ELSEVIER

Biophysical Chemistry 117 (2005) 207 – 215

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Melting temperatures of nucleic acids: Discrepancies in analysis

Richard Owczarzy*

Integrated DNA Technologies, 1710 Commercial Park, Coralville, IA-52241, USA

Received 2 February 2005; received in revised form 21 May 2005; accepted 21 May 2005
Available online 15 June 2005

Abstract

Melting temperature, T_m, is an important property of nucleic acid duplexes. It is typically determined from spectroscopic or calorimetric melting experiments. More than one analytical method has been used to extract T_m values from experimental melting data. Unfortunately, different methods do not give the same results; the same melting data can be assigned different T_m values depending upon which method is used to process that data. Inconsistencies or systematic errors between T_ms reported in published data sets can be significant and add confusion to the field. Errors introduced from analysis can be greater than experimental errors, ranging from a fraction of degree to several degrees. Of the various methods, the most consistent and meaningful approach defines melting temperature as the temperature at the transition midpoint where half of the base pairs are melted and standard free energy is zero. Assuming a two-state melting behavior, we present here a set of general equations that can be used to reconcile these analytical T_m differences and convert results to the correct melting temperatures at the transition midpoint. Melting temperatures collected from published sources, which were analyzed using different methods, can now be corrected for these discrepancies and compared on equal footing. The similar corrections apply to T_m differences between calorimetric and spectroscopic melting curves. New algorithm for selection of linear sloping baselines, 2nd derivative method, is suggested, which can be used to automate melting curve analysis.

Keywords: DNA; RNA; Melting curve; Melting profile; Thermodynamic analysis

1. Introduction

Melting temperature, T_m, is a physical property of nucleic acids that gives information about the stability of duplexes in a specified environment. T_m values are useful in a variety of fields ranging from practical assay design in molecular biology to theoretical biophysics. We and others [1] have been collecting and importing published thermodynamic data including melting temperatures into databases. Large, statistically significant sets of experimental melting data from independent sources are very useful for development and refinement of models that predict stability and hybridization energetics of nucleic acids. Unfortunately, published T_m values are not consistent; results vary between different laboratories. These discrepancies arise because of experimental errors in data collection and differences between methods of data analysis [2,3]. Different analytical methods provide different T_m values from the same experimentally measured melting curves, i.e., inherent systematic differences in reported melting temperatures exist. Building upon previous works [2–7], we present here a set of general formulas that allow correction of reported T_m values to the transition midpoint T_m where half of the base pairs are broken and allow different data set to be integrated. In addition, a novel method for automatic selection of sloping baselines and some advantages of plotting $d\theta/d(1/T)$ derivative melting curves are described.

2. Theoretical results

Melting profiles are usually collected using spectroscopic experiments. The intensity signal (absorbance, fluorescence, molar ellipticity) is measured as a function of temperature. In the following, we develop a theoretical model using an ultraviolet absorbance versus temperature profile as an

* Tel.: +1 319 626 8459.
E-mail address: science@owczarzy.net.
URLs: http://www.idtdna.com, http://www.owczarzy.net.

0301-4622/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2005.05.006

example. However, the same formalism will apply to other spectroscopic measurements, including circular dichroism, fluorescence or NMR melting experiments. Fig. 1 shows a typical ultraviolet melting profile and its derivatives for 10 base pair long DNA duplex [8]. Panel A displays dependence of absorbance at 268 nm on sample temperature. The upper and lower baselines are established from the linear regions before and after the transition. The fraction of melted base pairs, θ, is calculated from the standard formula, $\theta=(A-A_L)/(A_U-A_L)$, where A, A_L, and A_U are sample absorbance, absorbance of the lower baseline, and absorbance of the upper baseline, respectively [9]. The resulting melting profile (Fig. 1B) is usually smoothed by a digital filter [13] or by fitting a polynomial to a sliding window of data points. The correct definition of melting temperature, T_m(half), is the temperature at the midpoint of transition where $\theta=0.5$ and standard free energy of transition is zero. Assuming that spectroscopic signal is proportional to the number of melted base pairs, T_m(half) is a temperature where half of the base pairs are melted from the total number of base pairs in the nucleic acid sample. For the duplex on Fig. 1, T_m(half) of 55.4 °C was determined.

In common practice, alternative analytical methods are sometimes employed, which yield different T_m values [2,3]. These analytical methods are often available as a part of software package supplied with spectrophotometers. Derivatives of θ with respect to temperature or reciprocal of the temperature are calculated (panels C and D on Fig. 1). Melting temperatures, $T_m(\theta-T)$ and $T_m(\theta-1/T)$ are read as temperatures of peak height maxima of $d\theta/dT$ and $d\theta/d(1/T)$ derivative melting profiles, respectively. These two methods essentially determine melting temperatures from the inflection point of an S-shaped melting curve. Because melting curves are not exactly symmetrical, T_m(half), $T_m(\theta-T)$, and $T_m(\theta-1/T)$ are not identical. Differences among these values exist, typically $T_m(\theta-1/T)>T_m(\theta-T)>T_m$(half), as is illustrated in Fig. 1. When spectroscopic equipment is tested and validated, and proper experimental procedures are followed, experimental errors of melting temperatures that stem from data collection can be as low as 0.2 °C [3,8,10]. Differences among T_m(half), $T_m(\theta-T)$ and $T_m(\theta-1/T)$ are outside of these experimental errors, e.g., the difference $T_m(\theta-1/T)-T_m$(half) is equal to 1.5 °C for the melting profile shown in Fig. 1. Therefore, it would be useful to correct published $T_m(\theta-T)$ and $T_m(\theta-1/T)$ temperatures to T_m(half) values, so that data from different laboratories can be properly compared.

Consider bimolecular hybridization reaction of two different non-self-complementary oligomers S_1 and S_2 of the same concentrations $C_t/2$. The oligomers anneal to form a duplex D,

$$S_1 + S_2 \rightleftarrows D$$

Assuming that the transition proceeds in two-state fashion, i.e., no significant population of partially melted

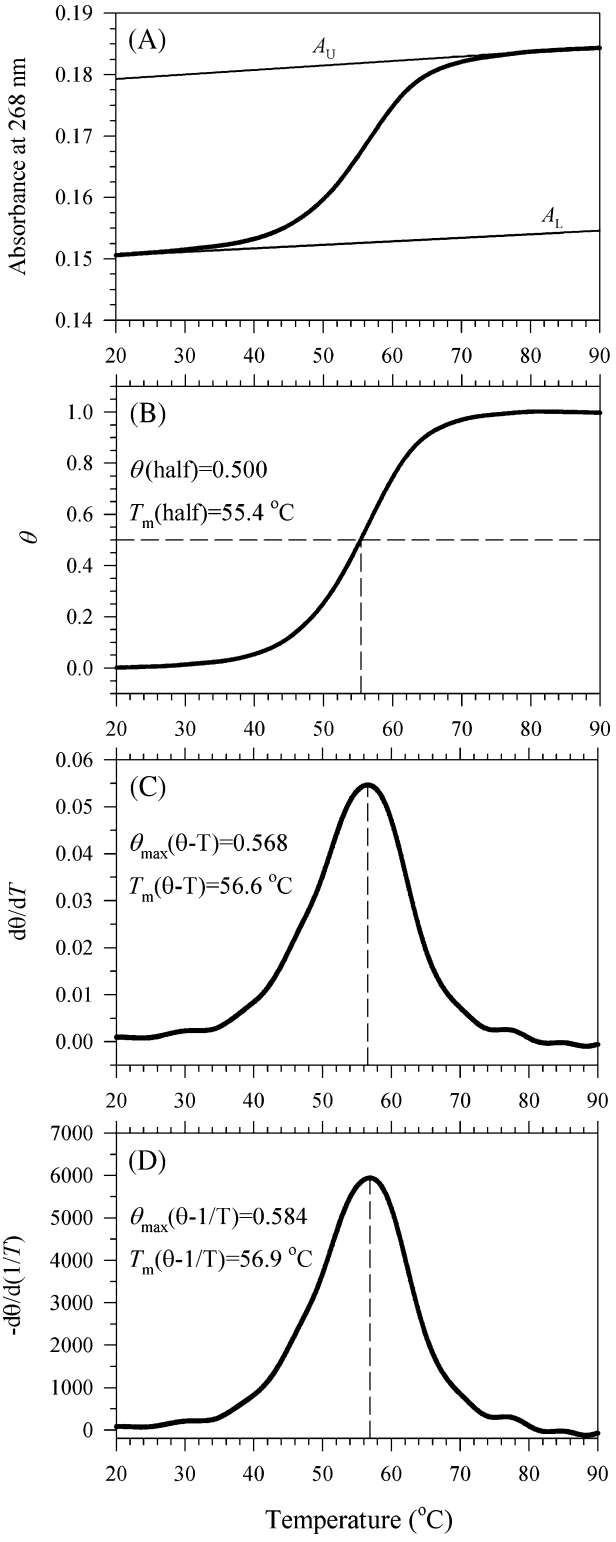

Fig. 1. Experimental ultraviolet melting profile of d(TGGCGAGCAC) duplex oligomer (A) was analyzed using three different methods: fraction of broken (melted) base pairs was plotted vs. temperature (B). First derivatives $d\theta/dT$ (C) and $d\theta/d(1/T)$ (D) were plotted vs. temperature. The DNA duplex was dissolved in 1M NaCl-phosphate buffer, at C_t=2 μM, pH 7.0 [8]. Melting profiles were averaged from four heating and four cooling melting curves.

duplexes is detected, the fraction of *melted base pairs* is equal to fraction of *melted duplexes* and can be written as,

$$\theta = \frac{[S_1] + [S_2]}{C_t} \quad (1)$$

where the total single strand concentration, $C_t=[S_1]+[S_2]+2[D]$. Note that the equations are sometimes cast in terms of fraction of intact duplexes [2,4,7], α, which is simply $1-\theta$. The equilibrium constant of annealing reaction [2,4,7] in terms of θ is,

$$K_a = \frac{[D]}{[S_1][S_2]} = \frac{2(1-\theta)}{\theta^2 C_t}. \quad (2)$$

Solution of this equation gives us the formula for prediction of melting curves (Fig. 1B),

$$\theta = \frac{\sqrt{1+2K_aC_t}-1}{K_aC_t}. \quad (3)$$

Substitution of K_a from the Eq. (2) into van't Hoff Eq. (4),

$$\frac{d\ln K_a}{dT} = \frac{\Delta H^\circ}{RT^2} \quad (4)$$

and further differentiation leads to the equation for the derivative melting curves, e.g., the graph in Fig. 1C,

$$\frac{d\theta}{dT} = -\frac{\Delta H^\circ}{RT^2}\frac{\theta(1-\theta)}{2-\theta}. \quad (5)$$

This expression relates $d\theta/dT$ to ΔH° at any θ. To find the fraction of melted base pairs at maximum of this derivative melting curve, $\theta_{max}(\theta-T)$, the second derivative is calculated,

$$\frac{d^2\theta}{dT^2} = \frac{\Delta H^\circ}{RT^3}\frac{\theta(1-\theta)}{2-\theta}\left(2+\frac{\Delta H^\circ}{RT}\frac{\theta^2+2(1-2\theta)}{(2-\theta)^2}\right). \quad (6)$$

Solution of the Eq. (6) for θ when $d^2\theta/dT^2=0$ and $0<\theta<1$ yields,

$$\theta_{max}(\theta-T) = 2-\sqrt{\frac{2}{1+2\frac{RT}{\Delta H^\circ}}} = 2-\sqrt{\frac{2}{1+2B}} \quad (7)$$

where B denotes the ratio of $RT/\Delta H^\circ$, which is often present in equations. Since negative ΔH° values are observed for strand annealing reactions, the values of B are negative.

Analogous procedure can be employed to get a relationship for $\theta_{max}(\theta-1/T)$ at the maximum of $d\theta/d(1/T)$ derivative melting curve. For bimolecular equilibria, the resulting equations are summarized in Table 1. Two important features are observed. (1) Comparing hybridization of a self-complementary strand with hybridization of non-self-complementary strands, equations for the equilibrium constant are different for each type of hybridizations (row 1 of Table 1). However, equations for derivative melting curves and their maxima are identical for both types of hybridizations (rows 3–8 of Table 1). In other words, derivative melting curves are described by the same relationships, in terms of θ, for hybridizations of both identical and non-identical single strands. (2) Furthermore, the value of θ at the maximum of $d\theta/d(1/T)$ derivative melting curve is independent of ΔH° and therefore is also independent of base sequence and oligomer length. The $\theta_{max}(\theta-1/T)$ is constant and equal to 0.5858 for bimolecular hybridizations when single strands are at the same concentration.

Table 1
Thermodynamic models of bimolecular equilibria between two single strands, S_1, S_2, and the duplex D, where both single strands have the same concentration

Reaction type	$S_1+S_2 \rightleftarrows D$ (non-self-complementary)	$2S \rightleftarrows D$ (self-complementary)
Equilibrium constant	$K_a = \frac{[D]}{[S_1][S_2]} = \frac{2(1-\theta)}{\theta^2 C_t}$	$K_a = \frac{[D]}{[S]^2} = \frac{(1-\theta)}{2\theta^2 C_t}$
Melting curve	$\theta = \frac{\sqrt{1+2K_aC_t}-1}{K_aC_t}$	$\theta = \frac{\sqrt{1+8K_aC_t}-1}{4K_aC_t}$
	$d\theta/dT$ derivative melting curves	
First derivative	$\frac{d\theta}{dT} = -\frac{\Delta H^\circ}{RT^2}\frac{\theta(1-\theta)}{(2-\theta)}$	
Second derivative	$\frac{d^2\theta}{dT^2} = \frac{\Delta H^\circ}{RT^3}\frac{\theta(1-\theta)}{(2-\theta)}\left(2+\frac{\Delta H^\circ}{RT}\frac{\theta^2+2(1-2\theta)}{(2-\theta)^2}\right)$	
$\theta_{max}(\theta-T)$[a]	$\theta_{max}(\theta-T) = 2-\sqrt{\frac{2}{1+2B}}$	
	$d\theta/d(1/T)$ derivative melting curves	
First derivative	$\frac{d\theta}{d(1/T)} = \frac{\Delta H^\circ}{R} \bullet \frac{\theta(1-\theta)}{2-\theta}$	
Second derivative	$\frac{d^2\theta}{d(1/T)^2} = \left(\frac{\Delta H^\circ}{R}\right)^2\frac{\theta(1-\theta)}{2-\theta}\frac{\theta^2+2(1-2\theta)}{(2-\theta)^2}$	
$\theta_{max}(\theta-1/T)$	$\theta_{max}(\theta-1/T) = 2-\sqrt{2}$	

[a] $(B=RT_m/\Delta H^\circ)$.

Equations from Table 1 can be used to calculate correct melting temperatures, T_m(half), from published $T_m(\theta-T)$, $T_m(\theta-1/T)$ values. Temperatures can be estimated from transition enthalpy, ΔH°, entropy, ΔS°, and equilibrium constants,

$$T_m(\text{half}) = \frac{\Delta H^\circ}{\Delta S^\circ - R\ln K_a(\text{half})} \quad (8)$$

$$T_m(\theta-T) = \frac{\Delta H^\circ}{\Delta S^\circ - R\ln K_a(\theta-T)} \quad (9)$$

where K_a(half) and $K_a(\theta-T)$ are the equilibrium constants of hybridization reaction at T_m(half) and $T_m(\theta-T)$ temperatures, respectively. Combining Eqs. (8) and (9) yields,

$$T_m(\text{half}) = \frac{\Delta H^\circ}{\frac{\Delta H^\circ}{T_m(\theta-T)} + R\ln\left[\frac{K_a(\theta-T)}{K_a(\text{half})}\right]} \quad (10)$$

which represents a general equation that allows calculation of T_m(half) from the temperature at maxima of derivative melting curve, $T_m(\theta - T)$. The analogous equation was found for the $d\theta/d(1/T)$ derivative melting curves,

$$T_m(\text{half}) = \frac{\Delta H^\circ}{\frac{\Delta H^\circ}{T_m(\theta - 1/T)} + R\ln\left[\frac{K_a(\theta - 1/T)}{K_a(\text{half})}\right]} \quad (11)$$

which can be applied to calculate T_m(half) melting temperatures from $T_m(\theta - 1/T)$ values. To complete the calculation of T_m(half), ΔH° and K_a must be known. If the value of ΔH° is not available experimentally, it can be estimated from the nearest-neighbor model [11]. The equilibrium constants, which enter into the Eqs. (10) and (11), are calculated from the $\theta_{max}(\theta - T)$, $\theta_{max}(\theta - 1/T)$ values and from the K_a relationships shown in row 1 of Table 1. For example, let us evaluate equilibrium constants for hybridization of non-self-complementary single strands. At T_m(half), the value of θ is 0.5 by definition. Therefore, the equilibrium constant at T_m(half) is evaluated as,

$$K_a(\text{half}) = \frac{2(1 - 0.5)}{0.5^2 C_t} = \frac{4}{C_t}. \quad (12)$$

Eq. (7) can be incorporated into Eq. (2) to give the equilibrium constant at $T_m(\theta - T)$,

$$K_a(\theta - T) = \frac{\sqrt{2(1+2B)} - (1+2B)}{C_t\left[3 + 4B - 2\sqrt{2(1+2B)}\right]} = \frac{1 - 2B - 8B^2 + \sqrt{2(1+2B)}}{C_t(1+4B)^2} \quad (13)$$

The expression for $\theta_{max}(\theta - 1/T)$ from Table 1 can be substituted into Eq. (2) to give the equilibrium constant at $T_m(\theta - 1/T)$,

$$K_a(\theta - 1/T) = \frac{1 + \sqrt{2}}{C_t} \quad (14)$$

The resulting K_a values from Eqs. (12)-(14) are then used with Eqs. (10) and (11) to calculate corrected melting temperatures, T_m(half). Note that the Eqs. (12)-(14) were derived assuming hybridization of non-self-complementary strands. If the annealing strands are self-complementary (identical), the appropriate equation for K_a given in row 1 of Table 1 must be used to obtain expressions for K_a(half), $K_a(\theta - T)$ and $K_a(\theta - 1/T)$.

Hybridization reactions of nucleic acids are not limited to bimolecular systems. Hairpins, 3-way junctions and 4-way junctions fold into structures, which consist of different numbers of single strands. The relationships between equilibrium constants and fractions of melted base pairs can be generalized for n-strand hybridizations [5,7]. Assuming that single strands are present in the same concentration, C_t/n, we derived general equations, which can be employed to predict derivative melting profiles, melting temperatures and their discrepancies. The equations for hybridization reactions of any molecularity are presented in the Table 2 and were derived from relationships between K_a and θ using procedures similar to that previously shown for bimolecular hybridizations. Comparison of Table 1 with Table 2 reveals that the expressions are consistent. The equations in Table 2 reduce to the equations of Table 1 when n equal to 2 is considered. Two features that were observed for bimolecular hybridizations are also seen for general n-strand hybridizations. Again, derivative melting curves are described by identical expressions for both self-complementary and non-self-complementary single strands. The equations in rows 2–7 of Table 2 hold for both cases. Further, the values of θ at maximum of $d\theta/d(1/T)$ derivative melting curves are independent of oligomer sequence and length. They depend solely on molecularity of the hybridization reaction. Corrected T_m(half) values are predicted in a manner similar to bimolecular hybridizations. General Eqs. (10) and (11) can be used, because they hold for hybridizations of any molecularity. Equilibrium constants, which enter the Eqs. (10) and (11), need to be evaluated from equations given in the first row of Table 2 and corresponding $\theta_{max}(\theta - T)$, $\theta_{max}(\theta - 1/T)$ values.

Table 2
Thermodynamic models of general n-molecular equilibria between the single strands, S_1, S_2, ... S_n, and the helical complex, D_n, when all single strands are either non-self-complementary or self-complementary

Reaction type	$S_1 + S_2 + \ldots S_n \rightleftarrows D_n$ (non-self-complementary)	$nS \rightleftarrows D_n$ (self-complementary)
Equilibrium constant[a]	$K_a = \frac{[D_n]}{[S_1][S_2]\ldots[S_n]} = \frac{n^{n-1}(1-\theta)}{\theta^n C_t^{n-1}}$	$K_a = \frac{[D_n]}{[S]^n} = \frac{(1-\theta)}{n\theta^n C_t^{n-1}}$
	$d\theta/dT$ derivative melting curves	
First derivative	$\frac{d\theta}{dT} = -\frac{\Delta H^o}{RT^2}\frac{\theta(1-\theta)}{\theta + n(1-\theta)}$	
Second derivative	$\frac{d^2\theta}{dT^2} = \frac{\Delta H^\circ}{RT^3}\frac{\theta(1-\theta)}{\theta+n(1-\theta)} \times \left(2 + \frac{\Delta H^\circ}{RT}\frac{\theta^2(n-1)+n(1-2\theta)}{\left[\theta+n(1-\theta)\right]^2}\right)$	
$\theta_{max}(\theta - T)$[b]	$\theta_{max}(\theta - T) = \frac{1}{n-1}\left[n - \sqrt{\frac{n}{1+2B(n-1)}}\right]$	
	$d\theta/d(1/T)$ derivative melting curves	
First derivative	$\frac{d\theta}{d(1/T)} = \frac{\Delta H^\circ}{R}\frac{\theta(1-\theta)}{\theta+n(1-\theta)}$	
Second derivative	$\frac{d^2\theta}{d(1/T)^2} = \left(\frac{\Delta H^\circ}{R}\right)^2\frac{\theta(1-\theta)}{\theta+n(1-\theta)} \times \frac{\theta^2(n-1)+n(1-2\theta)}{\left[\theta+n(1-\theta)\right]^2}$	
$\theta_{max}(\theta - 1/T)$	$\theta_{max}(\theta - 1/T) = \frac{1}{n-1}(n - \sqrt{n}) = 1 - \frac{1}{1+\sqrt{n}}$	

[a] Each strand has the same concentration of C_t/n. ($C_t = [S_1] + [S_2] + \ldots + [S_n] + n[D_n]$).
[b] ($B = RT_m/\Delta H^\circ$).

Table 3 shows expressions for two special cases of nucleic acid hybridizations. The left side of the table presents equations for hairpin-random coil equilibrium when a single strand folds into an intramolecular structure. The right side of Table 3 presents equations for a bimolecular reaction when one of two single strands (S_1) has a significantly higher concentration than the other strand (S_2). Under this condition, the expression for K_a of bimolecular reaction can be simplified and resembles the equation for K_a of hairpin-coil equilibrium (row 1 of Table 3). Further differentiation demonstrates that the expressions describing the derivative melting curves are same for both equilibria (row 3–8 of Table 3). It has been long recognized [3,4] that a plot of $d\theta/d(1/T)$ versus temperature for intramolecular hairpin-coil equilibrium has a maximum exactly at the midpoint of melting transition, i.e., $\theta_{max}(\theta-1/T)$ is equal to 0.5 and in turn $T_m(\theta-1/T)$ is equal to T_m(half). The analysis in Table 3 shows that this is also true for the bimolecular equilibrium when one of the single strands is in significant excess (at least 10× or more). Interestingly, this kind of equilibrium is seen in some important molecular biology applications, e.g., detection of gene targets using oligomeric probes. We conclude that in these two specific cases of nucleic acid equilibria, T_m(half) can be read directly as a temperature at the peak height maximum of $d\theta/d(1/T)$ derivative melting curve.

Table 3
Comparison of hairpin-melted coil equilibria to bimolecular equilibria where the analytical concentration (C_1) of the single strand S_1 is in significant excess to the concentration (C_2) of single strand S_2

Reaction type	$S \rightleftarrows D$ (melted coil-hairpin)	$S_1+S_2 \rightleftarrows D$ (non-self-complementary strands and $C_1 \gg C_2$)
Equilibrium constant	$K_a=\frac{[D]}{[S]}=\frac{1-\theta}{\theta}$	$K_a=\frac{[D]}{[S_1][S_2]}=\frac{1-\theta}{\theta\left[C_1-C_2(1-\theta)\right]}\approx\frac{1-\theta}{\theta C_1}$
Melting curve	$\theta=\frac{1}{1+K_a}$	$\theta\approx\frac{1}{1+C_1K_a}$
	$d\theta/dT$ derivative melting curves	
First derivative	$\frac{d\theta}{dT}=-\frac{\Delta H^\circ}{RT^2}\theta(1-\theta)$	
Second derivative	$\frac{d^2\theta}{dT^2}=\frac{\Delta H^\circ}{RT^3}\theta(1-\theta)\left[2+\frac{\Delta H^\circ}{RT}(1-2\theta)\right]$	
$\theta_{max}(\theta-T)$[a]	$\theta_{max}(\theta-T)=0.5+B$	
	$d\theta/d(1/T)$ derivative melting curves	
First derivative	$\frac{d\theta}{d(1/T)}=\frac{\Delta H^\circ}{R}\theta(1-\theta)$	
Second derivative	$\frac{d^2\theta}{d(1/T)^2}=\left(\frac{\Delta H^\circ}{R}\right)^2\theta(1-\theta)(1-2\theta)$	
$\theta_{max}(\theta-1/T)$	$\theta_{max}(\theta-1/T)=0.5$	

[a] $(B=RT_m/\Delta H^\circ)$.

2.1. Comparison of spectroscopic and calorimetric melting profiles

It is important to consider relationships between melting temperatures from spectroscopic and calorimetric experiments. Differential scanning calorimetry (DSC) provides direct, model-independent measurements of energetics for helix-random coil transitions [5,12]. Specifically, assumption of two-state transition is unnecessary to obtain meaningful thermodynamic data (ΔH and ΔS). DSC measures "excess" molar heat capacity, ΔC_p^{exc}, as a function of temperature for duplex to single strand transitions. The area below the excess heat capacity curve gives us calorimetric transition enthalpy, ΔH_{DSC}. This can be represented in differential form,

$$\Delta C_p^{exc}=\left(\frac{\partial \Delta H}{\partial T}\right)_p. \tag{15}$$

If we assume temperature-independent enthalpy and a strictly two-state nature of the transition, the enthalpy change of a DSC sample is proportional to fraction of melted duplexes,

$$\Delta C_p^{exc}=\left(\frac{\partial(\theta\times\Delta H_{DSC})}{\partial T}\right)_p=\Delta H_{DSC}\left(\frac{\partial\theta}{\partial T}\right)_p. \tag{16}$$

Eq. (16) reveals that the calorimetric melting temperature, which is typically read at maximum of ΔC_p^{exc} vs. temperature curve, is equivalent to $T_m(\theta-T)$. Therefore, if the two-state assumption is reasonable, DSC melting temperature could be corrected to T_m(half) using Eq. (10).

2.2. Automatic selection of melting curve baselines

To calculate the plot of θ versus temperature shown in Fig. 1, the lower baseline A_L and the upper baseline A_U must be drawn. Sloping baselines are caused by increased unstacking of bases in both single-stranded and double-stranded molecules [7,9]. Experimental errors contributing to sloping baselines include drift of lamp intensity, sample evaporation, cuvette expansion, and sensitivity of detector to temperature changes when the detector is in close proximity to Peltier cuvette holder. Sloping baselines are usually least-squared fit to linear regions of absorbance versus temperature plots (see Fig. 2). Since the beginning and the end of melting transition are not always apparent, selection of baselines is somewhat arbitrary and it is usually done manually. We propose new algorithm, the second derivative method, which can be programmed and executed automatically.

Because melting curves are linear in baseline regions, the first derivative of melting curve, dA/dT, is constant and the second derivative, dA^2/dT^2, is zero in these regions. In contrast, slopes of melting curves exhibit significant increases and decreases within melting transition where dA^2/dT^2 values are generally not zero. This difference in

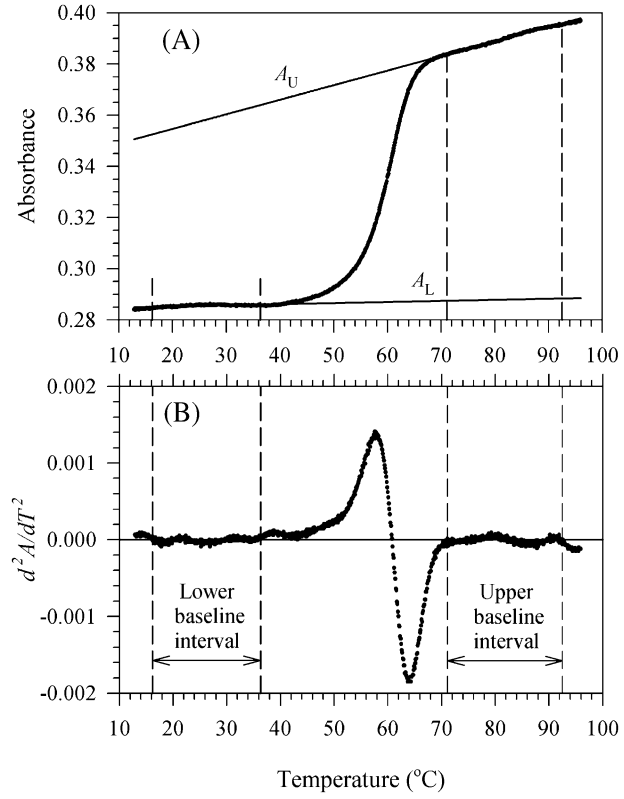

Fig. 2. (A) UV melting profile of d(TATGTATATTTTGTAATCAG) duplex in 1M NaCl-phosphate buffer [8], pH 7.0, C_t=2 μM. (B) Linear regions for baseline fit were selected using the second derivative plot.

values of the second derivative can be utilized for baseline selection. Fig. 2 shows that global maximum and minimum of the second derivative melting curve is observed within melting transition. The lower baseline is obtained from region early in the melting curve, that is, before the second derivative plot reaches global maximum. The region used for the lower baseline fit is linear with zero dA^2/dT^2 values. The upper baseline is analogously selected in the region after global minimum of the second derivative plot where dA^2/dT^2 values are zero.

Because absorbance values contain experimental errors, melting curves need to be smoothed before the second derivative can be calculated. Ideal smoothing procedure removes high-frequency noise present in absorbance values without disturbing overall melting profile. Noise frequency can be estimated from Fourier transformation of absorbance data. We routinely apply Nearly Equal Ripple Filter [13], which smoothes melting curves effectively. The filter is based on running average algorithm where weighting coefficients are adjusted based on noise bandwidth. Melting curves can be smoothed with parameters of 0.03–0.05, 0.05 and 50 db for passband edge, band width and stop band loss [13], respectively. Experimental errors and smoothing artifacts cause small deviations of dA^2/dT^2 values from zero in linear regions. Large deviations of dA^2/dT^2 from zero indicate either insufficient smoothing or nonlinearity. Linear regions for baseline fit are between the first and the last temperature where the second derivative of melting curve crosses zero value (see Fig. 2). This method provides good estimates of baselines and successfully resolves the beginning and the end of melting transition; however, it fails to determine baselines when a part of melting transition is missing.

2.3. Effects of sloping linear baselines on analysis of melting curves

Melting temperatures are occasionally estimated from maxima of dA/dT derivative melting curves, which are widely available in spectroscopic software. It would be useful to know the conversion function between dA/dT and $d\theta/dT$ melting curves. The following equations relate absorbance, fraction of broken base pairs, and linear baselines,

$$A = \theta(A_U - A_L) + A_L \tag{17}$$

$$A_L = a_L + b_L T \tag{18}$$

$$A_U = a_U + b_U T \tag{19}$$

where a_L, b_L, a_U, b_U are intercepts and slopes of lower and upper baselines, respectively.

Combining derivative of Eq. (17) with Eqs. (18) and (19) yields relationship between both derivative melting curves,

$$\frac{dA}{dT} = \frac{d\theta}{dT}\left[T(b_U - b_L) + a_U - a_L\right] + \theta(b_U - b_L) + b_L \tag{20}$$

The second derivative is calculated to find maximum of dA/dT versus temperature plot,

$$\frac{d^2A}{dT^2} = \frac{d^2\theta}{dT^2}\left[T(b_U - b_L) + a_U - a_L\right] + 2\frac{d\theta}{dT}(b_U - b_L) \tag{21}$$

If slopes of lower and upper baselines are same (b_U=b_L), then Eq. (21) reduces to,

$$\frac{d^2A}{dT^2} = \frac{d^2\theta}{dT^2}(a_U - a_L). \tag{22}$$

Eq. (22) indicates that maxima of dA/dT and $d\theta/dT$ melting curves are at identical temperature in this case. However, if slopes of baselines are different, the maxima of both derivative melting curves are not generally at the same temperature because the following relationship holds for dA/dT maximum,

$$\frac{d^2A}{dT^2} = \frac{d^2\theta}{dT^2}\left(T + \frac{a_U - a_L}{b_U - b_L}\right) + 2\frac{d\theta}{dT} = 0 \tag{23}$$

The Eqs. (20)–(23) show the importance of correct analysis of melting curves. Baselines should be subtracted, so that the proper melting temperature, T_m(half), is determined.

3. Experimental results

T_m discrepancies predicted from Eq. (7), Eqs. (10) and (11) were compared with experimental melting data. Melting profiles for duplex d(TGGCGAGCAC) were analyzed and are shown in Fig. 1. The transition enthalpy of this duplex is estimated to be −323 kJ/mol using the nearest-neighbor model and unified parameter set [11]. Therefore, ratio B has a value of $8.3115\times(273.15+56.6)/-323{,}000=-8.49\times10^{-3}$. Using this value with the equations from Table 1, maxima $\theta_{max}(\theta-T)$ and $\theta_{max}(\theta-1/T)$ of derivative melting curves are predicted to be 0.574 and 0.586, respectively. These values are in good agreement with measured experimental values of 0.568 and 0.584. Melting temperatures obtained from maxima of derivative melting curves, $T_m(\theta-T)$ and $T_m(\theta-1/T)$, are shown in Fig. 1C and D. These temperatures were corrected to the temperature of transition midpoint; T_m(half) values of 55.4 and 55.5 °C, calculated from the Eqs. (10) and (11), respectively, are consistent with the experimental value of 55.4 °C.

This analysis was repeated for a set of 14 duplex DNAs, which were studied in our laboratory [8]. Duplexes ranged in length from 15 to 40 base pairs and in G•C base pair content from 30% to 73%. Experimental T_m differences were compared with predicted values. Results are presented in Table 4. The predicted T_m differences (last column of Table 4) are in reasonable agreement with experimentally measured T_m differences (column 7 of Table 4). However, the predicted ΔT_m values are smaller by 0.3–0.4 °C on average. Although this residual inaccuracy is small and within expected experimental error of melting temperatures (0.3 °C), the result suggests that Eq. (10) may not account for the entire experimentally observed difference between $T_m(\theta-T)$ and T_m(half) values. The origin of this residual inaccuracy is uncertain. T_m corrections presented here were developed assuming two-state melting transitions. Differences between predicted and experimental ΔT_ms increase with N_{bp} (see Table 4). Because deviations of duplex melting transitions from the simple two-state model are known to increase with duplex

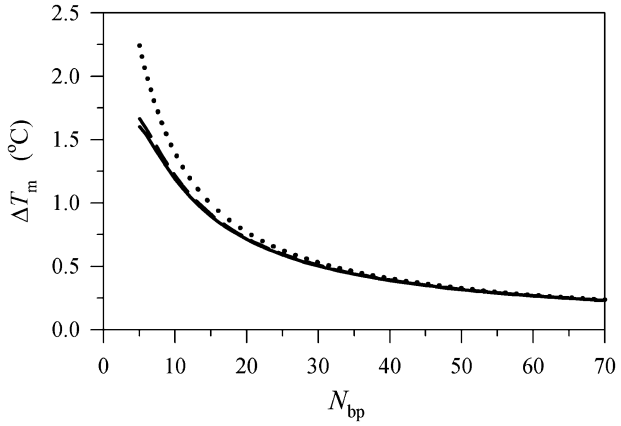

Fig. 3. Differences between the $T_m(\theta-T)$, $T_m(\theta-1/T)$, and T_m(half) temperatures, which were predicted from equations in Table 1 are plotted as a function of oligomer length, N_{bp}. The differences of $T_m(\theta-T)-T_m$(half) (solid line), $T_m(\theta-1/T)-T_m$(half) (dotted line) for non-self-complementary oligomers and the difference $T_m(\theta-T)-T_m$(half) for self-complementary oligomers (dash line) are shown. Results are calculated for a typical DNA duplex oligomer with 50% G•C base pair content, at C_t=1 μM and in 1M Na^+ buffer.

Table 4
Comparison of experimentally measured and predicted discrepancies in melting temperatures (°C)

DNA duplex sequence (5′ to 3′)	N_{bp}	f(G–C)	Predict. $\theta_{max}(\theta-T)$[a]	Exper. T_m(half)	Exper. $T_m(\theta-T)$	Exper. ΔT_m[b]	Predict. ΔT_m[c]
TTCTACCTATGTGAT	15	0.33	0.577	53.7	55.0	1.3	0.9
GCAGTGGATGTGAGA	15	0.53	0.577	63.3	64.4	1.1	0.9
CAGCCTCGTCGCAGC	15	0.73	0.578	72.0	72.9	0.9	0.9
TGATTCTACCTATGTGATTT	20	0.30	0.579	64.4	65.2	0.8	0.7
AGCTGCAGTGGATGTGAGAA	20	0.50	0.579	73.1	74.1	1.0	0.7
CAGCCTCGTTCGCACAGCCC	20	0.70	0.580	78.1	79.5	1.4	0.7
GTTCTATACTCTTGAAGTTGATTAC	25	0.32	0.581	67.7	68.9	1.2	0.6
CTGGTCTGGATCTGAGAACTTCAGG	25	0.52	0.581	75.6	76.6	1.0	0.6
CAGTGGGCTCCTGGGCGTGCTGGTC	25	0.72	0.581	83.4	84.9	1.5	0.6
CTTAAGATATGAGAACTTCAACTAATGTGT	30	0.30	0.581	71.8	72.8	1.0	0.5
AGTCTGGTCTGGATCTGAGAACTTCAGGCT	30	0.50	0.581	80.6	81.5	0.9	0.5
GACCTGACGTGGACCGCTCCTGGGCGTGGT	30	0.70	0.582	86.4	87.3	0.9	0.5
GCAATAGAAAGAGGAAATAATAGTTTTATATTCGACCTAG	40	0.30	0.583	75.4	76.2	0.8	0.4
AGCTGACGCCAAGTCCAAATCTAACCACATGCAAGACACG	40	0.50	0.583	84.5	85.3	0.8	0.4

[a] Predicted from the Eq. (7).
[b] Determined from experimental melting profiles, $\Delta T_m=T_m(\theta-T)-T_m$(half), in 1 M Na^+ buffer, C_t=2 μM [8].
[c] Predicted from the Eq. (10) and experimental $T_m(\theta-T)$. Transition enthalpies were estimated from the nearest-neighbor model and unified parameters [11].

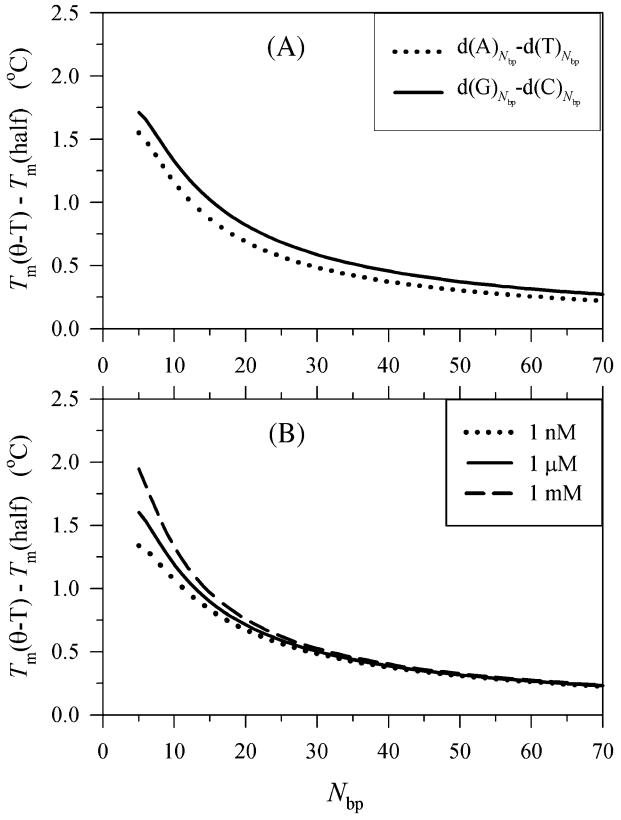

Fig. 4. Differences between the $T_m(\theta - T)$ and T_m(half) melting temperatures, which were predicted from the equations in Table 1 are plotted as a function of oligomer length, N_{bp}. Results are calculated for DNA duplexes in 1M Na^+ buffer using the nearest-neighbor model and unified parameters [11]. (A) Effects of G·C base pair content on T_m differences. Values of ΔT_m for $(dG)_{N_{bp}} - (dC)_{N_{bp}}$ duplexes (solid line) and $(dA)_{N_{bp}} - (dT)_{N_{bp}}$ duplexes (dotted line) are shown at $C_t = 1$ μM. (B) Effects of total single strand DNA concentrations from 1 nM to 1 mM on T_m differences calculated for a typical 50% G•C DNA duplex.

length, we speculate that residual inaccuracy in predicted ΔT_m values may be caused by minor deviations from two-state approximation.

It is useful to know the range of ΔT_m values that arise from use of different analysis methods and if these values vary with oligomer length, G–C base pair content, the type of hybridization, and nucleic acid concentration. We therefore predicted T_m discrepancies for most commonly studied bimolecular hybridizations using expressions from Table 1 as well as the Eqs. (10) and (11). Unified parameters [11] suggest that transition enthalpies and entropies of the representative DNA duplex with 50% G•C content depend on the number of base pairs, N_{bp}, in the following fashion, $\Delta H° = -34.4(N_{bp} - 1) + 10$ kJ/mol and $\Delta S° = -92.5(N_{bp} - 1) + 5.4$ J/(mol K). Using these relationships and assuming that both single strands are present in equal concentrations, T_m discrepancies were predicted. Results displayed in Fig. 3 show that $T_m(\theta - 1/T) - T_m$(half) difference (dotted line) is larger than $T_m(\theta - T) - T_m$(half) difference (solid line), in particular for DNA oligomers shorter than 20 base pairs. Comparing equilibrium of self-complementary oligomers (dash line) with equilibrium of non-self-complementary oligomers (solid line), $T_m(\theta - T) - T_m$(half) differences are very similar. In both kinds of hybridizations, significant dependence of ΔT_m on the length of duplexes is observed. The ΔT_m values are larger for shorter oligomers. They increase from 0.5 °C for a representative 30-mer to 1.5 °C for 6-mer. Fig. 4 shows effects of G•C base pair content and DNA concentration on ΔT_m values. Duplex DNAs with both high and low G•C contents exhibit similar ΔT_m values, i.e., the $T_m(\theta - T) - T_m$(half) difference is weakly dependent on oligomer f(G•C). Effects of DNA concentrations are significant for oligomers shorter than 20 base pairs. We observe that the $T_m(\theta - T) - T_m$(half) difference increases as DNA concentration increases.

4. Conclusions

General equations for correction of T_m values obtained using different methods of data analysis were derived and experimentally verified. For cases where a two-state model holds, these equations can be used to correct inconsistencies between published data sets, permitting improved analysis and collection of melting data from multiple sources. New method of melting curve analysis is suggested which allows automatic selection of sloping baselines.

Acknowledgements

I thank Mark Behlke and Yong You for their comments and review of this manuscript.

References

[1] W.L.A.K. Chiu, C.N. Sze, N.T. Ma, L.F. Chiu, C.W. Leung, S.C.F. Au-Yeung, NTDB: thermodynamic database for nucleic acids, version 2.0, Nucleic Acids Res. 31 (2003) 483–485.
[2] J. Gralla, D.M. Crothers, Free energy of imperfect nucleic acid helices: III. Small internal loops resulting from mismatches, J. Mol. Biol. 78 (1973) 301–319.
[3] J.-L. Mergny, L. Lacroix, Analysis of thermal melting curves, Oligonucleotides 13 (2003) 515–537.
[4] C.R. Cantor, P.R. Schimmel, Biophysical Chemistry: III. The Behavior of Biological Macromolecules, W.H. Freeman and Company, San Francisco, 1980, p. 1132.
[5] L.A. Marky, K.J. Breslauer, Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, Biopolymers 26 (1987) 1601–1620.
[6] G.E. Plum, A.P. Grollman, F. Johnson, K.J. Breslauer, Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex, Biochemistry 34 (1995) 16148–16160.
[7] J.D. Puglisi, I. Tinoco Jr., Absorbance melting curves of RNA, Methods Enzymol. 180 (1989) 304–325.

[8] R. Owczarzy, Y. You, B.G. Moreira, J.A. Manthey, L. Huang, M.A. Behlke, J.A. Walder, Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures, Biochemistry 43 (2004) 3537–3554.
[9] R.M. Wartell, A.S. Benight, Thermal denaturation of DNA molecules: a comparison of theory with experiment, Phys. Rep. 126 (1985) 67–107.
[10] R. Owczarzy, P.M. Vallone, R.F. Goldstein, A.S. Benight, Studies of DNA dumbbells: VII. Evaluation of the next-nearest-neighbor sequence dependent interactions in duplex DNA, Biopolymers 52 (1999) 29–56.
[11] J. SantaLucia Jr., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 1460–1465.
[12] J.M. Sturtevant, Biochemical applications of differential scanning calorimetry, Annu. Rev. Phys. Chem. 38 (1987) 463–488.
[13] J.F. Kaiser, W.A. Reed, Data smoothing using low-pass digital filters, Rev. Sci. Instrum. 48 (1977) 1447–1457.