J|A|C|S
ARTICLES
Published on Web 08/31/2004

Expanded-Size Bases in Naturally Sized DNA: Evaluation of Steric Effects in Watson−Crick Pairing

Jianmin Gao, Haibo Liu, and Eric T. Kool*

Contribution from the Department of Chemistry, Stanford University, Stanford California 94305-5080

Received March 15, 2004; E-mail: kool@stanford.edu

Abstract: We describe physicochemical properties in DNA of altered-size nucleobases that retain Watson−Crick analogous hydrogen-bonding ability. Size-expanded analogues of adenine and thymine (xA and xT, respectively, which are expanded by benzo-fusion) were incorporated into natural DNA oligonucleotides, and their effects on helix stability were measured. Base stacking studies revealed that the two stretched analogues stack much more strongly than do their naturally sized counterparts. In contrast to this, pairing studies showed that single substitutions of the new bases are destabilizing to the natural helix as compared to A or T in standard A−T pairs in the same context, unless multiple adjacent substitutions are used. Interestingly, the size-expanded bases displayed selective recognition of the hydrogen-bonding complementary partners, suggesting that Watson−Crick analogous pairs were still formed despite local backbone strain. In an attempt to compensate for the added size of the expanded adenine, we tested a formamide deoxynucleoside, which Leonard proposed as a shortened thymine analogue (F_o). Data showed, however, that this compound adopts a conformation unfavorable for pairing. On the basis of the combined thermodynamic data, we estimate the energetic cost of the 2.4 Å stretching of an isolated base pair in DNA at ca. +1 to 2 kcal/mol. Notably, during the pairing studies, the two size-expanded nucleobases were found to display significant changes in fluorescence on formation of stacked versus unstacked structures, suggesting possible applications in probing nucleic acid structures and biochemical mechanisms.

Introduction

Nearly three decades ago, Leonard began a series of studies probing the steric requirements of the active sites of ATP-dependent enzymes.[1] As a molecular strategy, he developed a linear benzo-fused analogue of the adenine base, increasing its dimensions by ca. 2.4 Å. Studies of the ribonucleoside triphosphate derivative of this compound with a number of ATP-processing enzymes were reported. In 1984, the deoxynucleoside derivative (here denoted dxATP) was prepared by deoxygenation of the ribonucleoside;[2] however, it was not incorporated into oligonucleotides.

Size-augmented nucleobases such as this would seem to be ideal candidates for evaluating steric effects in formation of duplex DNA structures and in enzymes that process DNA. Yet they have not been generally applied to such studies (save for one reported attempt at incorporation of dxATP into polymeric DNA).[2] Indeed, Leonard recognized that the xA base was too large for the natural duplex DNA or RNA backbone, and he proposed a "foreshortened" formamide analogue of uracil (denoted F_o here) that might compensate for this added size.[3] This compound also was never apparently tested for pairing, although a formamide deoxyriboside (dF_o) has been studied more recently in the context of its role as a possible product of DNA damage.[4]

In an attempt to address some of these issues, we embarked on a study of the properties of the dxA nucleoside and of a number of new size-altered analogues of DNA bases.[5] We developed an improved synthesis of dxA and designed a new analogue of thymidine (dxT) in which the nucleobase dimensions are stretched in a manner analogous to those of dxA. Interestingly, early studies have shown that dxA and dxT can be combined with the hydrogen-bonding complementary partners T and A (respectively) to form entire helices in which all base pairs are stretched.[5a] Such work is aimed at development of novel tools for biophysical studies, at possible biotechnological applications, and at development of novel genetic systems beyond the natural one. In addition, we expect that size-augmented nucleobases may also yield useful steric information about the active sites of enzymes involved in DNA replication and repair.[6,7]

(1) (a) Leonard, N. J.; Sprecker, M. A.; Morrice, A. G. *J. Am. Chem. Soc.* **1976**, *98*, 3987−3994. (b) Leonard, N. J.; Scopes, D. I. C.; Vanderlijn, P.; Barrio, J. R. *Biochemistry* **1978**, *17*, 3677−3685.
(2) Lessor, R. A.; Gibson, K. J.; Leonard, N. J. *Biochemistry* **1984**, *23*, 3868−3873.
(3) Czarnik, A. W.; Leonard, N. J. *J. Am. Chem. Soc.* **1982**, *104*, 2624−2631.
(4) Luo, Y.; Henle, E. S.; Linn, S. *J. Biol. Chem.* **1996**, *271*, 21167−21176.
(5) (a) Liu, H.; Gao, J.; Lynch, S. R.; Maynard, L.; Saito, D.; Kool, E. T. *Science* **2003**, *302*, 867−871. (b) Liu, H.; Gao, J.; Maynard, L.; Saito, D.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 1102−1109.
(6) (a) Summerer, D.; Marx, A. *Angew. Chem., Int. Ed.* **2001**, *40*, 3693−3695. (b) Strerath, M.; Cramer, J.; Restle, T.; Marx, A. *J. Am. Chem. Soc.* **2002**, *124*, 11230−11231. (c) Cramer, J.; Strerath, M.; Marx, A.; Restle, T. *J. Biol. Chem.* **2002**, *277*, 43593−43598.
(7) (a) Delaney, J. C.; Henderson, P. T.; Helquist, S. A.; Morales, J. C.; Essigmann, J. M.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4469−4473. (b) Kool, E. T. *Annu. Rev. Biochem.* **2002**, *71*, 191−219. (c) Matray, T. J.; Kool, E. T. *Nature (London)* **1999**, *399*, 704−708.

10.1021/ja048499a CCC: $27.50

Table 1. Free Energies ($-\Delta G^\circ_{37}$ (kcal/mol)) and Melting Temperatures (T_m (°C)) for DNA Duplexes Containing X−Y Pairs (ϕ = Tetrahydrofuran Abasic Analogue)

	X–Y	T_m[a]	$-\Delta G^\circ_{37}$[b] van't Hoff	$-\Delta G^\circ_{37}$[c] curve fits	$-\Delta G^\circ_{37}$[d] average		X–Y	T_m[a]	$-\Delta G^\circ_{37}$[b] van't Hoff	$-\Delta G^\circ_{37}$[c] curve fits	$-\Delta G^\circ_{37}$[d] average
1	A−T	40.7	9.3 ± 0.1	9.3 ± 0.1	9.3 ± 0.1	12	ϕ−xA	33.2	7.7 ± 0.1	7.5 ± 0.1	7.6 ± 0.1
2	T−A	40.4	9.1 ± 0.1	9.2 ± 0.1	9.2 ± 0.1	13	xT−A	35.2	8.1 ± 0.1	7.9 ± 0.2	8.0 ± 0.2
3	xA−T	35.8	8.1 ± 0.1	8.1 ± 0.1	8.1 ± 0.1	14	xT−T	31.3	7.1 ± 0.1	6.9 ± 0.1	7.0 ± 0.1
4	xA−G	29.3	6.8 ± 0.1	6.4 ± 0.1	6.6 ± 0.1	15	xT−G	28.2	6.0 ± 0.3	6.3 ± 0.2	6.2 ± 0.2
5	xA−C	29.9	6.8 ± 0.2	6.8 ± 0.1	6.8 ± 0.2	16	xT−C	25.5	5.8 ± 0.1	5.4 ± 0.3	5.6 ± 0.2
6	xA−A	27.8	6.5 ± 0.1	6.3 ± 0.1	6.4 ± 0.1	17	xT−ϕ	21.1	5.3 ± 0.2	4.9 ± 0.2	5.1 ± 0.2
7	xA−ϕ	29.6	7.0 ± 0.2	6.6 ± 0.2	6.8 ± 0.2	18	A−xT	34.0	7.8 ± 0.1	7.8 ± 0.1	7.8 ± 0.1
8	A−xA	29.5	6.7 ± 0.2	6.8 ± 0.1	6.8 ± 0.2	19	T−xT	33.7	7.6 ± 0.1	7.6 ± 0.1	7.6 ± 0.1
9	G−xA	31.1	6.9 ± 0.1	7.1 ± 0.1	7.0 ± 0.1	20	G−xT	29.3	6.7 ± 0.1	6.7 ± 0.1	6.7 ± 0.1
10	C−xA	31.8	7.2 ± 0.1	7.2 ± 0.1	7.2 ± 0.1	21	C−xT	28.7	6.3 ± 0.1	6.3 ± 0.1	6.3 ± 0.1
11	T−xA	36.2	8.2 ± 0.1	8.2 ± 0.1	8.2 ± 0.1	22	ϕ−xT	24.8	5.9 ± 0.2	5.8 ± 0.2	5.8 ± 0.2

[a] Conditions: 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES, pH = 7.0, 5.0 μM DNA, with DNA sequences AAG AA**X** GAA AAG and CTT TTC **Y**TT CTT. [b] Obtained by plotting $1/T_m$ vs $\ln(C_T)$ with data from five concentrations; standard deviations are shown. [c] Averaged from five denaturation curves; standard deviations are shown. [d] Average of ΔG° obtained by van't Hoff analysis and curve fits.

Before one can study the interactions of proteins and enzymes with modified DNA structures, it is important to understand the properties of the altered DNA structures alone. For example, there is the question of how sterically constraining the natural DNA backbone is to nucleobases that are larger than the surrounding natural pairs. To address this question, we undertook this study of two size-expanded nucleosides, dxA and dxT, in the context of the natural DNA double helix. The results provide new information about the flexibility of the natural DNA backbone and about the origins of DNA base stacking stability as well. We also report preliminary observations of changing fluorescence properties of dxT and dxA in the context of double- and single-stranded structures, a property that may be useful in future applications probing nucleic acid structures and DNA-recognizing enzymes.

Results

Pairing Near the Center of a Helix. Initial experiments were aimed at testing the ability of the natural DNA backbone to adapt to larger base-pair structures. The expanded nucleobases xA and xT (prepared as described previously)[5b] were used to probe this size effect (Figure 1). The base-pairing properties of nucleosides dxA and dxT were evaluated in short duplex DNAs by optically monitored thermal melting studies in a pH 7.0 buffer containing 100 mM Na^+ and 10 mM Mg^{2+}. A 12 base-pair context (with the variable pair centrally located) was used to measure any pairing preferences of xA and xT. To test the effects of local nearest-neighbor bases, we tested both expanded bases singly substituted in each strand of the duplex, with purine neighbors or pyrimidine neighbors, respectively. The pairing data are shown in Table 1.

Results show that when centrally substituted in a short duplex, both xA−T and xT-A pairs are destabilizing to the duplex as compared to natural A−T pairs in the same sequence contexts. For example, xA−T lowered the T_m by 4.2−4.9 °C and caused a loss of 1.0−1.2 kcal/mol of favorable free energy (37 °C) for duplex formation as compared to the A−T pair. Similarly, xT paired opposite A resulted in a loss of 5.2−6.7 °C and 1.2−1.5 kcal/mol relative to T−A. However, despite this destabilization, in both cases these expanded bases still showed a significant pairing preference for the expected H-bonded partner. For example, xA paired more stably opposite T as compared to opposite G, C, or A by 5.9−8.5 °C in T_m and 1.3−1.7 kcal/mol in free energy (entries 3−12, Table 1). The xT base, when incorporated into a purine strand, also showed a preference for being paired opposite A rather than T, C, or G, with a selectivity of 1.0−2.4 kcal/mol (entries 13−17, Table 1). However, with xT in the pyrimidine strand, A−xT and T−xT pairs showed similar stability (entries 18 and 19, Table 1). Taken together, the results suggest that, although destabilizing, these size-expanded bases are likely to exist as paired partners with their Watson−Crick complements. Also consistent with this interpretation is the results where an abasic sugar was present opposite the expanded bases. The data show that the abasic site was more destabilizing as a partner than the complementary base was (Table 1). For example, the xA−T pair displayed a T_m of 35.8 °C in this sequence, while the xA abasic pair gave a T_m of 29.6 °C. The added stability of the xA−T case suggests the positive contributions of stacking and hydrogen bonding by T within the duplex. Similar results were found for xT paired opposite the abasic analogue (Table 1).

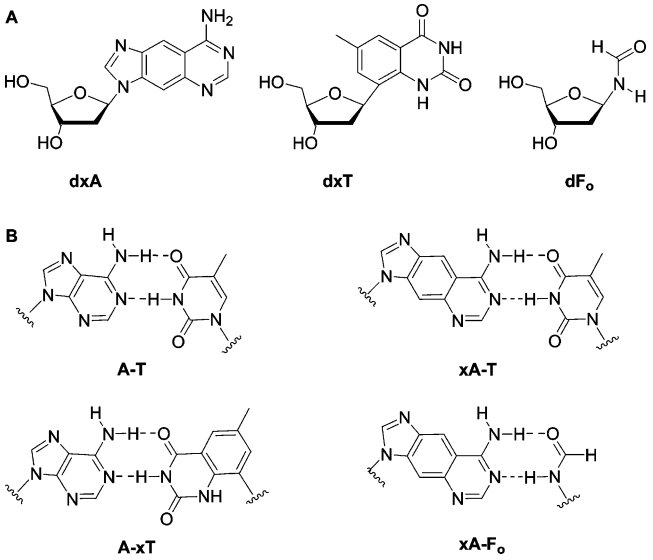

Figure 1. Structures of the size-expanded and shortened nucleoside analogues. (A) The free size-altered nucleosides. (B) Proposed structures of hydrogen-bonded base pairs involving extended or compressed bases, compared with natural A−T pairs.

Pairing and Stacking at the End of a Helix. Since the stability of a pair depends heavily on the stacking properties of the constituent bases, we explicitly measured stacking propensities for xA and xT in the context of a short self-complementary sequence designed to form a duplex with a single overhanging

Table 2. Free Energies ($-\Delta G^\circ_{37}$ (kcal/mol)) and Melting Temperatures (T_m (°C)) of Expanded Bases Stacking or Base-Pairing at the End of DNA Duplexes

sequence	T_m[a]	ΔT_m	$-\Delta G^\circ_{37}$[b] van't Hoff	$-\Delta G^\circ_{37}$[d] curve fits	$-\Delta G^\circ_{37}$[e] average	$-\Delta\Delta G^\circ_{37}$
CGCGCG[c]	41.7	−	8.1 ± 0.2	8.1 ± 0.1	8.1 ± 0.2	−
ACGCGCG[c]	51.6	9.9	10.1 ± 0.2	10.0 ± 0.4	10.0 ± 0.3	1.9 ± 0.3
TCGCGCG[c]	48.1	6.4	9.2 ± 0.2	9.2 ± 0.9	9.2 ± 0.5	1.1 ± 0.5
xACGCGCG	63.0	21.3	12.9 ± 0.3	12.1 ± 0.2	12.5 ± 0.2	4.4 ± 0.2
xTCGCGCG	57.7	16.0	11.2 ± 0.2	11.1 ± 0.3	11.2 ± 0.2	3.1 ± 0.2
CGCGCGT	44.7	3.0	8.8 ± 0.1	8.6 ± 0.1	8.7 ± 0.1	0.6 ± 0.2
ACGCGCGT	56.1	14.4	12.1 ± 0.3	11.0 ± 0.1	11.6 ± 0.2	3.5 ± 0.2
xACGCGCGT	62.8	21.1	12.3 ± 0.5	12.4 ± 0.2	12.4 ± 0.4	4.3 ± 0.4

[a] Conditions: 1.0 M NaCl, 10 mM NaH_2PO_4, pH = 7.0, 5.0 μM DNA strand concentration. [b] Obtained by plotting $1/T_m$ vs $\ln(C_T)$ with data from five concentrations; standard deviations are shown. [c] Data abstracted from ref 8f. [d] Average from five denaturation curves; standard deviations are shown. [e] Average of ΔG° obtained by van't Hoff analysis and curve fits.

base, putatively stacked at the end. The data (with comparisons to the natural nucleobases) are shown in Table 2, with comparison to the same duplex lacking an overhanging base. In this system, the added stabilization is expected to arise from stacking of the overhanging base on the neighboring duplex.[8]

Interestingly, dangling xA and xT bases were found to stabilize the core duplex much more strongly than their natural counterparts in this sequence. Assuming that this stabilization arises from stacking, then the xA case stacked more strongly than A by 2.5 kcal/mol (1.3 kcal per overhanging base). Similarly, xT was found to stack more strongly than T by 1.0 kcal/mol per base in this 5′ substitution. The bases are strongly stabilizing overall: xT, compared with no base, stabilized the helix by −1.6 kcal/mol per base, which is nearly as stabilizing as addition of a full T−A base pair (Table 2). The xA base stabilized the helix by −2.2 kcal/mol per base, which is stronger than any nucleobase analogue yet studied.[9]

To evaluate the effect of base-pair sterics at the end of a DNA sequence (as opposed to the center as described above), we also evaluated the stacking and pairing stabilities of xA at the terminus of this self-complementary helix. xA was tested alone and in combination with T, to allow for the possibility of pairing. The data show, interestingly, that the putative xA−T pair appears to be less disruptive at the end of a duplex than at the center. In fact, a single A−T pair added −1.8 kcal/mol of stability to this specific duplex, while the xA−T pair added −2.2 kcal of stabilization in the same context (Table 2).

However, the stacking data for xA and T suggest that this putative pair, if formed, still pays a significant energetic cost because of its excessive size: the total pair stabilization was less than the sum of the stacking of the component bases alone. The stacking of single xA and T summed to −2.5 kcal/mol (37 °C), suggesting a size penalty of roughly +0.4 kcal. By contrast, the stacking of A and T summed to −1.3 kcal/mol of stabilization, while the full pair added −1.8 kcal/mol, consistent with a small incremental (−0.5 kcal) benefit of H-bonding. If one assumes the same H-bonding contribution to the putative xA−T pair, this brings the estimated steric penalty for xA−T to +0.9 kcal/mol at the end of the duplex.

Table 3. Thermodynamic Parameters for DNA Duplexes Containing One, Two, or Three xA−T Pairs

sequence	T_m (°C)[a]	$-\Delta G^\circ_{37}$[b,c] van't Hoff	$-\Delta G^\circ_{37}$[b,d] curve fits	$-\Delta G^\circ_{37}$[b,e] average
d(AAGAAxAGAAAAG) ·d(TTCTTTCTTTTC)	35.8	8.1 ± 0.1	8.1 ± 0.1	8.1 ± 0.1
d(AAGAxAxAGAAAAG) ·d(TTCTTTCTTTTC)	35.5	8.1 ± 0.1	8.1 ± 0.1	8.1 ± 0.1
d(AAGxAxAxAGAAAAG) ·d(TTCTTTCTTTTC)	37.2	8.5 ± 0.1	8.4 ± 0.1	8.4 ± 0.1

[a] Conditions: 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES, pH = 7.0, 5.0 μM DNA. [b] $-\Delta G^\circ_{37}$ values in kcal/mol. [c] Obtained by plotting $1/T_m$ vs $\ln(C_T)$ with data from five concentrations; standard deviations are shown. [d] Average from five denaturation curves; standard deviations are shown. [e] Average of ΔG° obtained by van't Hoff analysis and curve fits.

Adjacent Substitutions of xA Base. The above results suggested that the natural phosphodiester backbone of duplex DNA experiences significant conformational strain in response to the added sizes of xA and xT. We tested this effect further by studying increasing numbers of adjacent substitutions of a stretched base. In the 12-base-pair duplex context described above, we compared the stabilities of one, two, or three xA bases consecutively placed opposite T bases in the sequence. These were compared to the sequence containing natural A at the same positions. The data are shown in Table 3. The second substitution of xA for A was found to have the same energetic penalty as a single substitution. Interestingly, the addition of a third xA residue showed an increase in stability, approaching the stability of the naturally substituted analogous duplex. These three xA-containing sequences were examined by CD spectroscopy to evaluate whether the DNA structure was perturbed broadly. The results showed (Figure S1, Supporting Information) that one and three substitutions gave spectra closely resembling that of the unsubstituted B-DNA helix. However, the doubly substituted case showed a significant difference, suggesting at least local distortion away from the natural helix structure.

Testing Size Compensation with a Shortened Base. The strain induced in the natural double helix might in principle be compensated for by shortening the complement of xA by the same length and direction that xA is extended. Thus, Leonard proposed the formamide nucleoside F_o (Figure 1) as a size-compensating partner for *lin*-benzo-A.[3] To test for such possible steric compensation, we prepared F_o and substituted it into duplexes opposite xA and opposite natural bases (Table S2,

(8) (a) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256−263. (b) Turner, D. H.; Sugimoto, N. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 167−192. (c) Senior, M.; Jones, R. A.; Breslauer, K. J. *Biochemistry* **1988**, *27*, 3879−3885. (d) Sugimoto, N.; Matsumura, A.; Hasegawa, K.; Sasaki, M. *Nucleic Acids Symp. Ser.* **1991**, *25*, 51−52. (e) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Paris, P. L.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182−8183. (f) 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*, 2213−2222.

(9) Nakano, S.; Uotani, Y.; Nakashima, S.; Anno, Y.; Fujii, M.; Sugimoto, N. *J. Am. Chem. Soc.* **2003**, *125*, 8086−8087.

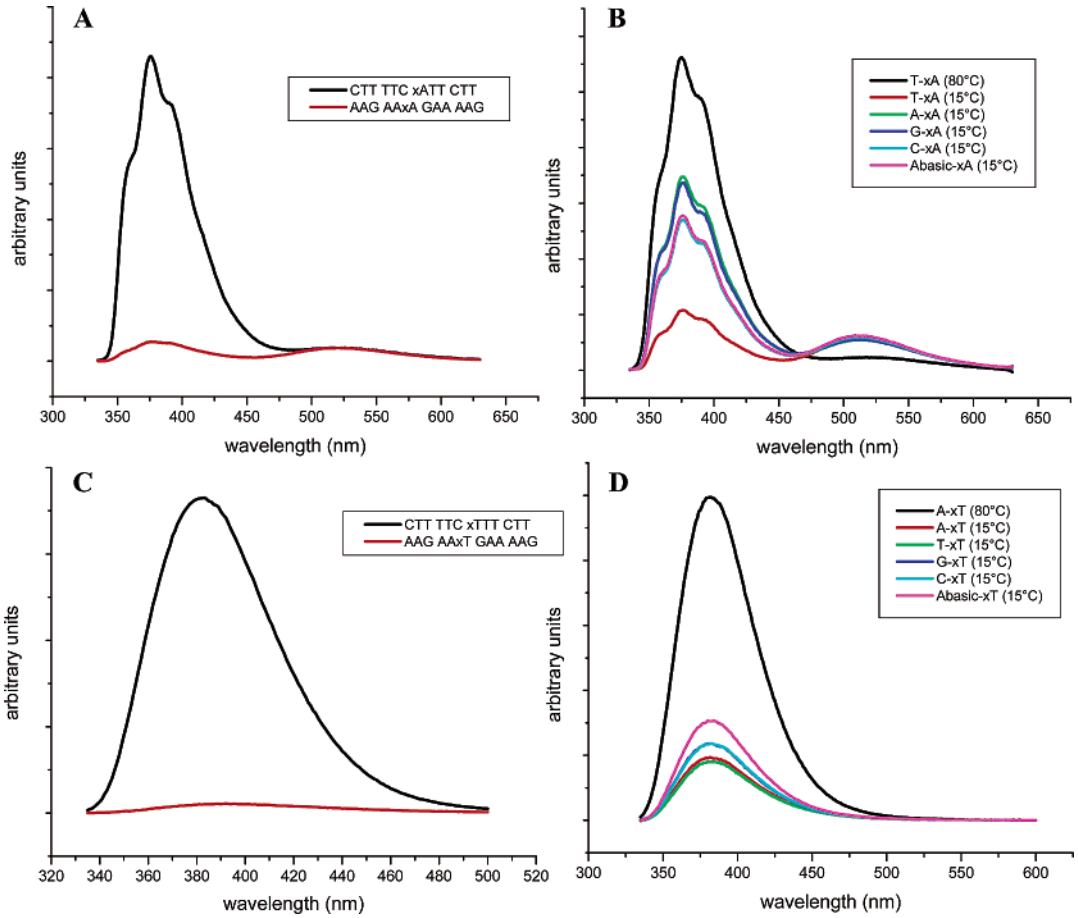

Figure 2. Fluorescence spectra of expanded base analogues in DNA, with excitation at 320 nm. (A) Fluorescence emission of single-stranded DNA containing xA. (B) Fluorescence emission of DNA duplexes containing xA. (C) Fluorescence emission of single-stranded DNA containing xT. (D) Fluorescence emission of DNA duplexes containing xT. All samples are dissolved in the buffer of 100 mM NaCl, 10 mM $MgCl_2$, 10 mM Na·PIPES, pH = 7.0. All plots are normalized with absorption at 320 nm. In (B) and (D), DNA duplexes used were d(AAG AA**X** GAA AAG)·d(CTT TTC **Y**TT CTT) with X−Y pairs listed on the plots.

Supporting Information). The results show that, compared with the abasic nucleoside described above, F_o gave no additional stabilization when paired opposite xA or other bases. Further examination of the free nucleoside of F_o and of dF_o in a trimer oligodeoxynucleotide by proton NMR revealed that F_o in these contexts exists as a mixture of two conformational isomers (data not shown and ref 10). This poor conformational complementarity may explain the unstable pairing by this compound (see Discussion).

Fluorescence Changes Report on Pairing and Stacking in DNA. Fluorescent nucleoside analogues have been widely useful as tools in the basic study of DNA structure and enzyme−DNA recognition.[11] The above two pairing and stacking contexts represent simple examples of DNA helical sequences similar to those that might be tested in such fluorescence studies. As a preliminary test of whether the xA and xT bases can respond to changes in local helicity by fluorescence, we characterized the emission properties of these bases in paired and unpaired (and presumably, well-stacked and less well-stacked) conformations. As the free nucleosides, these bases fluoresce with maxima near 385 nm and with quantum yields of 0.44 and 0.30, for xA and xT, respectively.[5b]

The results showed (Figure 2) that the fluorescence of both xA and xT strongly depend on their neighboring bases when incorporated into DNA. Adjacent purines quenched the fluorescence of xA five times more strongly than did the pyrimidines, and a >30-fold difference in quantum yield was observed with xT in a pyrimidine strand compared with a purine strand (Figure 2). When incorporated into pyrimidine strands, fluorescence of these expanded bases is also quenched upon duplex formation (Figure 2B,D). Therefore, both intrastrand and interstrand interactions between purines and these expanded bases can cause fluorescence quenching. Duplex formation of the dangling end sequence (5′-XCGCGCG-3′) also caused fluorescence quenching by 4- and 6-fold, respectively, for xT and xA (Table S3, Supporting Information). Presumably the well-stacked conformation in a duplex favors the interstrand interaction between expanded bases and guanine in the neighboring pair. Interestingly, when xA is incorporated into DNA, a new fluorescence emission band with maximum near 520 nm was observed (Figure 2A,B). Presumably this new emission band is due to electronic interactions between xA and the neighboring natural nucleobases. More interestingly, the ratio of 520-nm emission to xA monomer emission is dependent on the neighboring nucleobases as well. Quantum yields of xA and

(10) Guy, A.; Duplaa, A. M.; Ulrich, J.; Teoule, R. *Nucleic Acids Res.* **1991**, *19*, 5815−5820.
(11) For a review on pteridine nucleoside analogues, see: (a) Hawkins, M. E. *Top. Fluoresc. Spectrosc.* **2003**, *7*, 151−175. For recent work with 2-aminopurine, see: (b) Rachofsky, E. L.; Osman, R.; Ross, J. B. A. *Biochemistry* **2001**, *40*, 946−956 and references therein.

xT in single-stranded and duplex DNAs are given in Table S3 (Supporting Information).

Discussion

Our data are useful in analyzing how resilient the DNA backbone is to changes in steric size. This is an important issue not only for designed DNA structures, but also for its relevance to DNA replication and repair. The topic of how DNA repair enzymes recognize damage rapidly in a large genome has been under intense study recently.[12] Many of the known forms of DNA damage include adducts to nucleobases, often leading to increased size.[13] This is likely to lead not only to structural distortion of the helix, but also to energetic destabilization. The present results show that a 2.4 Å increase in size for an isolated base pair destabilizes the duplex by ca. 1 to 2 kcal/mol. In some respects this is a surprisingly small value, given the rigidity and persistence length of duplex DNA.[14] In the present case, this value is presumably diminished in part by the exceptionally strong stacking of the xA base. One known damaged base adduct, ethenoadenine,[13c] approaches the size of xA. We know of no available data on the effects of ethenoadenine on the stability of DNA helices.

Recently, Geyer et al. evaluated the energy penalty for DNA backbone distortion caused by incorporating a purine−purine pair into a standard DNA duplex.[15] Two pairing geometries are possible for a purine−purine pair: anti−anti and anti−syn. When both purines are anti, the Watson−Crick face is used, and the C1′−C1′ distance is longer than normal base pairs in DNA. 7-Deazaxanthine (dX), with the same Watson−Crick H-bonding arrangement as thymine and lacking N7 on the Hoogsteen face, was proposed to form an anti−anti pair with A. Compared with a T−A pair, the dX−A pair destabilized the DNA duplex by a drop of 4.4 °C in T_m and 1.0 kcal/mol in free energy. The value of energy cost for steric expansion for purine−purine pairs is close to our results obtained for an xT−A pair in the current study. This is not unexpected, since dX is similar in size to our dxT analogue. The authors also tested pairing stability of compressed base pairs composed of two pyrimidines. Compared with the G−C pair, the compressed iC−C pair destabilized a DNA duplex by ca. 2−3 kcal/mol. The data showed that compressed base pairs are more destabilizing than expanded base pairs of similar size alteration. This agrees with our hypothesis that the energy penalty caused by size expansion is partially compensated for by the stronger stacking contribution of the expanded base pairs.

Although our results are highly suggestive of pairing by xA with T and xT with A in natural DNA, the current data do not conclusively prove the existence of this Watson−Crick-like pairing of expanded bases in this context. Additional structural data will be needed to investigate this geometric question in detail. It is possible, for example, that the expanded bases might merely interdigitate themselves with their partners, in an intercalator-like mode, or alternatively, cause the pairing partner to be "flipped out" of the helix. However, we suggest that the present data are best interpreted by a paired structure for xA−T and xT−A. In both cases, there are favorable energetics observed with the expected partner relative to "mismatched" bases. It is difficult to interpret this in any way other than hydrogen-bonded pairing. For example, an interdigitated structure would be expected to yield greatest stability with the strongest-stacking combinations, i.e., xA with A (since A stacks more strongly than the other bases),[8d,f] but this is not observed. Finally, the CD data also support the idea of a relatively nonperturbed structure for DNAs containing single putative expanded pairs (Figure S1, Supporting Information). The final conclusions will await further structural information.

It is interesting to speculate on the structures of *mismatched* pairs of the expanded bases. The data show that there are essentially no differences in stability among the mismatched cases; for example, xA pairs equally poorly with C, T, or A. We suggest that one intriguing possibility is that the large size of the expanded bases, combined with their extremely stable stacking, may result in "flipped out" conformations for these mismatched partners. Our previously reported pyrene deoxynucleoside[16] (which also stacks with high proficiency[8e]) has been shown, in fact, to cause bases paired opposite to flip out of the helix.[17] Such a conformation is reminiscent of the intermediates in recognition of DNA by base excision repair enzymes. The pyrene nucleotide has been useful in the study of such enzymes, showing (for example) enhanced affinity of binding.[17] Studies are needed to see whether a similar effect may occur with the current size-expanded bases.

Our experiments with further successive substitutions of an expanded base show that they are not as destabilizing to DNA as a single expanded pair, and they begin to negate some of the destabilization cause by such a pair. Similar results have been reported recently for nucleobases designed to form four hydrogen bonds with complements in DNA.[18] Notably, recent data have shown that fully substituted helices can be formed with xT−A and xA−T pairs, yielding little discernible destabilization (in fact, the opposite is the case, with expanded helices showing considerably greater stability than natural DNA of the analogous sequence).[5a] Because fully expanded helices are stable, it is clear that the penalty for an expanded pair comes not from its size but rather from its *mismatch* in size relative to the surrounding DNA. Thus, one expects two destabilizing junctions between differently sized pairs embedded in a helix, and there should be two such penalties regardless of how many adjacent expanded pairs are present in natural DNA, as long as natural DNA surrounds the substituted bases. Our observation of smaller destabilization by xA paired opposite T at the end of a helix is also consistent with this notion.

Leonard proposed the interesting concept of compensating for the added size by pairing the base xA with the shortened formamide residue in RNA helices.[3] He successfully prepared a ribonucleotide dimer of rxA and rF_o, but did not characterize any pairing. Despite this, we felt that the concept was interesting

(12) Verdine, G. L.; Bruner, S. D. *Chem. Biol.* **1997**, *4*, 329−334.
(13) (a) Schärer, O. D. *Angew. Chem., Int. Ed.* **2003**, *42*, 2946−2974. (b) Marnett, L. J. *Carcinogenesis* **2000**, *21*, 361−370. (c) Marnett, L. J.; Plastaras, J. P. *Trends Genet.* **2001**, *17*, 214−221.
(14) (a) Jolly, D.; Eisenberg, H. *Biopolymers* **1976**, *15*, 61−95. (b) Frontali, C.; Dore, E.; Ferranto, A.; Gratton, E.; Bettini, A.; Pozzan, M. R.; Valdevit, E. *Biopolymers* **1979**, *18*, 1353−1373. (c) Schellman, J. A. *Biopolymers* **1974**, *13*, 217−226. (d) Millar, D. P.; Robbins, R. J.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 5593−5597.
(15) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure* **2003**, *11*, 1485−1498.
(16) Ren, R. X. F.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 7671−7678.
(17) Beuck, C.; Singh, I.; Bhattacharya, A.; Heckler, W.; Parmar, V. S.; Seitz, O.; Weinhold, E. *Angew. Chem., Int. Ed.* **2003**, *42*, 3958−3960.
(18) Minakawa, N.; Kojima, N.; Hikishima, S.; Sasaki, T.; Kiyosue, A.; Atsumi, N.; Ueno, Y.; Matsuda, A. *J. Am. Chem. Soc.* **2003**, *125*, 9970−9982.

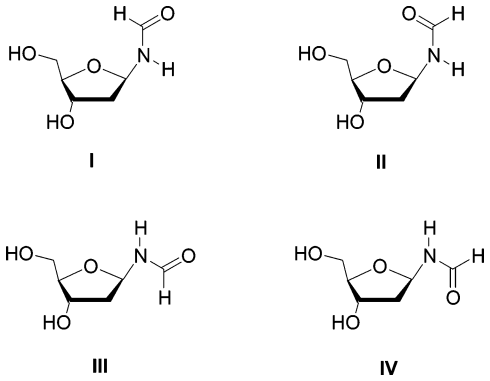

Figure 3. Possible conformational isomers of F_o. The conformation I is required for correct hydrogen bonding with xA in DNA.

enough to warrant an investigation in longer synthetic oligonucleotides, which were not available at the time of Leonard's work. The present data show that F_o pairs very poorly in DNA, and our data reveal multiple conformations for this nucleoside in DNA. In retrospect, the destabilization associated with F_o may not be surprising. A simple analysis predicts four conformational isomers for this compound, resulting from syn or anti glycosidic orientation combined with amide carbonyl−nitrogen rotamers orienting the carbonyl up or down (Figure 3). Only one of these (II, Figure 3) is expected to be disfavored sterically by standard nucleoside conformational analysis. For the desired pairing, the conformation I is needed; however, this one is expected to be oriented such that it stacks with the neighboring bases very little or not at all compared with other rotamers. It should be noted that although the present F_o base does not pair well, the present results do not preclude the possibility of other size-compensated bases forming stable pairs. They do suggest, however, that some degree of conformational analysis is warranted for new molecular tests of this strategy.

We expect that size-expanded pairs such as xT−A and xA−T may be useful as probes of steric tightness around DNA bases pairs in the active sites of DNA repair and DNA polymerase enzymes. We have proposed that the active site tightness is a primary factor in DNA replication fidelity, by closely enforcing complementary sizes and shapes of paired DNA bases, thus stabilizing the correct pair and sterically destabilizing the incorrect ones by several kcal/mol at the transition state for their formation.[7] Notably, work by Marx using C-4′ modified nucleosides suggests that such a mechanism may indeed be operable at the sugar backbone level.[6] The current compounds offer a way to examine this in the direct location of the base pair, where the DNA sequence information is stored.

Acknowledgment. This work was supported by the U.S. National Institutes of Health (GM52956/EB002059 and GM63587) and by Stanford Graduate Fellowships to J.G. and H.L.

Supporting Information Available: Experimental details of nucleoside and oligonucleotide synthesis and of thermodynamics methods. This material is available free of charge via the Internet at http://pubs.acs.org.

JA048499A