Y-DNA melting: a short tale of three scales

2009 J. Phys.: Condens. Matter 21 034115

(http://iopscience.iop.org/0953-8984/21/3/034115)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **21** (2009) 034115 (5pp) doi:10.1088/0953-8984/21/3/034115

Y-DNA melting: a short tale of three scales

A Dkhissi[1,2], G Renvez[2] and R Blossey[1]

[1] Biological Nanosystems, Interdisciplinary Research Institute, Lille University of Science and Technology, USR CNRS 3078, c/o IBL, Rue du Prof Calmette 1, 59021 Lille Cedex, France
[2] Laboratoire dAnalyse et dArchitecture des Systmes—CNRS, 7 Av du Colonel Roche 31077 Toulouse Cedex 04, France

E-mail: ralf.blossey@iri.univ-lille1.fr

Received 13 July 2008, in final form 7 October 2008
Published 17 December 2008
Online at stacks.iop.org/JPhysCM/21/034115

Abstract
We address some aspects of the thermal denaturation of Y-DNA—the three-way junction—on three different length scales: the effect of chain concentration on the extrinsic melting behaviour within a simple kinetic approach, the exponent of the loop entropy in intrinsic melting as it appears in statistical mechanics models, and the microscopics of stacking within the junction from molecular dynamics simulations. Our results suggest that a multiscale approach is needed to properly describe the denaturation properties of these systems. We propose experiments which can also shed light on the melting of short and long duplex DNA sequences.

1. Introduction

Y-DNA is, as its name suggests, a Y-shaped DNA hybrid made from three strands which form three double helices joining together in a three-way junction, which is its other name (Y-DNA is not to be confused with the Y chromosome). The structure of the molecule is sketched schematically in figure 1.

Three-way junctions arise naturally in several biological contexts such as: transients in strand-exchange (Reynaldo *et al* 2000), viral genomes (Kadrmas *et al* 1995), and more generally in RNA (Lescoute and Westhoff 2006). More recently, Y-DNA molecules have also found use as a building block or recognition element in nanotechnology (Li *et al* 2004, Feldkamp and Niemeyer 2006, Oleksi *et al* 2006). Melting of Y-DNA has been studied previously by several authors (Leontis *et al* 1991, Kadrmas *et al* 1995, Assenberg *et al* 2002). These studies predominantly focused on structural effects induced by few base pair changes of the nucleic acid composition.

The objective of the present study differs from these previous studies in that we are interested in conceptual features of the DNA melting process for which Y-DNA possesses some particular features not accessible with duplex DNA. Although our results are clearly not yet able to give a complete picture of the stability properties of these molecules, they are suggestive that a multiscale approach is needed in order to quantitatively address the properties of Y-DNA. The results also suggest experiments relevant for a better understanding of the melting of double-stranded DNA.

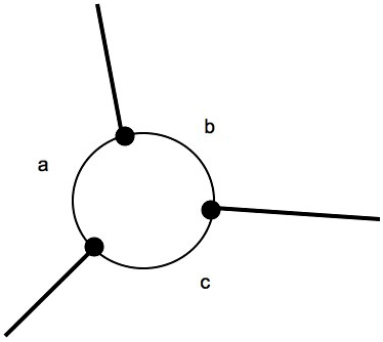

Figure 1. A schematic representation of a Y-DNA molecule. It is built from three single strands, denoted by *a*, *b* and *c*, which share partially complementary regions *a*–*c*, *a*–*b*, *b*–*c*. Double strands are indicated in boldface. In drawing this sketch we assume that both the length of the complementary regions at the extremities of the strands and the length of the non-complementary regions near the centre of the molecule are variable in length. In this way the size of the interior loop can be varied.

2. Extrinsic melting of Y-DNA

In DNA melting, one commonly distinguishes between *intrinsic* and *extrinsic* melting (Wartell and Benight 1985). The first refers to the breaking of bound pairs while the two DNA chains are still bound by at least one pair; the latter process is the complete dissociation of the two strands. Naturally, extrinsic melting dominates the melting process of short double-stranded DNA, which can, for most purposes,

Table 1. Base pair energetics and stacking. N_i are the combinations of stacked bases according to the singlet scheme of (Owczarcy *et al* 1997), with the free energies δG_i taken from Doktycz *et al* (1992).

Base pair	Hairpin	Duplex	Triplex	δG_i	δE^s (hp)	δE^s (d)	δE^s (t)
N_{AT}	4	10	12	—	—	—	—
N_{GC}	3	10	9	—	—	—	—
N_1	1	2	3	−190	−190	−380	−570
N_2	1	2	3	146	146	292	438
N_3	1	4	3	−28	−28	−112	−84
N_4	0	3	0	−240	0	−720	−0
N_5	2	4	6	−113	−226	−452	−678
N_6	1	4	3	−6	−6	−24	−18
N_7	3	9	9	−25	−75	−225	−225
N_8	5	11	15	−39	−195	−429	−585

even be considered as a two-state transition. By contrast, intrinsic melting dominates for long chains. In terms of the melting curves themselves, one finds that for temperatures $T < T_{\mathrm{m}}$ intrinsic melting dominates, while for $T > T_{\mathrm{m}}$ extrinsic melting takes over since the chains dissociate.

There is an ongoing discussion about double-stranded DNA of lengths around 100 bp, for which the distinction between extrinsic and intrinsic melting poses some problems; this is mainly due to the fact that at around this chain length, and depending on sequence composition, intrinsic and extrinsic melting effects are not easy to decouple. For Y-DNA, things can in fact get even worse. In this first part of this paper we address the question of the concentration dependence of the melting curve of Y-DNA.

The Y-DNA molecule exhibits a far greater option for variations than duplex DNA; e.g., by adjusting base pair composition, the length of its three arms can be tuned independently (also the angle between them). Since we want to look here at the concentration dependence only, it is best to select a situation in which this effect is presumably strongest, and which facilitates a comparison with duplex DNA.

We illustrate this here for the 20 bp sequence

$$5'\text{–CAATGGATCGCGATCCATTG–}3'$$

whose two halves are self-complementary: $a = s_a \bar{s}_a$, likewise for b and c. As a consequence, in principle, the sequence can hybridize in the form of hairpins, duplexes, Y-DNA, and even higher-order complexes. We consider only the three lowest possible structures and discuss their extrinsic melting curves within a two-state model (which is clearly an approximation, see below).

As a first step in this analysis we calculate the Gibbs free energy ΔG following the standard set-up used for DNA hairpins and duplexes. We use the expression

$$\Delta G = \delta S_{\mathrm{bp}}[N_{\mathrm{AT}}(T_{\mathrm{AT}} - T) + N_{\mathrm{GC}}(T_{\mathrm{GC}} - T)] + \sum_{i=1}^{8} \delta E_i^s \quad (1)$$

with the values for the entropy per base pair $\delta S_{\mathrm{bp}} = -24.85$ cal mol^{-1} K^{-1}, $N_{\mathrm{AT}}, N_{\mathrm{CG}}$ as the number or corresponding base pairs, $T_{\mathrm{AT}} = 338.36$ K and $T_{\mathrm{GC}} = 380.97$ K, all taken at a salt concentration of 0.115 M Na$^+$. The last term in the expression is the stacking energy, $\delta E_i^s = N_i \delta G_i$.

We assume that in the hairpin structure the six central bases of the sequence are unpaired. Further, we assume that this is also the case for Y-DNA, which means that $\Delta G_{\text{Y-DNA}} = 3\Delta G_{\text{hairpin}}$. This again is of course a (crude) assumption, and we will come to the true stacking of Y-DNA in the central junction below.

Table 1 lists the results from this computation. It can be seen that within our assumptions the stacking energies of the duplex DNA and DNA-Y are almost comparable. Since both sequences have a similar number of base pairs (20 versus 21), the melting curves should not be too different. As we will see, this is not the case.

To proceed we first recall how the extrinsic melting curve for duplex DNA is computed, assuming a two-state transition. This is done by considering the chemical equilibrium between single-and double-stranded DNA with equilibrium constant K_{D}, taking into account the fixed concentration of total DNA, $c_{\mathrm{T}} = c_1 + 2c_2$ (the factor two arises since each molecules contains two DNA strands). This yields for the single-chain fraction $\theta \equiv c_1/c_{\mathrm{T}}$ the result

$$\theta = \frac{1 + 4c_{\mathrm{T}}/K_{\mathrm{D}} - \sqrt{1 + 8c_{\mathrm{T}}/K_{\mathrm{D}}}}{4c_{\mathrm{T}}/K_{\mathrm{D}}}. \quad (2)$$

The same reasoning can be applied to the DNA-Y case, but we now have

$$c_{\mathrm{T}} = c_1 + \bar{c}_1 + 2c_2 + 3c_3 \quad (3)$$

with the equilibrium constants

$$K_{\mathrm{H}} = \frac{c_1}{\bar{c}_1}, \qquad K_{\mathrm{D}} = \frac{c_1^2}{c_2}, \qquad K_{\mathrm{Y}} = \frac{c_1^3}{c_3}. \quad (4)$$

In this case θ obeys the cubic equation

$$\theta^3 + \alpha\theta^2 + \beta(\gamma\theta - 1) = 0 \quad (5)$$

with

$$\alpha = \frac{2}{3}\frac{K_{\mathrm{Y}}}{K_{\mathrm{D}}}\frac{1}{c_{\mathrm{T}}}, \qquad \beta = \frac{1}{3}\frac{K_{\mathrm{Y}}}{c_{\mathrm{T}}^2}, \qquad \gamma = \left(1 + \frac{1}{K_{\mathrm{H}}}\right). \quad (6)$$

The cubic equation for θ can of course be solved by standard means. Figure 2 shows the result of the computations for our sequence, comparing hairpins, duplexes, DNA-Y and the equilibrium mixture of the three species.

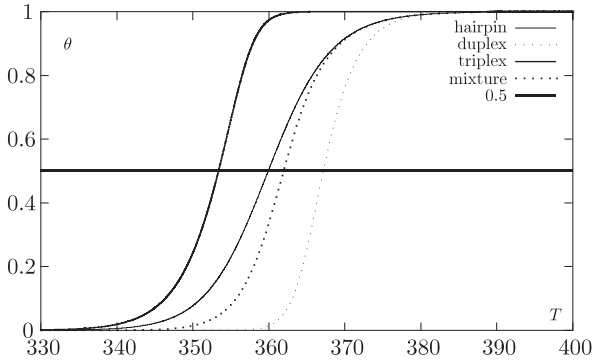

Figure 2. Computed extrinsic melting curves of the hairpin, duplex and Y-DNA, as well as the curve for the equilibrated mixture at a given concentration c_T, here shown for a value of 10^{-1} M.

The melting profiles of the hairpin, duplex, triplex molecules and the mixture have different shapes with different melting temperatures: one finds the progression

$$T_m^{\text{triplex}} < T_m^{\text{hairpin}} < T_m^{\text{mixture}} < T_m^{\text{duplex}}. \quad (7)$$

Hairpin melting curves have *no* concentration dependence since in the formation of the hairpin the number of DNA strands does not change; melting here reflects purely sequence properties. Due to the different nonlinearity of the quadratic and cubic equations for θ in the duplex and DNA-Y case, the melting curves of duplex DNA and DNA-Y differ quite distinctly in their temperature dependence. The overall shape of the melting curve is similar, which is a consequence of the similar values of the free energies of the chains; however, the extrinsic melting curve of DNA-Y has a considerably lower melting temperature than the duplex. This effect becomes more pronounced when the total concentration of DNA c_T is lowered. We finally note that the melting curve for the mixture of hairpins, duplexes and Y-DNA always resembles the (non-concentration-dependent) melting curve of the hairpins. In order to distinguish the concentration-dependence effect, the sequences must therefore be denatured separately in experiment (Bayer *et al* 2005).

Clearly, the general assumption of a two-state transition is only a very crude one, but it allows us to extract an interesting feature of DNA melting, testable in experiment. Note that the assumption of self-complementarity of the strands places the system at a peculiar symmetry point since the kinetics is assumed to be dominated by the simultaneous dissociation of the three strands; this might not be at all adequate for Y-DNA with markedly different base pair composition. The dissociation of the strands might therefore be only partial or occur in a multi-step process (e.g. one strand dissociating from the Y-DNA molecule, followed by a dissociation of the remaining dimer): again, a situation in principle easy to test experimentally.

3. Molecular melting

In our view a general mesoscopic theory of Y-DNA is difficult to achieve due to the coupling of sequence and structural

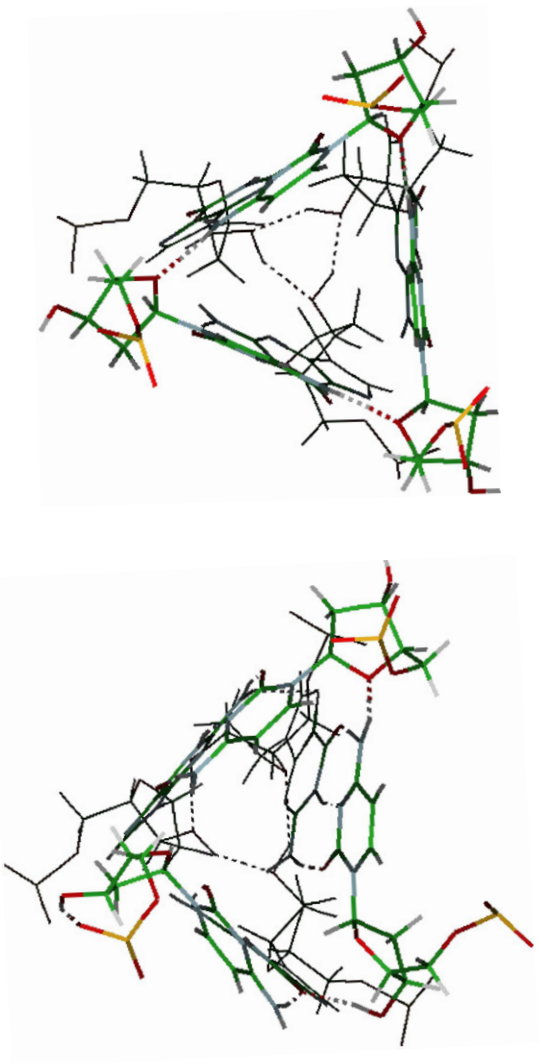

Figure 3. Top: initially optimized structure of G/CG/CG; bottom: after melting
(This figure is in colour only in the electronic version)

effects. Mesoscopic models e.g. require the knowledge of the stacking parameters. In order to get a better idea of the stacking within the three-way junction we descend down to molecular scales and address the question of the molecular unstacking or melting of the three-way junction. To address this issue we have performed molecular dynamics simulations of DNA-Y junctions.

The development of new force fields and representations of long-range effects make such techniques applicable to DNA as well. All our simulations used the AMBER 8 suite of programs (Case *et al* 2004). The *ff*99 force field was used to describe the DNA duplex (Wang *et al* 2000). Initial structures of each Y-junction were built first arbitrarily and were then minimized. After minimization MD simulations were started at 273 K. The temperature was raised to 373 K within 20 ps. The temperature was fixed using the Berendsen coupling algorithm (Berendsen *et al* 1984).

Our results are seen in figures 3 and 4. The structures in figure 3 (bottom) exhibits broken bonds and is about

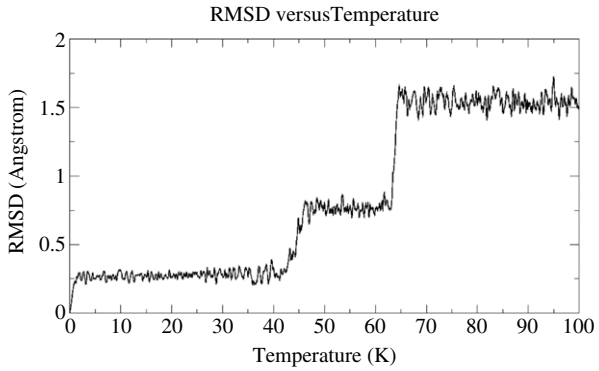

Figure 4. RMSD (root mean square deviation) versus temperature for the transition between the two structures.

Table 2. Force field energies. 1–4: non-bonded interactions, atoms separated by three covalent bonds.

Energy (kcal mol^{-1})	Before heating	After heating
$E_{\rm tot}$	−821.49	−753.50
$E_{\rm bond}$	11.7164	12.8197
$E_{\rm vdW}$	−55.6447	−59.6890
$E_{\rm vdW,1-4}$	43.6468	44.5041
$E_{\rm angle}$	26.4551	29.9132
$E_{\rm el}$	169.1941	219.7540
$E_{\rm el,1-4}$	−1110.9117	−1110.5910
$E_{\rm dihed}$	94.0584	109.7882

70 kcal mol^{-1} less stable than the structure shown at the top of that figure. The analysis of the energetics before and after the melting step indicated that this difference can be attributed to electrostatic interactions. Figure 4 shows the evolution in RMSD (root mean square deviation) over temperature. From 0 to 45 °C the RMSD is is almost constant, and then jumps up with a first structural change, and then, at about 65 °C, the melting step. Finally, table 2 lists a number of data (force field energies) recoded during the simulations.

4. Intrinsic melting of Y-DNA

We now look at intrinsic melting of Y-DNA. In order to increase its effect we consider a particular version of the molecule sketched in figure 1. Its composition is assumed to be such that each single strand has about 50 bp with a similar energetics of pairing. The interest is now in a control of the junction loop of size ℓ such that the intrinsic melting signal is dominated by its opening. The difference in equilibrium melting curves that either melt from the boundaries, or from an interior loop in a plasmid, has been studied in detail (Blossey and Carlon 2003).

The interest in such a thought experiment is the following. In the Poland–Scheraga model of DNA melting by Kafri *et al* (2002), the free energy of the chain is importantly influenced by the loop entropy. Loop entropy itself is controlled by the loop exponent c, for which a theory based on polymer networks exists (Duplantier 1986). The number of configurations in the thermodynamic limit is given by $\Gamma_{\mathcal{G}} \sim s^L L^{\gamma_{\mathcal{G}}-1}$ with

$$\gamma_{\mathcal{G}} = 1 - \nu d\mathcal{L} + \sum_{k\geqslant 1} n_k \sigma_k. \tag{8}$$

Here, s is a geometrical factor, L the polymer length, $\mathcal{L}$ the number of independent loops in the network, d the spatial dimension and ν the exponent of the radius of gyration of a self-avoiding random walk. Further, the n_k are the vertices of order k with the scaling dimensions σ_k.

For the case of the three-way junction, we have $\mathcal{L} = 1$, $d = 3$ so that

$$\gamma_{\rm loop} = 1 - 3\nu + 3\sigma_1 + 3\sigma_3. \tag{9}$$

In the limit $\ell/L \ll 1$, the partition sum decomposes into chain and loop elements such that

$$\Gamma \sim s^{\ell} \ell^{\gamma_{\rm loop}-\gamma} L^{\gamma-1} \tag{10}$$

with $\gamma = 1 + 3\sigma_1$ as the chain part. The loop exponent

$$c = \gamma - \gamma_{\rm loop} = 3\nu - 3\sigma_3. \tag{11}$$

In $d = 3$, $\nu \approx 0.588$ and $\sigma_3 \approx -0.175$ one has

$$c \approx 2.289 \tag{12}$$

which is about 13% larger than the duplex value (so it is not a small effect).

From this simple estimate it can be concluded than a systematic study of intrinsic melting, e.g. by varying the loop length inside the Y-DNA molecule, could shed additional light on the question of the adequate value of c to use also in DNA duplex calculations. (The argument ignores the value of the cooperativity parameter, which is assumed to not vary over such a series of experiments.) We expect that in this case an equilibrium differential melting curve $\mathrm{d}\theta/\mathrm{d}T$ might reveal information about the loop entropy exponent. In addition, loop entropy effects within Y-DNA could also be tested in the dynamics of bubble denaturation, as studied for double-stranded DNA (Fogedby and Metzler 2007).

5. Conclusions

In this short paper we have looked at aspects of Y-DNA melting at three different scales: macroscopic (=extrinsic melting), mesoscopic (=intrinsic melting), and microscopic. The specific nature of the three-way junction cannot be accessed by mesoscopic models, which for double-stranded DNA is the most useful approach to DNA melting. The sequence-dependent structural variability of Y-DNA requires a more microscopic analysis, standing in the way of a general theory.

This disadvantage can, however, be turned into an advantage by 'programming' Y-DNA to address specific issues of more general relevance, e.g. of interest also for duplex DNA. We have focused here on two aspects, extrinsic melting and the loop entropy. We believe that work along the lines sketched here could indeed be useful to elucidate some fundamental aspects of DNA melting.

Acknowledgment

A Dkhissi acknowledges support from a post-doctoral grant from the CNRS.

References

Assenberg R, Weston A, Cardy D L N and Fox K R 2002 *Nucleic Acids Res.* **30** 23
Bayer J, Rädler J O and Blossey R 2005 *Nano Lett.* **5** 497
Berendsen H J C *et al* 1984 *J. Chem. Phys.* **81** 3684
Blossey R and Carlon E 2003 *Phys. Rev.* E **68** 061911
Case D A *et al* 2004 AMBER 8, UC San Francisco
Doktycz M J *et al* 1992 *Biopolymers* **32** 849
Duplantier B 1986 *Phys. Rev. Lett.* **57** 941
Feldkamp U and Niemeyer C M 2006 *Angew. Chem. Int. Edn* **45** 2
Fogedby H C and Metzler R 2007 *Phys. Rev. Lett.* **98** 070601
Kadrmas J L, Ravin A J and Leontis N B 1995 *Nucleic Acids Res.* **23** 2212
Kafri Y, Mukamel D and Peliti L 2002 *Eur. Phys. J.* B **27** 135
Leontis N B, Kwok W and Newman J S 1991 *Nucl. Acids Res.* **19** 759
Lescoute A and Westhoff E 2006 *RNA* **12** 83
Li Y *et al* 2004 *Nat. Mater.* **3** 38
Oleksi A *et al* 2006 *Angew. Chem. Int. Edn* **45** 1227
Owczarcy R *et al* 1997 *Biopolymers* **44** 217
Reynaldo L P, Vologodskii A V, Neri B P and Lyamichev V I 2000 *J. Mol. Biol.* **297** 511
Wang J, Cieplak P and Kolman P A 2000 *J. Comput. Chem.* **21** 1049
Wartell R M and Benight A S 1985 *Phys. Rep.* **126** 67