DNA terminal base pairs have weaker hydrogen bonds especially for AT under low salt concentration

Izabela Ferreira,[1, a)] Tauanne D. Amarante,[1, b)] and Gerald Weber[1, c)]
Departamento de Física, Universidade Federal de Minas Gerais, 31270-901, Belo Horizonte-MG, Brazil

(Dated: Novermber 15, 2015)

DNA base pairs are known to open more easily at the helix terminal, a process usually called end fraying, the details of which are still poorly understood. Here, we present a mesoscopic model calculation based on available experimental data where we consider separately the terminal base pairs of a DNA duplex. Our results show an important reduction of hydrogen bond strength for terminal CG base pairs which is uniform over the whole range of salt concentrations, while for AT base pairs we obtain a nearly 1/3 reduction but only at low salt concentrations. At higher salt concentrations terminal AT pair have almost the same hydrogen bond strength than interior bases. The calculated terminal stacking interaction parameters display some peculiarly contrasting behavior. While there is mostly no perceptible difference to internal stacking, for some cases we observe an unusually strong dependence with salt concentration which does not appear follow any pattern or trend.

PACS numbers: 87.14.gk, 87.15.A-, 87.15.B-
Keywords: DNA mesoscopic models, Peyrard-Bishop model, DNA terminal effects, hydrogen bonds, stacking interaction

Terminal effects play a crucial role in the overall thermodynamic stability of DNA and RNA molecules. For example in viral DNA the fraying at the helix ends contributes to end recognition.[1] Understanding the opening at the helix terminals, known as end fraying, is also of technological interest such as for logical systems by controlling the end motion of DNA.[2] However, what exactly happens at the terminal base pair is still far from clear. Even less well understood are the influences of ionic interactions or hydration patterns at the duplex terminals.

There are relatively few experimental studies dedicated to terminal effects in DNA. Nonin, Leroy, and Gueron[3], for instance, estimated the dissociation constants from NMR experiments of terminal base pairs for two DNA duplexes and concluded for a wider separation of the terminal pairs and a larger dissociation constant for terminal AT. There are also some attempts of replacing terminal base pairs with non-natural analogues. Morales and Kool[4] use non-hydrogen bonding mimics of thymine and adenine to study what would happen to terminal pairs if they were not stabilized by hydrogen bonds. However, these experiments were carried out to evaluate proofreading rates and but provided not much insight about DNA terminal effects. Similarly, Nakano *et al.*[5] replaced terminal base pairs with A/T and C/G base analogs, which displayed a lower stability due to the loss of the hydrogen bonds.

Dangling ends are much better understood than blunt terminals as they have attracted much attention since the early days of oligonucleotide synthesis[6] and may give us some clues at what happens at the helix ends. Dangling ends tend to increase the duplex stability[7–9] and it is thought that this happens by shielding the hydrogen bonds of the terminal base pair from water.[10] Therefore, one would expect that the unprotected hydrogen bonds of blunt terminals would be much more susceptible to interaction with water. But in practice it is unknown how strong this effect is and what role the counter-ions would play.

How do theoretical models incorporate end effects? For molecular dynamics (MD) simulations understanding the terminal effects of DNA poses a very hard challenge as recently reviewed in by Zgarbová *et al.*[11]. Indeed, most researchers working with MD usually refrain from placing AT base pairs at the helix terminals to avoid end fraying during the simulations. This limitation probably stems from an incomplete knowledge of the force fields required for the simulations.[11]

Early uses of the nearest-neighbor (NN) models based on Gibbs free energy have struggled to provide acceptable melting temperature predictions, to the point that some authors expressed a rather negative view about this simplified model.[12] It was not until the introduction of initiation and terminal factors that those predictions started to improve considerably.[13,14] Still, quite a number of different model implementations exist for those factors, as reviewed by Guerra and Licínio[15], which complicates qual-

[a)] Electronic mail: izabelaferreira13@gmail.com
[b)] Electronic mail: tauamarante@gmail.com
[c)] Electronic mail: gweberbh@gmail.com

itative comparisons. But crucially, the resulting terminal factors, very much as all NN entropy and enthalpy parameters, offer very little insight about intra-molecular processes.

End fraying appears quite naturally in mesoscopic models such as the Peyrard-Bishop (PB) model[16] as well as in coarse-grained models such as by Šulc *et al.*.[17] This is largely due to the melting cooperativity along the helix which is well represented by those models. In particular, for the PB model the cooperativity is ensured by the coupling between adjacent bases in the Hamiltonian.[18] There is, in principle, no restriction in using different parameters for the terminal base pairs in this models. For instance, one could use different parameters for hydrogen bonds of base pairs located at the helix ends. Unfortunately, there is currently not sufficient knowledge to determine those terminal-related parameters.

The PB model however has the capability to overcome this difficulty. To our knowledge, this model is unique in that it is computationally feasible to reverse-engineer all parameters from melting temperatures in a very systematic way.[19] This provides us with the means of reinterpreting existing experimental data and extract new information about hydrogen bonds and stacking interactions for oligonucleotides.[19–22]

There is a growing interest in obtaining parameters for the PB model, for instance Dahlen and van Erp[23] recently investigated the parameter dependencies in the context of denaturation rates. One of the attractive properties of this model is that its Hamiltonian can be easily adapted to reflect a number of different experimental situations such as DNA overstretching[24] or to include solvent interactions.[25] Several theoretical approaches also rely on the PB model parameters as for example the path integral method[26,27] or the mesoscopic model to study the mechanical response of DNA recently proposed by Nisoli and Bishop.[28]

Here, we use the PB model to calculate the parameters for terminal DNA base pairs. This is accomplished by considering different potentials for the terminal bases and optimizing them separately from the internal bases. Effectively, this means almost tripling the number of parameters which makes it computationally much more difficult to perform the numerical minimization.

Here, we use a semi-empirical regression scheme which combines experimental melting temperature data and the Peyrard-Bishop (PB) model.[19,29] From the PB model and given a set of L parameters $P = \{p_1, p_2, \ldots, p_L\}$ we calculate an adimensional thermal index $\tau_i(P) = \omega_{\max}^{1/2}$ which can be used to predict the melting temperatures T_i of the ith sequence.[30] The parameter $\omega_{\max}^{1/2}$ is calculated from the classical partition function and represents a measure of sequence ordering combined with the base pair and stacking interaction. For details on the calculation of $\omega_{\max}^{1/2}$ and the related melting temperature regression see Ref. 30. We let the parameters P vary until we minimize the squared difference between the experimental and predicted temperatures

$$\chi_j^2 = \sum_{i=1}^{N} \left[T_i'(P_j) - T_i\right]^2 . \tag{1}$$

where P_j is the jth tentative set of parameters and N is the number of sequences in the dataset. The numerical parameter optimization is performed by a downhill simplex multidimensional minimization algorithm.[31]

The parameters of the PB model are those contained in the configurational part of the Hamiltonian[30]

$$U_{n,n+1} = \frac{k_{n,n+1}}{2}\left(y_n - y_{n-1}\right)^2 + D_n\left(e^{-y_n/\lambda_n} - 1\right)^2 , \tag{2}$$

describing the interaction of the nth base pair with its nearest-neighbors $n+1$. D_n is the depth and λ_n the width of Morse potential of the nth base pair which can be related to the hydrogen bond strength. The elastic constant $k_{n,n+1}$ describes the stacking interaction of the nearest-neighbors. The coordinate y represents the relative distance between the bases. Note that this simpler model is being preferred over more elaborate potentials such as including anharmonic stacking[32] or solvent interactions[25], since we have found that the simpler harmonic potential provides a better description of the melting temperatures of short DNA sequences[19]. The PB model, while lacking the abrupt melting transition characteristic of other models, still keeps the crucial base-pair cooperativity which causes the colaborative increase of the average base-pair opening with temperature, see for instance Fig. 1 of Ref. 33.

Usually, we would consider 2 Morse potentials D and 10 stacking parameters k to describe the canonical AT and CG base pairs for DNA. We will call this scheme the 'uniform' parameters for the remainder of this article. However, to distinguish terminal effects we need to consider additional parameters to describe the interactions at the end of the helix. First, let us establish our notation by considering the following example sequence

5′-AGCAAGTC-3′
3′-TCGTTCAG-5′

which we split into internal and terminal base pairs

$$\underbrace{\begin{matrix}\mathrm{A}\\ \mathrm{T}\end{matrix}}_{\text{terminal}} \cdot \underbrace{\begin{matrix}\mathrm{GCAAGT}\\ \mathrm{CGTTCA}\end{matrix}}_{\text{internal}} \cdot \underbrace{\begin{matrix}\mathrm{C}\\ \mathrm{G}\end{matrix}}_{\text{terminal}}$$

The terminal base pairs will be superscribed with *, in our example AT* at the 5′-side and CG* at the 3′-side. From the point of view of the Morse potential AT* and TA*, as well as CG* and GC* base pairs are symmetrical and share the same parameters D and λ, see Eq. (2). However, for the nearest-neighbor (NN) stacking parameter k we have a mixed notation of terminal and internal base pairs. For instance, the first NN pair of our example would be AT*pGC, that is a terminal AT* followed

by an internal GC. The AT*pGC pair is symmetric to CGpTA*,

$$\begin{array}{l} 5'\text{-}\mathbf{AG}\text{CAAGTC-}3' \\ 3'\text{-}\mathbf{TC}\text{GTTCAG-}5' \end{array} \leftrightarrow \begin{array}{l} 5'\text{-GACTTG}\mathbf{CT}\text{-}3' \\ 3'\text{-CTGAAC}\mathbf{GA}\text{-}5' \end{array}$$

Since both can be described by the same stacking parameter k, we retain only the one that precedes alphabetically, in this case AT*pGC. Therefore, the stacking parameter for ATpGC NN pairs is divided into three separate parameters: AT*pGC for terminal AT, ATpGC* for terminal CG and for internal NNs we keep the original notation ATpGC. In some cases, due to the symmetry of the NN pair, there will be only one additional parameter. For instance, CGpGC has only one terminal related NN CGpGC* since it is symmetric to CG*pGC. Considering all possible combinations of the canonical base pairs, the existing 10 stacking parameters are now complemented by additional 16 terminal related variables. Together with the 2 new terminal related Morse potentials D, the minimization searching space now extends from the original 12 to 30 parameters. From our previous calculations[19] we noticed that λ has a negligible effect on the merit function χ^2. Therefore, we decided to keep this parameter constant to the same values as for the uniform sequence calculation[19] and to avoid a further increase of the number of parameters to optimize.

To allow a direct comparison with our previous result for DNA,[19] we used the same melting temperature datasets and experimental uncertainties. We used all sequences of the high quality measurements by Owczarzy et al.[34] over 5 different salt concentrations and with 0.3 °C declared experimental uncertainty. Supplementary table S1[35] shows the number of occurrences of internal and terminal base pairs and NNs following the notation outlined previously.

We performed the parameter optimization, that is, the minimization of Eq. (1) in two separate rounds. First we run the minimization by varying the initial parameters randomly over an interval which averages to the values obtained for uniform Morse potentials.[19,21] In other words, the initial internal parameter p^i as well as the initial terminal parameter p^* are sampled between $0.5p^u$ and $1.5p^u$, where p^u is the uniform parameter calculated previously.[19] This is repeated 200 times and we calculate the average of the parameters with lowest χ^2 which is used as new fixed initial set of parameters for the second round of minimizations. However, this time we change the temperatures of the dataset by small random amounts such that the standard deviation between the original set and the modified set approaches the declared experimental uncertainty. This is again repeated 200 times and provides us with an estimate of the uncertainty over the calculated parameters. The results presented here are the averages over these minimizations. These minimizations were carried out independently for each of the 5 available salt concentrations and took a total of 20000 h processing time on 2 GHz processors. The overall reduction of χ^2, as compared to the uniform

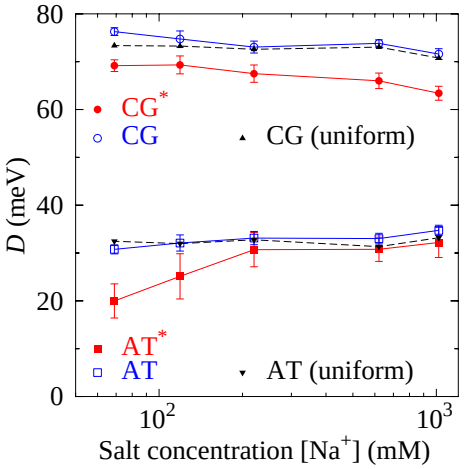

FIG. 1. Average DNA Morse potential D as function of salt concentration. Internal base pairs are shown as blue boxes (AT) and circles (CG). Terminal base pairs are shown as red filled boxes (AT*) and and bullets (CG*). The error bars are the calculated standard deviations. For comparison, base pairs calculated for uniform Morse potentials in Ref. 19 are shown as black triangles (error bars were omitted for clarity). Lines connecting the data points are intended as guides to the eye.

calculation[19] is between 13–25% depending on salt concentration (see supplementary Table S2[35]). This corresponds to average temperature deviations ΔT ranging between 0.73–0.80 °C, also shown in Tab. S2. Note that this is still larger than the ΔT of the order of 0.6 °C obtained for optimized Gibbs free energies,[36] which gives us confidence that increasing the number of parameters did not result in significant over-fitting. When separated into groups of same lengths, shown in supplementary Table S3[35], we notice that the poorest predictions are for the shortest sequences of length 10 bp. This is also the only group where we observe a continuous increase of ΔT with salt concentration. The best predictions are for the groups of intermediate lengths between 15–25 bp. For the longest sequences of size 30 bp the ΔT is somewhat larger but shows no discernible trend with salt concentration.

In Fig. 1 we show the calculated Morse potentials for DNA considering internal and terminal base pairs separately. Generally, we observe little difference between internal and uniform base pairs. This is expected as most base pairs are internal and therefore should have dominated the contribution for the uniform calculation. However, for terminal base pairs there are several important changes. Terminal CG base pairs present smaller Morse potentials when compared to internal base pairs. This smaller potential, around 5 meV less, shows little dependence with salt concentration. Terminal AT base pairs display an altogether more intriguing behavior. They start with a surprisingly large difference of 12 meV for low salt concentrations as shown in Fig. 1. This difference decreases rapidly with salt concentration and at 621 mM

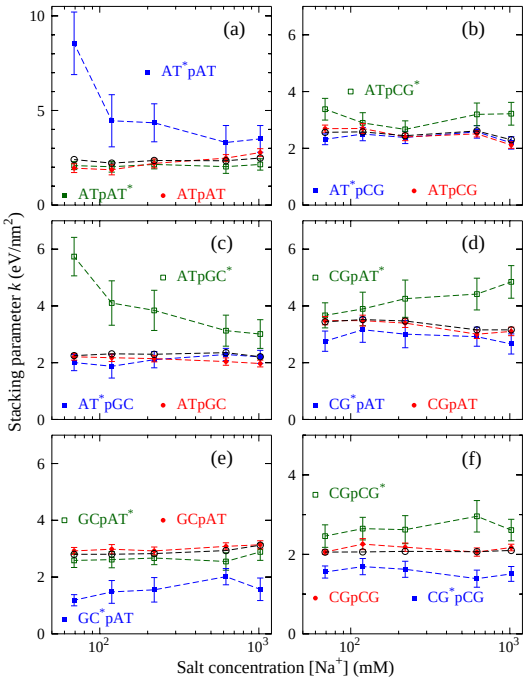

FIG. 2. Average DNA stacking parameter k for non-symmetrical NNs as function of salt concentration. Panels (a) show AT-AT, (b)–(e) AT-CG and (f) CG-CG nearest neighbors. The error bars are the calculated standard deviations. Red bullets refer to internal NN stacking parameters and green and blue boxes are for terminal NNs. For comparison, the black circles show the uniform parameters previously calculated in Ref. 19, where error bars were omitted for clarity. Vertical scales were adjusted for each panel to highlight differences and connecting dashed lines are intended as guides to the eye.

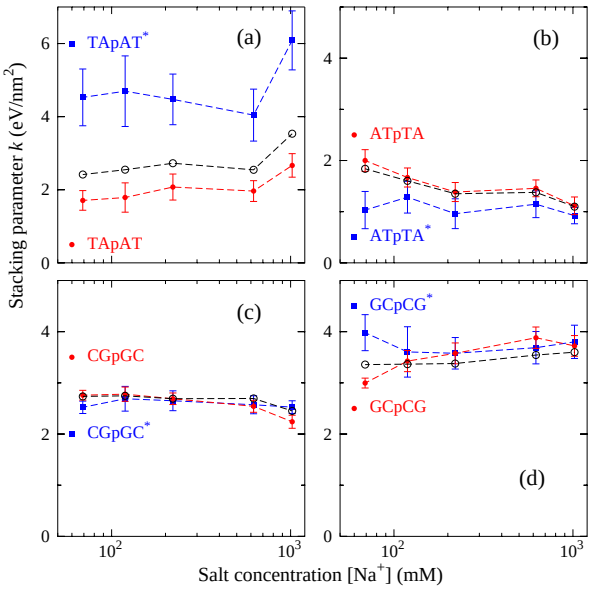

FIG. 3. Average DNA stacking parameter k for symmetrical NNs as function of salt concentration. Panels (a)–(b) show AT-AT and (c)–(d) CG-CG nearest neighbors. Red bullets refer to internal NN stacking parameters and blue boxes are for terminal NNs. Remaining figure elements are as in Fig. 2.

it almost vanishes. In other words, at high salt concentration the terminal AT Morse potentials are nearly as strong as their internal counterparts. This could perhaps explain why there is no observable increase in stability when AT terminals are capped either with polar or apolar carbohydrates,[37,38] since they are at their highest stability already.

The overall reduction of the Morse potential is likely due to the exposure of the hydrogen bonds to water.[39,40] But what could be the possible cause of the differences between the salt-dependent behavior of AT and CG Morse potentials? It is known that Na^+ binds predominantly to the minor groove of AT in DNA,[41] therefore salt concentration variations should be much more important for AT base pairs than for CG. This is indeed what is observed in Fig. 1. The lack of further changes for higher salt concentrations could perhaps indicate a saturation of the available sites for Na^+ binding.

The calculated terminal nearest neighbors (NN) stacking parameters are shown in Figs. 2 and 3, for non-symmetric and symmetric NNs, respectively. The largest stacking parameter is found for AT*pAT exceeding 8 eV/nm^2 as shown in Fig. 2a which rapidly drops to half of this value with increasing salt concentration. On the other hand, its terminal counterpart ATpAT* follows the internal NNs very closely. This contrasting behavior is observed for some other terminal NNs, such as ATpGC* in Fig. 2c However, for the symmetric CGpGC* and GCpCG*, Fig. 3c and 3d, there is hardly any difference. Clearly, the hydrogen bonding has a role in the variation of stacking interaction as demonstrated by the near absence of visible effects for CG-CG NNs in Fig. 3c and 3d, while the largest variations were observed for AT with AT NN Figs. 2a and 2c. However, what is not clear at all is why for instance AT*pAT has a large stacking parameter variation while its counterpart ATpAT* has virtually none. The NN stacking parameters that show important variation with salt concentration do not appear to fall into any discernible pattern such as pyrimidine/purine motifs or a 5′-end predominance over 3′ as commonly seen for dangling ends.[7–9] One possibility is that specific structural factors, such as propeller-twist angles,[42] may have some influence here. Unfortunately, there is currently not sufficient knowledge about end structure that would allow such a correlation, but the present results could serve as a starting point for additional experimental or theoretical studies.

The new terminal related parameters can be readily used with our free software implementation of the PB model[33] for the calculation of average opening profiles or to verify the accuracy of the present calculation. The most acurate predictions are for the range of 15 bp to 25 bp sequence lengths as shown in supplementary Table S3[35]. The variable Morse potentials could also be

used straightforwardly in modified a salt dependent PB Hamiltonian as proposed by Singh and Singh[43]. We believe that the new knowledge of the hydrogen bond related Morse potentials, especially in regard to salt concentration dependence could guide the planning of new experiments or be a target for adjusting molecular dynamics simulations[11] which aim to study the end effects of DNA.

This work was supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

[1]R. A. Katz, G. Merkel, M. D. Andrake, H. Roder, and A. M. Skalka, "Retroviral integrases promote fraying of viral DNA ends," Journal of Biological Chemistry **286**, 25710–25718 (2011).
[2]N. Kanayama, T. Takarada, M. Fujita, and M. Maeda, "DNA terminal breathing regulated by metal ions for colloidal logic gates," Chemistry-A European Journal **19**, 10794–10798 (2013).
[3]S. Nonin, J.-L. Leroy, and M. Gueron, "Terminal base pairs of oligodeoxynucleotides: imino proton exchange and fraying," Biochemistry **34**, 10652–10659 (1995).
[4]J. C. Morales and E. T. Kool, "Importance of terminal base pair hydrogen-bonding in 3'-end proofreading by the Klenow fragment of DNA polymerase I," Biochemistry **39**, 2626–2632 (2000).
[5]S.-i. Nakano, H. Oka, Y. Uotani, K. Uenishi, M. Fujii, and N. Sugimoto, "Stacking interaction in the middle and at the end of a DNA helix studied with non-natural nucleotides," Molecular BioSystems **6**, 2023–2029 (2010).
[6]S. M. Freier, D. Alkema, A. Sinclair, T. Neilson, and D. H. Turner, "Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helixes," Biochemistry **24**, 4533–4539 (1985).
[7]S. Bommarito, N. Peret, and J. J. SantaLucia, "Thermodynamic parameters for DNA sequences with dangling bonds," Nucl. Acids. Res. **28**, 1929–1934 (2000).
[8]B. G. Moreira, Y. You, M. A. Behlke, and R. Owczarzy, "Effects of fluorescent dyes, quenchers, and dangling ends on DNA duplex stability," Biochemical and Biophysical Research Communications **327**, 473–484 (2005).
[9]R. Dickman, F. Manyanga, G. P. Brewood, D. J. Fish, C. A. Fish, C. Summers, M. T. Horne, A. S. Benight, *et al.*, "Thermodynamic contributions of 5-and 3-single strand dangling-ends to the stability of short duplex dnas," Journal of Biophysical Chemistry **3**, 1 (2012).
[10]J. Isaksson and J. Chattopadhyaya, "A uniform mechanism correlating dangling-end stabilization and stacking geometry," Biochemistry **44**, 5390–5401 (2005).
[11]M. Zgarbová, M. Otyepka, J. Sponer, F. Lankas, and P. Jurečka, "Base pair fraying in molecular dynamics simulations of DNA and RNA," Journal of Chemical Theory and Computation **10**, 3177–3189 (2014).
[12]R. Kierzek, M. H. Caruthers, C. E. Longfellow, D. Swinton, D. H. Turner, and S. M. Freier, "Polymer-supported RNA synthesis and its application to test the nearest-neighbor model for duplex stability," Biochemistry **25**, 7840–7846 (1986).
[13]N. Sugimoto, S. Nakano, M. Yoneyama, and K. Honda, "Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes," Nucl. Acids. Res. **24**, 4501–4505 (1996).
[14]J. SantaLucia, Jr., H. T. Allawi, and P. A. Seneviratne, "Improved nearest-neighbour parameters for predicting DNA duplex stability," Biochem. **35**, 3555–3562 (1996).
[15]J. C. d. O. Guerra and P. Licínio, "Terminal contributions for duplex oligonucleotide thermodynamic properties in the context of nearest neighbor models," Biopolymers **95**, 194–201 (2011).
[16]Y.-L. Zhang, W.-M. Zheng, J.-X. Liu, and Y. Z. Chen, "Theory of DNA melting based on the Peyrard-Bishop model," Phys. Rev. E **56**, 7100–7115 (1997).
[17]P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. Doye, and A. A. Louis, "Sequence-dependent thermodynamics of a coarse-grained DNA model," The Journal of Chemical Physics **137**, 135101 (2012).
[18]M. Peyrard and A. R. Bishop, "Statistical mechanics of a nonlinear model for DNA denaturation," Phys. Rev. Lett. **62**, 2755–2757 (1989).
[19]G. Weber, J. W. Essex, and C. Neylon, "Probing the microscopic flexibility of DNA from melting temperatures," Nature Physics **5**, 769–773 (2009).
[20]G. Weber, "Finite enthalpy model parameters from DNA melting temperatures," Europhys. Lett. **96**, 68001 (2011).
[21]G. Weber, "Mesoscopic model parametrization of hydrogen bonds and stacking interactions of RNA from melting temperatures," Nucl. Acids. Res. **41**, e30 (2013).
[22]R. V. Maximiano and G. Weber, "Deoxyinosine mismatch parameters calculated with a mesoscopic model result in uniform hydrogen bonding and strongly variable stacking interactions," Chemical Physics Letters **631–632**, 87–91 (2015).
[23]O. Dahlen and T. S. van Erp, "Mesoscopic modeling of DNA denaturation rates: Sequence dependence and experimental comparison," The Journal of Chemical Physics **142**, 235101 (2015).
[24]A. E. B. Pupo, F. Falo, and A. Fiasconaro, "DNA overstretching transition induced by melting in a dynamical mesoscopic model," The Journal of Chemical Physics **139**, 095101 (2013).
[25]G. Weber, "Sharp DNA denaturation due to solvent interaction," Europhys. Lett. **73**, 806–811 (2006).
[26]M. Zoli, "Helix untwisting and bubble formation in circular DNA," The Journal of Chemical Physics **138**, 205103 (2013).
[27]M. Zoli, "Entropic penalties in circular DNA assembly," The Journal of Chemical Physics **141**, 174112 (2014).
[28]C. Nisoli and A. Bishop, "Thermomechanical stability and mechanochemical response of DNA: A minimal mesoscale model," The Journal of Chemical Physics **141**, 115101 (2014).
[29]G. Weber, N. Haslam, N. Whiteford, A. Prügel-Bennett, J. W. Essex, and C. Neylon, "Thermal equivalence of DNA duplexes without melting temperature calculation," Nature Physics **2**, 55–59 (2006).
[30]G. Weber, N. Haslam, J. W. Essex, and C. Neylon, "Thermal equivalence of DNA duplexes for probe design," J. Phys. Condens. Matter **21**, 034106 (2009).
[31]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1988).
[32]T. Dauxois, M. Peyrard, and A. R. Bishop, "Entropy-driven DNA denaturation," Phys. Rev. E **47**, R44–R47 (1993).
[33]G. Weber, "TfReg: Calculating DNA and RNA melting temperatures and opening profiles with mesoscopic models," Bioinformatics **29**, 1345–1347 (2013).
[34]R. Owczarzy, Y. You, B. G. Moreira, J. A. Manthey, L. Huang, M. A. Behlke, and J. A. Walder, "Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures," Biochem. **43**, 3537–3554 (2004).
[35]See supplemental material at [URL will be inserted by AIP] for the number of occurrences of base pairs and NN pairs contained in the dataset and the calculated temperature deviations.
[36]G. Weber, "Optimization method for obtaining nearest-neighbour DNA entropies and enthalpies directly from melting temperatures," Bioinformatics **31**, 871–877 (2015).
[37]R. Lucas, I. Gómez-Pinto, A. Aviñó, J. J. Reina, R. Eritja, C. González, and J. C. Morales, "Highly polar carbohydrates stack onto DNA duplexes via CH/π interactions," Journal of the American Chemical Society **133**, 1909–1916 (2011).
[38]R. Lucas, E. Vengut-Climent, I. Gómez-Pinto, A. Aviñó, R. Eritja, C. González, and J. C. Morales, "Apolar carbohydrates as DNA capping agents," Chemical Communications **48**, 2991–2993 (2012).

[39]E. Liepinsh, G. Otting, and K. Wüthrich, "NMR observation of individual molecules of hydration water bound to DNA duplexes: direct evidence for a spine of hydration water present in aqueous solution," Nucleic Acids Research **20**, 6549–6553 (1992).
[40]M. Becker, V. Lerum, S. Dickson, N. C. Nelson, and E. Matsuda, "The double helix is dehydrated: evidence from the hydrolysis of acridinium ester-labeled probes," Biochemistry **38**, 5603–5611 (1999).
[41]F. Cesare Marincola, V. P. Denisov, and B. Halle, "Competitive Na+ and Rb+ binding in the minor groove of DNA," Journal of the American Chemical Society **126**, 6739–6750 (2004).
[42]W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. H. U. Heinemann, X.-J. Lu, S. Neidle, Z. S. H. Sklenar, M. Suzuki, C.-S. Tung, E. W. C. Wolberger, and H. M. Berman, "A standard reference frame for the description of nucleic acid base-pair geometry," J. Mol. Biol. **313**, 229–237 (2001).
[43]A. Singh and N. Singh, "Effect of salt concentration on the stability of heterogeneous DNA," Physica A: Statistical Mechanics and its Applications **419**, 328–334 (2015).