JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Effect of Inert Tails on the Thermodynamics of DNA Hybridization

Lorenzo Di Michele,[†,∥] Bortolo M. Mognetti,[‡,∥,⊥] Taiki Yanagishima,[†,§] Patrick Varilly,[‡] Zachary Ruff,[†] Daan Frenkel,*[,‡] and Erika Eiser*[,†]

[†]Cavendish Laboratory, Department of Physics, University of Cambridge, J.J. Thomson Avenue, CB3 0HE Cambridge, United Kingdom

[‡]Deparment of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW Cambridge, United Kingdom

Supporting Information

ABSTRACT: The selective hybridization of DNA is of key importance for many practical applications such as gene detection and DNA-mediated self-assembly. These applications require a quantitative prediction of the hybridization free energy. Existing methods ignore the effects of non-complementary ssDNA tails beyond the first unpaired base. We use experiments and simulations to show that the binding strength of complementary ssDNA oligomers is altered by these sequences of non-complementary nucleotides. Even a small number of non-binding bases are enough to raise the hybridization free energy by approximately 1 kcal/mol at physiological salt concentrations. We propose a simple analytical expression that accounts quantitatively for this variation as a function of tail length and salt concentration.

DNA hybridization is the process by which a double stranded (ds) DNA is formed from two complementary sequences of single-stranded (ss) DNA. A quantitative description of DNA hybridization is necessary to predict the properties and functionalities of DNA. In particular, the reversibility and high selectivity of DNA hybridization has been utilized to design supramolecular interactions[1,2] in DNA-coated colloids,[3] DNA origami,[4] and biosensors.[5] The most common approach to predict the hybridization free energy ΔG of complementary sequences is based on the so-called "nearest-neighbor" rules.[6,7] In the unified framework of SantaLucia,[7] ΔG is computed using the tabulated free-energy contributions of all possible (oriented) nearest-neighbor pairs of bases, plus initiation parameters and contributions due to "dangling bases", which are defined as unpaired bases immediately adjacent to duplex regions.[7,8] These models do not explicitly treat the contribution from additional unpaired bases that are often present in DNA-nanotechnology experimental designs.[2,3,9,10]

In this Communication we present experiments and simulations showing that the hybridization free energy of DNA is significantly altered when the reactive sequences of length n_{hyb} are connected to inert tails of length n_{tail} in addition to the dangling base (Figure 1). The electrostatic repulsion between the tails and the hybridized section hinders the dimer formation. The binding weakens upon increasing n_{tail} and the ionic screening length. At physiological ionic strengths, the nonspecific free energy shift is enough to compensate the contribution of a complementary base pair. The observed effects can be captured by a simple empirical formula capable of predicting the tail corrections to the SantaLucia estimate of the hybridization free energies and melting temperatures. Beyond duplex formation, our findings can be used in other contexts, for instance to rationalize the influence of tails in toehold-mediated strand displacement.[9a]

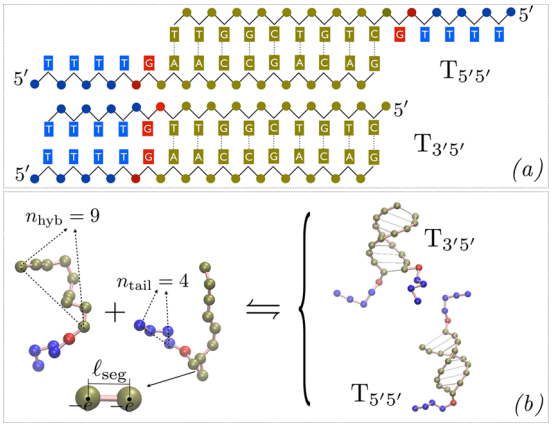

Figure 1. (a) Schematic view and (b) simulation snapshots of two hybridizing oligomers made of n_{hyb} complementary sequences (yellow), a dangling term (red), and n_{tail} inert bases (blue). The dots are a representation of the charged phosphate backbone.

Using UV absorbance spectroscopy,[11] we measure ΔG and the melting temperature for DNA duplexes with inert tails of different length (see SI, section S2). The results are summarized in Figure 2. We consider two topologies in which inert tails are either attached to the same side of the assembled dsDNA duplex ($T_{3'5'}$) or to opposite sides, namely to the 5′ termini of the DNA ($T_{5'5'}$) (Figure 1). We choose an effective monovalent salt concentration of I = 125 mM,[10] due to the presence of 100 mM NaCl and a 10 mM phosphate buffer at pH 7.5. In Figure 2a, we plot the shift in the hybridization free energy defined as

$$\delta\Delta G(n_{tail}) = \Delta G(n_{tail}) - \Delta G^0 \quad (1)$$

where $\Delta G^0 = \Delta G(n_{tail} = 0)$. We chose to measure $\delta\Delta G$ at $T_m{}^0 = T_m(n_{tail} = 0)$, where T_m is defined as the temperature where there are equal amounts of DNA in monomer and dimer form.

Received: January 14, 2014

   dx.doi.org/10.1021/ja500027v | J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

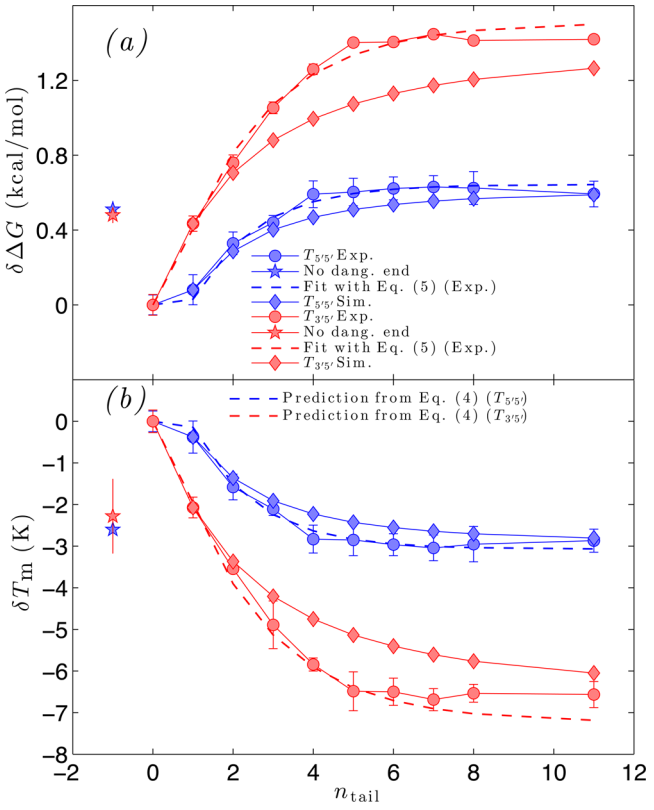

Figure 2. (a) Experiments (circles) and computer simulations (lozenges) of the free energy shift due to inert tails for strands with n_{hyb} = 9 with $T_{5'5'}$ (blue) and $T_{3'5'}$ (red) architectures (see Figure 1 and SI, section S2.1 for the sequences). For the $T_{5'5'}$ geometry, $\delta\Delta G$ is evaluated at $T_m(n_{tail})$. Dashed lines are fits with eq 5. (b) Melting temperature shift for the same strands. The simulated melting temperatures are extracted from $\delta\Delta G$ using eq 4 and the experimental estimates of T_m^0 and ΔH^0 tabulated in SI (Table S1). The dashed lines are predictions extracted from the experimental fits in panel a using eq 4. Stars indicate the sequences without the dangling ends. The concentration of both the ssDNA strands is ρ = 5 μM.

Adding a small number of inert bases substantially increases the hybridization free energy. However, for $n_{tail} \approx 5$, $\delta\Delta G$ saturates at the constant values of ~0.6 kcal/mol for the $T_{5'5'}$ strands and ~1.3 kcal/mol for the $T_{3'5'}$ strands. The effect of steric repulsion overcomes the contribution to ΔG due to the dangling terms (stars in Figure 2a). A similar qualitative behavior is observed in Figure 2b for the melting temperature shift

$$\delta T_m(n_{tail}) = T_m(n_{tail}) - T_m^0 \quad (2)$$

The plateaus correspond to $\delta T_m \approx -3$ and -6.5 K for $T_{5'5'}$ and $T_{3'5'}$, respectively.

We preform computer simulations based on a minimal model of screened charges, representing the backbone of the DNA, to highlight the role of electrostatics. Oligomers are mapped onto freely jointed chains with $n_{hyb} + n_{tail} + 1$ segments of length l_{seg} = 6.47 Å, corresponding to the distance between two backbone phosphates. In the dimerized strand, hybridized charges are located at the native phosphate positions of the dsDNA (see Figure 1 and SI, section S3.1). We compute $\Delta G(n_{tail})$ using Rosenbluth Monte Carlo simulations.[12] The experimental results for the $T_{5'5'}$ architecture (Figure 2a) show a kink at n_{tail} = 1 due to residual stacking between the first base of the tail and the dsDNA. Therefore, we compute the excess in ΔG,

$$\Delta G(n_{tail}) - \Delta G(n_{ref}) = \delta\Delta G(n_{tail}) - \delta\Delta G(n_{ref}) \quad (3)$$

caused by having n_{tail} bases with respect to a reference tail-length, using n_{ref} = 1 and the experimental value of $\delta\Delta G(n_{tail} = 1)$ as an input parameter (see SI, section S3.2). The results are presented in Figure 2a.

Overall the simulations reproduce the experimental results very well, and the agreement for the plateau value is almost quantitative. The simplicity of the coarse-grained model, along with the similarity between experimental and simulation results, supports the hypothesis that the electrostatic interaction between the tails and the dsDNA duplex is responsible for the shift. For the $T_{3'5'}$ architecture, the effect is further enhanced by tail−tail interactions.

From our knowledge of $\delta\Delta G$, we can predict the melting temperature shift as (SI, section S3.4)

$$\delta T_m(n_{tail}) \simeq \frac{T_m^0 \delta\Delta G}{\Delta H^0} \quad (4)$$

where $\Delta H^0 = \Delta H(n_{tail} = 0)$ is the hybridization enthalpy for n_{tail} = 0, which we measure experimentally (SI, section S2.2), but which could also be estimated using the nearest-neighbor rules.[7,8] Eq 4 would be exact (within the approximations of the two-state model of melting) if $\delta\Delta G$ were evaluated at $T = T_m(n_{tail})$ rather than at T_m^0. However, the differences between $\delta\Delta G(T_m^0)$ and $\delta\Delta G[T_m(n_{tail})]$ are much smaller than the statistical errors; therefore, both forms can be used (SI, section S2.2). The melting temperature shift derived from the simulated $\delta\Delta G$ using eq 4 is shown in Figure 2b. The agreement between the simulated and experimental values for δT_m is quite good and almost quantitative for the plateau value.

We now quantify the influence of the specific design of DNA oligomers on the tail effect. We focus on the most common $T_{5'5'}$ architecture. $T_{3'5'}$ would produce similar results, only with larger plateaus. In Figure 3a, we present experimental data showing the dependence of $\delta\Delta G$ on n_{hyb}. In a range of values usually employed when designing reversible supramolecular interactions, we can see that $\delta\Delta G$ is only weakly affected by n_{hyb}. This indicates that, for the $T_{5'5'}$ case, tail−tail interactions are negligible. Tail−tail interactions could play a role for very high values of n_{nail}, but, as explained in the SI (section S3.5), the resulting corrections should be negligible. We chose sequences (SI, section S2.1) for which changes in n_{hyb} affect T_m^0 by less than 3% (Table S2). Consequently, possible differences in $\delta\Delta G$ due to changes in T_m are well within experimental errors. In Figure 3b we show the experimentally determined δT_m. In this case, we notice a dependence on n_{hyb}, which is due to the n_{hyb} dependence of ΔH^0, as described by eq 4. In panels c and d, we show simulation results for $\delta\Delta G$ and δT_m, which are in good agreement with experiments. In Figure 3e−h we show the effect of changing ionic strength from I = 50 to 225 mM, at fixed n_{hyb} = 7. As expected, at high salt concentration the effect of the tails becomes small, while at low salt concentration the plateau value is larger.

The results in Figures 2a and 3a are quantitatively reproduced by the following interpolating formula,

$$\delta\Delta G_{fit}(n_{tail} \geq 1) = \delta\Delta G_{plat}\left[1 - \exp\left(-\frac{n_{tail}}{n_{plat}}\right)\right] - \delta\Delta G_{stack} \quad (5)$$

In I = 125 mM solutions we find, for the $T_{5'5'}$ architecture, $\delta\Delta G_{plat}$ = 1.15 kcal/mol, n_{plat} = 1.58, and $\delta\Delta G_{stack}$ = 0.51 kcal/

dx.doi.org/10.1021/ja500027v | J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

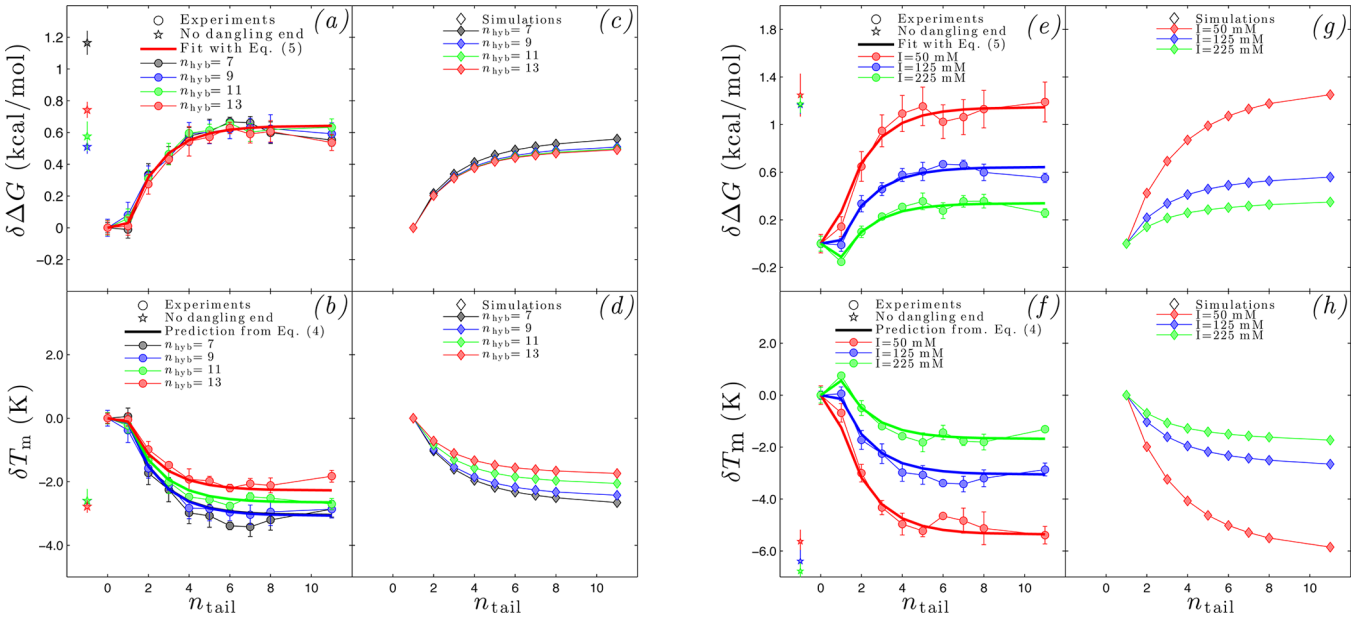

Figure 3. (a) Experimental free-energy shift for four different pairs of strands with n_{hyb} = 7, 9, 11, and 13 with $T_{5'5'}$ architecture. The red solid line is a fit with eq 5. (b) Experimental melting temperature shift for the strands in panel a. Solid lines are predictions calculated using the fit in panel a and eq 4. (c) Simulated $\delta\Delta G$ for the sequences in panel a. (d) Melting temperature shifts obtained using $\delta\Delta G$ from panel c and eq 4. (e) Experimental free energy shift for strands with n_{hyb} = 7 and different ionic strengths I. For I = 50, the free energy shifts is evaluated at $T_m(n_{tail})$. For full details see SI. Solid lines are fits with eq 5, in which the parameters n_{plat} and ΔG_{stack} are kept fixed to the values fitted in panel a. (f) Experimental melting temperature shifts for the strands in panel e. Solid lines are predictions calculated using the fit in panel e and eq 4. (g) Simulated $\delta\Delta G$ for the samples in panel e. (h) Melting temperature shifts obtained using $\delta\Delta G$ from panel g and eq 4. When applying eq 4 to extract δT_m, we use the experimental estimates of $T_m{}^0$ and ΔH^0 tabulated in Table S1.

mol (Figures 2a and 3a). For $T_{3'5'}$ we find $\delta\Delta G_{plat}$ = 1.75 kcal/mol, n_{plat} = 2.17, and $\delta\Delta G_{stack}$ = 0.24 kcal/mol (Figure 2a). The parameter $\delta\Delta G_{stack}$ describes the free-energy shift responsible for the kink in the experimental curves at n_{tail} = 1, which is almost absent for $T_{3'5'}$. Equations 4 and 5 can be used to predict $\delta T_m(n_{tail})$ for generic ssDNA sequences starting from the values of ΔH^0 and $T_m{}^0$ obtained using the nearest-neighbor rules.[7,8] In Figures 2b and 3b, we compare the melting-temperature shift predicted by this procedure with the experimental measurements: the agreement is excellent. Figure 3e illustrates that eq 5 can account for the ionic-strength dependence of $\delta\Delta G$. We find that $\delta\Delta G_{stack}$ and n_{plat} are rather insensitive to the ionic strength. We therefore keep these quantities fixed at the values fitted in Figure 3a and optimize $\delta\Delta G_{plat}$ to account for the change in the ionic strength. We obtain $\delta\Delta G_{plat}$ = 1.66 and 0.85 kcal/mol for I = 50 and 225 mM, respectively. The values of the fitting parameters $\delta\Delta G_{plat}$, $\delta\Delta G_{stack}$, and n_{plat} for all cases we tested are tabulated in the SI (Table S2). In Figure 3f we make use of eq 4 to predict δT_m from the fits in Figure 3e. The agreement with the experiments is good.

Importantly, eq 5 can be integrated into the nearest-neighbor phenomenological estimates of the two-state hybridization free energy.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Full details on experimental and simulation procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

df264@cam.ac.uk

ee247@cam.ac.uk

Present Addresses

[⊥]Physics of Complex Systems and Statistical Mechanics, Département de Physique, Université Libre de Bruxelles, Belgium

[§]Institute of Industrial Science, University of Tokyo, Japan

Author Contributions

[‖]L.D. and B.M.M. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported by the ERC Advanced Grant 227758, the Wolfson Merit Award 2007/R3 of the Royal Society of London, the EPSRC Programme Grant EP/I001352/1, the Marie Curie Initial Training Network ITN-COMPLOIDS grant no. 234810, the Ernest Oppenheimer Fund, and the Winton Programme for the Physics of Sustainability of the Cavendish Laboratory.

■ REFERENCES

(1) Kallenbach, N. R.; Ma, R.-I.; Seeman, N. C. *Nature* **1983**, *305*, 829.

(2) (a) Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. *Nature* **1996**, *382*, 607. (b) Alivisatos, A. P.; Johnsson, K. P.; Peng, X.; Wilson, T. E.; Loweth, C. J.; Bruchez, M. P.; Schultz, P. G. *Nature* **1996**, *382*, 609. (c) Macfarlane, R. J.; Lee, B.; Jones, M. R.; Harris, N.; Schatz, G. C.; Mirkin, C. A. *Science* **2011**, *334*, 204. (d) Maye, M. M.; Nykypanchuk, D.; van der Leile, D.; Gang, O. *J. Am. Chem. Soc.* **2006**,

dx.doi.org/10.1021/ja500027v | *J. Am. Chem. Soc.* XXXX, XXX, XXX−XXX

128, 14020. (e) Nykypanchuk, D.; Maye, M. M.; van der Leile, D.; Gang, O. *Nature* **2008**, *451*, 549.
(3) Di Michele, L.; Eiser, E. *Phys. Chem. Chem. Phys.* **2013**, *4*, 3158.
(4) (a) Rothemund, P. W. K. *Nature* **2006**, *440*, 297. (b) Dongran, H.; Suchetan, P.; Jeanette, N.; Zhengtao, D.; Yan, L.; Hao, Y. *Science* **2011**, *332*, 324.
(5) (a) Zhang, K.; Hao, L.; Hurst, S. J.; Mirkin, C. A. *J. Am. Chem. Soc.* **2012**, *134*, 16488. (b) Young, K. L.; Scoot, A. W.; Hao, L.; Mirkin, S. E.; Liu, G.; Mirkin, C. A. *Nano Lett.* **2012**, *12*, 3867.
(6) (a) Crothers, D. M.; Zimm, B. H. *J. Mol. Biol.* **1964**, *9*, 1. (b) De Voe, H.; Tinoco, I., Jr. *J. Mol. Biol.* **1962**, *4*, 500. (c) Dirks, R. M.; Bois, J. S.; Schaeffer, J. M.; Winfree, E.; Pierce, N. A. *SIAM Rev.* **2007**, *49*, 65.
(7) (a) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1960. (b) SantaLucia, J., Jr.; Hicks, D. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415.
(8) (a) Markham, N. R.; Zuker, M. *Nucleic Acid Res.* **2005**, *33*, W577. (b) Dimitrov, R. A.; Zuker, M. *Biophys. J.* **2004**, *87*, 215.
(9) (a) Srinivas, N.; Ouldridge, T. E.; Sulc, P.; Schaeffer, J. M.; Yurke, B.; Louis, A. A.; Doye, J. P. K.; Winfree, E. *Nucleic Acids Res.* **2013**, *41*, 10641. (b) Baker, B. A.; Mahmoudabadi, G.; Milam, V. T. *Soft Matter* **2013**, *9*, 11160.
(10) (a) Rogers, B.; Crocker, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *87*, 15687. (b) Mognetti, M.; Varilly, P.; Angioletti-Uberti, S.; Martinez-Veracoechea, F. J.; Dobnikar, J.; Leunissen, M.; Frenkel, D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E378.
(11) Mergny, J.-L-; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.
(12) Frenkel, D.; Smit, B. *Understanding Molecular Simulations, from Algorithms to Applications*; Acedemic Press: San Diego, 2002.
(13) Zhou, Z. Y.; Ou-Yang, Z. C. *Biophys. J.* **2001**, *81*, 1133.
(14) Hansen, J. P.; McDonald, I. R. *Theory of Simple Liquids*; Academic Press: New York, 1986.

dx.doi.org/10.1021/ja500027v | *J. Am. Chem. Soc.* XXXX, XXX, XXX–XXX