



# Thermodynamics of DNA: heat capacity changes on duplex unfolding

Anatoliy Dragan<sup>1</sup> · Peter Privalov<sup>2</sup> · Colyn Crane-Robinson<sup>3</sup>

Received: 24 June 2019 / Revised: 12 September 2019 / Accepted: 3 October 2019 / Published online: 5 November 2019  
© The Author(s) 2019

## Abstract

The heat capacity change,  $\Delta C_p$ , accompanying the folding/unfolding of macromolecules reflects their changing state of hydration. Thermal denaturation of the DNA duplex is characterized by an increase in  $\Delta C_p$  but of much lower magnitude than observed for proteins. To understand this difference, the changes in solvent accessible surface area ( $\Delta ASA$ ) have been determined for unfolding the B-form DNA duplex into disordered single strands. These showed that the polar component represents ~55% of the total increase in ASA, in contrast to globular proteins of similar molecular weight for which the polar component is only about 1/3rd of the total. As the exposure of polar surface results in a decrease of  $\Delta C_p$ , this explains the much reduced heat capacity increase observed for DNA and emphasizes the enhanced role of polar interactions in maintaining duplex structure. Appreciation of a non-zero  $\Delta C_p$  for DNA has important consequences for the calculation of duplex melting temperatures ( $T_m$ ). A modified approach to  $T_m$  prediction is required and comparison is made of current methods with an alternative protocol.

**Keywords** DNA · Heat capacity · Hydration · Solvent accessible surface area · Surface polarity

## Introduction

It is well established that the heat denaturation of globular proteins is accompanied by an increase in the heat capacity of the system as a consequence of the hydration of internal hydrophobic residues by weakly bound water molecules having a heat capacity greater than bulk water. In contrast, when polar residues become exposed on protein denaturation, solvating water molecules are more tightly bound than in free solution, so their heat capacity decreases (Makhatadze and Privalov 1990; Privalov and Makhatadze 1992; Spolar et al. 1992; Murphy and Friere 1992; Loladze et al. 2001).

The total heat capacity change,  $\Delta C_p(T)$ , is frequently formalized in equations of the type:

$$\Delta C_p(T) = \sum \Delta(ASA)^i \times \Delta C_p^i(T), \quad (1)$$

where the coefficients  $\Delta C_p^i(T)$  represent the heat capacity change per  $\text{Å}^2$  of surface of defined type  $i$  and  $\Delta(ASA)^i$  is the change (increase) in the accessible surface area of that type that becomes exposed upon unfolding. Coefficients have been derived by several authors for the polar, aliphatic and aromatic surface of proteins although the last two categories are often combined in a single apolar term. According to Makhatadze and Privalov (1995), the heat capacity effect of hydrating the apolar and polar groups of proteins can be expressed by the equation:

$$\Delta C_p(25^\circ\text{C}) = 2.14 \times \Delta ASA_{\text{apolar}} - 1.27 \times \Delta ASA_{\text{polar}}, \quad (2)$$

where  $\Delta ASA$ s represent the increase in accessible surface area of apolar and polar surface.

The heat capacity change,  $\Delta C_p$ , is an important parameter because it represents the temperature dependence of the enthalpy of the process:  $\Delta C_p = \delta(\Delta H)/\delta T$ . Knowledge of  $\Delta C_p$ , therefore, allows comparison of denaturation enthalpies at a standard temperature for proteins having very variable melting points (Privalov 2012). Equation (2)—but with opposite signs—represents the heat capacity change resulting from the dehydration of internal residues on folding and has also been of value in characterizing protein/DNA association interactions in terms of the interfacial surface area occluded. The observed heat capacity change on forming a

✉ Colyn Crane-Robinson  
colyn.crane-robinson@port.ac.uk

<sup>1</sup> Institute of High Technologies, Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine

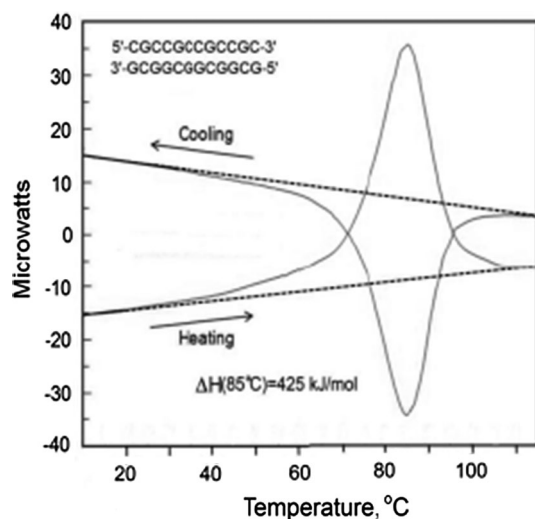
<sup>2</sup> Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA

<sup>3</sup> Biophysics Laboratories, School of Biology, University of Portsmouth, Portsmouth PO1 2DT, UK

protein/DNA complex was separated into the contribution from dehydration of protein surface, calculated for example using Eq. (2), and that from dehydration of the DNA surface to which the protein binds (e.g., Dragan et al. 2003). Application of this protocol to a substantial set of major and minor groove binding proteins led to the derivation of  $\Delta C_p^i$  coefficients for the dehydration of unit surface area in both grooves of the duplex (Dragan et al. 2019).

## The heat capacity of the DNA duplex

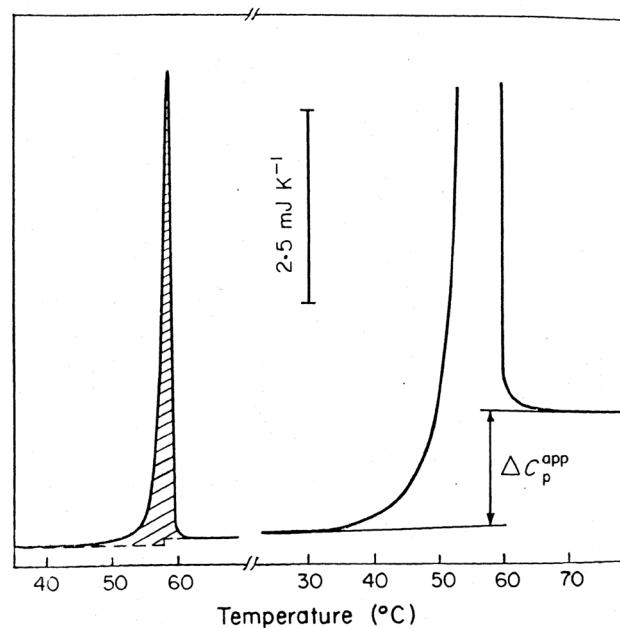
The situation with regard to unfolding the DNA duplex appears different from proteins as it is widely assumed that heat denaturation, i.e., strand separation, is not accompanied by any change in the heat capacity. This conclusion has been drawn from differential scanning calorimeter (DSC) studies of the heat denaturation process, as illustrated by the  $C_p/T$  function of a 12 bp all-CG duplex in Fig. 1. Linear extrapolation of the heat capacity function of the fully folded duplex (below 30 °C) coincides at high temperatures with the heat capacity of the fully unfolded state (above 100 °C). This appears to indicate that the heat capacities of the native duplex and the two separated strands are the same, a conclusion—if correct—of considerable convenience as it allows enthalpies determined for the melting of duplexes at different temperatures to be directly comparable. This simplification has resulted in tables of enthalpies/entropies of CG and AT pairs assumed valid for duplexes melting at different temperatures. Such tables are widely used, for example, in predicting the melting temperatures,  $T_m$ , of the primers and probes used in PCR reactions.



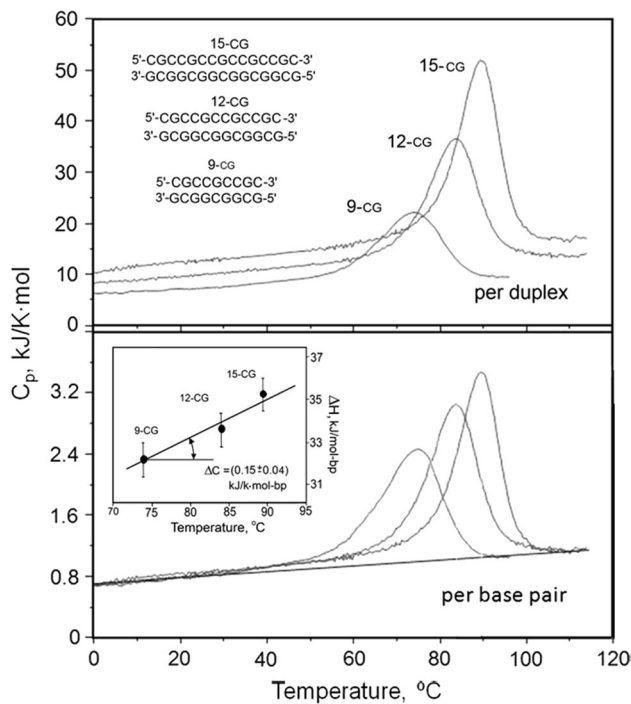
**Fig. 1** Original DSC recordings of the heat effect on heating and subsequent cooling at a constant rate of  $1 \text{ K min}^{-1}$  of a 12 bp all-CG DNA duplex (Privalov and Crane-Robinson 2018b)

The continuing use of these data tables is somewhat surprising bearing in mind the evidence for a significant increase in  $C_p$  on duplex denaturation. Filimonov and Privalov 1978 demonstrated a significant increment for the melting of long poly(A)•poly(U) molecules, and measured  $\Delta C_p$  as  $134 \pm 10 \text{ J K}^{-1} \text{ mol-bp}^{-1}$  (see Fig. 2). More recently, Chalikian et al. 1999, plotted the melting enthalpies of a broad range of double-stranded polynucleotides against their  $T_m$  values, to give a linear plot of slope  $\delta(\Delta H)/\delta T = 196 \text{ J K}^{-1} \text{ mol-bp}^{-1}$ . Holbrook et al. 1999 derived values of  $\Delta C_p$  between 240 and  $390 \text{ J K}^{-1} \text{ mol-bp}^{-1}$  for a 14 bp duplex of mixed composition, whilst Rouzina and Bloomfield 1999 indicated a range between 170 and  $420 \text{ J K}^{-1} \text{ mol-bp}^{-1}$  for a broad set of genomic DNAs. In the same year, Jelesarov et al. 1999 used DSC measurements of residual structures in the separated strands to correct ITC determinations of the enthalpies of duplex formation and thereby derived  $\delta(\Delta H)/\delta T$  functions over a wide temperature interval to yield  $\Delta C_p$  values of about  $200 \text{ J K}^{-1} \text{ mol-bp}^{-1}$ . The question is then whether these earlier estimates are borne out by more recent measurements.

A closer look at the melting of all-CG duplexes of different lengths, and thus melting temperatures, indicates that the assumption of a zero  $\Delta C_p$  is indeed incorrect (Fig. 3). The upper panel shows the molar  $C_p/T$  functions for 9, 12 and 15-bp duplexes. As expected, the absolute heat capacities



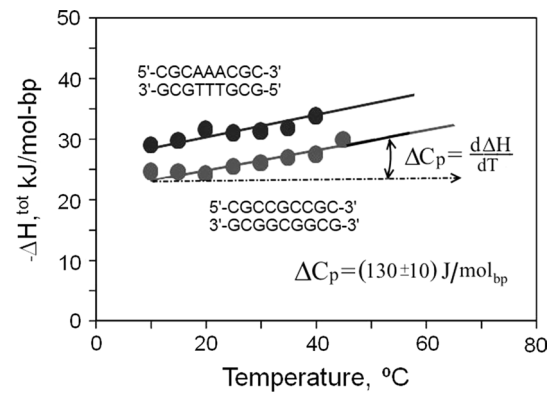
**Fig. 2** DSC recording of poly(A)•poly(U) melting at 0.3 mM concentration (left) and a fragment of a recording at 5.0 mM concentration (right) The hatched area corresponds to the apparent melting enthalpy; arrows indicate the observed heat capacity change. The NaCl concentration is 0.1 M. (Filimonov and Privalov 1978).  $d\Delta H^{(A-U)}/dT = 134 \pm 10 \text{ J K}^{-1} \text{ mol}^{-1}$



**Fig. 3** The partial heat capacity functions of three all-CG DNA duplexes calculated per mole of duplex (molar heat capacity, upper panel) and per mole of base pair (specific molar heat capacity, lower panel), all measured at the same molarity, 230  $\mu\text{M}$ , of the duplexes in 150 mM NaCl, 5 mM Na-phosphate, pH 7.4. Inset: the dependence of the excess enthalpy on the transition temperature, the slope of which gives an estimate of  $\Delta C_p$  (Privalov and Crane-Robinson 2018b)

of the fully folded duplexes (at low temperatures) are in proportion to their lengths, as are their denatured states at high temperature—and the  $T_m$  values also increase with the length, as expected. If the  $C_p/T$  functions are plotted per base pair (i.e., the specific heat capacities) and a baseline is drawn on the assumption of a zero  $\Delta C_p$  value (as in Fig. 1), it turns out that the total enthalpy increases somewhat with temperature—as seen in the inset to the lower panel of Fig. 3. The slope of the  $\Delta H/T$  plot, i.e.,  $\Delta C_p$ , is about  $0.15 \text{ kJ K}^{-1} \text{ mol-bp}^{-1}$ .

An alternative approach to verify the magnitude of  $\Delta C_p$ , without making any assumptions regarding the background appropriate for DSC scans, is to titrate one strand into its complement in the isothermal titration calorimeter (ITC). This has the advantage that experiments can be conducted over a wider temperature range, though the observed enthalpies require correction for residual structure in the individual stands at the temperature of each experiment (see Jelesarov et al. 1999; Vaitiekunas et al. 2015 for details). Figure 4 shows ITC-derived enthalpies for two 9-bp duplexes: one the 9-bp all CG duplex from Fig. 3 and the other of the same length but with the central 3 base pairs changed to AT. The enthalpies recorded for the AT-containing duplex



**Fig. 4** The ITC measured enthalpies of formation of two 9-bp DNA duplexes at various temperatures from 10 to 45  $^{\circ}\text{C}$ : one consisting only of CG base pairs, the other containing an A/T.A/T.A/T triplet (for more detail, see Vaitiekunas et al. 2015; Privalov and Crane-Robinson 2018b)

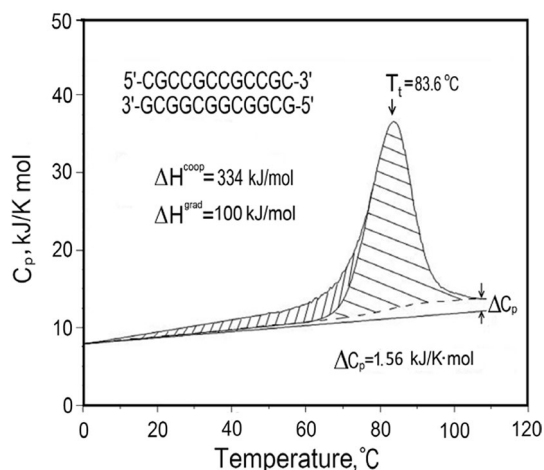
are somewhat greater, because the heat of denaturing AT pairs is significantly greater than for CG pairs (Vaitiekunas et al. 2015) but the slope of the two functions is the same at  $0.13 \text{ kJ K}^{-1} \text{ mol-bp}^{-1}$ —a value which can be taken as the magnitude of  $\Delta C_p$  for both AT and CG pairs.

The heat capacity increase on DNA dissociation is thus positive but much lower in magnitude than for proteins. For comparison: ubiquitin (M.Wt. = 8.6 kDa), has  $\Delta C_p^{25\text{C}} \sim +6 \text{ kJ K}^{-1} \text{ mol}^{-1}$ , whereas for the 12 bp DNA all-CG duplex (M.Wt. = 7.4 kDa),  $\Delta C_p^{25\text{C}}$  is measured as  $+(12 \times 0.13) = +1.56 \text{ kJ K}^{-1} \text{ mol}^{-1}$ , i.e., the *specific* heat capacity change on unfolding is much less for the DNA duplex than for the protein.

Determination of  $\Delta C_p$  for the base pairs of DNA has significant consequences for interpreting the  $C_p/T$  functions obtained in the scanning calorimeter. As explained in the caption to Fig. 5, it allows construction of a linear heat capacity function for the native, folded, state. This, in turn, allows the total excess heat to be deconvoluted into two components: the main peak that corresponds to a two-state cooperative dissociation process, preceded by a gradual accumulation of heat in the intact duplex structure. The enthalpy that characterizes the strand dissociation process, i.e., the melting, is that of the cooperative transition, not the total excess heat—as was previously assumed.

## Surfaces exposed upon DNA duplex dissociation

As the DNA duplex is a macromolecular complex with stacked aromatic bases located internally to the external polar phosphodiester chains, a situation not dissimilar to folded proteins, it is unsurprising that disruption of this



**Fig. 5** The observed heat capacity profile of a 12 bp all-CG duplex. The expected heat capacity of the fully folded DNA duplex is obtained by subtracting the heat capacity increment,  $\Delta C_p$ , ( $12 \times 0.13 = 1.56 \text{ kJ K}^{-1} \text{ mol}^{-1}$ ), from the heat capacity of DNA at 110 °C and linearly extrapolating back to the start of melting at 0 °C. The experimental excess heat effect is then deconvoluted into non-cooperative (gradual, vertical hatching) and cooperative (horizontal hatching) phases (Vaitiekunas et al. 2015)

structure is accompanied by an increase in the heat capacity—reflecting the exposure of apolar groupings to the solvent. Striking, however, is the fact that the magnitude of  $\Delta C_p$  is only about  $\frac{1}{4}$  that of a protein of similar weight. Understanding this difference comes from measuring the

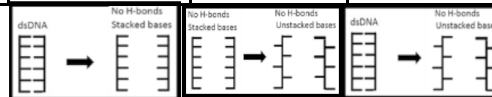
magnitude of the two types of surface exposed on duplex denaturation, i.e., the increase in the polar and apolar accessible surface areas ( $\Delta ASAs$  in Eq. 1).

To quantify the contributions of polar and apolar contacts in the DNA duplex, the increases in accessible surface areas on strand separation were assessed using the Naccess program with two categories of surface atoms: polar (N, O and P) and apolar (C and H), see Table 1. For the folded forms, four B-form duplexes with high-resolution structures available and having mixed composition and variable length were selected. The question then arises as to the state of the single-stranded oligonucleotides that result from the heat dissociation process: can they be taken as totally randomised with full solvent access to the bases or does some secondary/tertiary structure remain? This is answered by DSC scans of individual single strands that demonstrate the presence of intrinsic structures at low temperatures that melt to yield linear  $C_p/T$  functions above 80/90 °C. Measurement shows that the slopes of these linear  $C_p/T$  functions—that represent the intrinsic heat capacity of the disordered chains—and also the absolute values of the heat capacities at 80/90 °C are in direct proportion to the length of the oligonucleotides (Jel-sarov et al. 1999). This demonstrates that the heat denatured state is the same for all of such oligonucleotides—and is strongly suggestive that this is a fully disordered strand.

To model such randomly disordered and solvent-exposed strands, successive nucleotides in  $n$ -mer chains were spaced by  $(n - 1)$  abasic sites to ensure full access of the  $n$  bases to solvent. The Naccess programme does not

**Table 1** Water accessible surface areas (ASAs)—apolar and polar—and their changes for B-form DNA unfolding in two steps

DNA sequence	Polarity	ASA, Å <sup>2</sup>	ASA, Å <sup>2</sup>	ASA, Å <sup>2</sup>	ΔASA, Å <sup>2</sup>	ΔASA, Å <sup>2</sup>	ΔASA, Å <sup>2</sup>	ΔCp
		dsDNA	ssDNA	ssDNA				kj/K.mol-bp
		"folded"	"no H-bonds"	"unfolded"	"no H-bonds" - folded"	"unfolded" - "no H-bonds"	"unfolded" - folded"	"unfolded" - folded"
<sup>a</sup> 12 bp CGCGAATTCGCG	Apolar	2128	2235	4738	107	2503	2610	0.11
	Polar	2438	3479	5833	1041	2354	3395	
<sup>b</sup> 12 bp CGCAAATTTGCG	Apolar	2188	2341	4849	153	2508	2662	0.13
	Polar	2494	3466	5705	972	2239	3211	
<sup>c</sup> 16 bp ACTACAATGTTGCAAT	Apolar	2873	3108	6632	235	3524	3759	0.17
	Polar	3216	4400	7414	1184	3014	4198	
<sup>d</sup> 20 bp CCAATAATCGCGATTATTGG	Apolar	3475	3751	8193	276	4442	4718	0.15
	Polar	3826	5344	9380	1518	4036	5554	



*Step 1* native folded dsDNA to a hypothetical ssDNA state with broken H-bonds but the stacking within each strand preserved [LH cartoon]

*Step 2* the stacked ssDNA state to totally unfolded strands having complete solvent access to unstacked bases [central cartoon]. The RH cartoon depicts the aggregate of these two steps: the complete transition from folded dsDNA to totally unfolded strands. The total increases “unfolded–folded” ( $\Delta ASAs$ ) are used to calculate the corresponding changes in the heat capacities,  $\Delta C_p$ , using the equation of Makhatadze and Privalov (1990) and are given in the last column

dsDNA double-stranded DNA, ssDNA single-stranded DNA

<sup>a</sup>Drew et al. (1981), PDB:1BNA; <sup>b</sup>Woods et al. (2004), PDB:1S2R; <sup>c</sup>Narayana and Weiss (2009), PDB:3BSE. <sup>d</sup>Garcia et al. (2016), PDB:5F9I

recognize unnatural nucleotides, so the measured ASAs correspond to fully extended and exposed native oligodeoxynucleotide chains. Two checks of this assumption were made: (1) the abasic sites were removed manually—with no resulting change in ASA values—and (2) the chains were modelled by simple addition of the ASA values determined for the four individual mononucleotides: this resulted in an increase in the polar contribution by 1.4% and no change in the apolar contribution.

The duplex melting process was separated into two steps: Step 1 is separation of the two strands without altering their conformation, a process that represents breakage of the H-bonds (loss of pairing) without loss of base stacking. Table 1 shows that the increases in ASA in this step are largely polar (only ~14% apolar) as expected from exposure of the largely polar edges of the bases: this step represents loss of ~18% of the total contact area between the two strands. Step 2 is unstacking adjacent bases, i.e., randomising the separated polynucleotide chains, for which the total increase in ASA is much greater and is approximately equally divided between polar and apolar contributions: it represents the remaining 82% of the total contact area. The overall increase in accessible surface areas (“unfolded–folded”) averages to 45% apolar/55% polar. Previous measurements of changes in water accessible areas on DNA unfolding were reported by Holbrook et al. (1999) for a 14 bp duplex using the ANAREA programme. They noted that the base pairing interaction (‘helix to helix’) is almost completely polar—exactly as found here. However, the unstacking process (‘helix to disordered’) was found to be as much as 60% polar rather than the 47% polar measured here—although in both data sets, it is the unstacking process that makes the dominant contribution to the heat capacity change. When Holbrook et al.’s  $\Delta$ ASA values were substituted into their own heat capacity function (Spolar et al. 1992), the increased negative contribution to  $\Delta$ Cp from hydration of polar surface was sufficient to fully negate the lower positive effect from the apolar term, i.e., a net zero  $\Delta$ Cp was predicted.

In contrast to DNA, the total increase in ASA for the unfolding of ubiquitin (5780 Å<sup>2</sup>) is 67% apolar/33% polar, figures typical for small globular proteins. This comparison demonstrates a key difference between the unfolding of DNA and proteins: dissociation of the DNA strands results in a much greater exposure of polar surface than is the case for proteins. For proteins, the dominant apolar component leads to the well-known positive values of  $\Delta$ Cp, but with DNA this is very much reduced in magnitude by the substantial negative contribution to  $\Delta$ Cp from the large polar  $\Delta$ ASA—despite its smaller  $\Delta$ Cp<sup>i</sup> coefficient.

The last column of Table 1 gives the heat capacity changes calculated on the basis of the total apolar and polar  $\Delta$ ASAs using the above Eq. (2): the predicted values

are close to the observed value of 0.13 kJ K<sup>-1</sup> mol-bp<sup>-1</sup> (Vaitiekunas et al. 2015). This correspondence demonstrates that the equation derived on the basis of unfolding the polypeptide chain (Eq. 2) applies effectively to polynucleotides and also supports the assumption that the fully denatured state used to model the ASA of the unfolded DNA strands effectively corresponds to the heat denatured state at high temperature. It is clear, therefore, that the heat capacity increase on melting the DNA duplex, although positive as for proteins, is much less in magnitude as a result of the large negative contribution from exposure of internal polar surface, rather than dominated by the apolar surface as for folded proteins.

### The significance of a heat capacity increase in DNA duplex dissociation

A heat capacity increase of 0.13 kJ K<sup>-1</sup> mol-bp<sup>-1</sup>, i.e., 1.56 kJ K<sup>-1</sup> for a 12-mer duplex, may seem of little consequence when compared to that of comparable globular proteins but when applied, for example, to calculation of the melting temperatures of PCR primers—a widespread use of DNA thermodynamic data—it is of important significance. For example, the denaturation enthalpy of a CG pair is about 19 kJ mol<sup>-1</sup> at the standard temperature of 25 °C but extrapolation to, say, 75 °C (a typical primer  $T_m$ ) adds 6.5 kJ mol<sup>-1</sup> to this, which is an increase of more than 30%!

Bearing in mind that the usual protocols for calculating the  $T_m$  of primers and probes for real-time PCR assume no variation of the enthalpy/entropy with temperature—and are thus fundamentally flawed—what, in practical terms, are the consequences of the observed  $\Delta$ Cp for  $T_m$  prediction? At present, ‘universal’ tables of characteristic enthalpies and entropies for the ten possible base pair adjacencies (the ‘nearest neighbour—NN—interactions’) and valid at all temperatures, are used to predict a wide range of  $T_m$  values. With a finite  $\Delta$ Cp, however, a completely different prediction protocol is required. The central issue is as follows: the enthalpy/entropy values assigned to CG and AT pairs for the standard temperature of 25 °C must be extrapolated to  $T_m$ , the unknown we are attempting to calculate. This problem is best solved iteratively, rather than analytically: melting of short DNA duplexes takes place at temperatures between 50 and 95 °C, so to a first approximation a  $T_m$  of 75 °C can be used, i.e., 50° above the standard temperature, to calculate the enthalpies/entropies expected for that temperature—the ratio of which yields a postulated  $T_m$ . This value of  $T_m$  can then be used for a second extrapolation. Usually the second iteration gives a predicted value of  $T_m$  that is not changed by further cycles. This protocol is given in detail in Privalov and Crane-Robinson (2018a).



**Table 2** Comparison of melting temperatures predicted by OligoCalc (OC) and OligoAnalyser (OA)—with  $T_m$  values obtained by the iterative methodology of PLP/CCR (Privalov and Crane-Robinson 2018a)

DNA duplexes	$T_m^{\text{obs}}$ (°C) in 0.15 M NaCl	$\Delta H^{\text{OC}}$ (kJ mol <sup>-1</sup> )	$\Delta S^{\text{OC}}$ (J K <sup>-1</sup> mol <sup>-1</sup> )	$T_m^{\text{Predicted}}$ (°C) using OC OA	$\Delta H^{\text{PLP/CCR}}$ (kJ mol <sup>-1</sup> )	$\Delta S^{\text{PLP/CCR}}$ (J K <sup>-1</sup> mol <sup>-1</sup> )	$T_m^{\text{Predicted}}$ (°C) using PLP/CCR
5′-CGCCGCCCGCCCGC-3′ 3′-GCGGCCGCGCGCGC-5′	89.5	648	1632	86.3 91.2	407	1124	88.9
5′-CGCAAATTTAAACGC-3′ 3′-GCGTTTTAAATTTGCG-5′	64.8	512	1347	64.9 63.7	416	1236	63.4
5′-CGCACACACACACGC-3′ 3′-GCGTGTGTGTGTGCG-5′	75.9	554	1435	72.8 75.1	412	1189	73.3
5′-GCGAACCAATCGG-3′ 3′-CGCTTGTAGCC-5′	64.8	426	1109	63.8 64.7	320	940	67.2

All duplex concentrations = 283 μM. Observed melting temperatures in column 2; Predictions in columns 5 and 8

A matter of interest is then to establish if the use of more accurate and temperature-dependent enthalpies/entropies leads to increased precision in predictions of the melting temperatures of short duplexes, for example the primers and probes of PCR reactions. Comparison was, therefore, made, for four duplexes, of their melting temperatures predicted using: (1) two of the well-known online calculators (OligoCalc, OC, from Northwestern University and OligoAnalyser, OA, from the company IDT) based on the NN tables of temperature-independent enthalpies/entropies published by Breslauer et al. (1986), Sugimoto et al. (1996) and SantaLucia (1998), and (2) the iterative protocol outlined above using temperature-dependent single-valued enthalpies/entropies for CG and AT pairs.

Table 2 shows substantial differences in the enthalpy/entropy values between those used in the PLP/CCR protocol and those from the ‘historical’ data sets:  $\Delta H^{\text{PLP/CCR}}$  and  $\Delta S^{\text{PLP/CCR}}$  values are significantly lower. This is largely a result of using just the cooperative component of the total enthalpy/entropy and also results from the fact that the enthalpy/entropy of AT pairs are in fact greater than those of CG pairs, as explained in Vaitiekunas et al. (2015). However, it is the ratio of the enthalpy to the entropy that determines  $T_m$ , not their absolute values. Despite these large differences in the enthalpies and entropies, the predictive capacities of the two approaches do not differ greatly, with the notable exception of the 15-bp all-CG duplex, for which the PLP/CCR protocol is strikingly accurate: the difference of 0.4° in 362 K between the predicted and observed  $T_m$  represents an error of only 0.1%. The reason for such high precision is that the enthalpy/entropy for a CG pair is independent of its neighbour when that is also a CG pair, i.e.,

the predictive capability for all-CG duplexes is exceptionally high. This is not the case, however, for the enthalpy/entropy of an AT pair, which does depend on its neighbours, a variation that results from the presence of water tightly bound to AT pairs in the minor groove, but absent from all-CG duplexes. However, the current PLP/CCR protocol is based on single, unique values for the enthalpy/entropy of an AT pair, i.e., takes no account of NN interactions. The lack of correction for such interactions results in less accurate  $T_m$  predictions for the three AT-containing duplexes than for all-CG duplexes.

The remaining point of interest is that despite the very substantial inaccuracies in the historical enthalpy/entropy data sets and the neglect of their temperature dependence, the OC and OA predictive capability is reasonably good: why is this? Firstly, it must be recalled that only the enthalpy is measured experimentally for the calibrating duplexes and the entropy is then derived by dividing this by the melting temperature,  $T_m$ . As  $T_m$  prediction is the reverse process, errors in the predicted value are not great because inaccuracies in the calibrating enthalpies have been ‘compensated’ by corresponding errors in the entropies derived from them. Furthermore, since most of the predicted sequences melt at temperatures not so far from those of the calibrating duplexes, the importance of  $\Delta C_p$  for  $T_m$  prediction is reduced. It follows, therefore, that although the temperature-dependent enthalpies/entropies of the PLP/CCR protocol are much more appropriate than those in the historical data tables, further analysis of AT-containing duplexes is required to establish a precise NN interaction table and thereby bring the precision of all predictions up to the level already achieved for all-CG duplexes.

## Conclusions

Measurement of increases in solvent accessible surface areas (ASAs) as the B-form duplex dissociates into fully disordered single strands shows that these average to 45% apolar/55% polar surface. This distribution differs markedly from that of proteins for which the apolar surface exposed on denaturation amounts to about 2/3rd of the total. The core structure of the duplex is thus very much more dependent on polar interactions than is that of proteins. The immediate consequence of this is that the negative heat capacity effect of exposing the polar surface reduces the net increase in  $C_p$  for DNA to a low positive value. A finite  $\Delta C_p$  value for DNA implies that changes are required to the methods of calculating duplex melting temperatures.

**Acknowledgements** Initiation of this work was funded by the National Institutes of Health (NIH) Grant no. 105365 and subsequently supported by the Ministry of Education and Science of Ukraine Grant no. 0116U004757/16BΦ07-03 (D.A.I.).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Breslauer KJ, Frank R, Blocker H, Marky LA (1986) Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci USA* 83:3746–3750
- Chalikian TV, Volker J, Plum GE, Breslauer KJ (1999) A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques. *Proc Natl Acad Sci USA* 96:7853–7858
- Dragan AI, Liggins JR, Crane-Robinson C, Privalov PL (2003) The energetics of specific binding of AT-hooks from HMG1 to target DNA. *J Mol Biol* 327:393–411
- Dragan AI, Read CM, Crane-Robinson C (2019) Hydration differences between the major and minor grooves of DNA revealed from heat capacity measurements. *Eur Biophys J* 48:131–138
- Drew HR, Wing RM, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *Proc Natl Acad Sci USA* 78:2179–2183
- Filimonov VV, Privalov PL (1978) Thermodynamics of base interaction in (A)<sub>n</sub> and (AU)<sub>n</sub>. *J Mol Biol* 122:465–470
- García S, Acosta-Reyes FJ, Saperas N, Campos JL (2016) Crystal structure of an AT-rich 20mer DNA. (**Unpublished structure**)
- Holbrook JA, Capp MW, Saecker RM, Record MT 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
- Jelesarov I, Crane-Robinson C, Privalov PL (1999) Energetics of HMG box interactions with DNA: thermodynamic description of the target DNA duplexes. *J Mol Biol* 294:955–981
- Loladze VV, Ermolenko DN, Makhatadze GI (2001) Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci* 10:1343–1352
- Makhatadze GI, Privalov PL (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solutions: hydration effect. *J Mol Biol* 213:375–384
- Makhatadze GI, Privalov PL (1995) Energetics of protein structure. *Adv Prot Chem* 47:307–425
- Murphy KP, Friere E (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv Prot Chem* 43:313–361
- Narayana N, Weiss MA (2009) Crystallographic analysis of a sex-specific enhancer element: sequence-dependent DNA structure, hydration, and dynamics. *J Mol Biol* 385:469–490
- Privalov PL (2012) Microcalorimetry of macromolecules. The physical bases of biological structures. Wiley, New York
- Privalov PL, Crane-Robinson C (2018a) Translational entropy and DNA duplex stability. *Biophys J* 114:15–20
- Privalov PL, Crane-Robinson C (2018b) Forces maintaining the DNA double helix and its complexes with transcription factors. *Prog Biophys Mol Biol* 135:30–48
- Privalov PL, Makhatadze GI (1992) Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. *J Mol Biol* 224:715–723
- Rouzina I, Bloomfield VA (1999) Heat capacity effects on the melting of DNA. 1. General aspects. *Biophys J* 77:3242–3251
- SantaLucia J Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 95:1460–1465
- Spolar RS, Livingstone JR, Record MT Jr (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of non-polar and polar surface from water. *Biochemistry* 31:3947–3955
- Sugimoto N, Nakano S, Yoneyama M, Honda K (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res* 24:4501–4505
- Vaitiekunas P, Crane-Robinson C, Privalov PL (2015) The energetic basis of the DNA double helix: a combined microcalorimetric approach. *Nucleic Acids Res* 43:8577–8589
- Woods KK, Maehigashi T, Howerton SB, Sines CC, Tannenbaum S, Williams LD (2004) High-resolution structure of an extended A-tract: [d(CGCAAATTTGCG)]<sub>2</sub>. *J Am Chem Soc* 126:15330–15331

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.