JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Autonomously Pairing Cysteinyl-Linked Nucleotide Analogues with a Unique Architecture

Manuel Peifer and Andrea Vasella*

Laboratory for Organic Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland

Supporting Information

ABSTRACT: We report the efficient pairing in water of the first representative of oligonucleotide analogues in which the backbone is replaced by linking elements between the nucleobases. The architecture of the new analogue demonstrates that the structural differentiation of oligonucleotides into a contiguous backbone and nucleobases, as embodied by the natural nucleic acids and all nucleotide analogues analyzed to date, is not a prerequisite for pairing. UV and circular dichroism analyses of self-complementary and non-self-complementary octanucleotide analogues strongly suggest the fully reversible, sequence-specific association of our new analogues to form a left-handed double helix with an antiparallel strand orientation that is characterized by melting temperatures and free enthalpies higher than those of natural RNA and DNA of the same sequence. The linking element incorporates an L-cysteine moiety that allows a short and efficient synthesis of the monomeric building blocks and, through the choice of either L- or D-cysteine, gives access to either one of the enantiomeric oligomers and thus to left- or right-handed helices.

Extensive work by a number of research groups, including those of Eschenmoser,[1] Nielsen,[2] and Kool,[3] has addressed the synthesis and analysis of oligonucleotide analogues with altered backbones and nucleobases. In accord with the structure of naturally occurring nucleic acids and oligonucleotides, all of these analogues are characterized by the structural differentiation between a contiguous backbone and the nucleobases attached to it. Over the last several years, our group has described a series of oligonucleotide analogues that diverge from this common structural motif. Instead of a contiguous backbone, they possess linking elements directly between adjacent nucleobases. Such oxymethylene-,[4] ethynylene-,[5] (*Z*)-ethenylene-,[6] thiomethylene-,[7,8] sulfonyl- and sulfinylmethylene-,[9] and aminomethylene-linked[10] di- and tetranucleosides, called "oligonucleotide analogues with integrated bases and backbone" (ONIBs), pair efficiently in organic solvents, thus deepening our insight into the limits within which the structure of pairing nucleic acids can be varied.

Having shown that ONIBs are capable of pairing in organic solvents, we designed a new type of these analogues that allows their pairing properties in aqueous solution to be evaluated. As pairing in water requires longer oligomers, we chose an amide group as a key element of the new linker to allow assembly of the desired oligomers by one of the numerous established methods of amide coupling and deprotection in solution or on a solid support. We aimed at designing the new linker to be short enough to allow favorable twist angles but long enough to maintain sufficient flexibility, allowing the oligonucleotide analogues to adapt to the stereochemical requirements of pairing conformers. We realized this intent by incorporating an L-cysteinyl moiety (Figure 1). According to a conformational analysis based on vibrational spectra,[11] NMR data,[8] and crystal structures,[12] we conclude that the sulfur atom favors a first synclinal arrangement of $CH_2(1')$ and $CH(4')$ and that it further contributes to the flexibility of our new analogues on account of the low energy barriers for rotation about C−S bonds.[11] The polar side chain attached to C(4′) allows for a second crucial synclinal arrangement of S(2′) and NH(5′) and ensures the water solubility of the oligomers at neutral or basic pH. The preferred conformation of the aminoacetyl fragment attached to N(9) (as shown in Figure 1) is well-precedented by crystal and NMR structures of peptide nucleic acids (PNAs).[13] These conformational aspects suggest that the cysteinyl linker adopts a conformation similar to the one shown in Figure 1, where the nucleobases are positioned at a favorable distance for stacking and in the right orientation for hydrogen bonding in the Watson−Crick mode. In agreement with the common structural

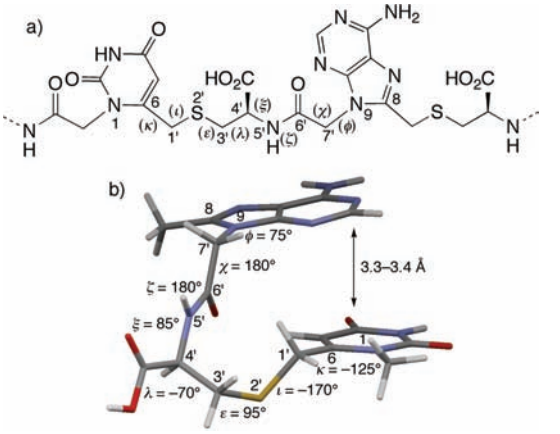


Figure 1. a) Definitions of the torsion angles and numbering of a cysteinyl-linked pyrimidine−purine (UA) dinucleotide analogue. (b) Conformation of the UA dinucleotide analogue with torsion angles $\kappa-\phi$ required for pairing [termini on N(1) and C(8) have been replaced by methyl groups for clarity].

Received: February 3, 2011
Published: March 08, 2011

dx.doi.org/10.1021/ja200829s | *J. Am. Chem. Soc.* 2011, 133, 4264–4267

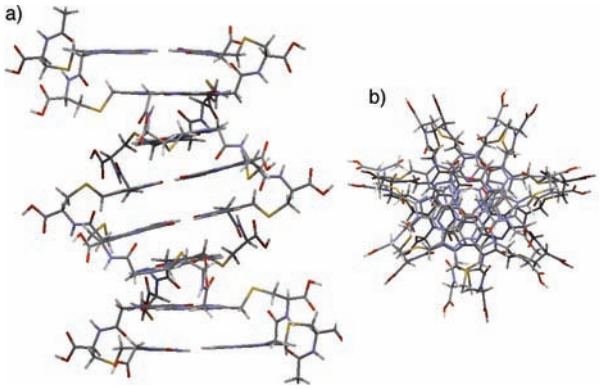


Figure 2. Model of the octanucleotide analogue $U_2C_2G_2A_2$ (**1**): (a) side view; (b) top view.

motif of ONIBs, the linker is attached to N(1) and C(6) of a pyrimidine base and to N(9) and C(8) of a purine base, thus fusing the nucleobases and the linker into a single linear structural unit.

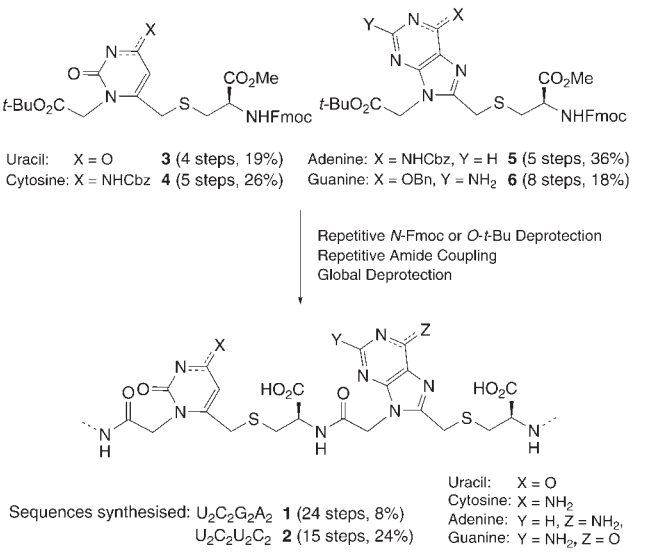
Amber* force field calculations[14] of cysteinyl-linked dimers suggest that the surface overlap of the nucleobases in the cysteinyl-linked analogues depends even more strongly on the sequence than in nucleic acids[14−16] and PNAs.[13] The surface overlap decreases in the order pyr·pur ≈ pyr·pyr ≈ pur·pur > pur·pyr (from the C- to the N-terminus, with pyr representing a pyrimidine base and pur representing a purine base). We thus synthesized the self-complementary octamer **1** having the sequence $U_2C_2G_2A_2$ and the non-self-complementary octamer **2** having the sequence $U_2C_2U_2C_2$ (naming from the C- to the N-terminus), which are devoid of a pur/pyr switch. Minimization of the conformational energy of **1** after releasing all of the constraints between the hydrogen bond donors and acceptors (Amber* force field calculation[14]) resulted in a left-handed double helix (Figure 2) with the left-handedness dictated by the L-cysteine moiety.

The helix is stabilized by Watson−Crick-type base pairing and base stacking, with the cysteinyl linker adopting a conformation having the values of the torsion angles $\kappa-\phi$ that we deduced for the pairing conformation shown in Figure 1. The helix is characterized by a rise of 3.3−3.4 Å per base pair (cf. B-DNA, 3.4 Å; A-RNA, 2.8 Å), a pitch of ∼26 Å (cf. 34 Å for B- and A-RNA), and a twist angle of ∼42° (cf. B-DNA, 36°; A-RNA, 31°).[14,15]

We synthesized the fully protected monomeric building blocks **3**−**5** in short sequences of four to five steps and overall yields of 19−36% from uracil, cytosine, and adenine, respectively. The fully protected monomeric building block **6** was obtained in eight steps and 18% overall yield from commercially available 2,6-diamino-4-chloropyrimidine (Scheme 1). Self-complementary octameric **1** and non-self-complementary octameric **2** were synthesized by convergent coupling of the *N*-Fmoc- or *N*-*tert*-butoxycarbonyl-deprotected building blocks **3**−**6** via di- and tetrameric intermediates and a final deprotection in 24 steps and 6% overall yield for **1** and 15 steps and 24% overall yield for **2** (see the Supporting Information).

The UV absorbance of the self-complementary octanucleotide analogue **1** (c = 10 μM in 10 mM sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA, pH 7.0) changed nonlinearly with linearly increasing temperature, evidencing cooperative stacking and thus pairing (Figure 3, top plot). The UV absorbance of the non-self-complementary octanucleotide

Scheme 1. Structures of the Monomeric Building Blocks 3−6 and Summary of the Synthesis of the Octanucleotide Analogues $U_2C_2G_2A_2$ (1) and $U_2C_2U_2C_2$ (2) of the Type Illustrated by the Depicted General Structure of a Dimeric Unit



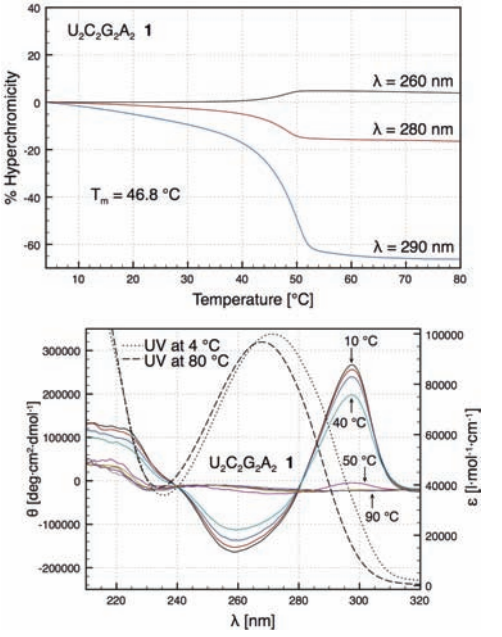


Figure 3. (top) UV melting curves and (bottom) temperature-dependent CD spectra (solid lines, 10 °C steps from 10 to 90 °C) and UV spectra (dotted line at 4 °C, dashed line at 80 °C) of self-complementary $U_2C_2G_2A_2$ (**1**) in 10 mM sodium phosphate buffer, 100 mM NaCl, and 0.1 mM EDTA at pH 7.0 (c = 10 μM, 1 cm cell).

analogue **2**, however, changed only linearly with linearly increasing temperature, evidencing the absence of cooperative stacking and thus suggesting the absence of pairing.[17] The self-recognition of **1** unambiguously proves sequence-specific hydrogen bonding between complementary nucleobases and thus an antiparallel orientation of the associating strands.

dx.doi.org/10.1021/ja200829s | J. Am. Chem. Soc. 2011, 133, 4264–4267

Similar to the UV melting curves of the octanucleotides $U_2(T_2)C_2G_2A_2$ of RNA and DNA (for the spectra of **2**, RNA, and DNA, see the Supporting Information), the UV melting curves resulting from heating (denaturation) and cooling (renaturation) experiments with self-complementary **1** were identical, evidencing fully reversible pairing and thermodynamic control of the association. The remarkable stability of the associate of **1** is reflected by its melting temperature of 46.8 °C (c = 10 μM, 10 mM sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA, pH 7.0), which is higher than those of RNA and DNA with the same sequence, as obtained under the same conditions (cf. RNA, 44.9 °C; DNA, 31.4 °C).

The transition of **1** in the UV melting profile is considerably sharper than those for RNA and DNA and is also reflected by the abrupt decrease of the strong circular dichroism (CD) signal of **1** precisely at the melting temperature of 47 °C (compare the CD spectra at 40 and 50 °C in the bottom panel of Figure 3), evidencing a high cooperativity during the pairing of **1**. The CD spectra of the non-self-complementary cysteinyl-linked analogue **2** at various temperatures, in contradistinction, show very small ellipticities and no temperature dependence (see the Supporting Information).

The UV spectrum of **1** at 4 °C shows a bathochromic shift relative to its UV spectrum at 80 °C (λ_{max} from 272 to 268 nm; see Figure 3, bottom panel, dashed and dotted lines). We attribute this observation to a pronounced stacking of the inner C_2G_2 fragment, as compared with the somewhat weaker stacking of the C-terminal U_2 and N-terminal A_2 fragments (λ_{max} for a C·G base pair, ~280 nm; λ_{max} for a T·A base pair, ~260 nm).[18] As a consequence, **1** shows only a small hyperchromic shift in the UV melting profile at a wavelength of 260 nm (5%), whereas one observes remarkable hypochromic shifts at λ = 280 nm (17%) and 290 nm (66%) (Figure 3, top panel). The melting temperatures calculated from the curves at 260 and 290 nm are very close (46.8 and 47.5 °C).

The stability of the associate of the cysteinyl-linked octamer **1** is expressed not only by the higher melting temperature but also by the high value of the negative free energy ($-\Delta G_{298}$) (cf. **1**, 13.4 kcal/mol; RNA, 9.8 kcal/mol; DNA, 8.0 kcal/mol) and of the negative energy of association ($-\Delta H_{298}$) (cf. **1**, 89.1 kcal/mol; RNA, 48.1 kcal/mol; DNA, 55.4 kcal/mol). The value of $-T\Delta S$ for **1** was calculated to be 75.7 kcal/mol (cf. RNA, 38.3 kcal/mol; DNA, 47.4 kcal/mol).[19] In fair agreement with the values obtained by Kool,[20] $-\Delta G_{298}$ per base pair for **1** amounts to 1.7 kcal/mol (cf. RNA, 1.2 kcal/mol; DNA, 1.0 kcal/mol; eight Watson−Crick base pairs, seven stacking interactions). The large value of $-\Delta H$ for **1** suggests a combination of stabilizing H-bonding and stacking,[20] possibly combined with a weaker strand repulsion of **1** relative to RNA and DNA. The large value for $-T\Delta S$ may be due to the conformational equilibria of the dissociated strands ($\lambda = -60°$ and $\lambda = 180°$ and/or $\xi = +90°$ and $\xi = -90°$) (Figure 1), which are "frozen" in the associated state ($\lambda = -70°$ and $\xi = +85°$). The hydrophobic effect could also have an influence on $-T\Delta S$, in that **1** could release less hydrated water upon association than RNA and DNA.[16]

The self-complementarity of **1**, its complex NMR spectra in water, and the fact that it could not be crystallized prevented us from unambiguously proving the stoichiometry of the associate of **1**. However, on the basis of a series of UV experiments, we are confident that our assumption of a bimolecular associate is correct. Specifically, we can exclude an intramolecular associate (monomolecular equilibrium), since the melting temperature of **1** depends on the concentration (cf. c = 5 μM, T_m = 44.7 °C; c = 10 μM, T_m = 46.8 °C; c = 100 μM, T_m = 55.4 °C; 10 mM sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA, pH 7.0). Triplex formation (trimolecular equilibrium) requires a homopyrimidine strand and a homopurine strand consisting of pyr·pur·pyr and/or pyr·pur·pur triads. $CG\cdot C^+$ and $TA\cdot T^+$ triads require protonation, while the stability of pyr·pur·pur triads depends on the presence of a bivalent metal cation.[21] However, **1** does not possess a pure homopyrimidine or pure homopurine sequence; moreover, its association does not require bivalent cations, and its T_m does not decrease with increasing pH (cf. pH 7.0, T_m = 46.8 °C; pH 8.0, T_m = 47.5 °C; pH 8.8, T_m = 47.0 °C; c = 10 μM in 10 mM sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA). Also, considering the single transition of **1** in its UV melting profile, the formation of a triplex by **1** is highly improbable. Association of **1** to form a G-quadruplex or an i-motif (tetramolecular equilibria) requires stabilization of the associate by mono- or bivalent cations [preferentially K^+, Rb^+, or Na^+ (150 mM−1 M) or Mg^{2+}, Ca^{2+}, Ba^{2+}, or Sr^{2+} (>10 mM)] or by protonation.[22] The octamer **1**, however, associates even more strongly in the presence of the only very weak G-quadruplex initializer Li^+ (cf. 10 mM sodium phosphate buffer + 100 mM NaCl, T_m = 46.8 °C; 10 mM lithium phosphate buffer + 100 mM LiCl, T_m = 50.9 °C at 10 μM concentration and pH 7.0) as well as at basic pH (see above). Also, the hysteresis characteristic of tetramolecular G-quadruplexes[22] was not observed. The progressive increase in the melting temperature of **1** with increasing salt concentration is in agreement with the findings for DNA and RNA duplexes (cf. 10 mM sodium phosphate buffer, no NaCl, T_m = 24.9°; 10 mM sodium phosphate buffer, 1 M NaCl, T_m = 63.7° at 10 μM concentration and pH 7.0).[23]

The analytically evidenced stacking, the pairing via hydrogen bonding (as evidenced by the self-recognition of **1** and the large values of $-\Delta H_{298}$ and $-\Delta G_{298}$), and the considerations regarding the stoichiometry provide evidence for the association of **1** in a double-helical structure similar to that proposed in Figure 2.

In conclusion, efficient pairing of the first oligonucleotide analogue with integrated bases and backbone (ONIB) in water has been demonstrated. Analytical data and molecular modeling suggest that L-cysteinyl-linked ONIBs associate to form a left-handed double helix. The chirality of the cysteine moiety allows the synthesis of either one of the enantiomeric cysteinyl-linked helical oligomers.

■ ASSOCIATED CONTENT

Supporting Information. Experimental procedures, spectral data, and additional UV and CD spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
vasella@org.chem.ethz.ch

■ ACKNOWLEDGMENT

We thank ETH Zürich and Syngenta (Basel, Switzerland) for generous support and Dr. Bruno Bernet for helpful discussions.

REFERENCES

(1) (a) Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R.; Eschenmoser, A. *Science* **1999**, *283*, 699–703. (b) Schoning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347–1351. (c) Eschenmoser, A. *Tetrahedron* **2007**, *63*, 12821–12843.

(2) (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500. (b) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895–1897. (c) Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677–9678.

(3) (a) Liu, H. B.; Gao, J. M.; Lynch, S. R.; Saito, Y. D.; Maynard, L.; Kool, E. T. *Science* **2003**, *302*, 868–871. (b) Lynch, S. R.; Liu, H. B.; Gao, J. M.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 14704–14711. (c) Wilson, J. N.; Teo, Y. N.; Kool, E. T. *J. Am. Chem. Soc.* **2007**, *129*, 15426.

(4) Matthews, A. J.; Bhardwaj, P. K.; Vasella, A. *Helv. Chim. Acta* **2004**, *87*, 2273–2295.

(5) (a) Zhang, X.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2006**, *89*, 2861–2917. (b) Zhang, X.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2007**, *90*, 792–819.

(6) Zhang, X.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2007**, *90*, 864–890.

(7) (a) Ritter, A.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2008**, *91*, 1675–1694. (b) Bernet, B.; Johar, Z.; Ritter, A.; Jaun, B.; Vasella, A. *Helv. Chim. Acta* **2009**, *92*, 2596–2630.

(8) Ritter, A.; Egli, D.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2008**, *91*, 673–714.

(9) Bogliotti, N.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2010**, *93*, 659–667.

(10) (a) Chiesa, K.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2010**, *93*, 1822–1843. (b) Chiesa, K.; Shvoryna, A.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2010**, *93*, 668–691.

(11) (a) Plusquellic, D. F.; Suenram, R. D.; Mate, B.; Jensen, J. O.; Samuels, A. C. *J. Chem. Phys.* **2001**, *115*, 3057–3067. (b) Vansteenkiste, P.; Pauwels, E.; Van Speybröck, V.; Waroquier, M. *J. Phys. Chem. A* **2005**, *109*, 9617–9626.

(12) (a) Kashino, S.; Ashida, T.; Kakudo, M. *Acta Crystallogr., Sect. B* **1974**, *30*, 2074–2076. (b) Voliotis, S.; Cordopatis, P.; Main, P.; Nastoroulos, V.; Germain, G. *Acta Crystallogr., Sect. C* **1986**, *42*, 1777–1780.

(13) (a) Rasmussen, H.; Sandholm, J. *Nat. Struct. Biol.* **1997**, *4*, 98–101. (b) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838–1841. (c) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410–413. (d) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777–780.

(14) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Caufield, C.; Lipton, M.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(15) Evaluation of the structural properties of DNA and RNA was performed according to the default settings for nucleic acids in *Macromodel 7.0*.

(16) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

(17) (a) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601–1620. (b) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515–537. (c) Owczarzy, R. *Biophys. Chem.* **2005**, *117*, 207–215.

(18) Puglisi, J. D.; Tinoco, I. *Methods Enzymol.* **1989**, *180*, 304–325.

(19) (a) SantaLucia, J., Jr.; Allawi, H. T.; Seneviratne, P. A. *Biochemistry* **1996**, *35*, 3555–3562. (b) Xia, T. B.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schröder, S. J.; Jiao, X. Q.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719–14735. (c) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460–1465.

(20) (a) Kool, E. T. *Annu. Rev. Biophys.* **2001**, *30*, 1–22. (b) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473–1487. (c) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990–1009. (d) Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **2006**, *34*, 564–574.

(21) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65–95.

(22) (a) Sen, D.; Gilbert, W. *Nature* **1990**, *344*, 410–414. (b) Venczel, E. A.; Sen, D. *Biochemistry* **1993**, *32*, 6220–6228. (c) Mergny, J. L.; Phan, A. T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74–78. (d) Hardin, C. C.; Perry, A. G.; White, K. *Biopolymers* **2000**, *56*, 147–194.

(23) Schildkraut, C.; Lifson, S. *Biopolymers* **1965**, *3*, 195–208.