BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

A New General Model for Predicting Melting Thermodynamics of Complementary and Mismatched B-Form Duplexes Containing Locked Nucleic Acids: Application to Probe Design for Digital PCR Detection of Somatic Mutations

Curtis Hughesman,[†] Kareem Fakhfakh,[†,‡,§] Roza Bidshahri,[†] H. Louise Lund,[†,∥] and Charles Haynes*[,†,‡,§,∥]

[†]Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada V6T 1Z4

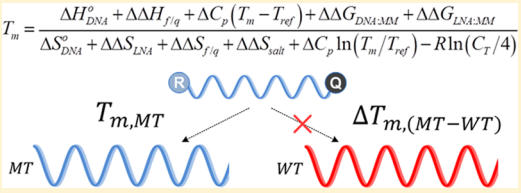
[‡]Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada V6T 1Z3

[§]RES'EAU Water Research Network, Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada V6T 1Z3

[∥]Genome Science and Technology (GSAT) program, University of British Columbia, Vancouver, BC, Canada V5Z 4S6

Supporting Information

ABSTRACT: Advances in real-time polymerase chain reaction (PCR), as well as the emergence of digital PCR (dPCR) and useful modified nucleotide chemistries, including locked nucleic acids (LNAs), have created the potential to improve and expand clinical applications of PCR through their ability to better quantify and differentiate amplification products, but fully realizing this potential will require robust methods for designing dual-labeled hydrolysis probes and predicting their hybridization thermodynamics as a function of their sequence, chemistry, and template complementarity. We present here a nearest-neighbor thermodynamic model that accurately predicts the melting thermodynamics of a short oligonucleotide duplexed either to its perfect complement or to a template containing mismatched base pairs. The model may be applied to pure-DNA duplexes or to duplexes for which one strand contains any number and pattern of LNA substitutions. Perturbations to duplex stability arising from mismatched DNA:DNA or LNA:DNA base pairs are treated at the Gibbs energy level to maintain statistical significance in the regressed model parameters. This approach, when combined with the model's accounting of the temperature dependencies of the melting enthalpy and entropy, permits accurate prediction of T_m values for pure-DNA homoduplexes or LNA-substituted heteroduplexes containing one or two independent mismatched base pairs. Terms accounting for changes in solution conditions and terminal addition of fluorescent dyes and quenchers are then introduced so that the model may be used to accurately predict and thereby tailor the T_m of a pure-DNA or LNA-substituted hydrolysis probe when duplexed either to its perfect-match template or to a template harboring a noncomplementary base. The model, which builds on classic nearest-neighbor thermodynamics, should therefore be of use to clinicians and biologists who require probes that distinguish and quantify two closely related alleles in either a quantitative PCR or dPCR assay. This potential is demonstrated by using the model to design allele-specific probes that completely discriminate and quantify clinically relevant mutant alleles (*BRAF* V600E and *KIT* D816V) in a dPCR assay.



Dual-labeled hydrolysis probes are widely used by clinics in real-time quantitative PCR (qPCR) assays, and more recently in digital PCR (dPCR) assays, to detect germline variations or somatic mutations associated with diseases, including various types of cancer, and thereby facilitate diagnosis and subtyping of disease, setting of treatment regimens, and monitoring of minimal residual disease levels following treatment.[1] These assays are often designed to detect a specific somatic mutation present in a rare population of abnormal or tumor cells dispersed within a tissue matrix abundant in normal cells carrying the homologous wild type (germline) allele.

The use of qPCR to distinguish between two alleles that differ greatly in abundance while differing in sequence by as little as a single base pair often requires the engineering of a probe that displays high affinity and specificity for the low-abundance mutant (MT) allele and little to no affinity for the closely related wild type (WT) gene at the qPCR annealing temperature (T_a). This in turn requires a sufficiently large difference (Δ) between $T_{m,MT}$, the melting temperature (T_m) of the perfectly matched duplex formed by the probe and its MT target, and $T_{m,WT}$, the T_m for the mismatched duplex formed by the probe and the more abundant WT gene. The

Received: July 24, 2014
Revised: January 14, 2015
Published: January 21, 2015

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

thermodynamics of bimolecular complexation equilibria provide the necessary conditions for achieving this goal:

$$\Delta T_{m,(MT-WT)} = T_{m,MT} - T_{m,WT} = \frac{\Delta H_{MT}}{\Delta S_{MT} - R \ln(C_{T,MT}/2)} - \frac{\Delta H_{WT}}{\Delta S_{WT} - R \ln(C_{T,WT}/2)} \quad (1)$$

where ΔH and ΔS are the enthalpy change and entropy change, respectively, for the helix-to-coil transition of a duplex formed from two single strands, each of which is non-self-complementary. ΔH and ΔS are both positive in value for the probe:MT or probe:WT duplex denaturation reaction; if the heat capacity change (ΔC_p) for the denaturation reaction is ignored, ΔH and ΔS may also be taken to be independent of temperature. C_T is the total strand concentration, probe plus template, and both $C_{T,MT}$ and $C_{T,WT}$ in eq 1 are equal to C_{probe} in assays of this type because the probe is present in great excess relative to either template at the start of the reaction. The concentration (C_T) is divided by 2 in eq 1 because the large excess of probe renders the hybridization reaction pseudo-first-order. As a result, $R \ln(C_{T,MT}/2)$ and $R \ln(C_{T,WT}/2)$ can, for the sake of simplicity, be denoted by a constant C that is set by C_{probe}. Achieving an acceptable $\Delta T_{m,(MT-WT)}$ therefore requires a probe for which $\Delta H_{MT}/(\Delta S_{MT} + C) > T_a > \Delta H_{WT}/(\Delta S_{WT} + C)$. This thermodynamic criterion is difficult to realize, as melting transitions occur across a range of temperatures and $\Delta T_{m,(MT-WT)}$ is often less than 6 °C, in part because of the relatively small proportional contribution of a single base pair to the overall thermodynamic stability of a duplex formed with a standard 20−25 bp pure-DNA probe. Current strategies for realizing a suitable $\Delta T_{m,(MT-WT)}$ therefore involve modifying oligonucleotides used as hydrolysis probes either by adding a 3′-terminal minor groove binder (MGB) ligand[2] or by replacing nucleotides within the probe with their corresponding locked nucleic acid (LNA) analogue.[3,4] The latter approach can increase the difference in MT and WT duplex stabilities through the ability of LNA substitutions to decrease probe length while maintaining $T_{m,MT}$,[4,5] and the generally greater energetic intolerance of LNAs to participate favorably in mismatched base pairs.[6] LNA-bearing probes have therefore been shown to offer specific advantages when applied to the detection of either germline variants, including single-nucleotide polymorphisms (SNPs),[4,7,8] or acquired variants, including somatic point mutations (SPMs).[9−11]

MGB- and LNA-modified probes can demonstrate comparable performance in qPCR genotyping assays.[3] For example, a 15-mer probe coupled to an MGB linker[12,13] and a 16-mer LNA-bearing[9] probe provide an equivalent limit of detection (LOD) of 1−2% (relative to the abundance of the WT allele in the specimen) when applied to the detection of the V617F somatic mutation within the human *JAK2* gene. The *JAK2* V617F mutation is often found in Philadelphia chromosome negative myeloproliferative neoplasm (MPN) patients,[14] including at high incidence in patients with polycythemia vera (PV), essential thrombocythemia (ET), or idiopathic myelofibrosis (IMF). Most cancer clinics therefore test for *JAK2* V617F on human chromosome 9 to identify MPN and leukemia patients and define appropriate treatment regimens.[15] Improving V617F mutational analysis could therefore provide for early or more sensitive diagnoses of PV, as well as related MPDs and leukemias.

Guidelines for designing probes bearing one or more LNA substitutions have been proposed.[6] Recently, we reported a molecular thermodynamic model that accurately predicts melting temperatures and thermodynamic changes under standard thermodynamic conditions (pH 7, 1 M NaCl) for short perfectly complementary duplexes containing any combination of LNA substitutions in one strand.[16] Other models have recently been proposed for this purpose, as well.[17−19] Here, we define differences between these models and when they become important and then show how our model may be extended to introduce two additional capabilities useful for designing probes bearing one or more LNA substitutions: namely, the ability to accurately predict under PCR solution conditions (1) the T_m value for any complementary duplex formed between a dual-labeled LNA-substituted probe and its target allele and (2) the value of $\Delta T_{m,(MT-WT)}$, which is needed to ensure that the probe hybridizes only to its target allele during qPCR amplification. The intent is to establish a simple, reliable method for accurately predicting T_m and $\Delta T_{m,(MT-WT)}$ for pure-DNA and LNA-substituted hydrolysis probes operating under standard and PCR-relevant conditions.

Melting thermodynamics data collected by thermal-ramp ultraviolet "melt" spectroscopy (UVM) are reported for a large set of pure-DNA and LNA-containing oligonucleotides that are either perfectly complementary to their duplexed template or form a mismatched DNA:DNA or LNA:DNA base pair with the template. The data are used to regress the model parameters needed to accurately predict from sequence information the melting thermodynamics of complementary or mismatched pure-DNA or LNA-substituted duplexes under standard thermodynamic conditions. In addition, we report UVM data for a large set of LNA-substituted oligonucleotides bearing a 5′-fluorescence reporter dye (e.g., FAM or HEX) and a 3′-fluorescence quencher [e.g., Black-Hole Quencher 1 (BHQ1)] duplexed either to their perfect complement or to a template containing a single mismatch. Those data are used to extend the model to allow accurate prediction of $\Delta T_{m,(MT-WT)}$ under PCR-relevant conditions.

We show the resulting model can be used to define LNA substitution patterns on a probe, as well as probe positions on a template, that create a $T_{m,MT}$ and $\Delta T_{m(MT-WT)}$ that permit efficient hybridization to a target MT allele while limiting probe duplexation with a nontarget WT allele under PCR conditions. This novel model-driven design strategy demands accurate prediction of hybridization thermodynamics, and we therefore validate results from our model against experiment for a large set of model-designed probes. Model-designed unlabeled and dual-labeled hydrolysis probes offering large $\Delta T_{m,(MT-WT)}$ values are created against a number of somatic point mutation "hot spots" for various genes (*BRAF*, *EGFR*, *JAK2*, *KIT* *PIK3CA*, *PTEN*, and *TP53*) that are clinically relevant in cancer monitoring and treatment. Finally, application of model-refined probes against *BRAF* V600E and *KIT* D816V in a droplet digital PCR (ddPCR) format is shown to result in complete segregation of data fields to permit quantitative detection of the mutant allele.

MATERIALS AND METHODS

Oligonucleotides. All DNA and LNA oligonucleotides used as templates or primers in UV melting (UVM)

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

experiments or in droplet digital PCR (ddPCR) assays were synthesized and purified through standard desalting by either Proligo/Sigma-Aldrich (St. Louis, MO) or Integrated DNA Technologies (IDT, Coralville, IA). All unlabeled and dual-labeled hydrolysis probes were purchased from either Exiqon Inc. (Vedbæk, Denmark) or IDT and purified by high-performance liquid chromatography (HPLC). Mini-gene plasmids were synthesized by IDT and used to represent the *BRAF* WT or V600E and *KIT* WT or D816V alleles. *BRAF* plasmids were ampicillin resistant and also of pUC origin and included a 280 bp fragment of *BRAF* spanning from intron 14 to 15, thru exon 15, and into introns 15 and 16. The *KIT* plasmids were kanamycin resistant and of pUC origin and contained a 335 bp segment of *KIT* spanning portions of introns 16 to 17 and 17 to 18, and all of exon 17.

Primer sets used for amplification of the *BRAF* WT and V600E plasmids included the forward primer (FP) 5′-ctactgttttccttttacttactacacctcaga-3′ and reverse primer (RP) 5′-agcctcaattcttaccatcca-3′, while those used for *KIT* WT and D816V plasmids included FP 5′-ctcctccaacctaatagtgtattcacag-3′ and (RP) 5′-gcagagaatgggtactcacg-3′. Both forward and reverse primer sets were designed to hybridize to appropriate intron/exon boundary regions within a mini-gene template sequence of the target gene displayed on a plasmid. Amplicon lengths for *BRAF* and *KIT* were 164 and 158 bp, respectively.

Monitoring Helix-to-Coil Transitions via UVM Spectroscopy. UVM profiles for the thermally induced melting of complementary or mismatched duplexes composed of either pure DNA or DNA containing LNA substitutions in one strand were collected at an absorbance wavelength of 260 nm using a Varian Inc. (Santa Clara, CA) Cary 1E spectrophotometer equipped with a 12-cell Peltier temperature controller. The two strands were combined at an equimolar concentration to reach a total strand concentration (C_T) sufficient for UVM-based melt analysis. For melting experiments conducted under standard thermodynamic solution conditions, perfect match or mismatched duplexes were resuspended to a C_T of 2, 50, or 100 μM in 10 mM Na_2HPO_4 (pH 7.0) buffer containing 1 M NaCl and 1 mM Na_2EDTA. UVM data were also collected under PCR conditions in a buffer containing New England Biolabs (Ipswich, MA) 1× Taq standard reaction buffer and $MgCl_2$ to form a solution with final $MgCl_2$ and KCl concentrations of 3 and 50 mM, respectively. UVM experiments involving dual-labeled hydrolysis probes and their corresponding template were performed at a C_T of 2, 5, or 10 μM in this buffer.

Duplex samples were annealed prior to UVM data collection by preheating the sample to a temperature between 75 and 95 °C, holding it for 3 min, and slowly cooling it to a temperature between 15 and 25 °C. The equilibrated samples were then loaded into quartz cuvettes with path lengths of 10 or 1 mm and sealed with fitted caps to prevent evaporation during heating. UVM experiments were conducted by collecting absorbance readings at 260 nm (A_{260}) every 0.5 °C as the temperature was increased from 5 to 95 °C at a rate of 0.5 °C/min to create a raw absorbance (A_{260}) versus temperature (T) profile, also known as a melt profile.

Raw UVM data were exported into Microsoft Excel and, in cases where a complete melting transition was recorded, used to compute the melting temperature and duplex-to-random coil state transition thermodynamics according to the procedures described by Hughesman et al.[16]

T_m values for the mismatched duplexes were often too low (<30 °C) to permit collection of a complete melting transition by the UVM method. In these cases, the T_{max} value, which could be determined accurately from the maximum of the dA_{260}/dT data obtained, was instead recorded.

Regression of $\Delta\Delta G_{LNA:MM}$ Parameters. Incremental $\Delta\Delta G_{LNA:MM_i}$ model parameters, which define the change in Gibbs energy of a duplex resulting from the introduction of mismatched LNA:DNA base pairs, were determined through model regression to UVM-derived T_{max} and ΔT_{max} data for those members among a set of 192 9-mer duplexes for which both values could be measured accurately (Table S1 of the Supporting Information). Although $T_{max} \neq T_m$, ΔT_{max} and $\Delta\Delta T_{max}$ values provide very good estimates of ΔT_m and $\Delta\Delta T_m$, respectively, and were applied in that manner.[20,21] $\Delta\Delta G_{LNA:MM_i}$ parameters could then be determined through standard χ^2 minimization based on model regression to the complete set of $\Delta\Delta T_m$ data

$$\chi^2 = \sum_i [\Delta\Delta T_{m(Pred)_i} - \Delta\Delta T_{m(Expt)_i}]^2 \quad (2)$$

where

$$\Delta\Delta T_{m_i} = \Delta T_{m(LNA-LNA:MM)_i} - \Delta T_{m(DNA-DNA:MM)_i}$$
$$= [T_{m(LNA)} - T_{m(LNA:MM)}]_i - [T_{m(DNA)} - T_{m(DNA:MM)}]_i \quad (3)$$

$\Delta\Delta T_{m(Expt)_i}$ was determined by collecting $\Delta\Delta T_{m(LNA-LNA:MM)_i}$ and $\Delta T_{m(DNA-DNA:MM)_i}$ data for each duplex sequence set *i*. The mean experimental errors for each ΔT_m value were determined from two upscans and two downscans and are reported in Table S1 of the Supporting Information. $\Delta\Delta T_{m(Pred)}$ was determined by using the specified model to calculate values for $T_{m(LNA)}$, $T_{m(LNA:MM)}$, $T_{m(DNA)}$, and $T_{m(DNA:MM)}$. The model developed in this work cannot solve for T_m explicitly (see below). A Newton−Raphson type iteration implemented in Excel was therefore applied. Convergence of the model-predicted T_m value (generally within 0.1 °C) was typically reached in two iterations, permitting straightforward determination of $\Delta\Delta G_{LNA:MM_i}$ parameters using the method described above.

Droplet Digital PCR Detection of *BRAF* V600E or *KIT* D816V. Standard qPCR assays directed against an SPM-bearing MT allele typically require two competing dual-labeled hydrolysis probes, with each targeting either the WT or MT allele, to both reduce cross hybridization effects and permit quantification of relative abundances.[22] In principle, however, the need for competing probes to detect rare MT alleles may be relaxed in dPCR because of the ability to partition WT and MT alleles into separate droplets.[23] Quantification of mutation frequencies may then be achieved using a FAM-labeled probe against the mutant allele that spans the site of variance. Within the same reaction, total copies of all forms of the target gene, MT + WT, may be quantified using a HEX-labeled consensus probe designed to hybridize to a highly conserved region within an exon of the gene. One then counts those droplets containing an end-point HEX amplification signal to compute the total number of copies of the target gene in the sample and then counts the segregated cluster of droplets containing both an end-point HEX and FAM signal to compute the mutation frequency.

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

To test this concept, plasmids harboring the *BRAF* or *KIT* template were linearized using restriction enzyme *Cla*I or *Sph*I, respectively, at a single junction outside the amplicon region prior to ddPCR amplification. The assay workflow followed methods described previously.[24] A 20 μL PCR mixture consisting of 10 μL of 2× dPCR Supermix for Probes (no dUTP) (Bio-Rad Inc., Hercules, CA), 2 μL of 10× primer and probe mix (to yield final concentrations of 900 nM primer and 200−250 nM probe), and 8 μL of template (to yield on average 300 *BRAF* copies or 100 *KIT* copies per microliter) was loaded into each sample well of an eight-channel disposable droplet generation cartridge (Bio-Rad); 60 μL of QX100 Droplet Generation Oil (Bio-Rad) was then loaded into the corresponding oil well. A droplet-containing emulsion was generated according to the manufacturer's instructions (Bio-Rad), and 40 μL droplets in excess oil were transferred manually into a 96-well PCR plate (Eppendorf, Hamburg, Germany). The plate was heat sealed (Eppendorf) with foil and amplified to end point in a CFX96 thermocycler (Bio-Rad). Thermocycling conditions included activation at 95 °C for 10 min (one cycle), denaturation at 94 °C for 30 s followed by annealing and/or extension at 60 °C for 60 s (50 cycles), and finally inactivation at 98 °C for 10 min. Individual droplet end-point fluorescence signals in channel 1 (FAM) and channel 2 (HEX) were read using a QX100 Droplet Reader (Bio-Rad) and then analyzed with Quantasoft analysis software (Bio-Rad).

THEORY

Predicting DNA Melting Thermodynamics Using a Nearest-Neighbor Type Model. Measurements of the melting thermodynamics of either complementary DNA duplexes or LNA/DNA heteroduplexes are typically conducted in a standard thermodynamic solution comprised of pH 7 background buffer and 1 M NaCl.[16,18,19,25] Conversely, solutions employed in PCR detection of somatic mutations typically contain 20−100 mM NaCl or KCl. In addition, magnesium (Mg^{2+}) is generally added to a concentration of 1−6 mM in PCR master mixes as a cofactor for Taq DNA polymerase. Both of these departures from the standard solution condition at which thermodynamic data are most often acquired affect duplex stability, as do other solution additives employed in amplification reactions. In particular, deoxyribonucleotide triphosphates (dNTPs) bind excess Mg^{2+}, while dimethyl sulfoxide (DMSO), which is added as a cosolvent to facilitate amplification of difficult templates, decreases T_m.[26] Finally, chemical modifications to oligonucleotides, including the addition of a reporter dye and a fluorescence quencher, affect T_m, as well.[27] A model capable of predicting probe properties under PCR-relevant conditions will therefore have to account for these various effects.

Recently, we reported a new model[16] that accurately predicts T_m values and melting thermodynamics for both complementary DNA homoduplexes and LNA/DNA heteroduplexes under standard conditions (pH 7, 1 M NaCl). Other models have also been described for this purpose.[17−19] They all compute the T_m of a short (<30 bp) complementary B-form duplex using the standard thermodynamic relation

$$T_m = \frac{\Delta H}{\Delta S - R \ln(C_T/4)} \tag{4}$$

where C_T is now divided by 4 because the two complementary strands are equimolar in concentration and the reaction must therefore be treated as bimolecular. Differences between these various models arise primarily through the manner in which the thermodynamic changes, ΔH and ΔS, for the helix-to-coil transition are computed. It therefore is useful to segregate the available models into those that assume heat capacity change ΔC_p for the melting transition is negligible (in part because accurate values for ΔC_p have only recently been acquired) and those that account for the temperature dependence of ΔH and ΔS through a non-zero ΔC_p. While either type of model can be extended to provide predictions of $\Delta T_{m(MT-WT)}$ for probe design, the approach, parametrization, and output of the two approaches differ as described below. Before proceeding to that discussion, we note that a feature common to these models is that they are generally applicable to melting thermodynamics of the canonical B-form helical structure and may not provide accurate predictions for duplexes containing DNA sequences known to have a propensity for alternative conformations such as B*-form and A-form helices associated with A-tract and G-tract sequence motifs, respectively.[28]

Predicting T_m Values of LNA/DNA Heteroduplexes Using a ΔC_p = 0 Type Model. Though other models have been described,[17,18] Owczarzy et al.[19] have arguably developed the most advanced zero-ΔC_p type model for predicting from sequence information the T_m of any short perfectly complementary duplex containing LNAs on one of the strands. Hereafter termed a type 1 model, it is based on the widely used unified nearest-neighbor thermodynamic (NNT) model of SantaLucia[29] and reduces to it when applied to pure-DNA duplexes. In type 1 models,[19,29] temperature-independent ΔH and ΔS values for complementary DNA/LNA heteroduplexes in standard thermodynamic solutions are computed as

$$\Delta H = \sum_{i=1}^{2} n_i \Delta H_{EE_i} + \sum_{j=1}^{10} n_j \Delta H_{DNA_j} + \sum_{k=1}^{48} n_k \Delta\Delta H_{LNA_k} \tag{5}$$

$$\Delta S = \Delta S^{\circ}_{sym} + \sum_{i=1}^{2} n_i \Delta S_{EE_i} + \sum_{j=1}^{10} n_j \Delta S_{DNA_j} + \sum_{k=1}^{48} n_k \Delta\Delta S_{LNA_k} \tag{6}$$

where ΔS°_{sym} is an exact result of statistical mechanics that applies only to self-complementary sequences and accounts for the fact that a bimolecular complex formed from two self-complementary strands has a 2-fold rotational symmetry that is not present in either of the single strands or any duplex formed from non-self-complementary strands.[30,31] In either equation, the first summation accounts for the unique energetics of terminal base pairs (end effects), with n_i specifying the number of terminal base pairs of type i in the duplex. The second summation accounts for the contribution to the melting energetics of the isosequential DNA duplex of each unique Watson−Crick nearest-neighbor base pair doublet of type j. There are 10 such unique doublets, and n_j counts the number of each in the duplex. The requirement for nearest-neighbor enthalpy (ΔH_{DNA_j}) and entropy (ΔS_{DNA_j}) parameters for each doublet is firmly supported by theory, most notably modified forms of the Poland−Sheraga Hamiltonian (see, for example, ref 32) that partition base pair formation and stacking interactions at the nearest-neighbor level to account for the unique nonlinear contributions of these two interactions to

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

duplex DNA stability. The third summation is more empirical in nature. It assumes that perturbations to duplex stability resulting from a DNA to LNA substitution also occur at the nearest-neighbor level, where n_k specifies the number of each of the 48 possible types of LNA-containing doublets that may be present in a duplex when LNA substitutions are restricted to only one of the two strands. The form of the summation therefore tacitly assumes that an LNA substitution perturbs not only base pair energetics but also base stacking interactions in a sequence-sensitive manner. While this assumption is certainly in keeping with the nearest-neighbor concept, it presents challenges in terms of both the physical meaning of the combined 96 new parameters (48 $\Delta H_{\mathrm{LNA}_k}$ and 48 $\Delta S_{\mathrm{LNA}_k}$ parameters) introduced into the model and the ability to regress a unique and reliable value for each of those parameters. Despite these concerns, type 1 models[19] are on average quite accurate in their T_m predictions for LNA/DNA heteroduplexes, as is the underlying unified NNT model of SantaLucia[29] when it is applied to pure-DNA homoduplexes.

Predicting T_m Values of LNA/DNA Heteroduplexes Using a Model Based on a Non-Zero ΔC_p. T_m values for complementary DNA homoduplexes and LNA/DNA heteroduplexes may also be predicted using a model that accounts for the temperature dependence of $\Delta H°$ and $\Delta S°$. The first such "type 2" model was described by Hughesman et al.[16] It accurately predicts melting thermodynamics of complementary duplexes containing any pattern of LNA substitutions on one of the strands. When applied to DNA homoduplexes under standard solution conditions, it differs from the unified NNT model by accounting for both a non-zero ΔC_p and the unique base pairing energetics of terminal 5′-ta groups. Temperature-dependent ΔH and ΔS values for either pure-DNA homoduplexes or DNA/LNA heteroduplexes are computed as

$$\begin{aligned}\Delta H(T_m) &= \sum_{i=1}^{2} n_i \Delta H^{\circ}_{\mathrm{EE}_i} + \lambda \Delta H^{\circ}_{5'\text{-ta}} \\ &\quad + \sum_{j=1}^{10} n_j \Delta H^{\circ}_{\mathrm{DNA}_j} + \Delta\Delta H_{\mathrm{LNA}} \\ &\quad + \Delta C_p (T_m - T_{\mathrm{ref}}) \\ &= \Delta H^{\circ}_{\mathrm{DNA}} + \Delta\Delta H_{\mathrm{LNA}} + \Delta C_p (T_m - T_{\mathrm{ref}}) \end{aligned} \tag{7}$$

$$\begin{aligned}\Delta S(T_m) &= \Delta S^{\circ}_{\mathrm{sym}} + \sum_{i=1}^{2} n_i \Delta S^{\circ}_{\mathrm{EE}_i} + \lambda \Delta S^{\circ}_{5'\text{-ta}} \\ &\quad + \sum_{j=1}^{10} n_j \Delta S^{\circ}_{\mathrm{DNA}_j} + \Delta\Delta S_{\mathrm{LNA}} + \Delta C_p \ln\left(\frac{T_m}{T_{\mathrm{ref}}}\right) \\ &= \Delta S^{\circ}_{\mathrm{DNA}} + \Delta\Delta S_{\mathrm{LNA}} + \Delta C_p \ln(T_m/T_{\mathrm{ref}}) \end{aligned} \tag{8}$$

where the parameters $\Delta H^{\circ}_{5'\text{-ta}}$ and $\Delta S^{\circ}_{5'\text{-ta}}$ correct the melting enthalpy and entropy, respectively, for each of the λ (λ = 0, 1, or 2) 5′-ta terminal nearest-neighbor doublets. When the model is applied to probe design, those terminal 5′-ta corrections are small and may be neglected. As a result, type 2 models primarily distinguish themselves from their type 1 counterpart through terms defining the temperature dependencies of ΔH and ΔS, as well as the manner in which energetic changes arising from LNA substitutions are treated. In eqs 7 and 8, all pure-DNA parameters, including the terminal base pair parameters ($\Delta H^{\circ}_{\mathrm{EE}_i}$ and $\Delta S^{\circ}_{\mathrm{EE}_i}$) and the NNT parameters ($\Delta H^{\circ}_{\mathrm{DNA}_j}$ and $\Delta S^{\circ}_{\mathrm{DNA}_j}$), are specific to a reference temperature, T_{ref}, which is taken to be 53 °C for reasons explained previously.[33] The values of ΔH and ΔS at the T_m of the duplex are then computed from a defined reference state (deg) of the isosequential pure-DNA duplex at T_{ref} and under standard thermodynamic solution conditions using perturbation terms for LNA substitutions ($\Delta\Delta H_{\mathrm{LNA}}$ and $\Delta\Delta S_{\mathrm{LNA}}$), as well as two temperature-dependent terms, where ΔC_p is given by $n\Delta C_p^{\mathrm{bp}}$; here, n is the number of base pairs in the duplex, and ΔC_p^{bp} is 42 cal mol^{-1} K^{-1} bp^{-1}.[33]

The perturbation term $\Delta\Delta S_{\mathrm{LNA}}$ is computed as

$$\Delta\Delta S_{\mathrm{LNA}} = \sum_{i=1}^{4} n_i \Delta\Delta S_{\mathrm{LNA}_i} \tag{9}$$

where n_i is the number of LNA/DNA base pairs of type i. The four required $\Delta\Delta S_{\mathrm{LNA}_i}$ parameters have been previously reported[16] and account for the incremental change in entropy for a helix-to-coil transition resulting from substitution of a given DNA nucleotide with its corresponding LNA nucleotide. They were determined by model regression to calorimetry-derived data sets (ΔG, ΔH, ΔS, and T_m) for 43 LNA/DNA heteroduplexes and their corresponding isosequential DNA homoduplexes. Taking into account their standard errors, those data sets indicate that changes in ΔS resulting from an LNA substitution are statistically significant at the individual base pair level, but not at the nearest-neighbor level. The data therefore indicate that, within experimental certainty, incremental changes in stacking interactions resulting from an LNA substitution do not depend on the type of nearest-neighbor base that stacks with the substituted LNA. For unlabeled oligonucleotides, LNA substitutions at terminal positions are generally ignored in calculating $\Delta\Delta S_{\mathrm{LNA}}$, in accordance with our data[16] and those of others[34−36] that show that an LNA substitution at a terminal position does not result in a significant change in duplex stability. However, in the case of labeled oligonucleotides such as hydrolysis probes, LNA substitutions made at the 5′- and/or 3′-end of the sequence are no longer considered terminal, as they are flanked by either a 5′-fluorescent dye or 3′-quencher; they are therefore included in calculating $\Delta\Delta S_{\mathrm{LNA}}$.

$\Delta\Delta H_{\mathrm{LNA}}$ is calculated from $\Delta\Delta H_{\mathrm{LNA}_i}$ parameters in an analogous manner. However, model regression to the 43-member UVM-derived data set yielded $\Delta\Delta H_{\mathrm{LNA}_i}$ parameters that are statistically insignificant, indicating that, within experimental certainty, $\Delta\Delta H_{\mathrm{LNA}_i}$ for all i may be assumed to carry zero value. Thus, if the temperature dependencies of ΔH and ΔS are taken into account through a non-zero ΔC_p, we find that the change in duplex stability arising from an LNA substitution is entropically driven.[16]

RESULTS AND DISCUSSION

Comparing Basic Properties of the Two Types of Models. Type 1 (zero ΔC_p) and type 2 (non-zero ΔC_p) models both provide accurate predictions of T_m for short complementary pure-DNA duplexes and LNA/DNA heteroduplexes. The difference is in how they achieve those results, in particular, the number of parameters they require (Table 1) and the values of ΔH and ΔS they compute to predict T_m using eq 4. To predict melting thermodynamics of complementary pure-

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

Table 1. Comparison of Parameters Required in Type 1 and Type 2 Models When They Are Applied to T_m Calculations for Complementary Pure-DNA Homoduplexes, Pure-DNA Homoduplexes Containing a Mismatched Base Pair, Complementary LNA:DNA Heteroduplexes, or LNA:DNA Heteroduplexes Containing a Mismatched Base Pair

duplex type	parameter	type 1 model	type 2 model
pure-DNA homoduplexes[a]	ΔC_p		1
	ΔS°_{sym}	1	1
	ΔH_{EE_i} or $\Delta H^\circ_{EE_i}$	2	2
	ΔS_{EE_i} or $\Delta S^\circ_{EE_i}$	2	2
mismatched DNA duplex	ΔH_{DNA_i} or $\Delta H^\circ_{DNA_i}$	10	10
	ΔS_{DNA_i} or $\Delta S^\circ_{DNA_i}$	10	10
	$\Delta H_{DNA:MM_i}$	48	
	$\Delta S_{DNA:MM_i}$	48	
	$\Delta G^\circ_{DNA_i}$		10
	$\Delta G^\circ_{DNA:MM_i}$		48
total no. of parameters for DNA duplexes		121	84
LNA:DNA heteroduplexes	$\Delta\Delta H_{LNA_i}$	48	
	$\Delta\Delta S_{LNA_i}$ *or* $\Delta\Delta S^\circ_{LNA_i}$	48	4
mismatched LNA:DNA duplex	$\Delta\Delta H_{LNA:MM_i}$	96	
	$\Delta\Delta S_{LNA:MM_i}$	96	
	$\Delta\Delta G^\circ_{LNA:MM_i}$		12
	$\Delta\Delta G^\circ_{3'NN\text{-}LNA}$/ $\Delta\Delta G^\circ_{5'NN\text{-}LNA}$		2
total no. of model parameters		409	102

[a]Parameters used to correct for 5′-ta sequences have not been included in this comparison as the correction is insignificant when these models are applied to dual-labeled hydrolysis probes.

DNA homoduplexes, the two model types require very similar parametrization as both are based on the unified NNT model of SantaLucia.[29] The type 2 model[16] utilizes one additional parameter, ΔC_p, which is required to compute the temperature dependencies of ΔH and ΔS. One consequence of this is illustrated in Table 2, which compares thermodynamic properties of an internal cc/gg doublet (within a pure-DNA homoduplex) computed as a function of duplex melting temperature using each type of model. In type 1 models, the enthalpy of the doublet in its duplexed state relative to that in its denatured state, $\Delta H_{cc/gg}$, is taken to be temperature-independent. The melting entropy change for the doublet, $\Delta S_{cc/gg}$, is likewise invariate with temperature. This is, however, physically implausable, which was demonstrated by Searle and Williams,[37] who combined single-molecule stretching data with Zimm−Bragg theory[38] to find that the two-state transition describing melting of short duplex DNA can be described only if the melting entropy of each doublet changes with temperature in accordance with a non-zero, positive-in-value ΔC_p.

Indeed, there are now compelling data from multiple independent studies, most notably calorimetry studies, showing that ΔC_p is non-zero and positive in value.[16,33,39−46] For short complementary duplexes melting above 45 °C, measured $\Delta C_p{}^{bp}$ values lie between 30 and 60 cal K^{-1} (mol of bp)$^{-1}$.[44] The positive ΔC_p is thought to arise in part because the denaturation of each antiparallel strand of a double helix proceeds via an overstretching transition that includes net changes in strand solvation.[37] Excellent correlations between changes in solvent-exposed surface area and the value have been described, most notably for proteins (for a review, see ref 47). Bases that remain stacked in the single strand following duplex denaturation are likely to produce relatively little net change in polar and nonpolar surface hydration, and thus little change in C_p, but those bases that do unstack undergo a concominant change in their states of hydration, contributing to a positive ΔC_p. Schwarz's lab[48,49] and Mrevlishvili et al.[50] have independently shown that single-stranded DNA exhibits broad changes in stacking structure with temperature, and that the duplex DNA denaturation reaction is accompanied by changes in hydration. Their findings align with a particularly elegant study by Spink and Chaires,[51] who combined UVM, circular dichroism, and vapor-pressure osmometry experiments to estimate that, on average, the melting of poly(A)-poly(T) duplex DNA results in the net exchange of four water molecules per base pair. Other contributions to ΔC_p are thought to include changes in the conformations of bases and strands resulting from counterion condensation and other electrostatic effects.[52] Though questions remain regarding the precise magnitude of these various contributions,[53] current theory and data both indicate that unstacking events provide important, if not dominant, contributions to ΔC_p, as summarized in the excellent review by Mikulecky and Feig.[44] Type 2 models exploit these findings through inclusion of a non-zero ΔC_p; ΔH and ΔS are therefore predicted to be dependent on temperature (Table 2), and we find that the inclusion of ΔC_p terms is most valuable when one is specifically concerned with accurate prediction of ΔH and ΔS.

This is because ΔG and T_m are less dependent on ΔC_p than are ΔH and ΔS. As a result, NNT models that assume zero ΔC_p accurately predict T_m values for pure-DNA duplexes, as was shown by SantaLucia in developing the unified NNT model.[29] Moreover, the two model types are equally accurate in their ability to predict T_m values for short LNA-containing heteroduplexes. However, by accounting for the temperature dependence of ΔH and ΔS, the type 2 model requires only four parameters (the set of $\Delta\Delta S_{LNA_i}$ values) to fully describe perturbations to duplex energetics arising from LNA substitutions (see Table 1).[16] A total of 96 parameters are required

Table 2. Comparison of Melting Thermodynamic Values under Standard Thermodynamic Conditions for the cc/gg Nearest-Neighbor Doublet Computed Using a Type 1 or Type 2 Model

	type 1 model ($\Delta C_p = 0$)			type 2 model ($\Delta C_p \neq 0$)[a]		
T (°C)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal mol^{-1} K^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal mol^{-1} K^{-1})
37	1.83	8.0	19.9	1.79	6.7	15.7
53	1.51			1.51	8.0	19.9
70	1.17			1.13	9.4	24.2

[a]Type 2 model calculations set ΔC_p = 42 cal (mol of bp)$^{-1}$ K^{-1} and T_{ref} = 53 °C.

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

for this purpose in type 1 models.[18,19] At the very least, this shows that equivalent predictive power can be captured in a much smaller set of model parameters when the temperature dependencies of ΔH and ΔS are addressed. More importantly, it argues that models of type 2, due to their significantly lower level of parametrization (see Table 1), are well suited for extension to permit quantitative prediction of $\Delta T_{m,(MT-WT)}$ values and the model-based design of dual-labeled hydrolysis probes for qPCR and dPCR genotyping assays.

Type 2 Model Extension To Predict Changes in Pure-DNA Duplex Stability Due to Mismatched Base Pairs, Added End Groups, and Changes in Solution Conditions. Equation 4 is derived from the fundamental relation for ΔG, the Gibbs energy change for the helix-to-coil transition of a complementary pure-DNA duplex at a total strand concentration C_T. That Gibbs energy function can be extended to describe helix-to-coil transition energies for either DNA homoduplexes or LNA heteroduplexes bearing a mismatched base pair, a terminal fluorophore, and/or a fluorescence quenching agent, as well as for any such duplex solubilized in a nonstandard solution, through the introduction of perturbation ($\Delta\Delta$) terms at the Gibbs energy level:

$$T_m = \left[\Delta H^\circ_{DNA} + \Delta\Delta H_{f/q} + \Delta C_p(T_m - T_{ref}) + \Delta\Delta G_{DNA:MM} + \Delta\Delta G_{LNA:MM}\right] / \left[\Delta S^\circ_{DNA} + \Delta\Delta S_{LNA} + \Delta\Delta S_{f/q} + \Delta\Delta S_{salt} + \Delta C_p \ln\left(\frac{T_m}{T_{ref}}\right) + R\ln\left(\frac{C_T}{4}\right)\right] \quad (10)$$

Computed as indicated in eqs 10 and 11, ΔH°_{DNA} and ΔS°_{DNA} apply to the fully complementary pure-DNA duplex under standard state conditions (T_{ref}, standard thermodynamic solution conditions). To this reference, a perturbation term ($\Delta\Delta H_{f/q}$) is applied to correct ΔH for the addition of a terminal fluorophore and fluorescence quenching agent. Perturbation terms are also applied to the melting entropy to correct for terminal fluorescence/quenching agents ($\Delta\Delta S_{f/q}$), LNA substitutions ($\Delta\Delta S_{LNA}$), and salts ($\Delta\Delta S_{salt}$). Values for $\Delta\Delta H_{f/q}$ and $\Delta\Delta S_{f/q}$ were reported by Moreira et al.,[27] while $\Delta\Delta S_{LNA_i}$ values needed to compute $\Delta\Delta S_{LNA}$ using eq 9 were reported by Hughesman et al.[16] A modification to the method of von Ahsen et al.[26]

$$\Delta\Delta S_{salt} = 0.847 n_{bp} \log_{10}[Na^+_{eq}] \quad (11)$$

that is consistent with the temperature dependencies of ΔH and ΔS is used to compute $\Delta\Delta S_{salt}$, where n_{bp} is the number of base pairs in the duplex and $[Na^+_{eq}]$ is the molar concentration of sodium ion equivalents, computed as

$$[Na^+_{eq}] = 0.1[\text{monovalent cations}] + 0.12\sqrt{[Mg^+] - [dNTPs]} \quad (12)$$

In eq 12, [dNTPs] is the total molar concentration of deoxyribonucleotide triphosphates in the sample.

The introduction of a mismatched DNA:DNA or DNA:LNA base pair into the duplex sequence is included as perturbation terms at the Gibbs energy level and computed as

$$\Delta\Delta G_{DNA:MM} = -\sum_{i=1}^{10} n_i \Delta G^\circ_{DNA_i} + \sum_{j=1}^{40} n_j \Delta G_{DNA:MM_j} \quad (13)$$

$$\Delta\Delta G_{LNA:MM} = \sum_{i=1}^{12} n_i \Delta\Delta G_{LNA:MM_j} + \Delta\Delta G_{5'NN\text{-}LNA} + \Delta\Delta G_{3'NN\text{-}LNA} \quad (14)$$

The first summation in eq 13 subtracts the Gibbs energy of each complementary nearest-neighbor DNA doublet i lost when the mismatch is formed. The second summation then adds to ΔG the free energy contributions of the corresponding nearest neighbors containing the mismatch, where $\Delta\Delta G_{DNA:MM_i}$ is the Gibbs energy of nearest-neighbor doublet j containing the mismatched DNA base pair. $\Delta\Delta G_{LNA:MM_i}$ quantifies the incremental change in the transition energy if an LNA is substituted into the mismatched base pair to form an LNA:DNA mismatch of type i. A final set of incremental corrections in eq 14 is applied only in cases (defined more fully in a later section) where a mismatch is flanked by a neighboring LNA substitution on either its 5′-side ($\Delta\Delta G_{5'NN\text{-}LNA}$) and/or 3′-side ($\Delta\Delta G_{3'NN\text{-}LNA}$). The set of required $\Delta G^\circ_{DNA_i}$ and $\Delta G_{DNA:MM_i}$ parameters used in eq 13 is provided by SantaLucia et al.[54] Type 2 model $\Delta\Delta G_{LNA:MM_i}$ parameters, all of which were determined in this work through regression of eq 13 to UVM data sets as described in Materials and Methods, are listed in Table 3. The values for $\Delta\Delta G_{5'NN\text{-}LNA}$ and $\Delta\Delta G_{3'NN\text{-}LNA}$ were likewise regressed and are −0.17 and −0.55 kcal mol^{-1}, respectively, indicating that the introduction of LNAs directly adjacent to a mismatch favors the denatured state.

In the extended type 2 model embodied in eqs 10−14, perturbations to the stability of the reference duplex arising from the introduction of either a DNA base pair mismatch or

Table 3. Regressed $\Delta\Delta G^\circ_{LNA:MM_i}$ Parameters and Associated Experimental $\Delta\Delta T_{m_i}$ Values (eq 3) for All Possible Mismatched LNA:DNA Base Pairs (LNA and DNA bases shown in uppercase and lowercase letters, respectively)

mismatch	$\Delta\Delta T_{m_i}$ (°C)	$\Delta\Delta G^\circ_{LNA:MM_i}$ (kcal/mol)
purine-purine		
A-a	2.7 ± 1.6	−0.50 ± 0.27
A-g	3.4 ± 3.1	−0.63 ± 0.33
G-a	6.7 ± 1.5	−1.18 ± 0.25
G-g	4.4 ± 1.6	−0.82 ± 0.27
pyrimidine-pyrimidine		
C-c	2.9 ± 1.7	−0.59 ± 0.38
C-t	2.1 ± 1.5	−0.44 ± 0.24
T-c	1.7 ± 0.9	−0.35 ± 0.15
T-t	1.4 ± 1.1	−0.32 ± 0.20
purine-pyrimidine		
A-c	0.4 ± 2.2	−0.11 ± 0.28
G-t	−2.5 ± 1.6	0.38 ± 0.28
pyrimidine-purine		
C-a	0.4 ± 1.2	−0.17 ± 0.41
T-g	1.6 ± 1.1	−0.28 ± 0.20

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

an LNA:DNA base pair mismatch are defined at the Gibbs energy level. In contrast, in corresponding type 1 models for mismatched pure-DNA homoduplexes[25,55−58] and for mismatched LNA-substituted heteroduplexes,[19] new values for ΔH and ΔS are computed for the mismatched duplex by replacing the enthalpic and entropic contributions of the lost complementary doublets with those of the corresponding doublets bearing the mismatch. As shown in Table 1, when applied to a pure-DNA homoduplex bearing a mismatched base pair, this approach requires an additional 48 enthalpic parameters and 48 entropic parameters to account for the contributions to ΔH and ΔS of all possible doublets containing a single mismatched base pair. While consistent with the NNT concept, many (~40%) of these added $\Delta H_{DNA:MM_i}$ and $\Delta S_{DNA:MM_i}$ parameters are not statistically significant, creating the possibility for anomolies in T_m predictions for mismatched duplexes.[25,55−58] This is reflected in Table 4, which reports as a function of temperature the melting thermodynamics of both complementary and mismatched pure-DNA doublets computed using a type 1 model. For the two perfectly matched doublets, that model predicts that at 37 °C the duplex state of the doublet is thermodynamically favored over the denatured state (i.e., $\Delta G_{37°C} > 0$, where Δ indicates the duplex-to-denatured state transition). The model also predicts, in accordance with all experimental data for duplex DNA melting, that the denatured state of the doublet becomes increasingly favored as the temperature is increased; that is, the model predicts $\Delta\Delta G_{70-37°C} < 0$. However, for several mismatched doublets, the type 1 model predicts a different trend that is at odds with available experiment. If the mismatched tc/gg doublet is taken as an example, the model captures the destabilizing effect the mismatch has on the duplexed doublet at 37 °C (i.e., $\Delta G_{37°C} < 0$). However, it erroneously predicts that the duplexed state of the doublet becomes more stable with an increase in temperature ($\Delta\Delta G_{70-37°C} > 0$), such that it is thermodynamically favored at 70 °C (i.e., $\Delta G_{70°C} = 0.27$ kcal mol^{-1}).

Table 4. Duplex-to-Denatured State Thermodynamic Changes under Standard Thermodynamic Conditions for Complementary and Mismatched Nearest-Neighbor Pure-DNA Doublets Computed as a Function of Temperature using a Type 1 Model[a]

doublet	ΔH	ΔS	$\Delta G_{37°C}$	$\Delta\Delta G_{70-37°C}$
		Complementary		
cg/cg	10.6	27.2	2.16	−0.90
cc/gg	8.0	19.9	1.83	−0.66
		Mismatched		
tg/ct	5.0	15.8	0.10	−0.52
ac/tt	−0.7	−0.2	−0.64	0.01
tc/gg	−3.3	−10.4	−0.07	0.34
cc/ga	−5.2	−14.2	−0.80	0.47
aa/tc	−7.6	−20.2	−1.33	0.67

[a]Nearest-neighbor parameters used for the calculation taken from ref 54 and references therein.

In contrast, all $\Delta G^{\circ}_{DNA_i}$ and most (89%) $\Delta\Delta G_{DNA:MM_i}$ parameters are statistically significant, with those $\Delta\Delta G_{DNA:MM_i}$ parameters that are not generally lying just inside their reported standard deviation.[25,55−58] As a result, we find that mismatch corrections are more quantitatively reliable when they are applied at the Gibbs energy level. This is reflected in Figure 1 and Table S2 (Supporting Information), which compare the performance of our extended type 2 model (eqs 10−14) to that of a corresponding type 1 model[25,55−58] when they are applied to a large set of pure-DNA duplexes containing one or two isolated mismatched base pairs. While the performance of the two models is similar for mismatched duplexes melting at a T_m of ≤50 °C, the type 1 model overpredicts T_m for more stable duplexes containing one or two mismatched base pairs due, at least in part, to the unusual temperature dependence that model ascribes to the stabilities of certain mismatched doublets (see Table 4).

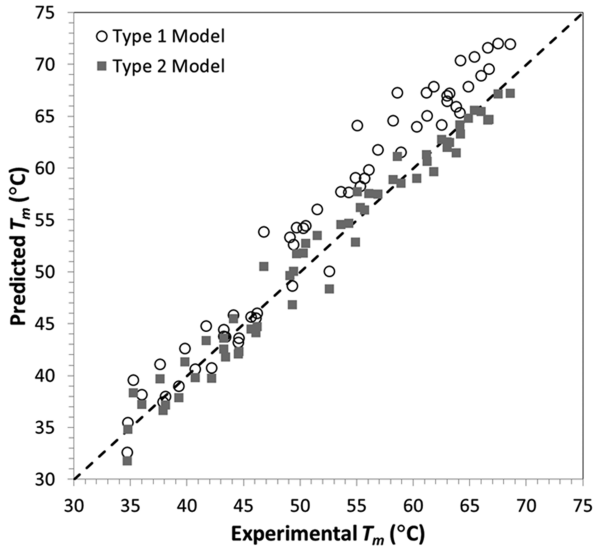


Figure 1. Comparison of T_m values predicted by a type 1[25,55−58] or type 2 model to corresponding experimental T_m data for 60 pure-DNA homoduplexes containing one or two mismatched base pairs.

Finally, in the application of eqs 10−14 to the prediction of T_m values for duplexes containing a mismatch, it is important to note that the calculation may be performed by specifying either of two different duplexes as the reference state. Figure S1 (Supporting Information) shows that our type 2 model predictions are independent of that choice.

Predicting Stabilities of Duplexes with One or More Mismatched LNA:DNA Base Pairs. Melting thermodynamic data (Table S1, Supporting Information) under standard conditions for a set of 384 duplexes containing either a central LNA:DNA base pair mismatch or the corresponding DNA base piar mismatch were collected and used to regress $\Delta\Delta G_{LNA:MM_i}$ parameters for all possible mismatched LNA:DNA base pairs (Table 3). The data were also used to compute mean $\Delta\Delta T_{m_i}$ [$=\Delta T_{m(LNA-LNA:MM)_i} - \Delta T_{m(DNA-DNA:MM)_i}$] values, as defined in eq 3, for all possible mismatched LNA:DNA base pairs. As noted previously,[6] LNAs are generally more energetically intolerant to mismatch formation, and this is reflected in the generally positive values of $\Delta\Delta T_{m_i}$, which indicate that nearly all LNA:DNA base pair mismatches are indeed more destabilizing to a duplex when compared to their corresponding DNA mismatch; the one exception, as noted by You et al.,[6] is the G-t mismatch. Our data show that this hyperdestabilization effect favoring the denatured state is largest for purine:purine LNA:DNA mismatches (Table 3).

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

You et al.[6] have reported data showing that the effect of a mismatched pure-DNA base pair (or LNA:DNA base pair) on duplex stability can be altered by a neighboring LNA substitution directly before (B; i.e., on the 5′-side) or after (A; 3′-side) the mismatch. Table 5 presents those data in a manner that shows that the destabilization of a duplex containing a pure-DNA mismatch is enhanced by an LNA substitution either directly before or after the mismatch, with the latter substitution having the larger destabilizing effect. Moreover, the data show that this hyperdestabilizing effect is additive in cases in which the mismatch is flanked on both sides by an LNA substitution. Most LNA substitutions within a mismatch are also seen to hyperdestabilize the duplex. Further destabilization is then observed for a neighboring LNA substitution on the 3′-side of the mismatch, but not on the 5′-side. Finally, Table 5 shows that these various destabilizing effects can be captured quantitatively in the type 2 model through the inclusion of two additional parameters, $\Delta\Delta G_{5'\text{NN-LNA}}$ and $\Delta\Delta G_{3'\text{NN-LNA}}$, and the provision that $\Delta\Delta G_{5'\text{NN-LNA}}$ not be applied to cases in which the mismatch and the 5′-neighboring base pair both contain an LNA substitution.

Table 5. Average Net Effect on Duplex Melting Temperatures of LNA Substitutions Directly before (B), Directly after (A), and/or within (M) a DNA:DNA (m) Base Pair Mismatch[a]

	ΔT_m (°C)[b]		
type of LNA substitution	experiment[c]	base model[d]	corrected model[e]
DNA mismatch (m)			
B-m	0.6 ± 1.3	−0.1 ± 0.0	0.4 ± 0.0
m-A	1.2 ± 1.1	−0.1 ± 0.0	1.5 ± 0.0
B-m-A	1.8 ± 1.5	−0.2 ± 0.0	1.9 ± 0.0
LNA mismatch (M)			
M	1.6 ± 1.6	1.2 ± 1.1	1.2 ± 1.1
B-M[f]	1.4 ± 2.0	1.1 ± 1.1	1.1 ± 1.1
M-A	2.8 ± 1.8	1.1 ± 1.1	2.7 ± 1.1
B-M-A[d]	3.0 ± 1.9	1.0 ± 1.1	2.6 ± 1.1

[a]Performance of eq 10 in capturing this effect when $\Delta\Delta G^\circ_{5'\text{NN-LNA}}$ and $\Delta\Delta G_{3'\text{NN-LNA}}$ parameters are included in eq 14 or not. Data used for analysis taken from ref 6. [b]All ΔT_m values reported are under standard thermodynamic conditions and are relative to the T_m of the corresponding mismatched pure-DNA duplex containing no B, M, or A type LNA substitutions. [c]For each type of LNA substitution, the mean ΔT_m and the range of ΔT_m values observed experimentally are reported. [d]Mean error and standard deviation of model predictions using eqs 10−14 with $\Delta\Delta G^\circ_{5'\text{NN-LNA}}$ and $\Delta\Delta G^\circ_{3'\text{NN-LNA}}$ parameters set equal to zero. [e]Mean error and standard deviation of model predictions using eqs 10−14 with $\Delta\Delta G^\circ_{5'\text{NN-LNA}}$ and $\Delta\Delta G^\circ_{3'\text{NN-LNA}}$ parameters set equal to −0.17 and −0.54 kcal mol^{-1}, respectively. [f]The $\Delta\Delta G_{5'\text{NN-LNA}}$ parameter is not applied to B-M type sequences.

Table 6 reports the performance of our extended type 2 model when it is applied to four different classes of duplexes: fully complementary pure-DNA duplexes, fully complementary duplexes containing one or more LNA substitutions, unsubstituted duplexes containing one or two mismatched base pair(s), and LNA-substituted duplexes containing one or two DNA:DNA and/or LNA:DNA mismatched base pair(s). None of the sequences included in this table were used to regress model parameters. All possible LNA:DNA and mismatched LNA:DNA base pairs are presented at least three times in the data sets. For each class of duplex, the results presented therefore represent a true test of model accuracy. In each column, the average experimental T_m and the range of experimental T_m values for the duplexes are reported along with the mean error and standard deviation of model predictions for the entire set of sequences within that class. In each case, model accuracy is very good, showing a low standard deviation and a mean error of <1 °C. As a result, the model is able to accurately predict $\Delta T_{m(\text{PM-MM})}$ for both pure-DNA duplexes and DNA:LNA heteroduplexes. In addition, the model accurately predicts $\Delta\Delta T_{m_i}$ (eq 3), thereby allowing one to reliably assess the impact that LNA substitutions at or near the site of variance will have on the specificity of a probe. Data and model results used to generate Table 6 are provided in Table S3 (Supporting Information).

Application of the Model to Predicting and Improving the $\Delta T_{m(\text{PM-MM})}$ of Allele-Specific Hydrolysis Probes through LNA Substitutions. As our type 2 NNT model is accurate in predicting melting thermodynamics under standard conditions for both pure-DNA duplexes and LNA-substituted duplexes (Table 6), we asked if the model might also be used to accurately predict T_m and $\Delta T_{m(\text{PM-MM})}$ values for LNA-substituted dual-labeled hydrolysis probes. Model corrections for PCR solution conditions and for the presence of both a fluorophore (reporter) and a fluorescence quenching agent are therefore required. In our model, eqs 11 and 12 are used to correct for changes in solution composition [salt concentrations, added magnesium ion (Mg^{2+})] away from standard

Table 6. Performance of Our Model (eqs 10−14) in Predicting T_m (degrees Celsius) under Standard Thermodynamic Conditions of Pure-DNA Duplexes and DNA/LNA Heteroduplexes Formed with Templates That Are Fully Complementary or Harbor One or Two Isolated (non-neighboring) Base Pair Mismatches

	DNA:PM	DNA:MM	LNA:PM	LNA:MM
no. of unique duplexed sequences	12	31	23	55
experimental T_m[a]	68.5 ± 6.9	54.1 ± 10.0	76.7 ± 9.9	63.0 ± 10.6
error in model predictions[b]	0.2 ± 1.0	−0.4 ± 1.4	−0.7 ± 1.2	−1.3 ± 2.4
experimental $\Delta T_{m(\text{DNA:PM-DNA:MM})}$ or $\Delta T_{m(\text{LNA:PM-LNA:MM})}$	12.2 ± 3.8		14.4 ± 4.3	
error in model predictions[b]	0.0 ± 1.6		0.4 ± 2.2	
experimental $\Delta\Delta T_{m_i}$ (eq 6)	2.2 ± 2.4			
error in model predictions[b]	0.0 ± 1.8			

[a]Experimental T_m values for each class of duplex are reported as the mean T_m values for all sequences ± the average range of T_m values recorded for the set of duplexes. All T_m data were collected under standard thermodynamic solution conditions (1 M NaCl in pH 7 buffer) and a C_T of 2 or 50 μM (Table S3 of the Supporting Information and data from refs 6 and 19). [b]Total errors in model predictions for a given class of duplex are reported as mean errors ± the standard deviation.

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

Table 7. Comparison of Experimental (UVM) and Model-Predicted Melting (duplex dissociation) Thermodynamics for Unlabeled and Dual-Labeled LNA-Substituted Probes under PCR-Relevant Solution Conditions

probe		experimentally measured by UVM[b]						model-predicted				
target	sequence (5′ to 3′)[a]	LNA (%)	c/g (%)	ΔG°_{37} (kcal/mol)	ΔH° (kcal/mol)	ΔS° (eu)[d]	T_m (°C)	$T_{m,PCR}$[c] (°C)	ΔG°_{37} (kcal/mol)	ΔH° (kcal/mol)	ΔS° (eu)[d]	T_m (°C)	$T_{m,PCR}$[c] (°C)
unlabeled probes													
EGFR WT exon 20	aTcACGcagcTca	38	54	18.0 ± 0.0	106.5 ± 0.6	282.5 ± 1.8	68.9 ± 0.2	65.4	18.7	108.5	286.5	71.0	67.5
EGFR T790M	aTcATGcagCTca	46	46	17.9 ± 0.1	103.9 ± 0.6	274.5 ± 1.8	69.3 ± 0.0	65.7	18.3	105.1	276.7	70.7	67.1
EGFR WT exon 21	agcAgtttggccAgcc	13	63	20.8 ± 0.3	131.1 ± 2.2	351.6 ± 6.2	71.4 ± 0.1	68.5	20.5	134.8	365.2	68.9	66.1
EGFR WT exon 21	ggcTggccaaacTgct	13	63	21.2 ± 0.2	132.4 ± 2.1	354.4 ± 6.2	72.2 ± 0.1	69.3	21.1	135.9	366.6	70.6	67.8
EGFR L858R	tggccCgCcca	18	82	17.9 ± 0.3	93.8 ± 3.2	241.6 ± 9.3	73.8 ± 0.1	69.7	17.2	91.0	235.5	71.3	67.1
EGFR L861Q	aCCcagcTgTTt	42	50	16.3 ± 0.1	93.2 ± 0.4	245.8 ± 0.9	66.3 ± 0.3	62.4	17.3	95.0	248.0	70.1	66.2
PIK3CA WT exon 10	ctGaaatCactGAgCa	31	44	19.1 ± 0.1	119.5 ± 1.2	320.3 ± 3.3	69.2 ± 0.1	66.0	20.2	131.5	355.6	68.9	66.0
PIK3CA WT exon 21	atgCaCATCatggt	36	43	18.5 ± 0.2	105.1 ± 2.1	275.7 ± 6.2	72.0 ± 0.1	68.4	18.8	112.3	298.4	70.0	66.6
PTEN WT exon 5	gggACGAACtgg	50	67	17.9 ± 0.3	94.5 ± 2.6	243.9 ± 7.4	73.5 ± 0.1	69.4	18.0	100.7	264.1	70.7	67.0
TP53 WT exon 7	cacACGCaCct	45	64	17.3 ± 0.4	89.8 ± 4.5	231 ± 13.1	72.6 ± 0.3	68.3	17.4	92.9	241.0	71.2	67.1
dual-labeled probes													
BRAF V600E	FAM-agatttCTCtGTagC-BHQ1	33	40	19.5 ± 0.3	121.3 ± 1.7	324.6 ± 4.3	71.2 ± 0.1	67.0	20.3	122.4	325.6	72.2	69.1
JAK2 V617F	FAM-ctccACagAAaCatac-BHQ1[e]	31	44	20.7 ± 0.4	136.1 ± 4.0	367.9 ± 11.8	72.6 ± 0.2	67.0	19.7	132.4	359.7	70.4	64.7
JAK2 V617F	FAM-acaGAAACATACT-BHQ1	77	31	19.7 ± 1.1	110.6 ± 10.0	288.8 ± 28.9	78.7 ± 0.6	71.6	19.2	106.0	275.1	79.3	71.9
KIT WT exon 17	HEX-ccagaGACATCaag-BHQ1	43	50	18.9 ± 0.6	112.8 ± 6.1	299.1 ± 17.6	73.2 ± 0.2	67.4	18.4	112.3	299.5	70.9	65.6
KIT D816V	FAM-ccagaGTCATC-BHQ1	55	55	17.4 ± 0.3	98.2 ± 0.7	258.0 ± 1.4	72.3 ± 0.4	65.4	16.3	88.8	231.2	70.8	64.1
KIT D816V	FAM-ccagaGTCaTCaag-BHQ1	36	50	19.8 ± 0.1	121.7 ± 1.0	324.7 ± 3.0	73.3 ± 0.0	67.9	19.0	115.8	308.8	71.7	66.6

[a]FAM = 6-carboxyfluorescein; HEX = hexachlorofluorescein; BHQ1 = Black Hole Quencher 1. LNA substitutions are capitalized. [b]Experimental data were measured by the UVM method in PCR buffer containing 50 mM K^+ and 3 mM Mg^{2+}; the data were collected at a C_T of 2 μM for all probes except for experiments involving *KIT* and *JAK2* probes that were performed at C_T values of 5 and 10 μM, respectively. [c]Experimental $T_{m,PCR}$ determined using eq 15 with a C_T of 0.2 μM. [d]eu = entropy unit (calories per mole per kelvin). [e]Previously reported probe designed by Markova et al.[9]

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

thermodynamic conditions. More comprehensive models are available for this purpose,[59] but we have found that eq 11, which we derived from the work of von Ahsen et al.,[26] provides a reliable estimate for the effect of solution composition on T_m. Similarly, the values of $\Delta\Delta H_{f/q}$ and $\Delta\Delta S_{f/q}$ used in eq 10 to correct T_m for reporter and fluorescence quenching groups were reported by Moreira et al.[27]

Table 7 reports LNA-substituted unlabeled and dual-labeled hydrolysis probes designed using our type 2 model to target known somatic point mutation "hot spots" in clinically relevant genes associated with various cancers, including *EGFR*, *PIK3CA*, *PTEN*, *TP53*, *BRAF*, *JAK2*, and *KIT*. These targets cover a broad range of sequence diversity (c/g content ranging from 31 to 82%), and the probes designed against them contain different LNA contents (ranging from 13 to 77%) as well as different LNA substitution patterns. Each probe was designed to contain LNA(s) at the site of variation(s) to take advantage of the greater thermodynamic penalty for an LNA:DNA mismatch (Table 5), as well as multiple LNA substitutions at other positions that collectively serve to decrease probe length while maintaining the T_m of the duplex formed by the probe and its perfect complement at a value between 65 and 70 °C under PCR conditions. In PCR, the initial probe concentration, and thus C_T (because the probe is in great excess), is typically set at (or near) 0.2 μM ($C_{T,PCR}$), a value that is too low to permit direct measurement of melting thermodynamics by the UVM method. However, the melting temperature for each probe:template duplex can be measured by UVM under PCR conditions and a somewhat higher C_T ($C_{T,Expt}$), and then corrected to $C_{T,PCR}$ using the rigorous two-state bimolecular melting thermodynamics relation

$$T_{m,PCR} = \frac{\Delta H_{Expt} + \Delta C_p(T_{m,PCR} - T_{m,Expt})}{\Delta S_{Expt} + \Delta C_p \ln(T_{m,PCR}/T_{m,Expt}) - R \ln(C_{T,PCR}/2)} \quad (15)$$

where ΔH_{Expt}, ΔS_{Expt}, and $T_{m,Expt}$ are the UVM-measured values at $C_{T,Expt}$. $C_{T,PCR}$ is divided by 2 in eq 15 because the probe is in great excess during PCR. For the set of 16 LNA-bearing probes, model-predicted T_m values at $C_{T,PCR}$ agree well with experiment, having a low total mean error and a standard deviation of −0.6 ± 2.0. Moreover, at $C_{T,Expt}$, the model not only accurately predicts T_m (−0.7 ± 2.0) but also provides accurate estimates of $\Delta G°_{37}$, $\Delta H°$, and $\Delta S°$ with average errors (predicted − experimental) determined to be 0 ± 4, 1 ± 5, and 1 ± 6% (mean error ± standard deviation), respectively.

To confirm the ability of the model to predict the discrimination [$\Delta T_{m,(PM-MM)}$] of single LNA/DNA base pair mismatches in LNA-bearing probes, UVM data were also collected on a total of 34 different mismatched probe:template duplexes. All 12 unique single LNA/DNA base pair mismatches were represented one or more times in this data set, and a good agreement between experimentally determined and model-predicted $\Delta T_{m,(PM-MM)}$ values was found (Figure 2), with a mean error of 0.9 ± 1.4 °C. Together, the results presented in Table 7 and Figure 2 demonstrate the general utility of the model in predicting melting thermodynamics of probes under PCR-relevant conditions and, more importantly, in designing LNA substitution patterns in allele-specific probes to achieve an effective $\Delta T_{m(MT-WT)}$ by increasing the destabilizing effect of a mismatch through an LNA-mediated decrease in probe length.

Quantitative Droplet Digital PCR Detection of the *BRAF* V600E or *KIT* D816V Mutation Using a Single

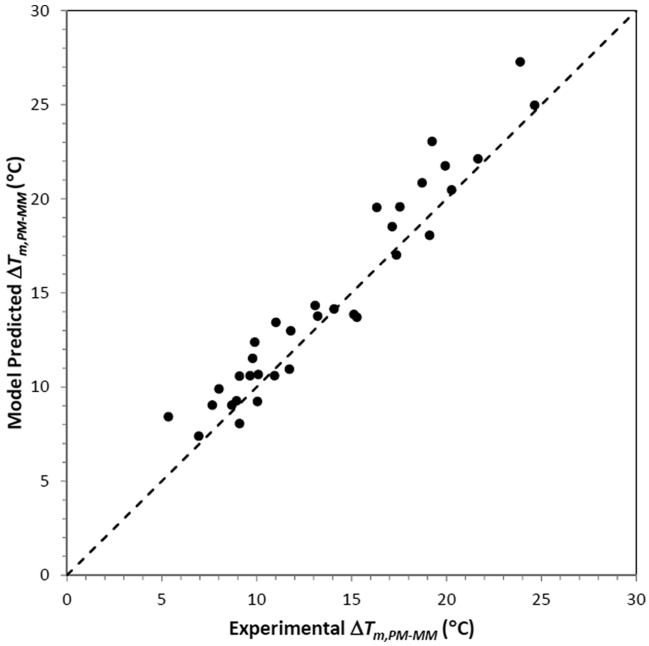


Figure 2. Comparison of experimental and model-predicted $\Delta T_{m,PM-MM}$ values for 34 mismatched probe:template duplexes.

Mutant-Specific Hydrolysis Probe. The ability of the model-designed LNA-substituted allele-specific (AS) probes to enhance the selective detection of a mutant allele was demonstrated in a ddPCR assay using the LNA-substituted probe directed against *BRAF* V600E. Quantitative ddPCR detection of the MT *BRAF* V600E (c. 1799T > A) allele in a background of WT *BRAF* requires complete segregation of output data fields (clusters) for the two alleles. Figure 3A shows that significant overlap of *BRAF* V600E and WT *BRAF* data clusters occurs when the ddPCR assay utilizes a conventional pure-DNA AS probe against *BRAF* V600E (5′-FAM-catcgag-atttctctgtagctagacc-IBFQ-3′) designed using our model. When targeting a MT allele harboring a single somatic point mutation, pure-DNA AS probes often offer a $\Delta T_{m(MT-WT)}$ of approximately ≤ 6 °C. Indeed, the pure-DNA probe against *BRAF* V600E provides experimental and predicted $\Delta T_{m(MT-WT)}$ values of 4.3 and 5.0 °C, respectively, which are too small for achieving complete segregation of data fields when T_a is set at a typical value of 60 °C.

When the LNA-substituted P15L6 probe against *BRAF* V600E, which offers experimental and predicted $\Delta T_{m(MT-WT)}$ values of 11.7 and 10.9 °C, respectively (Table S4, Supporting Information), is employed instead, complete segregation of MT and WT data fields is achieved at a T_a of 60 °C (Figure 3B), permitting quantitative detection of MT allele frequency within the sample.

Given that the pure-DNA AS probe against *BRAF* V600E does offer some thermodynamic discrimination, improved segregation of MT and WT data fields can be achieved by increasing T_a to a value closer to the $T_{m,PM}$ for that probe under PCR conditions. For example, at a T_a of 64 °C, reasonable segregation of MT and WT data clusters is achieved (Figure 4A). Nevertheless, it is unlikely that a cancer testing laboratory would adopt an assay in this form because of its strong sensitivity to small perturbations in T_a. In contrast, the performance of the LNA-substituted probe is relatively

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

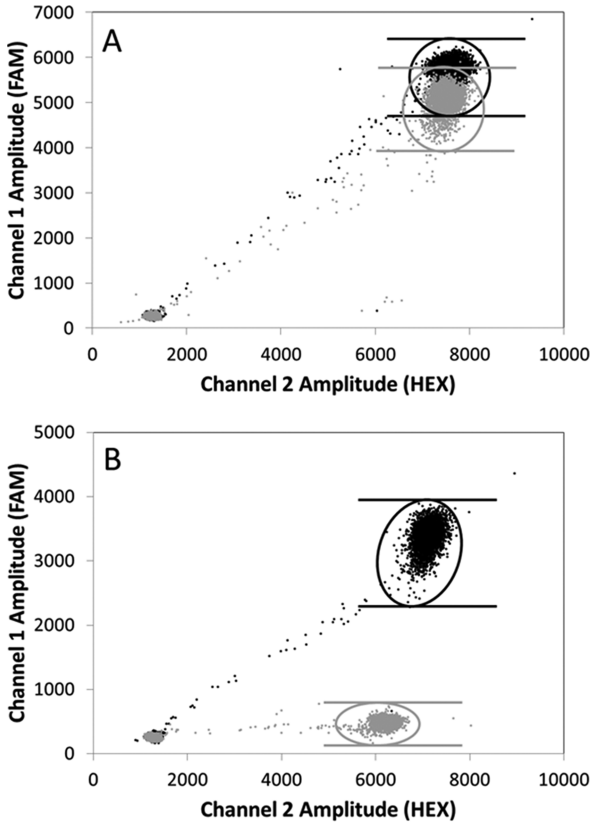


Figure 3. Detection and discrimination by ddPCR performed at a T_a of 60 °C of plasmids harboring either the WT *BRAF* (gray) or MT *BRAF* V600E (black) gene utilizing either (A) a *BRAF* V600E pure-DNA AS DNA probe (5′-FAM-catcgagatttctctgtagctagacc-IBFQ-3′) or (B) the *BRAF* V600E (5′-FAM-agatttCTCtGTagC-BHQ1-3′) LNA-substituted AS probe (Table 7). Horizontal lines indicate the upper and lower limits of each data cluster (circled). The cluster in the bottom right corner of each graph represents empty droplets.

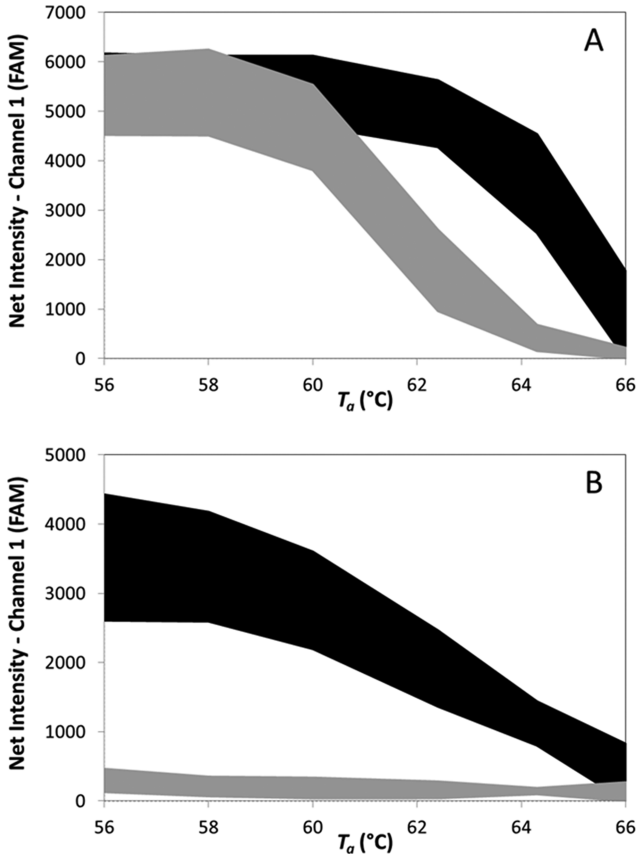


Figure 4. Ribbon diagram defining the discrimination of WT *BRAF* (gray) and MT *BRAF* V600E (black) data fields as a function of T_a in a ddPCR assay utilizing either (A) a *BRAF* V600E pure-DNA AS DNA probe (see Figure 2) or (B) the *BRAF* V600E LNA-substituted AS probe (Table 7). The upper and lower limits of each normalized data field (see Figure 2) define the span of each band shown. For either allele, normalized channel 1 (FAM) intensity values were determined by subtracting the upper and lower bounds of the empty droplet cluster from the corresponding upper and lower bounds of the data cluster.

insensitive to T_a, providing for complete segregation of data fields over a wide range of T_a values (Figure 4B).

Finally, we note that the model may be applied to the effective design of either pure-DNA or LNA-substituted AS probes. This is illustrated in Figure 5, which reports overlap of MT and WT data fields as a function of T_a for a ddPCR assay against *KIT* D816V utilizing either a model-designed pure-DNA AS probe (5′-FAM-ttggtctagccagagtcatcaagaa-IBFQ-3′) with experimental and predicted $\Delta T_{m(MT\text{-}WT)}$ values of 6.0 and 6.6 °C, respectively, or a model-designed LNA-substituted AS probe (5′-FAM-ccagaGTCaTCaag-BHQ1-3′) with experimental and predicted $\Delta T_{m(MT\text{-}WT)}$ values of 15.3 and 13.7 °C, respectively (Table S4, Supporting Information). While the LNA-substituted probe provides complete segregation of MT and WT data fields over a wide range of T_a values, the performance of the model-designed pure-DNA probe is sufficiently robust to yield an assay that potentially could satisfy clinical standards.

CONCLUSIONS

Real-time monitoring using dual-labeled hydrolysis probes has allowed the development of a wide range of PCR-based laboratory procedures and also improved the extent and reliability of information that can be obtained from patient specimens. Clinical applications of real-time PCR have grown exponentially, and the emergence of dPCR and useful modified probe chemistries, including LNA substitutions, is offering the potential to further expand diagnostic applications of PCR through their ability to improve quantification and differentiation of amplification products. However, fully realizing this potential will require model-based methods for designing probes and predicting their hydridization thermodynamics as a function of their sequence, chemistry, and template complementarity. Here, we have shown that formal accounting of the temperature dependence of ΔH and ΔS for the melting transition can be used to develop a nearest-neighbor thermodynamic model that accurately predicts the melting thermodynamics of a short oligonucleotide duplexed either to its perfect complement or to a template containing a mismatched base pair. The model may be applied to pure-DNA duplexes or to duplexes for which one strand contains any number and pattern of LNA substitutions. Perturbations to duplex stability arising from mismatched DNA:DNA or LNA:DNA base pairs are treated at the Gibbs energy level to maintain statistical significance in the regressed model parameters. As hydrolysis probe sequences are typically comprised of mixtures of all NNs, the model developed here is specifically designed to treat duplexes that are B-form in helix structure in their complementary unsubstituted (no LNA)

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

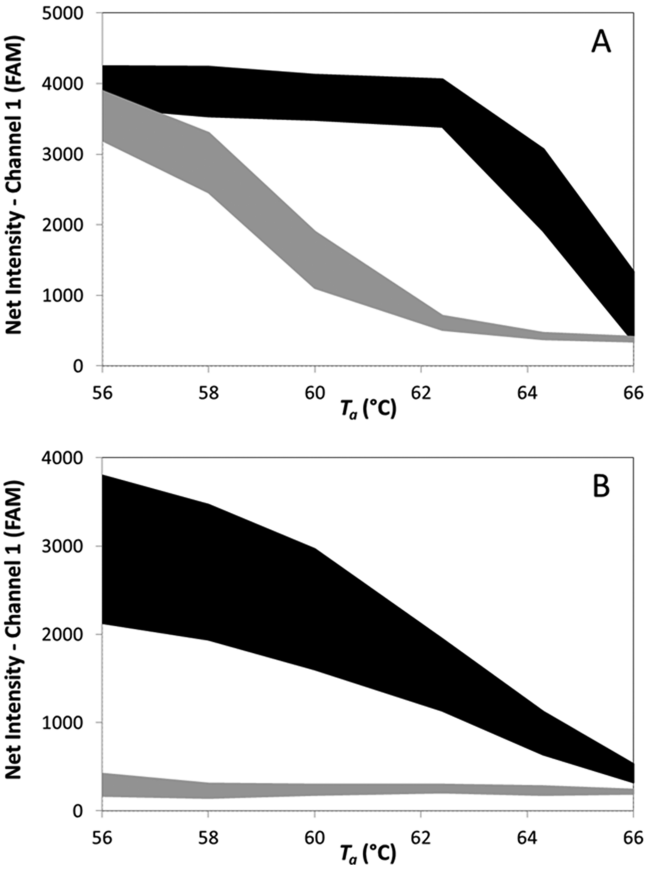


Figure 5. Ribbon diagram defining the discrimination of WT *KIT* (gray) and MT *KIT* D816V (black) data fields as a function of T_a in a ddPCR assay utilizing either (A) a *KIT* D816V pure-DNA AS DNA probe (5′-FAM-ttggtctagccagagtcatcaagaa-IBFQ-3′) or (B) the *KIT* D816V LNA-substituted AS probe (5′-FAM-ccagaGTCaTCaag-BHQ1-3′) (Table 7). The upper and lower limits of each normalized data field (see Figure 2) define the span of each band shown. For either allele, normalized channel 1 (FAM) intensity values were determined by subtracting the upper and lower bounds of the empty droplet cluster from the corresponding upper and lower bounds of the data cluster.

isosequential state. The model is therefore not specifically designed to treat duplexes containing substantial A-tract (i.e., consecutive aa/tt or at/at NNs) or G-tract (consecutive c/g base pairs) sequence motifs. Nevertheless, it predicts T_m and melting thermodynamics of short A-tract rich duplexes reasonably well, deviating from experimental T_m values, for example, by no more than ±3 °C for short duplexes containing up to six consecutive AA/TT NNs. This result is not surprising, as the B*-form helix associated with A-tract sequence motifs is structurally similar to the classic B-form helix. Poorer agreement is observed when the model is applied to short complementary duplexes high in G-tract content; deviations from experimental T_m values as large as ±6.6 °C are observed. This is consistent with the fact that G-tract sequences tend toward a modified A-form structural motif in which the sizes of the major and minor grooves of the duplex are reversed relative to those within the B-form helix. NNT parameters specific to B-form helix energetics and entropy are therefore not expected to be applicable to this significantly altered helix structure.

When appropriate equations are introduced to correct for changes in solution conditions and terminal addition of fluorescent dyes and quenchers, the model may be used to predict and thereby tailor the T_m of a pure-DNA or LNA-substituted hydrolysis probe when duplexed either to its perfect-match template or to a template harboring a non-complementary base. The model therefore should be of use to clinicians and biologists needing to design probes that unequivocally distinguish and quantify two closely related alleles in either a qPCR or a dPCR assay. This potential was demonstrated by showing that the model could be used to design AS probes that completely discriminate and quantify either *BRAF* V600E and WT *BRAF* templates or *KIT* D816V and WT *KIT* templates in a ddPCR assay.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional supplemental tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: israels@chbe.ubc.ca.

Funding

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR), and the Michael Smith Development Fund. C. Haynes receives salary support as a Canada Research Chair.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Tsongalis, G. J., and Coleman, W. B. (2006) Clinical genotyping: The need for interrogation of single nucleotide polymorphisms and mutations in the clinical laboratory. *Clin. Chim. Acta 363*, 127−137.

(2) Kutyavin, I. V., Afonina, I. A., Mills, A., Gorn, V. V., Lukhtanov, E. A., Belousov, E. S., Singer, M. J., Walburger, D. K., Lokhov, S. G., Gall, A. A., et al. (2000) 3′-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Res. 28*, 655−661.

(3) Letertre, C., Perelle, S., Dilasser, F., Arar, K., and Fach, P. (2003) Evaluation of the performance of LNA and MGB probes in 5′-nuclease PCR assays. *Mol. Cell. Probes 17*, 307−311.

(4) 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.

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

(6) 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.

(7) 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.

(8) Johnson, M. P., Haupt, L. M., and Griffiths, L. R. (2004) Locked nucleic acid (LNA) single nucleotide polymorphism (SNP) genotype analysis and validation using real-time PCR. *Nucleic Acids Res. 32*, e55.

(9) Markova, J., Prukova, D., Volkova, Z., and Schwarz, J. (2007) A new allelic discrimination assay using locked nucleic acid-modified nucleotides (LNA) probes for detection of JAK2 V617F mutation. *Leuk. Lymphoma 48*, 636−639.

(10) Pancrazzi, A., Guglielmelli, P., Ponziani, V., Bergamaschi, G., Bosi, A., Barosi, G., and Vannucchi, A. M. (2008) A sensitive detection method for MPLW515L or MPLW515K mutation in chronic

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

myeloproliferative disorders with locked nucleic acid-modified probes and real-time polymerase chain reaction. *J. Mol. Diagn. 10*, 435−441.
(11) Pinzani, P., Salvianti, F., Cascella, R., Massi, D., De Giorgi, V., Pazzagli, M., and Orlando, C. (2010) Allele specific Taqman-based real-time PCR assay to quantify circulating BRAFV600E mutated DNA in plasma of melanoma patients. *Clin. Chim. Acta 411*, 1319−1324.
(12) Bousquet, M., Le Guellec, S., Quelen, C., Rigal-Huguet, F., Delsol, G., and Brousset, P. (2006) Frequent detection of the JAK2 V617F mutation in bone marrow core biopsy specimens from chronic myeloproliferative disorders using the TaqMan polymerase chain reaction single nucleotide polymorphism genotyping assay: A retrospective study with pathologic correlations. *Hum. Pathol. 37*, 1458−1464.
(13) Poodt, J., Fijnheer, R., Walsh, I. B., and Hermans, M. H. (2006) A sensitive and reliable semi-quantitative real-time PCR assay to detect JAK2 V617F in blood. *Hematol. Oncol. 24*, 227−233.
(14) Bennett, M., and Stroncek, D. F. (2006) Recent advances in the bcr-abl negative chronic myeloproliferative diseases. *J. Transl. Med. 4*, 41.
(15) Bench, A. J. (2012) The role of molecular genetic analysis within the diagnostic haemato-oncology laboratory. *Int. J. Lab. Hematol. 34*, 21−34.
(16) Hughesman, C. B., Turner, R. F. B., and Haynes, C. A. (2011) Role of the Heat Capacity Change in Understanding and Modeling Melting Thermodynamics of Complementary Duplexes Containing Standard and Nucleobase-Modified LNA. *Biochemistry 50*, 5354−5368.
(17) 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.
(18) 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.
(19) Owczarzy, R., You, Y., Groth, C. L., and Tataurov, A. V. (2011) Stability and mismatch discrimination of locked nucleic acid-DNA duplexes. *Biochemistry 50* (43), 9352−9367.
(20) Mergny, J. L., and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides 13*, 515−537.
(21) Owczarzy, R. (2005) Melting temperatures of nucleic acids: Discrepancies in analysis. *Biophys. Chem. 117*, 207−215.
(22) Livak, K. J. (1999) Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. *Genet. Anal.: Biomol. Eng. 14*, 143−149.
(23) Vogelstein, B., and Kinzler, K. W. (1999) Digital PCR. *Proc. Natl. Acad. Sci. U.S.A. 96*, 9236−9241.
(24) Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A. J., Bright, I. J., Lucero, M. Y., Hiddessen, A. L., Legler, T. C., et al. (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem. 83*, 8604−8610.
(25) Allawi, H. T., and SantaLucia, J., Jr. (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry 36*, 10581−10594.
(26) von Ahsen, N., Wittwer, C. T., and Schütz, E. (2001) Oligonucleotide Melting Temperatures under PCR Conditions: Nearest-Neighbor Corrections for Mg^{2+}, Deoxynucleotide Triphosphate, and Dimethyl Sulfoxide Concentrations with Comparison to Alternative Empirical Formulas. *Clin. Chem. 47*, 1956−1961.
(27) 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.
(28) Hud, N. V., and Plavec, J. (2003) A unified model for the origin of DNA sequence-directed curvature. *Biopolymers 69*, 144−158.
(29) 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.
(30) Bailey, W. F., and Monahan, A. S. (1978) Statistical effects and the evaluation of entropy differences in equilibrium processes. Symmetry corrections and entropy of mixing. *J. Chem. Educ. 55*, 489.
(31) Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical chemistry, Part III: The Behavior of Biological Macromolecules*, W. H. Freeman, San Francisco.
(32) Dauxois, T., Peyrard, M., and Bishop, A. (1993) Entropy-driven DNA denaturation. *Phys. Rev. E 47*, R44−R47.
(33) Hughesman, C. B., Turner, R. F., 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.
(34) 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.
(35) 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.
(36) 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.
(37) 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.
(38) Zimm, B. H. (1960) Theory of "Melting" of the Helical Form in Double Chains of the DNA Type. *J. Chem. Phys. 33*, 1349−1356.
(39) 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.
(40) 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.
(41) 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.
(42) 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.
(43) 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.
(44) Mikulecky, P. J., and Feig, A. L. (2006) Heat capacity changes associated with nucleic acid folding. *Biopolymers 82*, 38−58.
(45) 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.
(46) 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.
(47) Prabhu, N. V., and Sharp, K. A. (2005) Heat Capacity in Proteins. *Annu. Rev. Phys. Chem. 56*, 521−548.
(48) Ramprakash, J., Lang, B., and Schwarz, F. P. (2008) Thermodynamics of single strand DNA base stacking. *Biopolymers 89*, 969−979.
(49) Zhou, J., Gregurick, S. K., Krueger, S., and Schwarz, F. P. (2006) Conformational changes in single-strand DNA as a function of temperature by SANS. *Biophys. J. 90*, 544−551.
(50) Mrevlishvili, G. M., Carvalho, A. P. S. M. C., and Ribeiro da Silva, M. A. V. (2002) Low-temperature DSC study of the hydration of ss-DNA and ds-DNA and the role of hydrogen-bonded network to the duplex transition thermodynamics. *Thermochim. Acta 394*, 73−82.

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352

(51) Spink, C. H., and Chaires, J. B. (1999) Effects of Hydration, Ion Release, and Excluded Volume on the Melting of Triplex and Duplex DNA. *Biochemistry 38*, 496−508.
(52) Madan, B., and Sharp, K. A. (2001) Hydration Heat Capacity of Nucleic Acid Constituents Determined from the Random Network Model. *Biophys. J. 81*, 1881−1887.
(53) Gallagher, K., and Sharp, K. (1998) Electrostatic Contributions to Heat Capacity Changes of DNA-Ligand Binding. *Biophys. J. 75*, 769−776.
(54) SantaLucia, J., Jr., and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct. 33*, 415−440.
(55) Allawi, H. T., and SantaLucia, J., Jr. (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: Sequence dependence and pH effects. *Biochemistry 37*, 9435−9444.
(56) Allawi, H. T., and SantaLucia, J., Jr. (1998) Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res. 26*, 2694−2701.
(57) Allawi, H. T., and SantaLucia, J., Jr. (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry 37*, 2170−2179.
(58) 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.
(59) Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A., and Walder, J. A. (2008) Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry 47*, 5336−5353.

DOI: 10.1021/bi500905b
Biochemistry 2015, 54, 1338−1352