Discrete Assembly of Synthetic Peptide−DNA Triplex Structures from Polyvalent Melamine−Thymine Bifacial Recognition

Yingying Zeng, Yaowalak Pratumyot, Xijun Piao, and Dennis Bong*

Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210, United States

Ⓢ Supporting Information

ABSTRACT: We have designed a 21-residue α-peptide that simultaneously recognizes two decadeoxyoligothymidine (dT_{10}) tracts to form triplexes with a peptide−DNA strand ratio of 1:2. The synthetic peptide side chain displays 10 melamine rings, which provide a bifacial thymine-recognition interface along the length of the 21-residue peptide. Recognition is selective for thymine over other nucleobases and drives the formation of ternary peptide·$[dT_{10}]_2$ complexes as well as heterodimeric peptide·$[dT_{10}C_{10}T_{10}]$ hairpin structures with triplex stems.
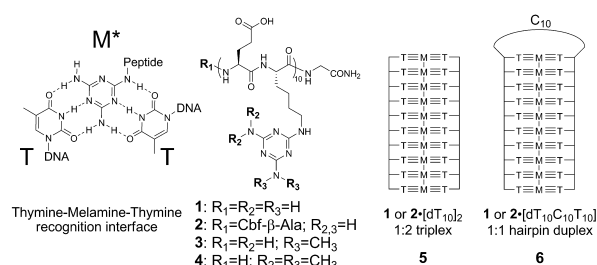
Figure 1. Bifacial melamine−thymine recognition. Synthetic peptides **1−4** present melamine and methylated melamine on derivatized lysine side chains. Peptide **2** is *N*-terminated with (5,6)-carboxyfluorescein (Cbf) and β-alanine (β-Ala). Putative peptide−DNA complex structures are shown: triplex **5** and hairpin **6**.

Nucleic acid triplex structures formed from native oligonucleotides are known to occur via purine Hoogsteen base-pairing of a third strand in the major groove of a Watson−Crick base-paired duplex.[1] This has been extensively developed by Dervan[2] and others[3] as a targeting concept using both native and artificial nucleobase recognition elements and backbones. McLaughlin[4] and Tor[5] have both reported elegant "Janus-wedge"[6] recognition of Watson−Crick interfaces using synthetic nucleobases on PNA and sugar−phosphate backbones, respectively. While these prior methods have largely sought to develop general strategies for sequence targeting of preformed oligonucleotide structures, there are fewer synthetic approaches to generate structure in single-stranded oligos or effect triplex formation without prior oligonucleotide duplex assembly; such structure-inducing recognition may be useful in design of synthetic regulators of transcription[7] or translation.[8] We have explored this notion using a Janus-wedge approach to address two *identical* interfaces: two oligothymidine (dT_{10}) DNA tracts were assembled on a peptide template via bifacial melamine recognition to form peptide−DNA triplex structures. Compounds closely related to melamine and its canonical hydrogen-bonding partner, cyanuric acid,[9] have been used to site-substitute for native nucleobases in PNA−DNA duplex recognition. Melamine itself is a well-known molecular recognition module in a number of contexts,[10] and Baranger and Zimmerman[11] have recently reported melamine targeting of thymine−thymine or uracil−uracil (T-T, U-U) mismatch sites in d(CTG) and r(CUG) repeats, assisted by acridine intercalation.[12] Indeed, the hydrogen-bonding pattern of melamine precisely complements the Watson−Crick face of thymine/uracil (Figure 1). We recently characterized the recognition of cyanuric acid and melamine derivatives at aqueous interfaces[13] as well as bulk solution[14] and found that binding was strongly dependent on the number of heterocycles

per scaffold: a trivalent system yielded robust binding while monovalent recognition was undetectable. Based on this prior work, we hypothesized that multivalent presentation of melamine heterocycles on a peptide backbone would bind two strands of dT_{10} into a peptide−DNA heterotrimeric bundle. Indeed, though dT_{10} has no detectable homooligomerization behavior, synthetic melamine-displaying peptide **1** induced assembly of a peptide−DNA triplex structures in heterotrimeric stem and heterodimeric hairpin systems with a peptide−DNA ratio of 1:2 and 1:1, respectively (Figure 1).

Prior studies from Eschenmoser and Krishnamurthy[9a,c] and Ghadiri[15] demonstrated the use of α-peptide backbones, instead of the traditional PNA backbone,[4,16] to recognize DNA. This method forms a recognition interface from alternate residues and was more synthetically convenient for our purposes. We introduced melamine through side-chain functionalization of Boc-Lysine with chlorodiaminotriazine, to yield a derivative we term melaminolysine (M*). Boc deprotection with TFA followed by reaction with Fmoc-OSu yielded Fmoc-M*, which was used in standard solid-phase peptide synthesis (SPPS) with a C-terminal glycinamide (Figure 1). Alternate residues were glutamic acid to provide water solubility and to avoid nonspecific electrostatic binding to DNA, yielding peptide **1**, $(EM^*)_{10}G$. This sequence was *N*-terminally capped with carboxyfluorescein (peptide **2**) to permit fluorescence-based binding analysis. Control peptides were also synthesized in which hydrogen-bonding sites were systematically blocked by methylation of the exocyclic amines on the melamine ring. Stepwise chloride displacement of

trichlorotriazine[17] with ammonia and/or dimethylamine and Boc-lysine yielded the dimethylated and tetramethylated melaminolysine derivatives, which were used in SPPS to provide peptides **3** and **4**, respectively.

Nanoparticle assembly from melamine and cyanuric acid derivatives results from noncovalent polymerization of the two-fold symmetric heterocycle recognition faces; thymine has only one recognition face complementary to melamine, and thus discrete assembly was anticipated. Gratifyingly, no large peptide−DNA aggregates were detectable by dynamic light scattering, consistent with the model of discrete triplex formation. Peptide-triggered base-stacking signatures were observed by UV absorbance changes, with a 1:2 stoichiometry between **1** and dT_{10}, consistent with bivalent melamine−thymine recognition and formation of triplex **5**; peptide alone did not induce such a signal (Figures 1 and 2, Supporting
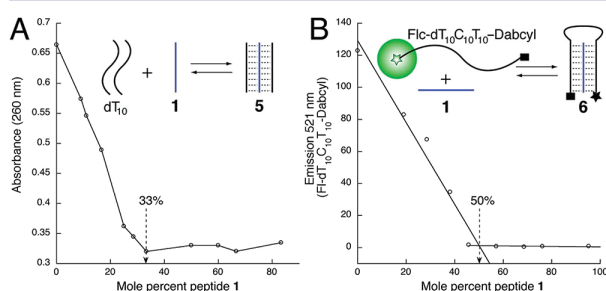

Figure 2. Peptide **1** titrated into (A) dT_{10} and (B) Flc-$dT_{10}C_{10}T_{10}$-Dabcyl, followed by UV absorbance (260 nm) and fluorescein emission (521 nm), respectively. Saturation is observed at 33 and 50 mol% peptide, indicating 1:2 and 1:1 peptide:DNA binding stoichiometries in (A) and (B), respectively.

Information (SI)). Unlike DNA and peptide−DNA[4] triplex structures which optimally form at high salt with divalent metal ions (1−2 M NaCl, 50 mM $MgCl_2$), robust assembly was observed under standard salt conditions (Dulbecco's phosphate buffered saline, DPBS), akin to the conditions used by Eschenmoser and Krishnamurthy[9c] for PNA−DNA duplex formation. Similarly, UV absorbance signatures indicated a 1:1 binding stoichiometry between **1** and $dT_{10}C_{10}T_{10}$, as would be expected if an intramolecular peptide−DNA triplex structure formed from the dT_{10} termini of $dT_{10}C_{10}T_{10}$. Indeed, binding of **1** to Flc-$dT_{10}C_{10}T_{10}$-Dabcyl resulted in maximal fluorescein quenching at a 1:1 peptide−DNA ratio, supportive of the formation of heterodimeric hairpin structure **6**, which would bring the 3′ and 5′ ends of the oligo in close proximity,[9c] resulting in efficient dabcyl quenching of fluorescein (Figure 2). Triplex and hairpin formation was further corroborated by circular dichroism, which indicated significant structuring of DNA upon addition of peptide, signified by the development of a negative CD signal at 260 nm at the expense of a positive CD signal at 280 nm, which we assign to the peptide complex and free DNA, respectively (Figure 3). Notably, while the peak at 280 nm is completely ablated in the triplex, there is a residual peak in the hairpin; this is consistent with the presence of an unstructured C_{10} loop found in hairpin **6** but not triplex **5**. Clean transformation of DNA to peptide−DNA complex bands on native polyacrylamide electrophoresis indicated discrete peptide−DNA recognition in both triplex and hairpin contexts (Figure 3).
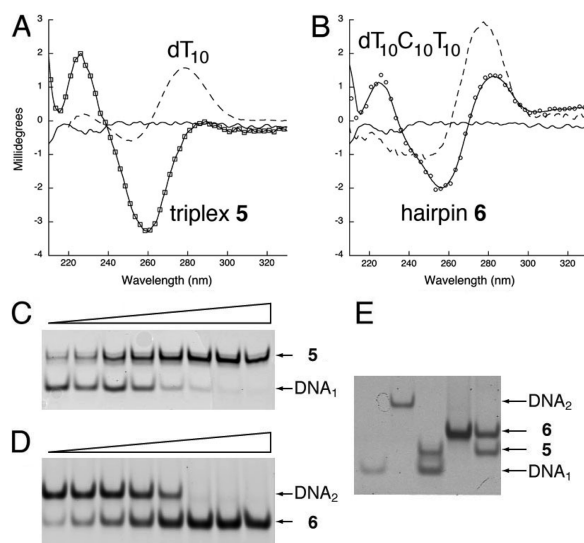

Figure 3. Circular dichroism spectra in DPBS, pH 7.4, of (A) **1** complexed with dT_{10} (□) vs dT_{10} alone (--) and (B) **1** complexed with $dT_{10}C_{10}T_{10}$ (○) vs $dT_{10}C_{10}T_{10}$ alone (--). Peptide **1** in both is at 5 μM concentration (—), while dT_{10} and $dT_{10}C_{10}T_{10}$ are maintained at 10 and 5 μM, respectively. Electrophoretic mobility shift assays imaged by Cy5 fluorescence for (C) Cy5-dT_{10} (DNA_1) and (D) Cy5-$dT_{10}C_{10}T_{10}$ (DNA_2) at 20 nM in each lane, with increasing peptide **1** concentration from left to right. (E) Relative electrophoretic mobilities of the free DNA oligos and their peptide complexes, a mixture of complex **5** and DNA_1 shown in the central lane. See SI for further details.

Cooperative melting was observed for triplex **5** and hairpin **6** by both UV and fluorescence dequenching and differential scanning calorimetry (DSC) (Figure 4, SI). The hairpin
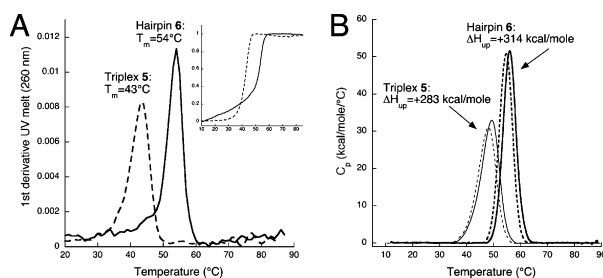

Figure 4. (A) First-derivative plot of melting transitions of triplex **5** (--) and hairpin **6** (—) followed by UV absorbance (260 nm). Normalized absorbance change is shown inset. (B) DSC upscan traces of triplex **5** (—) and hairpin **6** (—), with downscan traces shown as dashed regular and bold lines. Peptide−DNA ratios used in triplex **5** and hairpin **6** experiments were 1:2 and 1:1, respectively. DSC and UV experiments were performed in DPBS, pH 7.4 at peptide concentrations of 25 μM (DSC) and in UV experiments, 1 μM (**5**) and 2.5 μM (**6**).

heterodimer structure **6** was more thermally stable ($T_m = 54$ °C) than the heterotrimeric triplex **5** ($T_m = 43$ °C). For comparison, a dA_{10}-T_{10} duplex has $T_m = 35$ °C, and a dA_{10}-$(T_{10})_2$ triplex has $T_m = 17$ °C in the presence of 50 mM $MgCl_2$.[18] It is calculated that a dA_{10}-T_{10} duplex will have $T_m = 22.5$ °C under similar salt conditions. Reversible peptide−DNA complexation was supported by observation of endothermic melting and exothermic cooling peaks by DSC; though triplex **5**

transition temperature was elevated (48 °C) relative to the other methods, the hairpin T_m was very similar (55 °C). Assembly of complexes **5** and **6** was highly exothermic and reproducible over several thermal cycles ($\Delta H_{triplex}$ = −285 kcal/mol peptide, $\Delta H_{hairpin}$ = −314 kcal/mol), consistent with base-stacking-driven assembly.[19] The large enthalpy values (−28.5 to −31.4 kcal/mol per peptide:DNA triplet layer) are somewhat higher than, but similar to, enthalpies observed for $(TAT)_n$ intramolecular DNA triplexes,[20] which exhibit −21 kcal/mol per triplet stack. Peptide-displayed melamine−DNA binding appears to have exothermic assembly profiles and T_m values similar to those of DNA−DNA and melamine−cyanuric acid recognition, suggestive of similar driving forces, perhaps to be expected given the similarity of melamine and native nucleobases. Given that peptides **1** and **2** and DNA are all anionic polyelectrolytes, we anticipate negligible nonspecific electrostatic interactions. Indeed, partial and full methylation of each melamine ring on the peptide to respectively yield peptides **3** and **4** abolished all detectable binding to dT_{10} or $dT_{10}C_{10}T_{10}$, indicating the importance of the melamine recognition interface (SI). As methylation increases hydrophobicity, it is clear that assembly depends on recognition of donor−acceptor patterning and not simply nonspecific hydrophobic collapse. Notably, peptides **1** and **2** did not show any binding signatures when annealed with dA_{10}, dC_{10}, and dG_5A_{10}, indicating a selectivity for thymine over the other native nucleobases (SI). Thus, melamine peptide recognition is relatively specific for the hydrogen-bonding pattern of oligothymidine, as predicted. Taken together with the observed stoichiometry of binding with dT_{10} and $dT_{10}C_{10}T_{10}$, this is strongly supportive of the expected bivalent thymine recognition by melamine (Figure 1).

We quantified association using fluorescein-tagged peptide **2** binding to unlabeled DNA (Figure 5). Complex formation was
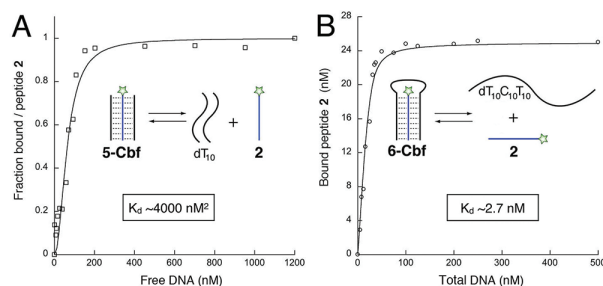

Figure 5. Binding isotherms in DPBS, pH 7.4, of (A) peptide **2** binding to dT_{10} followed by fluorescein quenching upon binding and (B) **2** binding to $dT_{10}C_{10}T_{10}$ followed by fluorescence anisotropy. Solid lines show fits to (A) 1:2 binding (fraction bound peptide = $[DNA]^2/K_d + [DNA]^2$, $R^2 \geq 0.96$) and (B) 1:1 binding model, corrected for fluorescence quenching, $R^2 \geq 0.98$. Peptide concentration is constant at 25 nM.

followed by both fluorescence anisotropy and quenching, which accompanies assembly; similar results were obtained with the two signatures. Using the stoichiometries established by titration (Figure 2), we fit the binding curves to a trimer−monomer model (SI) to obtain an apparent $K_d \approx 4000$ nM^2 (by quenching or anisotropy) for dissociation of triplex **5**. A second transition was undetectable by fitting to a two-step model or by observation of assembly processes, suggesting that the peptide−dT_{10} heterodimer that forms initially must react

rapidly with another strand of dT_{10} to form the heterotrimer. Indeed, the coefficient derived from a Hill-type binding isotherm was essentially 2, supportive of a cooperative 1:2 peptide−DNA assembly process. A hypothetical equivalent two-step dissociation process with an overall apparent K_d = 4000 nM^2 would have K_d = 63 nM for both trimer−dimer and dimer−monomer dissociation steps, which is the observed free DNA concentration at half-saturation (Figure 5). A good fit to a 1:1 binding model was obtained from the binding isotherm of peptide **2** to $dT_{10}C_{10}T_{10}$, revealing an approximate K_d = 2.7 nM for dissociation of hairpin **6**.

Overall, these data support the model of bivalent melamine−thymine recognition yielding formation of novel discrete peptide−DNA triplex structures with robust binding affinities. Notably, prior work from Eschenmoser and Krishnamurthy using similar diaminotriazine nucleobase mimics derived from aspartic and glutamic acids resulted in *duplex* formation with oligothymidines rather than triplex.[9a] Diaminotriazine has two exocyclic hydrogen bond donor sites compared to three on melamine, which yields two potential thymine recognition interfaces in melaminolysine and just one in diaminotriazine derivatives. Thus, it appears that the assembly state pivots from dimer to trimer based on the addition of a single hydrogen bond donor, with the increased entropic cost likely paid for by increased base stacking. Polyvalent melamine−thymine recognition between peptides **1** and **2** with dT_{10} tracts is unique in that structure is induced in unstructured single-stranded oligothymidines to yield novel triplex and hairpin peptide−DNA hybrid structures, thus expanding the range of non-native nucleobase-derived structures already known.[21] Though melamine recognition of preformed T-T/U-U mismatch sites has been previously reported,[11] the ability of polyvalent melamine peptides to broker assembly with two non-interacting oligothymidine strands to form peptide−DNA triplex structures is non-obvious. We anticipate that this design element may be used to manipulate the structure and function of thymine- or uracil-rich targets.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional CD, fluorescence, UV, and fluorescence anisotropy data, fitting procedures, experimental details, and detailed synthetic procedures and compound characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
bong@chem.osu.edu

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) (a) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123. (b) Francois, J. C.; Saison-Behmoaras, T.; Hélène, C. *Nucleic Acids Res.* **1988**, *16*, 11431. (c) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645.

(2) Dervan, P. B.; Doss, R. M.; Marques, M. A. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 373.

(3) Simon, P.; Cannata, F.; Concordet, J.-P.; Giovannangeli, C. *Biochimie* **2008**, *90*, 1109.

(4) Chen, H.; Meena.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2008**, *130*, 13190.

(5) Shin, D.; Tor, Y. *J. Am. Chem. Soc.* **2011**, *133*, 6926.

(6) Branda, N.; Kurz, G.; Lehn, J.-M. *Chem. Commun.* **1996**, *1996*, 2443.

(7) (a) Mapp, A. K.; Ansari, A. Z. *ACS Chem. Biol.* **2007**, *2*, 62. (b) Payankaulam, S.; Li, L. M.; Arnosti, D. N. *Curr. Biol.* **2010**, *20*, R764.

(8) (a) DeiganK. E.Ferré-D'AmaréA. R.Acc. Chem. Res.2011(b) Ogawa, A. *RNA* **2011**, *17*, 1. (c) Zhang, J.; Lau, M. W.; Ferré-D'Amaré, A. R. *Biochemistry* **2010**, *49*, 9123. (d) Suess, B.; Weigand, J. E. *RNA Biol.* **2008**, *5*, 24.

(9) (a) Mittapalli, G. K.; Reddy, K. R.; Xiong, H.; Munoz, O.; Han, B.; De Riccardis, F.; Krishnamurthy, R.; Eschenmoser, A. *Angew. Chem., Int. Ed.* **2007**, *46*, 2470. (b) Vysabhattar, R.; Ganesh, K. N. *Tetrahedron Lett.* **2008**, *49*, 1314. (c) Mittapalli, G. K.; Osornio, Y. M.; Guerrero, M. A.; Reddy, K. R.; Krishnamurthy, R.; Eschenmoser, A. *Angew. Chem., Int. Ed.* **2007**, *46*, 2478.

(10) (a) Ai, K.; Liu, Y.; Lu, L. *J. Am. Chem. Soc.* **2009**, *131*, 9496. (b) Kawasaki, T.; Tokuhiro, M.; Kimizuka, N.; Kunitake, T. *J. Am. Chem. Soc.* **2001**, *123*, 6792. (c) Prins, L. J.; De Jong, F.; Timmerman, P.; Reinhoudt, D. N. *Nature* **2000**, *408*, 181. (d) Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem., Int. Ed.* **2001**, *40*, 2382. (e) Kimizuka, N.; Kawasaki, T.; Hirata, K.; Kunitake, T. *J. Am. Chem. Soc.* **1998**, *120*, 4094. (f) Mathias, J. P.; Simanek, E. E.; Seto, C. T.; Whitesides, G. M. *Macromol. Symp.* **1994**, *77*, 157. (g) Zerkowski, J. A.; MacDonald, J. C.; Seto, C. T.; Wierda, D. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 2382. (h) Seto, C. T.; Whitesides, G. M. *J. Am. Chem. Soc.* **1993**, *115*, 905. (i) Zerkowski, J. A.; Seto, C. T.; Whitesides, G. M. *J. Am. Chem. Soc.* **1992**, *114*, 5473. (j) Seto, C. T.; Whitesides, G. M. *J. Am. Chem. Soc.* **1990**, *112*, 6409.

(11) Arambula, J. F.; Ramisetty, S. R.; Baranger, A. M.; Zimmerman, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16068.

(12) (a) Rapireddy, S.; He, G.; Roy, S.; Armitage, B. A.; Ly, D. H. *J. Am. Chem. Soc.* **2007**, *129*, 15596. (b) Fechter, E. J.; Olenyuk, B.; Dervan, P. B. *Angew. Chem., Int. Ed.* **2004**, *43*, 3591. (c) Fechter, E. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8476. (d) Durand, M.; Maurizot, J. C.; Asseline, U.; Thuong, N. T.; Hélène, C. *Bioconjugate Chem.* **1993**, *4*, 206. (e) Bentin, T.; Nielsen, P. E. *J. Am. Chem. Soc.* **2003**, *125*, 6378.

(13) (a) Ma, M.; Bong, D. *Langmuir* **2011**, *27*, 1480. (b) Ma, M.; Gong, Y.; Bong, D. *J. Am. Chem. Soc.* **2009**, *131*, 16919. (c) Ma, M.; Paredes, A.; Bong, D. *J. Am. Chem. Soc.* **2008**, *130*, 14456. (d) Ma, M.; Bong, D. *Org. Biomol. Chem.* **2011**, *9*, 7296.

(14) Ma, M.; Bong, D. *Langmuir* **2011**, *27*, 8841.

(15) Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. *Science* **2009**, *325*, 73.

(16) (a) Nielsen, P. E. *Chem. Biodiversity* **2007**, *4*, 1996. (b) Porcheddu, A.; Giacomelli, G. *Curr. Med. Chem.* **2005**, *12*, 2561.

(17) Zhang, W.; Nowlan, D. T.; Thomson, L. M.; Lackowski, W. M.; Simanek, E. E. *J. Am. Chem. Soc.* **2001**, *123*, 8914.

(18) (a) Pilch, D. S.; Brousseau, R.; Shafer, R. H. *Nucleic Acids Res.* **1990**, *18*, 5743. (b) Pilch, D. S.; Levenson, C.; Shafer, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1942.

(19) (a) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *30*, 1. (b) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213. (c) McKay, S. L.; Haptonstall, B.; Gellman, S. H. *J. Am. Chem. Soc.* **2001**, *123*, 1244. (d) Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **2006**, *34*, 564. (e) SantaLucia, J. Jr.; Hicks, D. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415.

(20) Soto, A. M.; Loo, J.; Marky, L. A. *J. Am. Chem. Soc.* **2002**, *124*, 14355.

(21) (a) Krueger, A. T.; Kool, E. T. *Chem. Biol.* **2009**, *16*, 242. (b) Krueger, A. T.; Kool, E. T. *Curr. Opin. Chem. Biol.* **2007**, *11*, 588.