Determination of melting temperature and temperature melting range for DNA with multi-peak differential melting curves

Dmitri Y. Lando\textsuperscript{a,b,*}, Alexander S. Fridman\textsuperscript{a}, Chun-Ling Chang\textsuperscript{b}, Inessa E. Grigoryan\textsuperscript{c}, Elena N. Galyuk\textsuperscript{a}, Oleg N. Murashko\textsuperscript{d}, Chun-Chung Chen\textsuperscript{b}, Chin-Kun Hu\textsuperscript{b,*}

\textsuperscript{a} Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, 220141 Minsk, Belarus
\textsuperscript{b} Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan
\textsuperscript{c} Department of Physics, Yerevan State University, 0025 Yerevan, Armenia
\textsuperscript{d} Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, 220141 Minsk, Belarus

ARTICLE INFO

Article history:
Received 21 August 2014
Received in revised form 9 January 2015
Accepted 22 January 2015
Available online 30 January 2015

Keywords:
DNA melting temperature
DNA temperature melting range
High-resolution melting profiles
Plasmid DNA
Calf thymus DNA

ABSTRACT

Many factors that change the temperature position and interval of the DNA helix–coil transition often also alter the shape of multi-peak differential melting curves (DMCs). For DNAs with a multi-peak DMC, there is no agreement on the most useful definition for the melting temperature, \(T_m\), and temperature melting width, \(\Delta T\), of the entire DNA transition. Changes in \(T_m\) and \(\Delta T\) can reflect unstable variation of the shape of the DMC as well as alterations in DNA thermal stability and heterogeneity. Here, experiments and computer modeling for DNA multi-peak DMCs varying under different factors allowed testing of several methods of defining \(T_m\) and \(\Delta T\). Indeed, some of the methods give unreasonable “jagged” \(T_m\) and \(\Delta T\) dependences on varying relative concentration of DNA chemical modifications \((r_0)\), [Na\textsuperscript{+}], and GC content. At the same time, \(T_m\) determined as the helix–coil transition average temperature, and \(\Delta T\), which is proportional to the average absolute temperature deviation from this temperature, are suitable to characterize multi-peak DMCs. They give smoothly varying theoretical and experimental dependences of \(T_m\) and \(\Delta T\) on \(r_0\), [Na\textsuperscript{+}], and GC content. For multi-peak DMCs, \(T_m\) value determined in this way is the closest to the thermodynamic melting temperature (the helix–coil transition enthalpy/entropy ratio).

© 2015 Published by Elsevier Inc.

In general, the temperature melting range, \(\Delta T\), is related to the size of the temperature interval spanned by DNA helix–coil transition, and the melting temperature, \(T_m\), is somewhere in its middle. For the majority of the definition methods, approximately 80% of base pairs are melted out in the temperature interval \((T_m - \Delta T)/2, T_m + \Delta T/2\). When the composition of solution used for DNA melting experiments is changed, or DNA is chemically modified, the shift of the melting temperature characterizes the impact of these factors on the thermal stability of the double helix and the temperature melting range reflects their influence on DNA thermal heterogeneity.

This study was carried out because there is no agreement on the most useful definition for the melting temperature, \(T_m\), and temperature melting width, \(\Delta T\), whose dependences on various factors can be properly defined if DNA differential melting curves (DMCs)\textsuperscript{1} have multi-peak fine structure. Indeed, theoretical and experimental studies demonstrate that the increasing concentration of DNA chemical modifications that alter \(T_m\) and \(\Delta T\) also strongly changes the shape of multi-peak DMC [1–4]. Therefore, a change in \(T_m\) and \(\Delta T\) also reflects this alteration in the shape besides the relevant change in the thermal stability and thermal heterogeneity of the double helix. However, there was no theoretical or experimental evidence supporting or disproving this viewpoint. Here, we have carried out such a study.

The \(T_m\) and \(\Delta T\) dependences on GC content, [Na\textsuperscript{+}], and per nucleotide concentration of chemical modifications \((r_0)\) randomly distributed along DNA have been calculated in several ways described below. Indeed, some of the ways give unreasonable “jagged” dependences of the \(T_m\) and \(\Delta T\) on \(r_0\), [Na\textsuperscript{+}], and GC content. At the same time, \(T_m\) taken as the helix–coil transition average temperature and \(\Delta T\) that is proportional to the average

\textsuperscript{1} Abbreviations used: DMC, differential melting curve; RMSEA, root mean square error of approximation; SD, standard deviation.

\textbf{Keywords:} DNA melting temperature, DNA temperature melting range, High-resolution melting profiles, Plasmid DNA, Calf thymus DNA

\textbf{Article info}:
Received 21 August 2014
Received in revised form 9 January 2015
Accepted 22 January 2015
Available online 30 January 2015

\textbf{Abstract}:
Many factors that change the temperature position and interval of the DNA helix–coil transition often also alter the shape of multi-peak differential melting curves (DMCs). For DNAs with a multi-peak DMC, there is no agreement on the most useful definition for the melting temperature, \(T_m\), and temperature melting width, \(\Delta T\), of the entire DNA transition. Changes in \(T_m\) and \(\Delta T\) can reflect unstable variation of the shape of the DMC as well as alterations in DNA thermal stability and heterogeneity. Here, experiments and computer modeling for DNA multi-peak DMCs varying under different factors allowed testing of several methods of defining \(T_m\) and \(\Delta T\). Indeed, some of the methods give unreasonable “jagged” \(T_m\) and \(\Delta T\) dependences on varying relative concentration of DNA chemical modifications \((r_0)\), [Na\textsuperscript{+}], and GC content. At the same time, \(T_m\) determined as the helix–coil transition average temperature, and \(\Delta T\), which is proportional to the average absolute temperature deviation from this temperature, are suitable to characterize multi-peak DMCs. They give smoothly varying theoretical and experimental dependences of \(T_m\) and \(\Delta T\) on \(r_0\), [Na\textsuperscript{+}], and GC content. For multi-peak DMCs, \(T_m\) value determined in this way is the closest to the thermodynamic melting temperature (the helix–coil transition enthalpy/entropy ratio).

© 2015 Published by Elsevier Inc.
absolute temperature deviation from this temperature give smoothly varying dependences even if DMCs are multi-peak, and the same factors cause a strong alteration in their multi-peak shape. For multi-peak DMCs, $T_m$ value determined in this way is the closest to the thermodynamic melting temperature (the ratio of the helix–coil transition enthalpy and entropy).

**Materials and methods**

**Experiment**

Calf thymus DNA from Sigma–Aldrich was used after additional purification. High-resolution melting profiles of DNA were obtained using a model CSC 6300 NanoDSC differential scanning calorimeter (Calorimetry Sciences, USA) with a cell volume of 0.3 ml. In the differential scanning calorimetry experiments, further primary processing, and determination of the enthalpy and entropy of the helix–coil transition, we followed standard procedures [5–7].

In the melting studies at various [Na$^+$], DNA concentration was 1 mg/ml. Melting media included buffer used in our earlier studies (0.05 mM ethylenediaminetetraacetic acid [EDTA] and 1 mM Na$_2$CO$_3$, pH 7.0 [48]), and NaCl was added to obtain [Na$^+$] from the interval 10 to 210 mM. We have shown that replacement of 1 mM Na$_2$CO$_3$ with 5 mM cacodylate or with any other concentrations of Na$_2$CO$_3$ from 0.05 mM to 5 mM does not alter DNA melting profiles if total [Na$^+$] and neutral acidity are conserved.

For the study of platination on the melting parameters, DNA at a concentration of 1.2 mg/ml was incubated in 10 mM NaClO$_4$ for 48 h at 37°C in the dark at approximately pH 6.0 with cisplatin (Sigma–Aldrich). Pt/nucleotide molar ratio ($r_0$) was 0.001 to 0.05. The melting was carried out in the same buffer at [Na$^+$] = 210 mM. In these ionic conditions, the destabilizing influence of cisplatin on melting temperature is the strongest [8].

**Computer modeling**

We examined a DNA molecule of $N_{bp}$ base pairs (bp) that can be unmodified or include $\omega$ chemically modified sites located at base pairs with numbers $n_1$, $n_2$, ..., $n_{\omega}$ that corresponds to the per nucleotide concentration of the modifications $r_0 = \omega/(2N_{bp})$. It is supposed that each modification locally changes the free energy of the helix–coil transition by $\delta G(T)$. A slightly modified conventional Poland–Fixman–Freire procedure [9–12] was used for calculation of DNA melting curves of unmodified and chemically modified DNA.

The following DNA parameter values were used for calculation: total number of base pairs, $N_{bp} = 10^4$ bp; fraction of GC base pairs, GC = 0.25 to 0.75; loop entropy factor for a loop of $L$ base pairs formed between boundaries of internal melted region, $f(L) = (L + 1)^{-1.7}$ [11,12]; factor of cooperativity or statistical weight assigned to the boundaries of an internal melted region bordered by helical ones, $\sigma_{coop} = 4.5 \times 10^{-5}$ [11]; strand association parameter, $\beta = \sigma_{coop}$ [13]; concentration of base pairs $C_{bp} = 3 \times 10^{-5}$ M that corresponds to $C_{DNA} = 0.02$ mg/ml and to the concentration of DNA strands $C_{str} = 2C_{bp}/N_{bp} = 6 \times 10^{-8}$ M; melting temperatures of AT and GC base pairs and the enthalpy of the helix–coil transition, $T_{AT} = 65.2$ °C and $T_{GC} = 107.8$ °C, $\Delta H_{AT} = 8.408$ kcal/(mole bp), $\Delta H_{GC} = 9.467$ kcal/(mole bp) [11]. Using these parameter values, the melting curve $\psi(T)$ (the fraction of melted base pairs) and differential melting curve $\psi'(T)$ with pronounced fine structure were calculated. Then the melting temperature and temperature melting range were determined using various definitions.

For the sake of clarity and ease of analyzing the results of calculation, we have used a model that does not include stacking heterogeneity. For a specified shape and temperature location of DMC, the $T_m$ and $\Delta T$ are independent of the melting model.

For a given GC content, DNA sequences of AT and GC base pairs were produced with a random number generator. For a random generated sequence or for a known real sequence of $N_0$ base pairs, 2$r_0N_{bp}$ random sites of chemical modification were generated for a given per nucleotide concentration ($r_0$).

As an example of a real sequence, plasmid pBR 322 (4361 bp) linearized with EcoRI was used in calculations.

**Various definitions of DNA melting temperature $T_m$**

In general, the temperature melting range $\Delta T$ is related to the size of the temperature interval spanned by the DNA helix–coil transition, and the melting temperature $T_m$ is somewhere in its middle. In Fig. 1, the illustration of various definitions of melting temperature ($T_m$) is shown in the picture of melting curve $\psi(T)$ and differential melting curve $\psi'(T)$. If DMC demonstrates a single peak, then the position of the maximum can be taken as the melting temperature ($T_m = T_{max}$) [5,6,14,15] (Fig. 1). If DMC includes several peaks [7,16–18] and they do not change their relative position under the influence of a factor that shifts $T_m$ [16–18] (see also Fig. 2), then the position of the highest peak can be considered as the melting temperature, which is determined with Eq. (1):

$$\psi'(T_{max}) = \max[\psi'(T)].$$

(1)

A strong change in $T_m$ without a change in the shape of DMC occurs, for example, under alteration in Na$^+$ ion concentration [16,17] (see also Fig. 2) or in the concentration of formamide [18]. For some melting experiments, it is convenient to define $T_m$ as the temperature position of the highest point of DMC if the end of melting occurs at a temperature too high to be recorded due to instrumental limitations. In this case, other definitions of $T_m$ cannot be used because the whole melting curve is required for their determination.

The most popular definition of DNA melting temperature is the temperature that corresponds to the half of melted base pairs

![Fig1](image-url)
\( T_m = T_{0.5} \) or to a half-change of the parameter used for registration of the helix–coil transition:
\[
\psi(T_{0.5}) = 0.5.
\]

The melting temperature can be determined as the average temperature of the helix–coil transition or rather the temperature mean of DMC \( (T_{\text{int}}) \):
\[
T_{\text{int}} = \int_{T_1}^{T_2} T \cdot \psi'_f(T) \, dT,
\]
where \( \psi'_f(T) \, dT \) is considered as the fraction of base pairs melted out in the temperature interval \( [T_1, T_2] \) and \( T_1 \) and \( T_2 \) are the temperatures that correspond to the start and end of DNA melting, respectively.

\( T_{\text{int}} \) was used to characterize asymmetric DMCs \[19–21\]. For \( T_{\text{int}} \) determination, DMC is considered as a density function of temperature and the melting curve is considered as the corresponding cumulative function. The definition is especially helpful for evaluation of the DNA average GC content \[19,22\].

If DMC includes a single symmetric peak, or the asymmetry is negligible, then all \( T_m \) definitions give very close values (Fig. 1B). It is not true for multi-peak DMCs.

The most general thermodynamic definition of the melting temperature \[6\] is given by Eq. (4):
\[
T_{\text{therm}} = \Delta H/\Delta S,
\]
where \( \Delta H \) and \( \Delta S \) are enthalpy and entropy of the helix–coil transition measured with differential scanning calorimetry or calculated (see comments on Eqs. (16) and (17) in Ref. [6]). This definition is independent of the fine structure of DMC but requires knowledge of the enthalpy and entropy.

**Definitions of temperature melting range \( \Delta T \)**

In general, the temperature melting range \( \Delta T \) is related to the size of the temperature interval spanned by the DNA helix–coil transition. The most obvious definition of the temperature melting range is the difference between the temperatures that correspond to the end \( (T_e) \) and start \( (T_s) \) of DNA melting (Fig. 1A). However, defining those two temperatures for an experimental melting curve is not always straightforward. In the experiment, different representations of melting curve give different positions of \( T_s \) and \( T_e \). It is well seen from comparison of Fig. 1A and B. In theoretical and some experimental studies, the positions of \( T_s \) and \( T_e \) are strongly shifted to lower and higher temperatures because of low temperature opening of the double helix and high temperature single-stranded stacking \[23,24\]. Moreover, in long DNAs of higher organisms, a very small fraction of base pairs is included in long regions (~1000 bp) in which nearly all base pairs are \( AT \) or \( GC \) \[25\]. Strictly speaking, it means that \( T_s \) is close to \( T_{AT} \) and that \( T_e \) is close \( T_{GC} \). Because the start and end of a helix–coil transition can be ambiguous, the temperature melting range \( \Delta T \) is often considered as a temperature interval \( (T_m - \Delta T/2, T_m + \Delta T/2) \) where the majority of base pairs is melted out. Those definitions are based on the properties of the central part of the melting curve and not on the location of the extreme points \( T_s \) and \( T_e \).

In early studies \[26\], \( \Delta T \) was defined as \( \Delta T_{\text{ng}} \), the distance between intersections with horizontal lines \( \psi = 0 \) and \( \psi = 1 \) of the tangent to melting curve at the point of melting temperature (Eq. (5)) and Fig. 1A):
\[
\Delta T_{\text{ng}} = 1/\psi'(T_{0.5}).
\]

However, this definition method demonstrated a strong ruggedness of concentration dependences of \( \Delta T_{\text{ng}} \) in computer modeling of melting behavior of DNA–ligand complexes \[27\]. The ruggedness is much weaker for \( \Delta T_{25–75} \), which is twice the difference between temperatures \( T_{0.75} \) and \( T_{0.25} \) corresponding to the fractions of melted base pairs 0.75 and 0.25, respectively:
\[
\Delta T_{25–75} = 2(T_{0.75} - T_{0.25}).
\]

In the case of smooth single-peak DMCs, \( \Delta T_{25–75} \) is slightly larger than \( \Delta T_{\text{ng}} \) (see below).

Another definition of the temperature melting range \( \Delta T \) often used in various studies is the full width at half maximum of the cumulative DMC \( (\Delta T_{\text{hh}}) [5,6] \). For a multi-peak DMC, we have defined \( \Delta T_{\text{hh}} \) as the difference between maximal \( (T_{1/2\text{max}}) \) and minimal \( (T_{1/2\text{min}}) \) temperatures corresponding to the half of the maximal height of DMC (Fig. 1B):
\[
\Delta T_{\text{hh}} = (T_{1/2\text{max}}) - (T_{1/2\text{min}}).
\]

The temperature standard deviation of DMC was also used to measure a temperature melting range \[20\]:
\[
\Delta T_{\text{sd}} = \left( \int_{T_1}^{T_2} (T - T_{\text{int}})^2 \cdot \psi'_f(T) \, dT \right)^{1/2}.
\]

We made use of another definition of the temperature melting range for DNA complexes with platinum compounds, which is proportional to the average absolute deviation from the average temperature \[8\]:
\[
\Delta T_{\text{ana}} = \int_{T_1}^{T_2} |T - T_{\text{int}}| \cdot \psi'_f(T) \, dT.
\]

If DMC is described by the Gaussian function, then all definitions of the melting temperature give the same value (as for any symmetrical DMC) and
\[
\psi'_f(T) = (\sqrt{2\pi} \cdot \sigma)^{-1} \cdot \exp\{-0.5[(T - T_m)/\sigma]^2\}
\]
and
\[
\psi'_f(T) = \Delta T_{\text{ng}}^{-1} \cdot \exp\{-\pi \cdot (|T - T_m|/\Delta T_{\text{ng}})^2\}.
\]
Expressing all $\Delta T$ definitions in terms of $\sigma$, one obtains:

\[
\Delta T_{\text{int}} = 0.7979 \sigma. \tag{16}
\]

The first three $\Delta T$ definitions give close values (Eqs. (12)–(14)) and the temperature interval ($T_m - \Delta T/2$, $T_m + \Delta T/2$) spanning the region where the majority of base pairs is melted out (Fig. 1). The deviation from $\Delta T_{\text{int}}$ is less than 10% for $\Delta T_{25-75}$ and $\Delta T_{\text{th}}$. However, the values of $\Delta T_{\text{sa}}$ and $\Delta T_{\text{int}}$ are small. Therefore, we multiplied $\Delta T_{\text{sa}}$ and $\Delta T_{\text{int}}$ by the factors $(2\pi)^{1/2}$ and $\pi$, respectively, to make them equal to $\Delta T_{\text{int}}$ given by Eq. (12).

Thus, the final expressions for $\Delta T_{\text{sa}}$ and $\Delta T_{\text{int}}$ are the following:

\[
\Delta T_{\text{sa}} = (2\pi)^{1/2} \left[ \int_{T_l}^{T_r} (T - T_{\text{int}})^2 \cdot \varphi(T) \cdot dT \right]^{1/2} \tag{17}
\]

\[
\Delta T_{\text{int}} = \pi \cdot \int_{T_l}^{T_r} |T - T_{\text{int}}| \cdot \varphi(T) \cdot dT. \tag{18}
\]

Quantitative evaluation of suitability of different $T_m$ and $\Delta T$ definitions

A factor that alters the melting temperature ($T_m$) and temperature melting range ($\Delta T$) often also strongly changes the fine multi-peak structure of the DMC [1–4]. Definitions of $T_m$ and $\Delta T$ can be used to characterize a change in DNA thermal stability and thermal heterogeneity caused by this varying factor ($x$) if they produce smoothly varying dependences $T_m(x)$ and $\Delta T(x)$ independent of alteration in the fine structure. Varying per nucleotide concentration of chemical modifications ($x$), GC content, $[\text{Na}^+]$, time or dose of irradiation, and the like are such factors. Each point of the dependences $T_m(x)$ and $\Delta T(x)$ is obtained from the corresponding DMC.

In Fig. 3, all $T_m$ and $\Delta T$ dependences on $[\text{Na}^+]$ (except $\Delta T_{\text{sa}}([\text{Na}^+])$) can be considered as smoothly varying. In Fig. 5B (see Results and discussion), $T_{\text{int}}([\text{GC}])$ is smoothly varying, but $T_{\text{max}}([\text{GC}])$ is “jagged.” A “smooth” arrangement of points $T_m(x_i)$ is characterized with low deviation from their polynomial approximant $T_{\text{mapp}}(x_i)$. In this case, the root mean square error of approximation (RMSE) given by Eq. (19) must be small:

\[
\text{RMSEA}[T_m(x)] = \left( \frac{1}{N_p} \sum_{i=1}^{N_p} [T_m(x_i) - T_{\text{mapp}}(x_i)]^2 \right)^{1/2}, \tag{19}
\]

where $N_p$ is the number of points $T_m(x_i)$ or, more exactly, the number of DMCs registered at different $x_i$ and used for determining $T_m(x_i)$.

Thus, the RMSEA value can be a measure of the “quality” of a given definition. If RMSEA is much lower than experimental error, as for $T_{\text{int}}([\text{GC}])$ in Fig. 5B (RMSEA = 0.0039 °C), then the quality is high. At the same time, RMSEA = 0.59 °C for $T_{\text{max}}([\text{GC}])$; that is, it is 100 times higher. Thus, in the considered case, $T_{\text{max}}$ demonstrates low “quality” and is not suitable for definition of the melting temperature. Here, linear polynomial was used for RMSEA calculation. Depending on the type of $T_m$ and $\Delta T$ dependences, various polynomial approximants were used.

In some cases, the dependence of $\Delta T(\text{GC})$ was weak. Therefore, to compare different definitions of $\Delta T(\text{GC})$, the mean of $\Delta T(\text{GC})$ and standard deviation ($\Delta T_{\text{sa}} \pm SD$) were used instead of RMSEA.

For all definitions of $T_m$, the comparison with the conventional expression given by Eq. (20),

\[
T_{\text{exp}}(\text{GC}) = T_{\text{AT}} + (T_{\text{GC}} - T_{\text{AT}}) \cdot \text{GC}, \tag{20}
\]

was also carried out for calculated DMCs. The mean and standard deviation of the difference between $T_m$ and $T_{\text{exp}}$ denoted by $[\text{GC}] - [\text{exp}]$ and $T_{\text{exp}}([\text{GC}])$ were calculated. A deviation from linearity for $T_m$ is caused by a change in the shape of the DMC as $[\text{Na}^+]$ decreases from 0.011 to 0.01 M. Peak 3, which is higher than the half of the maximal height of DMC, becomes lower than this value and is omitted from the $\Delta T_{\text{int}}$ melting range (see Figs. 1B and 2A).

Results and discussion

Influence of $[\text{Na}^+]$ on melting parameters of calf thymus DNA

As an example of various definitions of $T_m$ and $\Delta T$, let us consider the influence of ionic strength and chemical modifications formed by the antitumor drug cisplatin on DMC, melting temperature, and temperature melting range of calf thymus DNA. DNA thermograms after subtraction of buffer baseline, sample baseline, and normalization to the area are shown in Fig. 2. It is seen that Na$^+$ concentration strongly influences the position of melting curve with a minimal change in the shape if $[\text{Na}^+] > 0.011$ M. When $[\text{Na}^+]$ becomes smaller than 0.011 M, a change in the shape occurs (Fig. 2A). At the same time, cisplatin suppresses the fine structure (Fig. 2B).

The melting temperature and temperature melting range calculated with different methods for the DMCs from Fig. 2A are represented in Fig. 3. It is seen that both parameters give smoothly varying $[\text{Na}^+]$ dependences for any determination method. In general, the melting temperature monotonously increases with $[\text{Na}^+]$, and the temperature melting range decreases (Fig. 3A). The only case of non-monotony for $\Delta T_{\text{int}}([\text{Na}^+])$ (Fig. 3B) is caused by a strong change in the shape of the curve $\psi(T)$ registered at 0.01 M.
Na" relative to all other curves (0.011 ≤ [Na"] ≤ 0.21 M). Peak 3, which is higher than the half of the maximal height of DMC, becomes lower than this value at [Na"] = 0.01 M and stops to influence ΔT
th (Fig. 2). At the same time, the [Na"] dependences of ΔT
int, ΔT
25-75, ΔT
int, and ΔT
r are smoothly varying, and ΔT
int and ΔT
r almost coincide. Thus, all definitions of the melting temperature and temperature melting range except ΔT
th give parameter values that are changed with [Na"] in the same smoothly varying and monotonous way.

Although the difference between various melting temperature determinations reaches 1.5 °C (Fig. 3A), the shift of melting temperature caused by a change in [Na"] from one given value to another is independent of the determination method (Fig. 3A). All of them are smoothly varying. Choosing the T
m definition that gives values closer to T
therm helps us to make a choice of the most suitable among them. T
int is much closer to T
therm in comparison with other definitions. The average difference between them is smaller at 0.05 °C.

For platinated DNA, the relationship between different definitions of T
m and ΔT are the same (Fig. 4). The r
b dependences of melting temperatures are linear; T
int(r
b) and T
therm(r
b) are very close (Fig. 4A). A strong non-monotony of ΔT
int(r
b) is also caused by a decrease of peak 3 as r
b changes from 0.005 to 0.01 (Fig. 4B).

Melting parameters for random sequences with approximately the same GC ~ 0.5

So long as the dependence of T
int(GC) was measured in various DNA thermodynamic studies (see Ref. [16] and references therein), let us consider which of the definitions give a correct dependence T
int(GC). This issue was considered before on the basis of experimental data [19,22]. In this part of the work, we have obtained 10 sequences of N
bp = 10
8 bp in the narrow range of GC = 0.4906 to 0.5067. The multi-peak shapes of calculated DMCs are different (Fig. 5A).

As follows from Fig. 5B, the curve T
int(GC) is located lower and parallel to the line corresponding to the conventional expression defined by Eq. (20), and their mean difference is almost constant (−0.070 ± 0.0047 °C) (Fig. 5B).

At the same time, T
int(GC) is located slightly higher than T
exp(GC), and there is no parallelism. In contrast to T
int(GC), the SD of the difference is larger than the mean (0.066 ± 0.096). For T
max, the excess over T
exp and the deviation from parallelism are even stronger (0.62 ± 0.59). The maximal difference reaches 1.43 °C. The RMSEA[T
int(GC)] values for linear approximation shown in Table 1 correspond to these results and are equal to 0.084, 0.0039, and 0.38 for T
b, T
int, and T
max, respectively.

This example demonstrates that distortions of T
int(GC) and especially T
exp(GC) caused by varying fine structure of DMCs neither disturb monotony nor exceed errors of experimental studies. At the same time, a change in multi-peak DMC shape strongly influences T
max(GC) (Fig. 5B). Only T
int can be used in theoretical studies to elucidate various weak effects. Here, the dependence T
int(GC) reliably demonstrates the lowering in T
int(GC) relative to T
exp(GC) that is caused by strand separation after full DNA melting. Indeed, the prohibition of strand separation in the model of DNA melting strongly reduces the absolute value of the difference T
int(GC) – T
exp(GC) (not shown). At the same time, T
int(GC) and T
max(GC) are not changed under such prohibition because their values depend on the part of the melting curve that corresponds to ϑ < 0.7. For long DNA, the strand separation usually begins at ϑ > 0.9. Calculations demonstrate that a decrease in N
bp from 10
4 to 5·10
4 bp causes 2-fold strengthening of the destabilizing effect of the strand separation on T
int (not shown). In contrast to the current case, strand dissociation influences T
b of shorter DNAs with N
bp > 1000 bp [28–30].

In the considered very small range of GC, ΔT values are independent of GC, but dependent on sequence and corresponding multi-peak shape of DMCs. For ΔT
int and ΔT
r, the SDs are approximately 10% of the means. For other definitions of ΔT, the diversity is much larger and SD exceeds 30% for ΔT
int (Table 1). Nevertheless, all definitions of ΔT except ΔT
th give the close mean values.

Thus, the influence of the shape of multi-peak DMC is minimal for T
int(GC), ΔT
int(GC), and ΔT
r (GC). Although RMSEA[T
int(GC)] is 20 times larger than RMSEA[T
int(GC)], its value (as well as deviation from T
exp) does not exceed experimental error (Table 1).
Dependences \( T_m(GC) \) and \( \Delta T(GC) \): Simple case

To avoid influence of GC content on DMC fine structure in the previous part of our study, \( T_m(GC) \) and \( \Delta T(GC) \) were calculated for very close GC values. In the current part, 45 random sequences with 0.25 \( \leq \) GC \( \leq 0.75 \) were generated, and DMCs were calculated to study the dependences \( T_m(GC) \) and \( \Delta T(GC) \). To our surprise, the results shown in Fig. 6 and Table 1 demonstrate general similarity with the previous part. The mean differences between \( T_m(GC) \), \( T_{av}(GC) \), and \( T_{max}(GC) \) and \( T_{exp}(GC) \) are equal to \( -0.065 \pm 0.0062 \), \( 0.111 \pm 0.0822 \), and \( 0.68 \pm 0.44 \) °C, respectively (Table 1). Corresponding RMSEAs are also close. As in the previous case, deviations of \( T_{av}(GC) \) and \( T_{inf}(GC) \) from \( T_{exp}(GC) \) are smaller than 0.11 °C (Table 1); that is, they do not exceed experimental error. Thus, a strong change in fine structure of DMC does not influence the dependence of \( T_{av}(GC) \) and \( T_{inf}(GC) \).

At the same time, the mean values of \( \Delta T \) are approximately 10% lower than in the previous case of narrow GC range (GC \( \approx 0.5 \)) (Table 1). This decrease in the temperature melting range is caused by a decrease in sequence heterogeneity as GC content deviates from 0.5. However, in whole, \( \Delta T \) is independent of GC for simple random sequences that include one random block.

Dependences \( T_m(GC) \) and \( \Delta T(GC) \): Complex case

For the two types of sequences described above, the distortions found for melting temperatures measured as \( T_m \) and \( T_{int} \) are negligible (0.1 °C) when it comes to experimental studies. Therefore, we should analyze the suitability of these definitions for more complex sequences and corresponding more complex structures of DMCs when both \( T_m \) and \( \Delta T \) are strongly dependent on the GC content. Such a complex fine structure can be produced by sequences that include two random regions (blocks) of \( m_{GC1} \) and \( m_{GC2} \) base pairs that are different in GC content (GC1 < GC2). The total average GC content (GCav) of whole sequence is calculated using Eq. (21):

\[
GC_{av} = [m_{GC1} \cdot GC_1 + m_{GC2} \cdot GC_2]/N_{bp} = NGC/N_{bp}.
\]

The two sequences with \( GC_{av} = GC_1 \) or \( GC_{av} = GC_2 \) include only one random block. The sequences with other GC values include the two random blocks. In our calculations, GC1 = 0.25 and GC2 = 0.75. Three of the DMCs calculated for \( GC_{av} \) values 0.25 (\( m_{0.25} = 10^3 \) bp), 0.60 (\( m_{0.60} = 3000 \) bp), and 0.75 (\( m_{0.75} = 0 \) bp) are shown in Fig. 7.

As follows from Fig. 7, the sequence that includes two blocks demonstrates a two-component DMC, and each component includes several narrow peaks (curve \( GC_{av} = 0.6 \)). The results of calculation for three \( T_m \) definitions are shown in Fig. 8A as their dependence on \( GC_{av} \). It is seen that only \( T_m \) (Eq. (3)) demonstrates linear dependence on \( GC_{av} \) and can be used for a complex multi-peak shape of DMC. For linear approximation of \( T_{inf}(GC_{av}) \), the value of RMSEAt\( T_{inf}(GC_{av}) \) is equal to 0.051 °C. It is 10 times worse than for the two types of sequences described above, but it is lower than experimental error in \( T_m \) determination. At the same time, the difference between \( T_{int} \) and \( T_{exp} \) cannot be determined with the same high validity for this type of sequence (SD is comparable to the mean of the difference [Table 1]). The parameters \( T_{int} \) and \( T_{max} \) are not suitable in this case because the function \( T_{inf}(GC_{av}) \) does not demonstrate linear behavior for two-component curves, and \( T_{max}(GC_{av}) \) poorly reflects a change in DNA thermal stability with \( GC_{av} \) (Fig. 8A).

The use of \( T_m \) (Eq. (3)) solves the problem of representing \( T_m \) dependence on the concentration of irresponsibly bound large ligands such as basic polypeptides and histones that cover long DNA regions [31–33]. They give rise to the two-component melting

![Fig. 6. GC dependences for various definitions of the melting temperature for 45 sequences with GC content from 0.25 to 0.75 that include one random block. In the plot, the curves \( T_{inf}(GC) \) and \( T_{exp}(GC) \) = \( T_{int} \) + (\( T_{exp} \) – \( T_{int} \)) GC look as coinciding ones. The difference between them is \( -0.065 \pm 0.0062 \) °C. Primary points are shown for \( T_{max}(GC) \) only because the points of the dependences \( T_{av}(GC) \) and \( T_{inf}(GC) \) are very close to \( T_{exp}(GC) \).](image-url)
and well approximated with cubic polynomial \( \text{RMSEA}[\Delta T_{\text{int}}(GC_{\text{av}})] = 0.71 \) °C. For \( \Delta T_T(GC_{\text{av}}) \), the smoothness is worse \( \text{RMSEA}[\Delta T_T(GC_{\text{av}})] = 1.67 \) °C. Thus, in the case of complex shape of DMC, when multi-peak fine structure appears at background of two-component transition, the use of \( \Delta T_{\text{int}} \) is more appropriate than \( \Delta T_T \), although they are equivalent in all cases that were described above and are reported below. It would seem that the obtained RMSEAs are too high. However, the \( \Delta T \) values are also high in the case of the two-component transition, and the ratios of the RMSEAs to the maximal values of \( \Delta T_{\text{int}} \) and \( \Delta T_T \) are not large at 0.022 and 0.061, respectively.

Modeling of melting parameters for chemically modified DNA

A similar problem of correct determination of the melting temperature and temperature melting range arises in thermodynamic studies of DNAs chemically modified with antitumor platinum compounds [34] that strongly change the fine structure of DMC [1–4,8] and other DNA binders [35,36].

As in the previous case, computer modeling allows us to test the suitability of different definitions of the melting parameters for DNA that includes chemical modifications. We have considered the dependences of \( T_{\text{m}} \) and \( \Delta T \) on relative per nucleotide concentration (\( r_b \)) of chemical modifications randomly distributed along DNA with a given primary structure.

For the modeling of impact of chemical modifications on DNA melting, the sequence of the EcoRI-cut pBR322 plasmid DNA was used. It was supposed that a chemical modification increases the free energy of the helix–coil transition by 2.5 kcal per mole of modifications. The calculated DMCs are shown in Fig. 9. It is seen that an increase in the melting temperature and in the temperature melting range with \( r_b \) is accompanied by a strong change in the fine multi-peak structure of the pBR322 DMC. As follows from Fig. 10A and Table 1, \( T_{\text{int}}(r_b) \) demonstrates the most smoothly varying dependence with the minimal RMSEA (0.03 °C). The RMSEA of \( T_{0.5}(r_b) \) is nearly 10 times higher, but it is not larger than usual experimental errors. Therefore, \( T_{0.5} \) can also be used to study the influence of various chemical modifications on DNA thermal stability [37].

Using Eq. (19), we have calculated the values of the RMSEA for the quadratic approximation of \( \Delta T(r_b) \). The results of calculation for the temperature melting range of chemically modified EcoRI-cut pBR322 are shown in Fig. 10B and Table 1. As for other cases, the dependences \( \Delta T_{\text{int}}(r_b) \) and \( \Delta T_{T}(r_b) \) are the most smoothly varying, and their RMSEAs are minimal and comparable to experimental errors.

For the temperature melting range \( \Delta T \), only \( \Delta T_{\text{int}} \) and \( \Delta T_T \) demonstrate smoothly varying \( GC_{\text{av}} \) dependences for both one- and two-component helix–coil transition (Fig. 8B and Table 1). The dependence \( \Delta T_{\text{int}}(GC) \) is the most smoothly varying curve because of formation of two types of blocks of different stability. Formally, their DMCs are similar to that shown in Fig. 7 (curve \( GC_{\text{av}} = 0.6 \)). However, low- and high-temperature components correspond to free (uncovered) DNA and DNA covered with these irreversibly bound ligands. The dependence of the melting temperature \( T_{0.5} \) on ligand concentration given by Eq. (2) has a sigmoid shape and is poorly informative (Fig. 8A). In contrast, \( T_{\text{int}} \) (Eq. (3)) shows gradual linear increase with concentration the same as with \( GC_{\text{av}} \) in Fig. 8A.

For the temperature melting range \( \Delta T \), only \( \Delta T_{\text{int}} \) and \( \Delta T_T \) demonstrate smoothly varying \( GC_{\text{av}} \) dependences for both one- and two-component helix–coil transition (Fig. 8B and Table 1). The dependence \( \Delta T_{\text{int}}(GC) \) is the most smoothly varying curve because of formation of two types of blocks of different stability. Formally, their DMCs are similar to that shown in Fig. 7 (curve \( GC_{\text{av}} = 0.6 \)). However, low- and high-temperature components correspond to free (uncovered) DNA and DNA covered with these irreversibly bound ligands. The dependence of the melting temperature \( T_{0.5} \) on ligand concentration given by Eq. (2) has a sigmoid shape and is poorly informative (Fig. 8A). In contrast, \( T_{\text{int}} \) (Eq. (3)) shows gradual linear increase with concentration the same as with \( GC_{\text{av}} \) in Fig. 8A.

**Fig. 7.** Three DMCs for DNAs with sequences of \( N_{\text{av}} = 10^6 \) bp obtained with a random number generator. The curves \( GC_{\text{av}} = 0.25 \) and \( GC_{\text{av}} = 0.75 \) correspond to DNA sequences that include a single random block of \( 10^6 \) bp. The sequence corresponding to curve \( GC_{\text{av}} = 0.6 \) consists of two random blocks. For the first and second blocks, \( GC_1 = 0.25 \) (\( m_{\text{int}} = 3000 \) bp) and \( GC_2 = 0.75 \) (\( m_{\text{int}} = 7000 \) bp), respectively.

**Fig. 8.** \( GC \) dependences for various definitions of the melting temperature (A) and the temperature melting range (B) for the sequences that include the two random blocks with \( GC = 0.25 \) and 0.75. For \( \Delta T_{\text{int}} \) and \( \Delta T_T \), the lines show the approximation with cubic polynomial: RMSEA(\( \Delta T_{\text{int}} \)) = 0.71 °C and RMSEA(\( \Delta T_T \)) = 1.67 °C. The ratios of RMSEA to the maximal values of \( \Delta T_{\text{int}} \) and \( \Delta T_T \) are 0.022 and 0.061, respectively.

**Fig. 9.** DMCs (\( S(T) \)) for unmodified (\( r_b = 0 \)) and chemically modified DNA. The \( r_b \) is per nucleotide concentrations of randomly distributed chemical modifications that increase the free energy of the helix–coil transition by 2.5 kcal per mole of modification. Calculation was carried out for EcoRI-cut pBR322 DNA.
Dependence of the melting temperature $T_m(n)$ (A) and the temperature melting range $\Delta T_m$ (B) on the per nucleotide concentration of randomly distributed chemical modifications that increase the free energy of the helix–coil transition by 2.5 kcal per mole of modifications. The curves $T_m(n)$ (Eq. (3)), $\Delta T_m(n)$ (Eq. (18)), and $\Delta T_b(n)$ (Eq. (17)) are the most smoothly varying and monotonic relative to other definitions of $T_m(n)$ and $\Delta T(n)$. Calculation was carried out for Escoff-cut pBR322 DNA.

**Conclusion**

This study was carried out to check a viewpoint that the melting temperature and temperature melting range cannot be properly calculated for multi-peak differential melting curves. Here, we analyzed various definitions of those parameters and studied their dependences on the factors that also influence fine structure of DMC: varying GC content, $[\text{Na}^+]$, and the relative per nucleotide concentration of chemical modifications ($r_b$). It was shown that the average temperature of the helix–coil transition or rather the temperature mean of DMC ($T_{\text{int}}$) and the average absolute deviation from this temperature ($\Delta T_{\text{int}}$) are the best for various studies. They are the least sensitive to a change in fine structure of DMC and always give reasonable smoothly varying dependences on varying $r_b$, $[\text{Na}^+]$, and GC. Besides, $T_{\text{int}}$ value is the closest to the thermodynamic melting temperature $T_{\text{thermo}}$, which is the ratio of the helix–coil transition enthalpy and entropy. The melting temperature calculated as a temperature corresponding to the half of melted base pairs $T_{b,5}$ as well as the temperature melting range $\Delta T_{b,5}$ defined as the temperature standard deviation of DMC multiplied by $\sqrt{2}\pi$ can also be used in the majority of cases. The distortion of $T_{b,5}$ caused by changes in DMC fine structure is usually less than 0.3 °C for all considered cases except two-component DMCs.

**Acknowledgments**

We are grateful to T. Chalikian, J. Völker, and R. Wartell for fruitful discussions of the case in point. This work was supported by the Ministry of Science and Technology of the Republic of China (Taiwan) under Grants MOST 103-2811-M-001-031 (D.Y.L.) and 103-2811-M001-074, the National Center for Theoretical Sciences (NCTS) in Taiwan, and the Belarusian Republican Foundation for Fundamental Research (X13-068). We are grateful to the Biophysics Core Facility, Scientific Instrument Center of Academia Sinica, for use of the CSC 6300 NanoDSC differential scanning calorimeter.

**References**


