM}(T_m)}{\Delta S^\circ_{UVM}(T_m) - R\ln(C_T/4)} \quad (4)$$

and the solution iteration is continued with eqs 2 and 3 until a best fit with the experimental data is obtained using Microsoft Excel Solver.

Error Analysis. To establish the accuracy of the DSC experiment, we conducted repeated melting transition studies for several duplexes to obtain three to six complete and independent data sets for each duplex. From these data, the average errors in the regressed thermodynamic parameters ΔC_p, ΔH°, ΔS°, and T_m were determined to be 25, 3, and 3% and 0.5 °C, respectively. All UVM studies were performed in triplicate and the combined data also used to estimate errors in reported thermodynamic data.

Regression of SBT Model Parameters. Incremental thermodynamic parameters ($\Delta\Delta H_i^\circ$ and $\Delta\Delta S_i^\circ$) within the SBT model for each possible LNA substitution i (where i = A, T, G, C, D, or H) were determined through global regression to the set of incremental thermodynamic data (63 modified sequences identified as j = 1−63) reported in Tables 1 and 2. For example, the set of incremental enthalpy parameters $\Delta\Delta H_i^\circ$ was computed by minimizing the error-weighted squares of the residual (χ^2) as shown in eqs 5 and 6

$$\chi^2 = \sum_j \frac{\Delta\Delta H^\circ_{j(pred)} - \Delta\Delta H^\circ_j}{\sigma_j^2} \quad (5)$$

$$\Delta\Delta H^\circ_{j(pred)} = \sum_i n_i \Delta\Delta H^\circ_i \quad (6)$$

where $\Delta\Delta H^\circ_{j(pred)}$ and $\Delta\Delta H^\circ_j$ represent the model-predicted and experimental values, respectively, and n_i is the number of LNA substitutions of type i within duplex sequence j. The same approach was used to determine $\Delta\Delta S_i^\circ$ and $\Delta\Delta G_i^\circ{}_{37}$ parameters for each possible base substitution. Parameter regression was completed using Microsoft Excel with the add-in software XLSTAT version 2009.4.03.

Model-predicted ΔT_m values for LNA-substituted duplexes were determined and used to compare and evaluate the performance of the SBT model relative to other available methods. For the OligoDesign method,[20] the web-based algorithm accessed

at http:\\lna-tm.com was used to predict ΔT_m [$=T_{m(LNA)} - T_{m(DNA)}$] values, hereafter denoted $\Delta T_{m(pred)}$ to explicitly recognize that the value is model-derived, by computing the T_m of both the LNA-substituted duplex [$T_{m(LNA)}$] and the isosequential DNA reference duplex [$T_{m(DNA)}$] with all input fields including the total strand concentration (C_T) set to match experimental conditions. $\Delta T_{m(pred)}$ values predicted using the LNA NNT model of McTigue et al. were obtained as described in ref 21.

RESULTS

The central goal of this work was to conduct UVM and high-sensitivity DSC experiments that serve to establish an improved understanding of the thermodynamic basis for the enhanced duplex stability provided by replacement of a nucleotide with its LNA analogue and likewise provide the first detailed study of the thermodynamic properties of oligodeoxynucleotide duplexes incorporating the modified nucleobases LNA-2-aminoadenine (D) and LNA-2-thiothymine (H). Note that the thermodynamic data reported here (typically, T_m, $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and ΔC_p values) are for the helix-to-coil transition, and thus, their signs are opposite to those in some previously published reports of DNA and LNA thermodynamics.[21,22] These data are used to compute incremental thermodynamic changes ($\Delta\Delta G°_{37}$, $\Delta\Delta H°$, and $\Delta\Delta S°$) at the chosen reference temperature, $T_{m(DNA)}$, the melting temperature of the corresponding isosequential DNA duplex under the same solution conditions (P = 1 atm, same C_T, 1 M NaCl, 10 mM Na_2HPO_4, 1 mM Na_2EDTA, and pH 7.0).

Thermodynamic data determined by DSC for 43 duplexes containing A, T, G, and/or C substitutions (identified as ULx) and for 23 unmodified DNA duplexes used as references (identified as UDx) are listed in Table 1. Similarly, thermodynamic data for 20 duplexes containing D and/or H substitutions (identified as USLx) are listed in Table 2. The thermodynamic values reported in Tables 1 and 2 were computed as the average of those determined by model-dependent and model-independent analyses (see Material and Methods). For all 86 duplexes, there was very good agreement in the thermodynamic data obtained by these two methods, with the value of $\Delta H_{2\text{-st}}/\Delta H_{cal}$ lying between 0.97 and 1.03 for all samples. Average values of the two-state (model-dependent) and calorimetric (model-independent) analysis could therefore be used to ameliorate any very minor biases that might be introduced by either data analysis method. We note that Zhou et al.[42] have shown that agreement between these two methods is a necessary but not completely sufficient condition to prove that the melting reaction occurs as a two-state transition.

There is also excellent agreement between our DSC-derived thermodynamic data for the unmodified DNA reference duplexes and those predicted using our recently published model for DNA thermodynamics that accounts for a non-zero ΔC_p.[34] Absolute average differences in $\Delta H°$, $\Delta S°$, and T_m values predicted by the model and those obtained by DSC were 5 ± 6% (% mean error ± % standard deviation), 5 ± 6%, and 0.5 ± 1.8 °C, respectively. Moreover, six of the DNA reference duplexes investigated by DSC were previously studied by UVM.[21,43] Good agreement between these independent methods was also observed, with average differences in $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ of 2 ± 5, 3 ± 7, and −2 ± 1%, respectively.

Each of the 43 duplexes containing one or more A, T, G, and/or C substitutions was more stable than its isosequential DNA reference sequence, with ΔT_m values ranging from 1.1 to 15.9 °C. When the same sequences contained D and/or H substitutions, even larger stability increases (ΔT_m = 1.7−14.9 °C) were observed. Finally, through DSC analysis (a typical thermogram is shown in Figure 2) of all 86 duplexes reported in Tables 1 and 2, we determined (Table 3) an average heat capacity change per base pair (ΔC_p^{bp}) of 42 ± 11 cal mol^{-1} K^{-1} bp^{-1} for the helix-to-coil transition. This value is in excellent agreement with a ΔC_p^{bp} recently reported for melting of duplex DNA[34] and also lies within the consensus range (30−60 cal mol^{-1} K^{-1} bp^{-1}) of all previously reported ΔC_p^{bp} values for duplex DNA.[29] We also note that over the range of melting temperatures analyzed (~42 °C < T_m < ~80 °C), the experimental value of ΔC_p^{bp} decreases slightly but not appreciably with temperature.

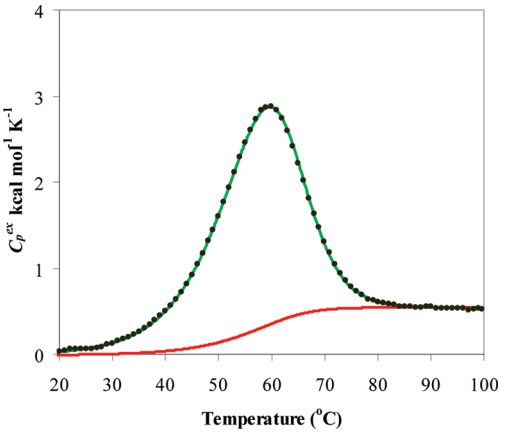


Figure 2. Representative DSC thermogram for the helix-to-coil transition for duplex UD1 showing the raw data (●), the two-state fit (green), and the baseline-derived ΔC_p (red).

Table 3. Average ΔC_p^{bp} and T_m Values Determined by DSC for Duplexes with and without LNA Substitutions

source	group	ΔC_p^{bp} (cal mol^{-1} K^{-1} bp^{-1})	T_m (°C)
Table 1	UD	47 ± 10	59 ± 9
Table 1	UL	41 ± 10	65 ± 8
Table 2	USL	40 ± 10	64 ± 4
Tables 1 and 2	all	42 ± 11	63 ± 8

Accounting for ΔC_p Shows That the Increase in Duplex Stability Resulting from LNA Substitutions Is Predominantly Driven by a Favorable Entropy Change. A non-zero value of ΔC_p indicates that $\Delta H°$ and $\Delta S°$ are functions of temperature.[31,35] Before determining incremental enthalpy $\Delta\Delta H°$ and $\Delta\Delta S°$ changes, we therefore used the measured ΔC_p for each modified duplex to compute $\Delta H°$ and $\Delta S°$ for that duplex at the T_m of the corresponding DNA reference sequence. Nearly identical results were obtained when the average ΔC_p^{bp} value of 42 cal mol^{-1} K^{-1} bp^{-1} was used for this computation. However, while the two methods yield similar results, we find that the errors in reported $\Delta\Delta H°$ and $\Delta\Delta S°$ values (as well as in incremental Gibbs energy changes $\Delta\Delta G°_{37}$) decrease slightly when the experimental ΔC_p value for the modified duplex is employed. For each sequence, the values of $\Delta H°$ and $\Delta S°$ as well as values of $\Delta\Delta H°$, $\Delta\Delta S°$, and $\Delta\Delta G°_{37}$ are reported in Tables 1 and 2. Average ΔC_p^{bp} values for each data set are reported in Table 3,

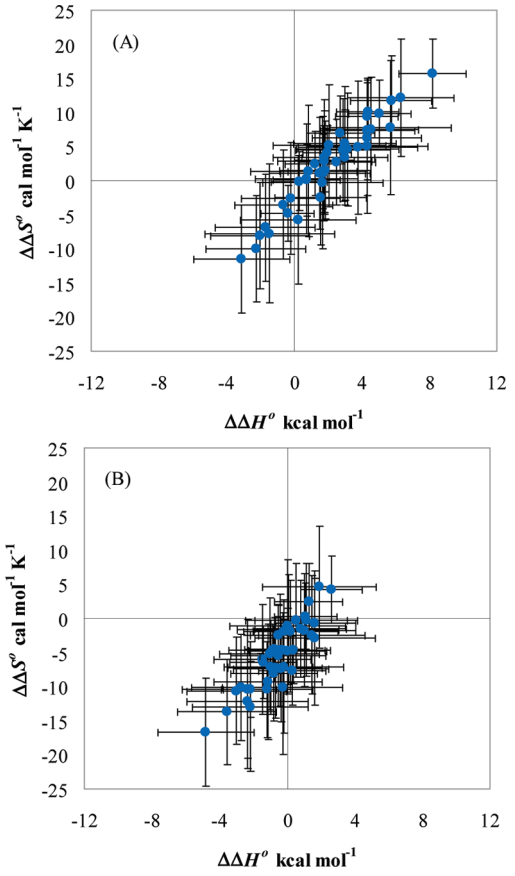


Figure 3. Experimental $\Delta\Delta H°$ and $\Delta\Delta S°$ data for the helix-to-coil transition of 43 duplexes with standard LNA substitutions determined under the assumption that $\Delta C_p = 0$ (A) or determined at the T_m of the isosequential DNA sequence using the ΔC_p of 42 cal mol^{-1} K^{-1} bp^{-1} as measured by DSC (B).

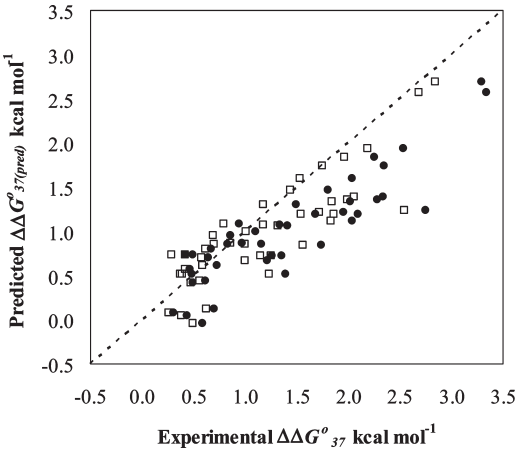


Figure 4. Comparison of $\Delta\Delta G°_{37}$ values predicted using the LNA NNT model of McTigue et al.[21] to experimental helix-to-coil $\Delta\Delta G°_{37}$ data for the 43 duplexes with standard LNA substitutions reported in Table 1. Both the raw experimental data (●) and data corrected to the T_m of the isosequential DNA reference sequences using ΔC_p (□) are shown.

and we find no statistical difference in the average value of $\Delta C_p{}^{bp}$ for LNA-modified duplexes relative to that for unmodified DNA duplexes, indicating that there is no appreciable $\Delta\Delta C_p$ associated with LNA substitution(s). $\Delta\Delta H°$, $\Delta\Delta S°$, and $\Delta\Delta G°_{37}$ values are therefore thought to be temperature-independent.

These results were used to assess the impact of a non-zero ΔC_p on our fundamental understanding of melting thermodynamics for short complementary duplexes containing LNA substitutions. For all standard LNA (A, T, C, or G)-substituted duplexes studied in this work, Figure 3A plots $\Delta\Delta H°$ versus $\Delta\Delta S°$ for the case in which both incremental changes are computed assuming ΔC_p equals zero. When applied to our data set, this assumption, which has been widely adopted in the DNA melting thermodynamics literature, leads to a result that suggests that the observed stability enhancement for LNA-substituted duplexes can be driven enthalpically, entropically, or by a combination thereof. These findings agree with previous thermodynamic studies of LNA-containing duplexes[6,7,21,23,24,35] and represent the primary justification made by McTigue et al.[21] for including both incremental enthalpy and entropy correction terms in their LNA NNT model.

However, when $\Delta H°_{LNA}[T_{m(DNA)}]$ and $\Delta S°_{LNA}[T_{m(DNA)}]$ are first computed from the $\Delta H°_{LNA}[T_{m(LNA)}]$, $\Delta S°_{LNA}[T_{m(LNA)}]$, and ΔC_p data for the modified duplex, we find, within experimental error, that a favorable incremental entropy change is the predominant duplex stabilization mechanism (Figure 3B), with average $\Delta\Delta H°$ and $\Delta\Delta S°$ values for the helix-to-coil transition for all duplexes studied of -0.5 ± 1.5 kcal mol^{-1} and -5.3 ± 4.8 cal mol^{-1} K^{-1}, respectively. Thus, with respect to duplex stability, a single LNA substitution is on average an athermal (or enthalpically slightly unfavorable) process that succeeds in stabilizing the duplex by reducing the entropy gain that accompanies the helix-to-coil transition. Together, panels A and B of Figure 3 show that the erroneous assumption that $\Delta C_p = 0$ can lead to significant errors in the interpretation of both DNA and LNA melting thermodynamics, including the introduction of nonrandom errors that result in an overestimation of any entropy–enthalpy compensation associated with LNA substitutions.

This motivated us to ask if any advances in the thermodynamic modeling of LNA-bearing oligonucleotides could be achieved by recognizing the non-zero value of ΔC_p. The existing LNA NNT model of McTigue et al.[21] can be used to predict incremental thermodynamic parameters. To identify areas in which model improvement efforts are justified, we therefore compared incremental thermodynamic values predicted by that model [$\Delta\Delta H°_{(pred)}$, $\Delta\Delta S°_{(pred)}$, and $\Delta\Delta G°_{37(pred)}$] with those determined from experimental data for the 43 standard LNA-containing duplexes reported in Table 1. $\Delta\Delta G°_{37(pred)}$ values computed by the LNA NNT model of McTigue et al. correlate somewhat (R^2 = 0.78) with $\Delta\Delta G°_{37}$ values derived from experiment assuming $\Delta C_p = 0$ [Figure 4 (●)]. A similar correlation (R^2 = 0.73) was obtained when LNA NNT model predictions were compared to $\Delta\Delta G°_{37}$ values corrected for the non-zero value of ΔC_p [Figure 4 (□)]. However, many significant outliers are observed in either correlation. Moreover, $\Delta\Delta H°_{(pred)}$ and $\Delta\Delta S°_{(pred)}$ values predicted by the LNA NNT model do not correlate at all (R^2 = 0) with either the corresponding uncorrected or temperature-corrected (using ΔC_p) experimental data, indicating that the LNA NNT model, which assumes ΔC_p equals zero, is unable to accurately quantify the enthalpic and entropic changes that drive the stability enhancements observed for LNA substitutions. As a result, the fundamental significance of the large set of LNA NN enthalpic and entropic parameters generated and required for that model is not clear. This then represents

Table 4. General SBT Model Parameters for the Helix-to-Coil Transition of Duplexes Containing Standard and Nucleoside-Modified LNA Substitutions

LNA-DNA base pair	$\Delta\Delta G_i^\circ{}_{37}$ (kcal mol^{-1})	$\Delta\Delta H_i^\circ$ (kcal mol^{-1})	$\Delta\Delta S_i^\circ$ (cal mol^{-1} K^{-1})
A-t	0.58 ± 0.06	−0.1 ± 0.3	−2.3 ± 0.9
T-a	0.88 ± 0.06	−0.1 ± 0.3	−3.2 ± 1.0
G-c	0.83 ± 0.10	0.1 ± 0.5	−2.5 ± 1.4
C-g	1.14 ± 0.09	−0.3 ± 0.4	−4.8 ± 1.3
D-t	1.05 ± 0.06	0.4 ± 0.3	−2.2 ± 0.9
H-a	1.20 ± 0.06	−1.5 ± 0.3	−8.7 ± 0.9
Effect of Nucleoside Modification in LNA			
D-t to A-t	0.47 ± 0.06	0.5 ± 0.3	0.1 ± 0.9
H-a to T-a	0.32 ± 0.06	−1.4 ± 0.3	−5.5 ± 1.0

one area in which useful advances in the understanding and modeling of LNA hybridization thermodynamics may be realized.

We therefore used the incremental thermodynamic values ($\Delta\Delta G^\circ_{37}$, $\Delta\Delta H^\circ$, and $\Delta\Delta S^\circ$) reported in Table 1 to regress a set of thermodynamic parameters (Table 4), identified as $\Delta\Delta G_i^\circ{}_{37}$, $\Delta\Delta H_i^\circ$, and $\Delta\Delta S_i^\circ$, that quantify the average incremental contribution (i.e., the excess contribution relative to that provided by the corresponding DNA nucleotide) to duplex stability provided by a single LNA of type i within its complementary base pair (A-t, T-a, G-c, or C-g). Consistent with Figure 3B, these average incremental changes indicate that the improved stability observed for LNA substitutions is entropically driven. Similar results were obtained when incremental thermodynamic values from either the two-state (model-dependent) or the calorimetric (model-independent) analysis were instead used to regress the parameters, but in either case, this resulted in a slightly larger uncertainty (standard error) in the parameters. On average, no statistically significant incremental enthalpy change $\Delta\Delta H_i^\circ$ is observed for any of the four standard LNA substitutions. A simple model describing the melting thermodynamics of LNA bearing duplexes may therefore be realized by setting all $\Delta\Delta H_i^\circ$ equal to 0 so that $\Delta\Delta G_i^\circ{}_{37} = -310\Delta\Delta S_i^\circ$. Differences in stability enhancement provided by a particular LNA substitution are then ascribed to the fact that the magnitude of $\Delta\Delta S_i^\circ$ is LNA base specific (Table 5).

Base Classification and Pairing Explain Differences in the $\Delta\Delta S_i^\circ$ Value and Stability Enhancement Offered by Different LNAs. Structural studies of single-stranded and duplexed oligonucleotides containing one or more LNA substitutions indicate that introduction of an LNA preorganizes (i.e., lowers the entropy of) the LNA-containing single strand.[9−12] This serves to improve duplex stability. In particular, the introduction of a 2′-O,4′-C-methylene bridge forces the nucleoside into its N-type conformation in the single strand. It remains fixed in its N-type conformation within the duplex, which lowers the entropy of the duplex and thereby reduces duplex stability, albeit likely to a lesser extent because of the inherently greater degree of organization of the duplex structure. These opposing entropic effects associated with LNA-nucleoside structure are consistent with our thermodynamic data. They further suggest that the differences in $\Delta\Delta S_i^\circ$ values for the four possible LNA nucleosides arise from differences in the extent of compensation between the two opposing entropy changes. To improve our understanding of this, we compared incremental entropy changes as a function of LNA base classification (i.e., purine vs pyrimidine) and pair (i.e., a-t base pair versus c-g base pair). Table 5 provides a summary of this analysis, where the $\Delta\Delta S_i^\circ$ values have now been calculated directly from our more accurate $\Delta\Delta G_i^\circ{}_{37}$ values reported in Table 4 assuming $\Delta\Delta H_i^\circ$ is 0 for each of the four standard LNA base substitutions. The results show that either possible LNA-pyrimidine substitution (t to T or c to C) results in a $-\Delta\Delta S_i^\circ$ (the incremental entropy change for the hybridization reaction) that is 1.0 ± 0.3 cal mol^{-1} K^{-1} more favorable than that of the corresponding LNA-purine substitution, suggesting that pyrimidinic LNA substitutions provide a comparatively greater increase in the level of preorganization of single strands. We also find that $-\Delta\Delta S_i^\circ$ values increase in the following order: C > T ≥ G > A (Table 5). This is in good agreement with the average stability enhancement observed for each of these substitutions as determined both by us and by others.[21]

Table 5. $\Delta\Delta S_i^\circ$ Parameters and Differences in Them for the Helix-to-Coil Transition of Duplexes Containing Standard LNA bases[a]

	$\Delta\Delta S_i^\circ$ (cal mol^{-1} K^{-1})		
LNA-DNA base pair	parameter	base-type difference	base pair difference
A-t	−1.9 ± 0.2	−1.0 ± 0.2	−0.8 ± 0.3
T-a	−2.9 ± 0.2		
G-c	−2.7 ± 0.3	−1.0 ± 0.3	
C-g	−3.7 ± 0.3		

[a] $\Delta\Delta S_i^\circ$ parameters were determined from $\Delta\Delta G_i^\circ{}_{37}$ parameters listed in Table 4 assuming $\Delta\Delta H^\circ = 0$.

The entropy gain arising from an LNA substitution likewise depends on base pair type, with $-\Delta\Delta S_i^\circ$ for substitution within a g-c base pair being 0.8 ± 0.3 cal mol^{-1} K^{-1} larger than that when the corresponding nucleotide is replaced with an a-t base pair. In this case, however, the effect is likely related to differences in the degrees of freedom available to the DNA base pair within the reference duplex. The unified NNT model of Santa Lucia[22] estimates the average thermodynamic contributions of a g-c base pair within duplex DNA to be as follows: $-\Delta G^\circ_{37} = -1.8$ kcal mol^{-1}, $-\text{-}\Delta H^\circ = -9.0$ kcal mol^{-1}, and $-\Delta S^\circ = -23.1$ cal mol^{-1} K^{-1}. Those for an a-t base pair are estimated to be −1.0 kcal mol^{-1}, −7.8 kcal mol^{-1}, and −21.6 cal mol^{-1} K^{-1}, respectively. On average then, the additional hydrogen bond in a c-g base pair acts to enhance the stability of a duplex by approximately −0.8 kcal mol^{-1} relative to that of an a-t base pair. As a result of this stronger interaction, the entropy of each base within the g-c pair is ∼0.8 cal mol^{-1} K^{-1} lower than that of the corresponding base within an a-t pair. Given this more organized base structure within a g-c pair, the average entropic penalty associated with locking either base into its N-type conformation is reduced when compared to that when the LNA substitution occurs within an a-t base pair.

Terminal 5′ and 3′ LNA Substitutions Are Much Less Stabilizing Than Internal LNA Substitutions. The thermodynamic values reported in Tables 1−5 were regressed from DSC data for duplexes that were intentionally designed to avoid artifacts that could be attributed to the duplex termini. However, development of a general and accurate model for predicting the T_m and melting thermodynamics of LNA-containing duplexes will require proper accounting of the influence of terminal LNA substitutions on duplex stability. Although it was not generally

Table 6. DSC-Derived Thermodynamic Data for Duplexes Used To Study the Effect of LNA Substitution at the 3′ or 5′ Terminus[a]

name	sequence	ΔC_p (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta\Delta G°_{37}$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$\Delta\Delta S°$ (cal mol^{-1} K^{-1})	ΔT_m (°C)
C1	ctacgcattcc	462	12.18	81.7	224.0	59.2				
L5C-C1	Ctacgcattcc	492	12.28	82.8	227.2	59.4	0.03	1.0	3.4	0.1
L5G-C1	Ggaatgcgtag	405	12.26	81.8	224.2	59.6	0.18	0.0	0.5	0.3
L3C-C1	ctacgcattcC	634	12.16	81.0	222.1	59.4	0.16	−0.7	−1.7	0.1
L3G-C1	ggaatgcgtaG	449	12.46	82.5	226.0	60.2	0.58	0.4	3.3	1.0
T1	ttcatagccgt	353	11.98	79.3	216.9	59.1				
L5T-T1	Ttcatagccgt	430	11.73	78.2	214.2	58.3	−0.43	−0.7	−3.7	−0.8
L5A-T1	Acggctatgaa	260	11.64	77.2	211.4	58.2	−0.13	−1.8	−6.3	−1.0
L3T-T1	ttcatagccgT	414	12.05	79.7	218.2	59.3	0.07	0.4	1.5	0.2
L3A-T1	acggctatgaA	475	12.08	80.3	220.0	59.2	0.00	1.0	3.3	0.1
T5	tactccgcatt	478	11.95	73.9	199.7	60.7				
LT5-T5	Tactccgcatt	357	11.58	71.5	193.1	59.7	−0.33	−2.0	−7.6	−1.0
LA3-T5	aatgcggagtA	280	12.03	74.1	200.1	61.0	0.09	0.1	0.7	0.3

[a] Measured thermodynamic values ($\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m) are for the helix-to-coil transition and were determined as reported in Materials and Methods. For the mean experimental values reported, the estimated errors (standard deviation) for ΔC_p, $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_m were determined to be 25, 2, 3, and 3% and 0.5 °C, respectively. $\Delta H°$ and $\Delta S°$ are determined at the T_m of the sample, while $\Delta\Delta H°$ and $\Delta\Delta S°$ are determined at the T_m of the reference isosequential DNA duplex and computed using the average ΔC_p^{bp} of 42 cal mol^{-1} K^{-1} bp^{-1}. C_T = 75 μM for all duplexes listed here.

their focus, previous investigations found that LNA substitutions at the 5′ or 3′ end of duplex DNA are slightly destabilizing;[15,23,44] however, conflicting findings have been reported,[24] and the available information is not sufficient for reliable model development. We therefore designed and studied 10 duplexes with a single LNA substitution at the 5′ or 3′ end. All four possible base substitutions at each position were tested, and our results are listed in Table 6 together with reference data for the associated isosequential DNA duplexes.

We find that, on average, a single LNA substitution at the 5′ or 3′ terminal end results in no significant stability change in the duplex (ΔT_m = − 0.1 ± 0.7 °C, and $\Delta\Delta G°_{37}$ = −0.02 ± 0.28 kcal mol^{-1}). Duplexes with 5′ terminal substitutions exhibited an average ΔT_m of −0.5 ± 0.6 °C, with three of the modified sequences showing a reduction in T_m of >0.5 °C. One LNA substitution at the 3′ terminus resulted in a statistically significant increase in stability (ΔT_m > 0.5 °C). However, on average, 3′ substitutions were also found to provide no significant stability improvement, exhibiting an average ΔT_m of 0.3 ± 0.4 °C. These end effects are not described well by either the LNA NNT model or the incremental thermodynamic changes reported in Table 5 for LNA substitutions within the interior of the duplex. Instead, the results, particularly the finding that on average $\Delta\Delta G°_{37}$ for terminal substitutions is −0.02 ± 0.28 kcal mol^{-1}, indicate that to a good approximation single isolated LNA substitutions at oligonucleotide termini can be ignored for the purpose of stability prediction.

A New Model for Predicting the Melting Thermodynamics of LNA-Substituted Duplexes. Our results suggest that accurate calculation of $\Delta T_{m(pred)}$ and melting thermodynamics for complementary duplexes containing standard LNAs might be achieved with a new model, which we call the single-base thermodynamic (SBT) model, that corrects predictions for the isosequential DNA duplex made by new model described by Hughesman et al.[34] The SBT model uses only the experimentally determined average ΔC_p^{bp} value and the set of four incremental entropy parameters ($\Delta\Delta S_i°$) reported in Table 5. The model is therefore considerably simpler in structure than the existing 64-parameter LNA NNT model.[21] Moreover, as the incremental thermodynamic parameters reported in Table 5 specifically characterize the individual LNA base substitution and not the associated nearest neighbor, the SBT model can in principle be applied to highly substituted duplexes containing one or more adjacent LNAs within one of the duplex-forming oligonucleotides.

In its most general form, the SBT model can be applied to oligonucleotides substituted with standard LNAs and the nucleobase-modified LNAs D and H, which, unlike standard LNAs, provide both entropic and enthalpic stability enhancement (see below). Thermodynamic changes for the helix-to-coil transition at temperature T are computed as

$$\Delta H°_{LNA}(T_m) = \Delta H°_{DNA}(T_{ref}) + \Delta\Delta H°_{LNA} + \Delta C_p(T_m - T_{ref}) \quad (7)$$

$$\Delta S°_{LNA}(T_m) = \Delta S°_{DNA}(T_{ref}) + \Delta\Delta S°_{LNA} + \Delta C_p \ln(T_m/T_{ref}) \quad (8)$$

where $\Delta H°_{DNA}(T_{ref})$ and $\Delta S°_{DNA}(T_{ref})$ are determined using the aforementioned new model of Hughesman et al. for duplex DNA that corrects for both a non-zero ΔC_p and the unique base-paring energetics of terminal 5′-ta groups.[34] Equations 7 and 8 show that the SBT model treats $\Delta H°_{LNA}$ and $\Delta S°_{LNA}$ as temperature-dependent functions, with their values at the T_m of the LNA-containing duplex computed using ΔC_p and the predicted melting thermodynamics of the isosequential DNA duplex at the chosen reference temperature of 53 °C. A T_{ref} of 53 °C was identified as described previously.[34] In particular, regression of the DNA thermodynamics model of Hughesman et al. to melting data for 128 unmodified duplex DNA sequences resulted in minimal model error when T_{ref} = 53 °C. In eqs 7 and 8, ΔC_p is given by $n\Delta C_p^{bp}$, where n is the total number of base pairs in the duplex and ΔC_p^{bp} is given by our average experimental value, 42 cal mol^{-1} K^{-1} bp^{-1}.

$\Delta\Delta H°_{LNA}$ and $\Delta\Delta S°_{LNA}$ are computed as

$$\Delta\Delta H°_{LNA} = \sum n_i \Delta\Delta H°_i \quad (9)$$

$$\Delta\Delta S°_{LNA} = \sum n_i \Delta\Delta S°_i \quad (10)$$

dx.doi.org/10.1021/bi200223s |Biochemistry 2011, 50, 5354–5368

where n_i is the number of standard and base-modified LNA substitutions of type i (not including those at either terminus) and the values of $\Delta\Delta H°_i$ and $\Delta\Delta S°_i$ are those reported in Table 4. The value of T_m for a duplex in which one strand contains one or more internal LNA substitutions is then given by

$$T_m = \frac{\Delta H°_{LNA}(T_m)}{\Delta S°_{LNA}(T_m) - R\ln(C_T/4)} \quad (11)$$

where C_T is the total molar concentration of single-stranded oligonucleotides. Solving for T_m using eq 11 requires iteration, but convergence is rapid if the initial estimate for T_m is found by setting ΔC_p equal to 0 in eqs 7 and 8.

For duplexes containing only standard LNA base substitutions, we have shown that all $\Delta\Delta H°_i$ values are statistically zero. A simplified form of the SBT model can then be obtained by setting $\Delta\Delta H°$ equal to 0 and calculating $\Delta\Delta S°$ using the $\Delta\Delta S°_i$ parameters listed in Table 5.

The SBT Model Predicts T_m Values for Standard LNA-Containing Mixmer Duplexes with Accuracy Similar to That of More Complex NNT Models. The performance of our simplified SBT model was first tested by determining errors in model-predicted $T_{m(LNA)}$ and $\Delta T_{m(pred)}$ values through comparison with experimental data for 206 mixmer duplexes containing one or more interspersed (nonadjacent) standard LNA substitutions. The 206 test sequences included the 43 duplexes from Table 1 that were used to regress the SBT model parameters, along with 163 mixmer duplexes for which experimental $T_{m(LNA)}$ and ΔT_m values have been reported previously.[14,21,44] Many of these 163 sequences were part of the basis set used to regress the parameters of the LNA NNT model.[21] All experimental ΔT_m values were measured at or very near 1 M Na^+, eliminating the need to correct for differences in salt concentration. SBT model-predicted $\Delta T_{m(pred)}$ values were computed as $T_{m(LNA)} - T_{m(DNA)}$, with both required T_m values determined by eq 11 ($\Delta\Delta H°$ and $\Delta\Delta S°$ = 0 for the unmodified reference duplex calculations).

As shown in Table 7, when applied to mixmers, the four-parameter SBT model calculates $\Delta T_{m(pred)}$ with an accuracy similar to that of the two methods previously described.[20,21] Average $\Delta T_{m(error)}$ values for duplexes not used for parameter regression are 0.4 ± 1.5 °C for the SBT model, −0.2 ± 1.4 °C for the LNA NNT model, and 0.0 ± 1.9 °C for the OligoDesign algorithm. However, though all of these methods show similar performance in correctly estimating ΔT_m values of LNA mixmer duplexes, we found for duplexes highly substituted with LNA that the SBT model provides ΔT_m values that are more accurate and more precise because of the temperature dependence of $\Delta H°$ and $\Delta S°$ in our model. In particular, for eight 15-mer LNA mixmer duplexes studied by You et al., each containing seven LNAs on the modified strand, there is a significant improvement in the $\Delta T_{m(error)}$ when the SBT model is used (0.0 ± 1.2 °C) as compared to either the OligoDesign (−2.2 ± 1.9 °C) or LNA NNT (−0.4 ± 2.3 °C) model.

Finally, independent data from Levin et al.[44] provide a reasonable confirmation of our correction for the effect of single LNA substitutions at terminal ends. For the four duplexes they studied that contain both interspersed internal LNA(s) and a 5′ terminal LNA substitution, the $\Delta T_{m(error)}$ values decreased from 3.7 °C when the appropriate $\Delta\Delta S°_i$ parameter of the SBT model was applied to the 5′ LNA to 0.2 °C when it was not. T_m data reported by Lattora et al.[45] for duplex primers containing terminal 3′ LNA substitutions show similar behavior, supporting our finding that an LNA substitution at the duplex termini provides no significant change in duplex stability relative to the DNA isosequence when adjacent base pairs in the duplex are not populated with an LNA.

Table 7. Errors in ΔT_m (°C) Values Predicted Using the Specified Model for Standard LNA (A, T, G, and/or C)-Substituted Mixmer Duplexes Not Used for SBT Model Parameter Regression[a]

source	count	Exiqon OligoDesign	LNA NNT model[b]	SBT NNT model[b]
this study[c]	43	0.1 ± 2.1	−0.6 ± 1.2	
McTigue[d]	100	0.3 ± 1.7		0.5 ± 1.3
You[e]	36	−1.3 ± 1.4	−0.1 ± 1.1	0.1 ± 1.2
Levin[f]	27	0.2 ± 2.1	0.1 ± 2.1	0.5 ± 2.2
subtotal		0.0 ± 1.9	−0.2 ± 1.4	0.4 ± 1.5

[a] Error values [denoted as $\Delta T_{m(errors)}$ in the text] give the mean error ± standard deviation for model-predicted $\Delta T_{m(pred)}$ values relative to experimental ΔT_m values. [b] $\Delta T_{m(pred)}$ values predicted by the LNA NNT and SBT models ignore 5′ terminal LNA substitutions. [c] Data from Table 1. [d] Data from ref 21. [e] Data from ref 14. [f] Data from ref 44.

Testing the Validity of SBT Model Assumptions. By measuring and accounting for the non-zero value of ΔC_p, we have shown that stability improvements resulting from standard LNA substitutions are entropically driven, likely due to a preorganization of the specific nucleotide involved. In developing the SBT model, we have assumed, but not yet proven, that these preorganization effects manifest themselves within the modified nucleotide, making incremental energy and entropy corrections additive at the individual base pair level. Here, we test the validity of this potential insight and associated model assumption.

All parameters of the SBT model were regressed from DSC data for mixmer duplexes. However, if our underlying thermodynamic analysis is correct, we should be able to successfully apply the model to gapmer-type duplexes incorporating tandem (neighboring) LNAs. To test this, we designed and studied by UVM a set of 16 gapmer duplexes (Table 8), each containing one of the 16 possible tandem LNA sequences. Melting of the three associated reference DNA duplexes was likewise studied both by UVM and DSC, and DSC was also used to confirm thermodynamic values obtained by UVM for seven of the 16 tandem LNA-substituted sequences. There was excellent agreement between DSC and UVM data for all 10 duplexes studied by both techniques, with the following average differences between the two methods: $\Delta H°$ = 0 ± 1%, $\Delta S°$ = 1 ± 1%, $\Delta G°_{37}$ = 0 ± 1%, and T_m = 0.3 ± 0.2 °C.

Table 8 reports helix-to-coil transition data determined by UVM for these sequences, as well as the experimental incremental thermodynamic values $\Delta\Delta G°_{37}$, $\Delta\Delta H°$, $\Delta\Delta S°$, and ΔT_m. On average, $\Delta\Delta H°$ is −1.6 ± 2.4 kcal mol^{-1} and $\Delta\Delta S°$ is −9.9 ± 5.3 cal mol^{-1} K^{-1}, in accordance with our finding that the stability enhancement is an entropically driven process within experimental error. Incremental thermodynamic values for the 16 duplexes were also computed using the four-parameter SBT model, and good agreement was observed between model predictions and experiment, with average model errors of −0.2 ± 0.3 kcal mol^{-1}, −1.3 ± 1.7 kcal mol^{-1}, −3.6 ± 5.4 cal mol^{-1} K^{-1}, and 0.0 ± 1.5 °C for $\Delta\Delta G°_{37(error)}$, $\Delta\Delta H°_{(error)}$, $\Delta\Delta S°_{(error)}$, and $\Delta T_{m(error)}$, respectively. As these sequences were not part of the data set used to regress SBT model parameters, the

Table 8. UVM-Derived Thermodynamic Data for Duplexes with Tandem LNA Substitutions[a]

	sequence									
name	strand 1	strand 2	ΔG°_{37} (kcal mol^{-1})	ΔH° (kcal mol^{-1})	ΔS° (cal mol^{-1} K^{-1})	T_m (°C)	$\Delta\Delta G^\circ_{37}$ (kcal mol^{-1})	$\Delta\Delta H^\circ$ (kcal mol^{-1})	$\Delta\Delta S^\circ$ (cal mol^{-1} K^{-1})	ΔT_m (°C)
T1	ttcatagccgt	acggctatgaa	11.96 ± 0.13	79.2 ± 2.6	216.9 ± 7.6	59.0 ± 0.4				
LCT	ttcatagccgt	acggCTatgaa	14.70 ± 0.37	81.7 ± 4.9	215.9 ± 14.1	70.6 ± 0.4	2.3	−2.9	−16.8	11.6
LGG	ttcatagccgt	acGGctatgaa	13.90 ± 0.34	81.5 ± 3.5	218.1 ± 10.0	67.0 ± 0.6	1.6	−1.4	−9.8	8.0
LTA	ttcatagccgt	acggcTAtgaa	13.83 ± 0.28	81.9 ± 4.3	219.4 ± 12.5	66.6 ± 0.2	1.6	−0.8	−7.9	7.5
LTG	ttcatagccgt	acggctaTGaa	14.26 ± 0.21	84.3 ± 0.9	225.8 ± 2.9	67.5 ± 0.6	2.0	1.1	−2.8	8.5
LAG	ttcatAGccgt	acggctatgaa	13.19 ± 0.06	76.9 ± 2.7	205.3 ± 8.0	65.6 ± 0.3	1.0	−5.4	−20.6	6.6
LCA	ttCAtagccgt	acggctatgaa	13.55 ± 0.21	81.9 ± 3.7	220.4 ± 10.9	65.3 ± 0.1	1.4	−0.2	−5.1	6.2
LCC	ttcatagCCgt	acggctatgaa	14.38 ± 0.26	81.3 ± 2.1	215.7 ± 6.0	69.3 ± 0.5	2.0	−2.7	−15.3	10.3
C1	ctacgcattcc	ggaatgcgtag	12.08 ± 0.05	80.6 ± 1.6	221.0 ± 4.8	59.1 ± 0.1				
LAT	ctacgcATtcc	ggaatgcgtag	12.98 ± 0.27	80.6 ± 3.5	217.9 ± 10.2	63.2 ± 0.5	0.8	−1.9	−8.9	4.1
LCG	ctaCGcattcc	ggaatgcgtag	14.23 ± 0.09	83.4 ± 2.0	222.9 ± 5.7	67.8 ± 0.2	1.9	−1.3	−10.1	8.6
LGC	ctacGCattcc	ggaatgcgtag	13.97 ± 0.18	82.4 ± 2.8	220.8 ± 8.2	66.9 ± 0.2	1.6	−1.8	−11.1	7.8
LTT	ctacgcaTTcc	ggaatgcgtag	14.08 ± 0.07	84.6 ± 2.2	227.4 ± 6.7	66.6 ± 0.3	1.8	0.5	−4.1	7.5
LAA	ctacgcattcc	ggAAtgcgtag	12.82 ± 0.08	80.3 ± 2.0	217.5 ± 5.9	62.5 ± 0.2	0.7	−1.9	−8.4	3.4
C2	ctaacggatgc	gcatccgttag	12.00 ± 0.05	80.0 ± 0.3	219.2 ± 0.7	59.0 ± 0.5				
LGT	ctaacggatgc	gcatccGTtag	12.79 ± 0.18	80.6 ± 0.6	218.5 ± 1.7	62.3 ± 0.3	0.7	−1.0	−5.3	3.4
LTC	ctaacggatgc	gcaTCcgttag	13.97 ± 0.24	81.1 ± 3.7	216.4 ± 10.6	67.5 ± 0.3	1.6	−2.8	−14.4	8.5
LAC	ctaACggatgc	gcatccgttag	13.85 ± 0.06	83.8 ± 0.8	225.4 ± 2.2	65.9 ± 0.4	1.6	0.6	−3.2	6.9
LGA	ctaacgGAtgc	gcatccgttag	13.70 ± 0.11	80.2 ± 1.9	214.5 ± 5.6	66.6 ± 0.3	1.4	−3.3	−15.1	7.6

[a] The reported UVM data (ΔG°_{37}, ΔH°, ΔS°, and T_m) are for the helix-to-coil transition and were determined as reported in Materials and Methods. The reported errors refer to the standard deviation of triplicate runs. ΔH° and ΔS° are determined at the T_m of the sample, while $\Delta\Delta H^\circ$ and $\Delta\Delta S^\circ$ are determined at the T_m of the reference isosequential DNA duplex and computed using the average ΔC_p^{bp} of 42 cal mol^{-1} K^{-1} bp^{-1}. C_T = 75 μM for all duplexes listed here.

agreement of model predictions with experiment provides good evidence that the assumptions upon which the SBT model was derived are sound. Moreover, the SBT model accurately predicts ΔT_m values for 19 additional gapmer duplexes (Table 9) studied previously.[14] The prediction accuracy for the combined set of 35 duplexes [$\Delta T_{m(error)}$] was 0.1 ± 1.4 °C (Table 9), similar to that reported for mixmer duplexes (Table 7).

As an extreme test of SBT model performance, predicted ΔT_m values were compared with experimental data for five heteroduplexes in which the modified ss bore an LNA at every position. For each of these fully substituted sequences, the experimental ΔT_m is very large, falling between 31.1 and 40.4 °C. $\Delta T_{m(error)}$ values for this prediction are listed in Table 9, and the results show that, even for these extreme cases, the predicted ΔT_m values differ from the corresponding experimental ΔT_m by no more than 15%. A bias is observed, as the SBT model underpredicts the melting temperature for the five fully substituted LNA-DNA heteroduplexes. Our studies suggest that this bias arises, at least in part, because terminal LNA can contribute to the net stability enhancement, albeit to a lesser extent when compared to an internal LNA, when they are flanked by additional LNA. Indeed, we generally find that $\Delta T_{m(pred)}$ values for fully substituted sequences differ from experiment by no more than ±6% (Table 9) when the $\Delta\Delta S^\circ_i$ for each terminal LNA is taken to be half that reported in Table 5 for the corresponding internal LNA. However, additional studies are required to establish a firm understanding of the contribution to duplex stability made by terminal LNA(s) in highly substituted gapmer sequences such as those found in fully substituted LNA-DNA heteroduplexes.

Table 9. Errors [denoted $\Delta T_{m(error)}$] in ΔT_m Values Predicted Using the SBT Model for LNA-Substituted Gapmer Duplexes and Fully Modified LNA-DNA Heteroduplexes[a]

source	no. of sequences	$\Delta T_{m(error)}$ (°C)
LNA-Substituted Gapmer Duplexes		
this study[b]	16	0.0 ± 1.5
You[c]	19	0.2 ± 1.3
subtotal	35	0.1 ± 1.4
Fully Substituted LNA-DNA Heteroduplexes[d]		
CTTGTGA[e]		−4.1 (1.1)
TCACAAG[e]		−2.5 (1.8)
CGTATAGT[f]		−6.9 (−0.5)
CGTDTAGT[f]		−6.1 (0.2)
CGTDTDGT[f]		−5.6 (0.7)

[a] $\Delta T_{m(error)}$ reported as the mean error ± the standard deviation between the model-predicted $\Delta T_{m(pred)}$ and the experimental ΔT_m. [b] Data from Table 8. [c] Data from ref 14. [d] $\Delta T_{m(pred)}$ values predicted for fully modified LNA-DNA heteroduplexes were calculated either ignoring terminal LNA substitutions or assuming that terminal LNA substitutions contribute half the stability of those for internal LNA (shown in parentheses). [e] ΔT_m values estimated from ΔT_{max} as the high stability of the LNA-DNA heteroduplex prevented complete thermodynamic analysis. [f] Data from ref 38.

Substitution of Standard LNAs with Modified LNA Nucleosides Provides Further Stability Increases. There are very few thermodynamic data available for duplexes containing the

dx.doi.org/10.1021/bi200223s |Biochemistry 2011, 50, 5354–5368

base-modified LNA nucleotides D and H. Table 4 reports average incremental thermodynamic parameters for D and H substitutions determined from our experiments and a non-zero value of ΔC_p. When compared to its standard LNA analogue (A or T, respectively), either substitution further stabilizes the duplex. However, as the chemical modification is now within the base of the LNA, it is intended by design to alter the energetics (enthalpy) of complementary base pairing. Therefore, $\Delta\Delta H°_i$ values are no longer athermal and indicate a change in the strength of base pairing, but not necessarily in base stacking because stacking is believed to be primarily related to the surface area of the base, which is not changed appreciably by the base modification.[2]

The incremental entropy change for the modified LNA base D is similar to that for its standard LNA homologue A. The D-t base pair is then further stabilized as compared to an A-t pair by a favorable incremental enthalpy change for the hybridization reaction ($-\Delta\Delta H°_i$) of -0.5 ± 0.3 kcal mol^{-1}, which is consistent with previously reported values for the contribution to DNA duplex stability of a single hydrogen bond.[46,47] This result is also consistent with previous studies of DNA-2-aminoadenine (d) substitutions that indicate that the increased duplex stability occurs through the creation of a third hydrogen bond between the bases, formed between the 2-amine group introduced into adenine and the existing 2-keto oxygen atom in thymine.[48−50]

In contrast, the enhancement of duplex stability provided by H substitution when compared to T substitution is found to be due to a more favorable incremental entropy change, which in this case is partially compensated by an unfavorable incremental enthalpy change. Previous studies of DNA-2-thiothymine (h) suggest the replacement of the 2-keto oxygen with 2-keto sulfur may act to shift the base pair into a more stable A-type form.[51] Other work has suggested that the stability increase could be caused by improved stacking and/or a change in solvation or cation interactions.[52] The signs and magnitudes of our regressed incremental thermodynamic parameters for the T to H substitution appear to be more consistent with the first argument and also suggest a greater degree of preorganization of both the altered base and the sugar ring of H within the single strand.

The incremental thermodynamic parameters for D and H reported in Table 4 can be combined with our general SBT model to accurately determine ΔT_m values for mixmer and gapmer duplexes containing both standard and base-modified LNAs. Average errors in predicted ΔT_m and $T_{m(LNA)}$ values for the 20 duplexes in Table 2 bearing D and/or H substitutions were -0.5 ± 1.3 and -0.3 ± 2.1 °C, respectively.

D-H Base Pairs Demonstrate Pseudocomplementary Properties. Two duplexes from Table 1 (UD1 and UD11) were chosen to further study all possible complementary base pairs between adenine (a, d, A, and D) and thymine (t, h, T, and H) analogues. The results of thermal denaturation experiments with these duplexes are compiled in Table 10. We find that duplexes containing d-h, D-h, and/or d-H base pairs have lower T_{max} values than the corresponding isosequential reference duplex containing a-t base pairs. Duplexes containing D-H base pairs have a higher T_{max} than their reference duplex (a-t) but still exhibit net lower stability compared to the stabilities of isosequences containing their three LNA-LNA base pair analogues (A-T, D-T, and A-H). The D-H base pair (as well as the d-h, d-H, and D-h base pairs) is therefore pseudocomplementary because of steric hindrance (Figure 1) between the 2-amino group in d or D and the 2-thioketo group in h or H that serves to reduce base pair stability.[37,39,41,53] This could prove to be useful in certain applications, as D-T and H-A base pairs are very stable (Table 10) because of the absence of these steric hindrance effects. As a result, oligonucleotides containing A and T, A and H, and/or D and T must be designed carefully, especially in regions of self-complementary, to ensure that highly stable secondary products are not formed. Our results suggest that when these undesired secondary products are observed, it may be advantageous to redesign the required oligonucleotides exclusively with D and H substitutions and avoid the use of A and T.

Table 10. UVM-Derived ΔT_{max} Values for the Helix-to-Coil Transition of Duplexes with Base Pairs Formed between a, d, A, or D and t, h, T, or H[a]

	X = a	X = d	X = A	X = D
		gcccXgcg/cgcYgggc		
Y = t	–	3.5	1.8	8.3
Y = h	2.8	−8.4	4.8	−3.3
Y = T	4.7	7.9	12.0	20.5
Y = H	9.4	−0.4	17.1	9.8
		aacatagXttacat/atgtaaYctatgtt		
Y = t	–	0.7	2.0	3.4
Y = h	0.3	−6.6	3.8	−2.7
Y = T	3.8	4.9	6.5	9.1
Y = H	5.6	−2.8	10.0	4.0

[a] Estimated error in ΔT_{max} of ±1.0 °C. For both sequences tested, the T_{max} of the duplex containing a-t base pair at the X-Y position was used as the reference to determine ΔT_{max}.

DISCUSSION

We have previously demonstrated that modification of the unified DNA NNT model to account for the temperature dependence of $\Delta H°$ and $\Delta S°$ can improve $T_{m(DNA)}$ predictions, especially at temperatures above 70 °C.[34] Here, we report DSC and UVM data that show that $\Delta H°$ and $\Delta S°$ for the helix-to-coil transition in duplexes containing LNA substitutions are also temperature-dependent. By measuring and recognizing the non-zero value of ΔC_p, we are then able to provide important insights into the manner by which LNA substitutions serve to stabilize a duplex. In particular, previous studies, which treat melting data assuming $\Delta C_p = 0$, have concluded that the stability enhancement in duplexes substituted with standard LNAs is enthalpically driven when ΔT_m becomes large.[21] Indeed, incremental enthalpy $\Delta\Delta H°$ and entropy $\Delta\Delta S°$ changes (Figure 3A) computed from our DSC data assuming $\Delta C_p = 0$ are consistent with this observation. However, when the temperature dependencies of $\Delta H°$ and $\Delta S°$ are properly taken into account, we find that a negative $\Delta\Delta S°$ provides the stability enhancement, and that this stabilization effect is localized at the level of the individual base pair. Our thermodynamic data therefore suggest that neighboring bases play little role in the net change in stability accompanying the substitution of a standard LNA into a duplex. Though we cannot categorically rule out the possibility that LNA nucleobase substitutions promote subtle changes in the base pair bonding interaction and/or stacking interactions, our results clearly show that $\Delta\Delta H°$ equals zero within the experimental error of current measurement techniques, including high-sensitivity DSC equipment.

Structural studies of duplexes containing one or more LNA substitutions have been conducted to investigate possible differences in base pairing and stacking between LNA and DNA nucleotides.[9−12] Regrettably, these studies have generally been conducted at temperatures (25−27 °C) at which stacking in the single-stranded and double-stranded states occurs much more strongly than near the T_m of the duplex. As a result, values of ΔC_p are known to be significantly higher at 25 °C than at T_m.[28] This makes it difficult to relate structural data collected near room temperature to thermodynamic data determined at T_m. However, even when structural studies are conducted at these lower temperatures, which should promote improved base pair and stacking interactions, Jensen et al.[9] found that although LNA substitutions do alter the conformation of the sugar within the LNA nucleotide, they do not promote significant changes in the base stacking when compared to that of an unsubstituted DNA duplex. Though limited, these structural data are consistent with our finding that the stability enhancement mainly occurs at the level of the individual base pair.

In accordance with both the structural data of Jensen et al.[9] and our extensive set of DSC- and UVM-derived melting data, we assumed and then demonstrated that the melting thermodynamics of mixmer (Table 7) and gapmer (Table 9) duplexes can be accurately predicted with a simple group contribution-type model that is additive at the level of the individual base pair. We note that this model structure is challenged by one study of RNA duplexes (LNA can also be introduced into RNA) that suggests that a saturation of ΔT_m occurs with an increasing number of LNA substitutions.[10] Other groups have likewise argued that the stability enhancement within duplex DNA is saturated with an increasing number of LNA substitutions.[11−13] These arguments are largely based on a limited set of data that show that the average increase in ΔT_m per LNA substitution appears to decrease with increasing substitution percentage. Our SBT model, which is restricted to short complementary oligonucleotides that show two-state melting behavior, does not predict a saturation effect and instead predicts that ΔT_m will continue to increase with each successive internal LNA substitution by an amount specific to the type of LNA substituted. To address this discrepancy, we measured ΔT_m for a set of gapmer and LNA-DNA heteroduplexes offering increasing LNA content (Table 9). The data show that when the reduced contribution of terminal LNAs to stability is taken into account, ΔT_m is not saturated as LNA content increases, a result that is predicted well by the SBT model. The SBT model accurately predicts the ΔT_m of gapmer duplexes containing two, three, or six neighboring LNA substitutions. More importantly, we find that the experimental ΔT_m values for the five fully substituted LNA-DNA heteroduplexes are slightly larger than those predicted by the additive SBT model, a result that is completely inconsistent with the idea that ΔT_m is saturated (levels off) with increasing LNA content. The saturation effect noted in a few previous studies therefore appears not to be a general property of LNA-substituted duplexes but may arise in certain specific situations, including terminal LNA substitutions due to their limited contribution to duplex stability.

Finally, we have shown that the use of D and H in place of A and T in a-t base pair rich DNA oligonucleotides may provide an effective strategy for improving the design of more challenging oligonucleotides for which formation of stable hairpins or homodimers can be problematic[6,17] because of the exceptionally high stability of the A-T base pair (Table 10).

AUTHOR INFORMATION

Corresponding Author

*Telephone: (604) 822-5136. Fax: (604) 822-2114. E-mail: israels@chbe.ubc.ca.

Funding Sources

C.A.H. receives salary support as the Canada Research Chair in Interfacial Biotechnology. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

ABBREVIATIONS

LNA, locked nucleic acid (2′-O,4′-C-methylene-β-D-ribofuranosyl nucleotide); ΔC_p, heat capacity change for the helix-to-coil transition; $\Delta H°$, standard state enthalpy change for the helix-to-coil transition; $\Delta S°$, standard state entropy change for the helix-to-coil transition; T_m, melting temperature; ΔT_m, difference in T_m for an LNA-modified duplex and its unmodified isosequential DNA duplex; DSC, differential scanning calorimetry; UVM, UV spectroscopy-monitored melting transition data; NNT, nearest-neighbor thermodynamics; a, t, g, and c, DNA bases adenine, thymine, guanine, and cytosine, respectively; d, DNA-2-amino-adenine; h, DNA-2-thiothymine; A, LNA-adenine; T, LNA-thymine; G, LNA-guanine; C, LNA-cytosine; D, LNA-2-amino-adenine; H, LNA-2-thiothymine; $\Delta\Delta G°_{37}$, $\Delta\Delta H°$, $\Delta\Delta S°$, and ΔT_m, experimental incremental thermodynamic changes for LNA substitutions; $\Delta\Delta G°_{37(pred)}$, $\Delta\Delta H°_{(pred)}$, $\Delta\Delta S°_{(pred)}$, and $\Delta T_{m(pred)}$, predicted incremental thermodynamic changes for LNA substitutions; $\Delta\Delta G°_{37(error)}$, $\Delta\Delta H°_{(error)}$, $\Delta\Delta S°_{(error)}$, and $\Delta T_{m(error)}$, errors (mean error ± standard deviation) in model-predicted thermodynamic values for melting of LNA-substituted duplexes.

REFERENCES

(1) Bergstrom, D. E. (2001) Unnatural nucleosides with unusual base pairing properties. *Current Protocols in Nucleic Acid Chemistry*, Chapter 1, Unit 1, p 4, Wiley, New York.

(2) Kool, E. T. (1997) Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides. *Chem. Rev. 97*, 1473–1488.

(3) Singh, S. K., Koshkin, A. A., Wengel, J., and Nielsen, P. (1998) LNA (locked nucleic acids): Synthesis and high-affinity nucleic acid recognition. *Chem. Commun.*, 455–456.

(4) Koshkin, A. A., Singh, S. K., Nielsen, P., Rajwanshi, V. K., Kumar, R., Meldgaard, M., Olsen, C. E., and Wengel, J. (1998) LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron 54*, 3607–3630.

(5) Obika, S., Nanbu, D., Hari, Y., Morio, K.-i., In, Y., Ishida, T., and Imanishi, T. (1997) Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering. *Tetrahedron Lett. 38*, 8735–8738.

(6) 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.

(7) Obika, S., Nanbu, D., Hari, Y., Andoh, J.-i., Morio, K.-i., Doi, T., and Imanishi, T. (1998) Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2′-O,4′-C-methyleneribonucleosides. *Tetrahedron Lett. 39*, 5401–5404.

(8) Wengel, J., Koshkin, A., Singh, S. K., Nielsen, P., Meldgaard, M., Rajwanshi, V. K., Kumar, R., Skouv, J., Nielsen, C. B., Jacobsen, J. P.,

Jacobsen, N., and Olsen, C. E. (1999) LNA (Locked Nucleic Acid). *Nucleosides, Nucleotides Nucleic Acids 18*, 1365–1370.

(9) 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.

(10) 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.

(11) 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.

(12) Petersen, M., Nielsen, C. B., Nielsen, K. E., Jensen, G. A., Bondensgaard, K., Singh, S. K., Rajwanshi, V. K., Koshkin, A. A., Dahl, B. M., Wengel, J., and Jacobsen, J. P. (2000) The conformations of locked nucleic acids (LNA). *J. Mol. Recognit. 13*, 44–53.

(13) Singh, S. K., and Wengel, J. (1998) Universality of LNA-mediated high-affinity nucleic acid recognition. *Chem. Commun.,* 1247–1248.

(14) 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.

(15) Di Giusto, D. A., and King, G. C. (2004) Strong positional preference in the interaction of LNA oligonucleotides with DNA polymerase and proofreading exonuclease activities: Implications for genotyping assays. *Nucleic Acids Res. 32*, e32.

(16) Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hokfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., Ossipov, M., Koshkin, A., Jakobsen, N., Skouv, J., Oerum, H., Jacobsen, M. H., and Wengel, J. (2000) Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. U.S.A. 97*, 5633–5638.

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

(18) Kaur, H., Babu, B. R., and Maiti, S. (2007) Perspectives on chemistry and therapeutic applications of Locked Nucleic Acid (LNA). *Chem. Rev. 107*, 4672–4697.

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

(20) 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.

(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) 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.

(23) Christensen, U., Jacobsen, N., Rajwanshi, V. K., Wengel, J., and Koch, T. (2001) Stopped-flow kinetics of locked nucleic acid (LNA)-oligonucleotide duplex formation: Studies of LNA-DNA and DNA-DNA interactions. *Biochem. J. 354*, 481–484.

(24) 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.

(25) Chalikian, T. V., Volker, J., Plum, G. E., and Breslauer, K. J. (1999) A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. *Proc. Natl. Acad. Sci. U.S.A. 96*, 7853–7858.

(26) Holbrook, J. A., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1999) Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: Interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry 38*, 8409–8422.

(27) Jelesarov, I., Crane-Robinson, C., and Privalov, P. L. (1999) The energetics of HMG box interactions with DNA: Thermodynamic description of the target DNA duplexes. *J. Mol. Biol. 294*, 981–995.

(28) Mikulecky, P. J., and Feig, A. L. (2006) Heat capacity changes associated with DNA duplex formation: Salt- and sequence-dependent effects. *Biochemistry 45*, 604–616.

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

(30) Tikhomirova, A., Beletskaya, I. V., and Chalikian, T. V. (2006) Stability of DNA duplexes containing GG, CC, AA, and TT mismatches. *Biochemistry 45*, 10563–10571.

(31) Tikhomirova, A., Taulier, N., and Chalikian, T. V. (2004) Energetics of nucleic acid stability: The effect of ΔC_p. *J. Am. Chem. Soc. 126*, 16387–16394.

(32) Wu, P., Nakano, S., and Sugimoto, N. (2002) Temperature dependence of thermodynamic properties for DNA/DNA and RNA/DNA duplex formation. *Eur. J. Biochem. 269*, 2821–2830.

(33) Rouzina, I., and Bloomfield, V. A. (1999) Heat capacity effects on the melting of DNA. 1. General aspects. *Biophys. J. 77*, 3242–3251.

(34) Hughesman, C. B., Turner, R. F. B., and Haynes, C. (2011) Correcting for Heat Capacity and 5′-TA Type Terminal Nearest Neighbors Improves Prediction of DNA Melting Temperatures Using Nearest-Neighbor Thermodynamic Models. *Biochemistry 50*, 2642–2649.

(35) Bruylants, G., Boccongelli, M., Snoussi, K., and Bartik, K. (2009) Comparison of the thermodynamics and base-pair dynamics of a full LNA:DNA duplex and of the isosequential DNA:DNA duplex. *Biochemistry 48*, 8473–8482.

(36) Koshkin, A. A. (2004) Syntheses and base-pairing properties of locked nucleic acid nucleotides containing hypoxanthine, 2,6-diaminopurine, and 2-aminopurine nucleobases. *J. Org. Chem. 69*, 3711–3718.

(37) Lahoud, G., Arar, K., Hou, Y. M., and Gamper, H. (2008) RecA-mediated strand invasion of DNA by oligonucleotides substituted with 2-aminoadenine and 2-thiothymine. *Nucleic Acids Res. 36*, 6806–6815.

(38) Rosenbohm, C., Pedersen, D. S., Frieden, M., Jensen, F. R., Arent, S., Larsen, S., and Koch, T. (2004) LNA guanine and 2,6-diaminopurine. Synthesis, characterization and hybridization properties of LNA 2,6-diaminopurine containing oligonucleotides. *Bioorg. Med. Chem. 12*, 2385–2396.

(39) Compagno, D., Lampe, J. N., Bourget, C., Kutyavin, I. V., Yurchenko, L., Lukhtanov, E. A., Gorn, V. V., Gamper, H. B., Jr., and Toulme, J. J. (1999) Antisense oligonucleotides containing modified bases inhibit in vitro translation of *Leishmania amazonensis* mRNAs by invading the mini-exon hairpin. *J. Biol. Chem. 274*, 8191–8198.

(40) Gamper, H. B., Jr., Arar, K., Gewirtz, A., and Hou, Y. M. (2006) Unrestricted hybridization of oligonucleotides to structure-free DNA. *Biochemistry 45*, 6978–6986.

(41) Kutyavin, I. V., Rhinehart, R. L., Lukhtanov, E. A., Gorn, V. V., Meyer, R. B., Jr., and Gamper, H. B., Jr. (1996) Oligonucleotides containing 2-aminoadenine and 2-thiothymine act as selectively binding complementary agents. *Biochemistry 35*, 11170–11176.

(42) Zhou, Y., Hall, C. K., and Karplus, M. (1999) The calorimetric criterion for a two-state process revisited. *Protein Sci. 8*, 1064–1074.

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

(44) Levin, J. D., Fiala, D., Samala, M. F., Kahn, J. D., and Peterson, R. J. (2006) Position-dependent effects of locked nucleic acid (LNA) on DNA sequencing and PCR primers. *Nucleic Acids Res. 34*, e142.

(45) Latorra, D., Arar, K., and Hurley, J. M. (2003) Design considerations and effects of LNA in PCR primers. *Mol. Cell. Probes 17*, 253–259.

(46) SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science 256*, 217–219.

(47) Searle, M. S., and Williams, D. H. (1993) On the stability of nucleic acid structures in solution: Enthalpy-entropy compensations, internal rotations and reversibility. *Nucleic Acids Res. 21*, 2051–2056.

(48) Chazin, W. J., Rance, M., Chollet, A., and Leupin, W. (1991) Comparative NMR analysis of the decadeoxynucleotide d-(GCATT-AATGC)$_2$ and an analogue containing 2-aminoadenine. *Nucleic Acids Res. 19*, 5507–5513.

(49) Cheong, C., Tinoco, I., Jr., and Chollet, A. (1988) Thermodynamic studies of base pairing involving 2,6-diaminopurine. *Nucleic Acids Res. 16*, 5115–5122.

(50) Gryaznov, S., and Schultz, R. G. (1994) Stabilization of DNA:DNA and DNA:RNA duplexes by substitution of 2′-deoxyadenosine with 2′-deoxy-2-aminoadenosine. *Tetrahedron Lett. 35*, 2489–2492.

(51) Connolly, B. A., and Newman, P. C. (1989) Synthesis and properties of oligonucleotides containing 4-thiothymidine, 5-methyl-2-pyrimidinone-1-β-D-(2′-deoxyriboside) and 2-thiothymidine. *Nucleic Acids Res. 17*, 4957–4974.

(52) Sintim, H. O., and Kool, E. T. (2006) Enhanced base pairing and replication efficiency of thiothymidines, expanded-size variants of thymidine. *J. Am. Chem. Soc. 128*, 396–397.

(53) Sismour, A. M., and Benner, S. A. (2005) The use of thymidine analogs to improve the replication of an extra DNA base pair: A synthetic biological system. *Nucleic Acids Res. 33*, 5640–5666.