Nucleic Acids Research, 2005, Vol. 33, No. 22 ***7111–7119***
doi:10.1093/nar/gki1018

DNA base flipping by a base pair-mimic nucleoside

Shu-ichi Nakano[1], Yuuki Uotani[2], Kazuya Uenishi[3], Masayuki Fujii[3,4] and Naoki Sugimoto[1,2,*]

[1]Frontier Institute for Biomolecular Engineering Research (FIBER) and [2]Department of Chemistry, Faculty of Science and Engineering, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan, [3]Molecular Engineering Institute and [4]Department of Environmental and Biological Chemistry, Kinki University, 11-6 Kayanomori, Iizuka, Fukuoka 820-8555, Japan

Received October 27, 2005; Revised and Accepted November 29, 2005

ABSTRACT

On the basis of non-covalent bond interactions in nucleic acids, we synthesized the deoxyadenosine derivatives tethering a phenyl group (X) and a naphthyl group (Z) by an amide linker, which mimic a Watson–Crick base pair. Circular dichroism spectra indicated that the duplexes containing X and Z formed a similar conformation regardless of the opposite nucleotide species (A, G, C, T and an abasic site analogue F), which was not observed for the natural duplexes. The ΔG^0_{37} values among the natural duplexes containing the A/A, A/G, A/C, A/T and A/F pairs differed by 5.2 kcal mol^{-1} while that among the duplexes containing X or Z in place of the adenine differed by only 1.9 or 2.8 kcal mol^{-1}, respectively. Fluorescence quenching experiments confirmed that 2-amino purine opposite X adopted an unstacked conformation. The structural and thermodynamic analyses suggest that the aromatic hydrocarbon group of X and Z intercalates into a double helix, resulting in the opposite nucleotide base flipping into an unstacked position regardless of the nucleotide species. This observation implies that modifications at the aromatic hydrocarbon group and the amide linker may expand the application of the base pair-mimic nucleosides for molecular biology and biotechnology.

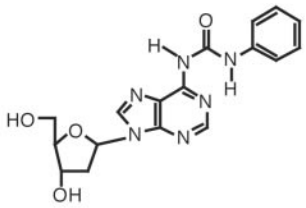
INTRODUCTION

Hydrogen bonds play a central role in the formation of Watson–Crick A/T and G/C base pairs. In addition to the interstrand hydrogen bonds, base stacking is important for the integrity and stabilization of the double helix structure of DNA and RNA. The stabilization energy provided by the base stacking is estimated to be 1.8–0.5 kcal mol^{-1} (1 kcal = 4.18 kJ) in ΔG^0_{37}, which is comparable with the energy provided by hydrogen bond formation (1–3). On the basis of the non-covalent interactions in nucleic acids, artificial nucleosides have been examined from the standpoints of evaluation of the base-pairing energy (1,3), recognition of a target DNA sequence and dNTPs by polymerases (4–6), expansion of the genetic alphabet (7,8) and development of functional molecules (9–16), where planar aromatic molecules are often examined because compounds having a contiguous π-system have a high potential for stacking on a nucleotide base and can intercalate between two adjacent base pairs in a duplex.

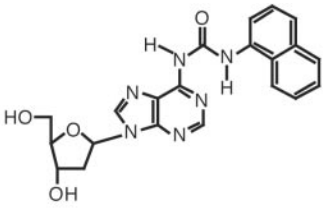
We have synthesized the deoxyadenosine derivatives tethering a phenyl group [N^6-(*N*′-phenylcarbamoyl)-2′-deoxyadenosine, **X**] and a naphthyl group [N^6-(*N*′-naphthylcarbamoyl)-2′-deoxyadenosine, **Z**] by an amide linker (Scheme 1), where the aromatic hydrocarbon group is expected to stack with a double helix when it orients in an appropriate position. Our previous work demonstrated that the single deoxyadenosine derivatives at the 5′ dangling ends stabilized the self-complementary DNA duplexes of 5′-ATGCGCAT-3′ and 5′-TGCGCA-3′ more than a single adenine, and the stabilization energy was as much as or greater than the formation of the Watson–Crick A/T base pair (17). The strong stabilization suggests that the deoxyadenosine derivatives stack on the adjacent terminal base pairs efficiently, probably by adopting a Watson–Crick base pair-mimic geometry.

In this study, the non-self complementary DNA duplexes containing **X** and **Z** in the middle of a sequence were investigated in order to estimate the stacking interaction by the deoxyadenosine derivatives in a duplex. We measured circular dichroism (CD) spectra, ultraviolet (UV) melting curves and fluorescence emission spectra of the DNA duplexes. It was demonstrated that the duplexes containing **X** or **Z** adopted a similar conformation regardless of the opposite nucleotide species, and it was suggested that **X** and **Z** could force the target base to flip into an unstacked position by intercalating

*To whom correspondence should be addressed. Tel: +81 78 435 2497; Fax: +81 78 435 2539; Email: sugimoto@konan-u.ac.jp



N^6-(*N*'-phenylcarbamoyl)-2'-deoxyadenosine (**X**)



N^6-(*N*'-naphthylcarbamoyl)-2'-deoxyadenosine (**Z**)

Scheme 1. Chemical structures of the deoxyadenosine derivatives used in this study.

their aromatic hydrocarbon group in a duplex. The lack of the pairing selectivity of the deoxyadenosine derivatives would have an advantage for the site-selective base flipping in a target sequence.

MATERIALS AND METHODS

Syntheses of the deoxyadenosine derivatives

Synthesis of the phosphoramidite derivatives of **X** and **Z** was started with 2′-deoxyadenosine (the details are described in the Supplementary Data). All chemicals were purchased from Aldrich and Wako Chemicals Co., Ltd, and the solvents were freshly distilled just before use. ^{1}H NMR spectra were measured at 400 MHz with a Varian INOVA 400 NMR spectrometer, and electro spray ionization mass spectra were measured with a Finnigan Mat LCQ mass spectrometer. Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectra were measured with a PE Biosystems Voyager DE MALDI-TOF mass spectrometer using 3-hydroxypicolinic acid as a matrix.

Synthesis of the DNA oligomers

Oligonucleotides were synthesized on a solid support using phosphoramidite chemistry by automated DNA synthesis (Applied Biosystems Model 391) and purified with reversed-phase high-performance liquid chromatography on a C18 column (TOSOH) after the removal of the protection groups, followed by desalting with a C18 cartridge column (18–20). All oligonucleotides were confirmed by MALDI-TOF mass spectra (PE Biosystems Voyager).

CD measurements

CD spectra of the DNA duplexes were obtained with a JASCO J-820 spectropolarimeter equipped with a temperature controller. All spectra were measured at 20 μM in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4 (pH 7.0) and 1 mM Na_2EDTA at 4°C. The DNA samples were heated to 90°C and cooled at a rate of 2°C min^{-1} before use.

Measurements and analysis of the UV melting curves

The UV absorbance was measured with a Shimadzu 1650 or 1700 spectrophotometer equipped with a temperature controller. The melting curves were monitored at 260 nm or 325 nm in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4 (pH 7.0) and 1 mM Na_2EDTA. The extinction coefficients for the single-strand oligonucleotides were calculated based on the nearest-neighbor approximation (21), and those containing the deoxyadenosine derivatives were assumed to be the same as those of the deoxyadenosine.

The thermodynamic parameters (ΔH^0, ΔS^0 and ΔG^0 at 37°C) for duplex formations were calculated from the T_m (melting temperature) obtained at different total strand concentrations (C_ts) using the T_m^{-1} versus $\log(C_t/4)$ plot and the fit of the melting curves with a theoretical equation, as reported previously (22,23).

Fluorescence measurements

Fluorescence emission spectra of the 2-amino purine-containing DNA strands were measured with a JASCO FP-6500 spectrofluorometer equipped with a temperature controller. Since 2-amino purine and the naphthyl group of **Z** have overlapping absorption and emission bands, reliable emission spectrum of 2-amino purine cannot be obtained for the DNAs containing **Z**. The emission spectra were measured at 20 μM duplex or 10 μM single-strand DNA in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4 (pH 7.0) and 1 mM Na_2EDTA at 4°C. The DNA samples were heated to 90°C before use. The emission spectra were obtained with an excitation wavelength at 305 nm, and the band path for the excitation and emission was set at 5 nm.

RESULTS

CD spectra of the DNA duplexes containing the deoxyadenosine derivatives

Scheme 1 indicates the structures of the deoxyadenosine derivatives of **X** (phenylurea derivative) and **Z** (naphthylurea derivative). The artificial nucleotides tether a phenyl group or a naphthyl group with the N^6 of deoxyadenosine by an amide linker, where the aromatic hydrocarbon group is expected to stack with a double helix when it orients in an appropriate position. The DNA duplexes of 5′-GTGTC$\underline{W_1}$CTGTC-3′/5′-GACAG$\underline{W_2}$GACAC-3′ form identical Watson–Crick base pairs except the W_1/W_2 pair, where W_1 is A (A-series), **X** (X-series) or **Z** (Z-series), and W_2 is A, G, C, T or F (tetrahydrofuran abasic site analogue) (Figure 1). When W_1 is A and W_2 is T, the duplex has the Watson–Crick base pairs exclusively. In contrast, the duplexes formed by the DNA strands containing A as W_1 and A, G, C or F as W_2 have a single mismatched pair of A/A, A/G, A/C or A/F, respectively. When W_1 is **X** or **Z**, the duplex contains the **X**/W_2 or **Z**/W_2 pair, respectively.

Figure 2 compares the CD spectra of the duplexes of 5′-GTGTC$\underline{A}$CTGTC-3′/5′-GACAG$\underline{W_2}$GACAC-3′ (A-series),

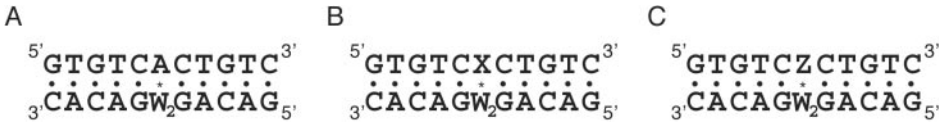


Figure 1. The sequences of (**A**) the A-series, (**B**) the X-series and (**C**) the Z-series DNA duplexes forming the A/W_2, X/W_2 or Z/W_2 pair, respectively, where W_2 is A, G, C, T or F.

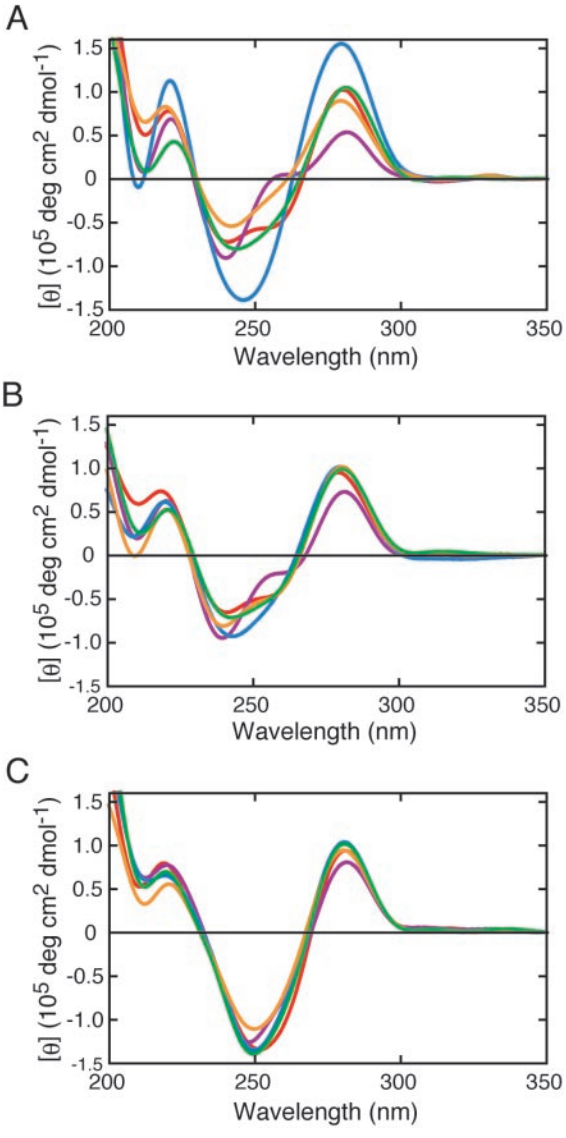


Figure 2. CD spectra of the (**A**) A-series, (**B**) X-series and (**C**) Z-series duplexes [W_2 = A (red), G (purple), C (blue), T (green) or F (orange)]. All measurements were done in the 1 M NaCl-phosphate buffer at 4°C.

5′-GTGTCXCTGTC-3′/5′-GACAGW_2GACAC-3′ (X-series) and 5′-GTGTCZCTGTC-3′/5′-GACAGW_2GACAC-3′ (Z-series). All the duplexes showed a typical B-form signal with positive peaks around 280 and 220 nm and negative peaks around 245 and 210 nm. However, the intensity of the peaks and the overall shape of the spectra of the A-series duplexes differed substantially depending on the W_2 nucleotide (Figure 2A), indicating that the A/A, A/G, A/C, A/T and A/F pairs adopt a different conformation, as has been reported for the duplexes containing a single mismatch (24–29). In contrast, differences in the CD spectra of the X-series duplexes (Figure 2B) were less than those of the A-series duplexes. Figure 2C also indicates nearly identical spectra for all the Z-series duplexes. These observations are suggestive of similar duplex structures regardless of W_2 nucleotide in the **X**/W_2 or **Z**/W_2 pair.

Thermodynamic parameters for duplex formations containing the deoxyadenosine derivatives

Figure 3A–C shows the typical melting curves of the duplexes monitored at 260 nm. The UV melting curves are consistent that all the duplexes denatured in a two-state transition. The T_m values of the A-series duplexes were in the range from 55.4 to 38.5°C ($\Delta T_m = T_{m,max} - T_{m,min} = 16.9$°C), while those of the X-series and Z-series duplexes were from 52.2 to 46.6°C ($\Delta T_m = 5.6$°C) and from 58.7 to 53.2°C ($\Delta T_m = 5.5$°C), respectively, consistent with the fact that the duplexes containing **X** or **Z** adopt a similar structure. The melting curves of the Z-series duplexes monitored at 325 nm are also shown in Figure 3D. The absorption at 325 nm by the naphthyl group decreased above 50°C, consistent with the T_m determined from the absorption change monitored at 260 nm. The finding suggests the naphthyl group of **Z** stacking on nucleotide bases in a duplex and melting simultaneously with the dissociation of nucleotide bases. In addition, the facts that the extinction coefficients of the Z-series duplexes are close ($8.2–7.0 \times 10^4$ cm^{-1} M^{-1} at 325 nm, determined from the absorption at 0°C) and the hypochromicities at 325 nm owing to heat denaturation are similar imply that the naphthyl group assumes a similar environment in a duplex regardless of the type of W_2 nucleotide.

Table 1 compares the thermodynamic parameters for the duplex formations obtained from the UV melting curves. The ΔG^0_{37} of the duplexes differed depending on the W_1/W_2 pair. In the A-series DNAs, the fully matched duplex (W_1 = A and W_2 = T) presented the greatest stability (−12.6 kcal mol^{-1} in ΔG^0_{37}), and the duplexes containing the A/A, A/G and A/C mismatches showed a lower stability (−8.3 to −10.2 kcal mol^{-1}). The ΔH^0 and ΔS^0 values of these duplexes also differed considerably (−62.2 to −89.0 kcal mol^{-1} and −174 to −248 cal mol^{-1} K^{-1}, respectively), suggesting different interactions near the A/W_2 pair. In addition, the duplex containing the abasic site analogue showed −7.4 kcal mol^{-1} in ΔG^0_{37}, which was the lowest stability owing to the lack of the nucleotide base (30,31).

It is remarkable that the ΔG^0_{37} values among the X-series duplexes differed by only 1.9 kcal mol^{-1} (−8.8 to −10.7 kcal mol^{-1}) while those among the A-series duplexes differed by 5.2 kcal mol^{-1} (−7.4 to −12.6 kcal mol^{-1}). The differences in the ΔH^0 and ΔS^0 values among the X-series duplexes (16.8 kcal mol^{-1} and 50 cal mol^{-1} K^{-1}, respectively)

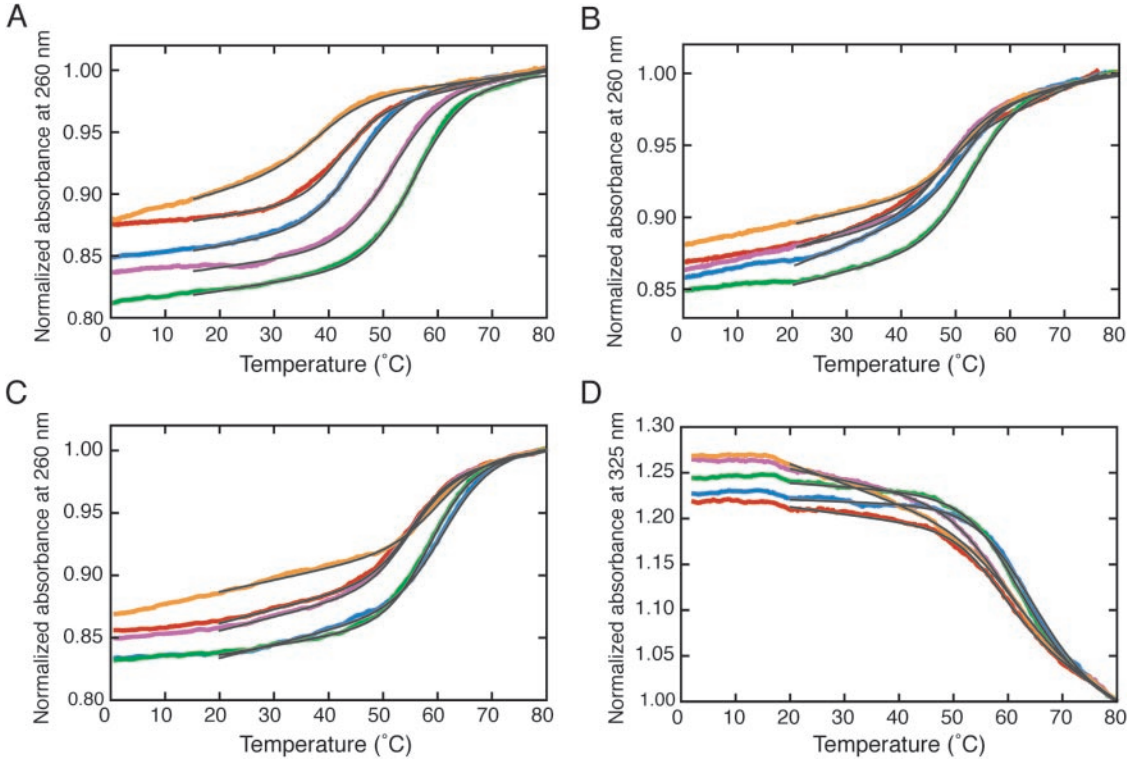


Figure 3. Melting curves of the (**A**) A-series, (**B**) X-series and (**C**) Z-series duplexes monitored at 260 nm [W_2 = A (red), G (purple), C (blue), T (green) or F (orange)]. All measurements were carried out at ~40 μM in the 1 M NaCl-phosphate buffer. The T_m values are as follows: A/A (42.5°C), A/G (51.3°C), A/C (44.5°C), A/T (55.4°C), A/F (38.5°C), **X**/A (46.6°C), **X**/G (48.0°C), **X**/C (49.3°C), **X**/T (52.2°C), **X**/F (50.8°C), **Z**/A (53.4°C), **Z**/G (53.2°C), **Z**/C (58.5°C), **Z**/T (57.4°C) and **Z**/F (58.7°C). (**D**) Melting curves of the Z-series duplexes monitored at 325 nm. The fits to a two-state model are indicated in a black line.

are also considerably smaller than those among the A-series duplexes (26.8 kcal mol^{-1} and 74 cal mol^{-1} K^{-1}, respectively). It is also striking that the thermodynamic parameters of the duplexes containing **X** opposite the purine base are quite similar (−57.5 and −56.6 kcal mol^{-1} in ΔH^0, −157 and −153 cal mol^{-1} K^{-1} in ΔS^0, and −8.8 and −9.2 kcal mol^{-1} in ΔG^0_{37}), and those duplexes containing **X** opposite the pyrimidine base are identical (−73.1 and −72.8 kcal mol^{-1} in ΔH^0, −202 and −200 cal mol^{-1} K^{-1} in ΔS^0 and −10.7 and −10.5 kcal mol^{-1} in ΔG^0_{37}). The same trends were also found in the parameters of the Z-series duplexes.

Comparison of the thermodynamic parameters of the duplexes having different adjacent base pairs

Table 1 demonstrates the results of the duplexes containing **X** next to two adjacent C/G base pairs (C**X**C/GW_2G, as a trinucleotide expression). In order to investigate influences of the adjacent base pairs on the duplex stability, different sequence sets were further examined. Table 2 summarizes the thermodynamic parameters of the duplexes containing the **X**/A pair next to two adjacent A/T, G/C or T/A base pairs (A**X**A/TAT, G**X**G/CAC or T**X**T/AAA, respectively) and the corresponding natural duplexes. As demonstrated for CAC/GW_2G and C**X**C/GW_2G in Table 1, the displacement of A at W_1 by **X** increased the duplex stability. The increases in ΔG^0_{37} for A**X**A/TAT and G**X**G/CAC were greater (1.7 and 1.5 kcal mol^{-1}, respectively) than those for C**X**C/GAG and T**X**T/AAA (0.5 and 1.1 kcal mol^{-1}, respectively). It is therefore likely that the stabilization energy provided by the displacement of adenine by **X** is context-dependent, affected by the two adjacent base pairs.

Fluorescence emission spectra of a fluorescent base opposite the deoxyadenosine derivative

To reveal the nucleotide conformation opposite the deoxyadenosine derivative, the DNA strands containing the fluorescent base, 2-amino purine (P), were examined. The 2-amino purine base can form a base pair with thymine by forming two hydrogen bonds in the same geometry as the Watson–Crick base pairs, where the fluorescence emission is efficiently quenched by the base stacking (32).

The fluorescence emission spectra were measured for the duplexes containing 2-amino purine opposite A, T and **X**, in addition to the 2-amino purine nucleotide forced into a single bulge and that in a single strand state (Figure 4). Figure 5A demonstrates the fluorescence emission spectra of 2-amino purine adjacent to two cytosines, 5′-GACAC<u>P</u>CACAC-3′. The fluorescence emission around 370 nm was efficiently quenched when thymine or adenine was located opposite the 2-amino purine, resulting from base stacking. In contrast, the duplex containing the **X**/P pair demonstrated a significant fluorescence emission, comparable with the 2-amino purine nucleotide forced into a bulge and that in a single strand state. This observation suggests that the 2-amino purine opposite **X** adopts an unstacked conformation similar to the single bulge nucleotide and the nucleotide in a single-strand DNA.

Table 1. Thermodynamic parameters for the DNA duplex formations with different W_1/W_2 pairs[a]

Sequence	ΔH^0 (kcal mol^{-1})	ΔS^0 (cal mol^{-1}K^{-1})	ΔG^0_{37} (kcal mol^{-1})	T_m (°C)
A-Series				
5′-GTGTCACTGTC-3′ 3′-CACAGAGACAG-5′	−62.2 ± 2.2	−174 ± 5	−8.3 ± 0.1	45.9
5′-GTGTCACTGTC-3′ 3′-CACAGGGACAG-5′	−69.7 ± 1.4	−192 ± 4	−10.2 ± 0.3	54.7
5′-GTGTCACTGTC-3′ 3′-CACAGCGACAG-5′	−72.4 ± 1.9	−203 ± 3	−8.9 ± 0.2	48.8
5′-GTGTCACTGTC-3′ 3′-CACAGTGACAG-5′	−89.0 ± 1.4	−248 ± 4	−12.6 ± 0.1	57.4
5′-GTGTCACTGTC-3′ 3′-CACAGFGACAG-5′	−69.5 ± 2.0	−199 ± 5	−7.4 ± 0.2	42.9
X-series				
5′-GTGTC**X**CTGTC-3′ 3′-CACAGAGACAG-5′	−57.5 ± 2.1	−157 ± 5	−8.8 ± 0.1	50.4
5′-GTGTC**X**CTGTC-3′ 3′-CACAGGGACAG-5′	−56.6 ± 2.5	−153 ± 7	−9.2 ± 0.4	52.6
5′-GTGTC**X**CTGTC-3′ 3′-CACAGCGACAG-5′	−73.1 ± 2.4	−202 ± 9	−10.7 ± 0.5	55.2
5′-GTGTC**X**CTGTC-3′ 3′-CACAGTGACAG-5′	−72.8 ± 3.0	−200 ± 8	−10.5 ± 0.3	55.8
5′-GTGTC**X**CTGTC-3′ 3′-CACAGFGACAG-5′	−73.4 ± 2.2	−203 ± 5	−10.4 ± 0.4	54.9
Z-series				
5′-GTGTC**Z**CTGTC-3′ 3′-CACAGAGACAG-5′	−67.8 ± 1.8	−184 ± 5	−10.5 ± 0.1	57.3
5′-GTGTC**Z**CTGTC-3′ 3′-CACAGGGACAG-5′	−63.5 ± 1.5	−172 ± 5	−10.0 ± 0.2	54.8
5′-GTGTC**Z**CTGTC-3′ 3′-CACAGCGACAG-5′	−84.7 ± 2.1	−232 ± 6	−12.8 ± 0.2	60.9
5′-GTGTC**Z**CTGTC-3′ 3′-CACAGTGACAG-5′	−80.2 ± 1.7	−219 ± 7	−12.3 ± 0.6	60.9
5′-GTGTC**Z**CTGTC-3′ 3′-CACAGFGACAG-5′	−76.5 ± 2.1	−209 ± 5	−11.5 ± 0.2	60.2

[a]All experiments were carried out in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA at pH 7.0. Underline indicates the nucleotides forming the W_1/W_2 pair. T_m was calculated at the total strand concentration of 100 μM.

Table 2. Thermodynamic parameters for the DNA duplex formations with different adjacent base pairs[a]

Sequence	ΔH^0 (kcal mol^{-1})	ΔS^0 (cal mol^{-1}K^{-1})	ΔG^0_{37} (kcal mol^{-1})	$\Delta\Delta G^0_{37}$ (kcal mol^{-1})[b]	T_m (°C)
5′-GTGTAAATGTC-3′ 3′-CACATATACAG-5′	−55.9 ± 1.9	−160 ± 4	−6.6 ± 0.1	0	35.5
5′-GTGTA**X**ATGTC-3′ 3′-CACATATACAG-5′	−61.1 ± 1.6	−170 ± 4	−8.3 ± 0.1	−1.7	46.9
5′-GTGTGAGTGTC-3′ 3′-CACACACACAG-5′	−73.4 ± 2.4	−208 ± 8	−9.1 ± 0.1	0	47.4
5′-GTGTG**X**GTGTC-3′ 3′-CACAC**A**CACAG-5′	−69.4 ± 4.6	−189 ± 13	−10.6 ± 0.2	−1.5	56.3
5′-GTGTCACTGTC-3′ 3′-CACAGAGACAG-5′	−62.2 ± 2.2	−174 ± 5	−8.3 ± 0.1	0	45.9[c]
5′-GTGTC**X**CTGTC-3′ 3′-CACAGAGACAG-5′	−57.5 ± 2.1	−157 ± 5	−8.8 ± 0.1	−0.5	50.4[c]
5′-GTGTTATTGTC-3′ 3′-CACAAAAACAG-5′	−66.0 ± 2.2	−188 ± 8	−6.9 ± 0.3	0	41.9
5′-GTGTT**X**TTGTC-3′ 3′-CACAAAAACAG-5′	−62.9 ± 1.7	−175 ± 6	−8.0 ± 0.2	−1.1	47.0

[a]All experiments were carried out in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA at pH 7.0. Underline indicates the nucleotides forming the W_1/W_2 pair. T_m was calculated at the total strand concentration of 100 μM.
[b]$\Delta\Delta G^0_{37}$ indicates the difference between the ΔG^0_{37} values of the duplexes containing **X**/W_2 and A/W_2 with the same adjacent base pairs.
[c]The data are derived from Table 1.

Figure 5B shows the fluorescence emission spectra of 2-amino purine adjacent to two guanines, 5′-GACAGPGACAC-3′. Unlike the DNA strand containing 2-amino purine adjacent to two cytosines, the fluorescence emission of the single strand was quenched, probably by partial stacking on the adjacent guanines. Nevertheless, the facts that the emission was efficiently quenched when thymine or adenine was opposite the 2-amino purine and that the duplex containing the **X**/P pair showed a large fluorescence emission comparable with those of the 2-amino purine nucleotide forced into a single

bulge also suggest that the 2-amino purine assumes an unstacked form when locating opposite **X**.

DISCUSSION

Conformation of the DNA duplex containing the deoxyadenosine derivative

We have synthesized the deoxyadenosine derivatives tethering a phenyl group (**X**) and a naphthyl group (**Z**) to deoxyadenosine by means of an amide linker. We have previously reported that the single **X** and **Z** nucleoside at the 5′ ends of a self-complementary DNA duplex increased the duplex stability by 1.8–3.3 kcal mol^{-1} in ΔG^0_{37} which was equal to or more than the formation of Watson–Crick A/T base pairs (17). This observation is suggestive of a strong stacking interaction of the aromatic hydrocarbon group at a helix terminus and adoption of a geometry analogous to the Watson–Crick base pairs. Here, we investigated the structure and thermodynamic parameters of the duplexes containing **X** and **Z** in the middle of a DNA sequence in order to estimate the stacking interaction by the base pair-mimic nucleosides in a duplex.

The conformation of single mismatches in a DNA duplex has been investigated in previous studies: the geometry of the A/G mismatch in a DNA duplex is context-dependent, and it can form a base pair with two hydrogen bonds, differing from Watson–Crick base pairs, which may also interconvert to other pairing geometries (25). The A/A mismatches have a high stacking potential and can form a base pair with one hydrogen bond in anti–anti conformation by disrupting the backbone structure (24,27). The A/C mismatches are unlikely to form a stable base pair at neutral pH, although a pair may be formed with one hydrogen bond depending on the sequence (26). The abasic site perturbs the canonical helix structure and induces additional flexibility (28). Therefore, it can be concluded that most of the differences in the CD spectra in Figure 2A result from the differences in the A/W_2 pair geometry. The existence of different interactions among the A-series duplexes is also suggested by the thermodynamic parameters in Table 1 that demonstrate different ΔH^0, ΔS^0 and ΔG^0_{37} values.

Contrary to the observations for the A-series duplexes, the CD spectra in Figure 2B and C suggests that all the X-series or Z-series duplexes adopt a similar geometry at the W_1/W_2 pair. The CD spectrum of the fully matched duplex (W_1 = A and W_2 = T) was similar to those of the X-series duplexes and the Z-series duplexes. A relatively large negative peak around 250 nm of the Z-series duplexes, probably partly originating from the naphthyl group, implies that the naphthyl group adopts a stacked conformation regardless of the opposite nucleotide species. One probable explanation for the similarity in the duplex conformations containing **X** or **Z** regardless of the type of the W_2 nucleotide is that the aromatic hydrocarbon group intercalates into the double helix and causes the W_2 nucleotide base to flip into an unstacked position.

The fluorescence emission spectra in Figure 5 indicate that the environment of the 2-amino purine opposite **X** differs considerably from that opposite adenine or thymine. The fluorescence emission intensity of the **X**/P pair was comparable with that of the 2-amino purine nucleotide forced into a single bulge and that in a single-strand DNA. Therefore, it is likely that the **X** opposite 2-amino purine causes the nucleotide base to flip into an unstacked position.

In our previous study, we demonstrated that the duplexes of 5′-**X**TGCGCA-3′ and 5′-**Z**TGCGCA-3′ showed a higher stability (−10.5 and −10.9 kcal mol^{-1}, respectively) than those of 5′-**X**TGCGCAT-3′ and 5′-**Z**TGCGCAT-3′ with terminal **X**/T and **Z**/T pairs (−9.8 and −9.9 kcal mol^{-1}, respectively), implying that the deoxyadenosine derivatives did not form a base pair with the opposing thymine (17). We have also found that the duplexes of an RNA strand associated with the DNA strand containing **X** showed a similar thermal stability regardless of the ribonucleotide species (A, G, C or U) opposite **X**

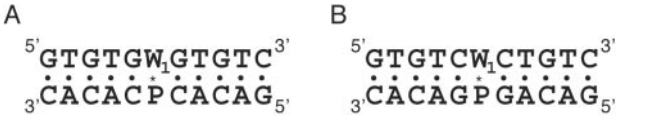


Figure 4. The duplexes containing the 2-amino purine nucleotide (P) opposite W_1 (A, T, X or •), forming the A/P, T/P or X/P pair or a single P-bulge, respectively. The 2-amino purine nucleotide is adjacent to (**A**) two cytosines or (**B**) two guanines.

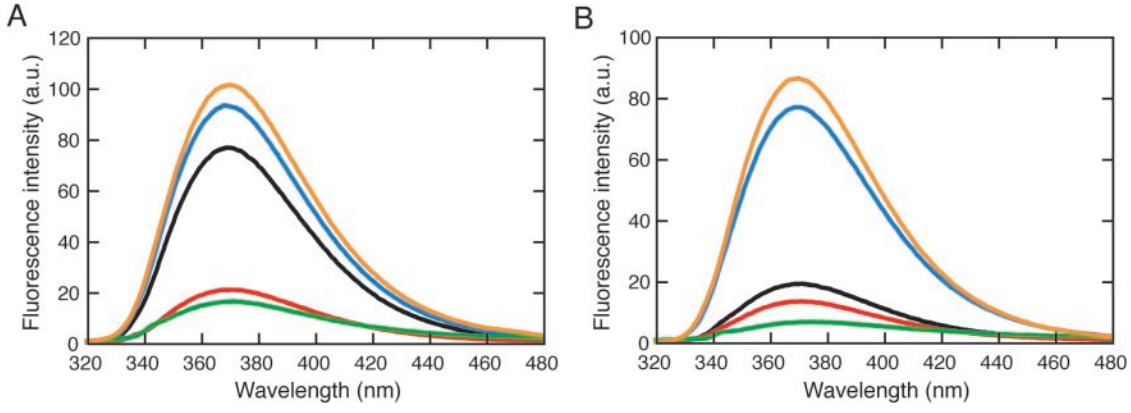


Figure 5. Fluorescence emission spectra of the duplexes containing 2-amino purine adjacent to two cytosines (**A**) or two guanines (**B**), opposite A (red), T (green) or **X** (blue) in a duplex, in addition to a duplex forming a 2-amino purine bulge (orange) and the single-stranded DNA (black). All measurements were carried out in the 1 M NaCl-phosphate buffer at 4°C.

and that the phosphorodiester bond of the RNA strand on the 3′-side of the ribonucleotides opposite **X** was prone to be hydrolysed (33). These observations also support that **X** induces unstacking of the opposite base regardless of the nucleotide species.

Conformation of the duplexes from the standpoint of the energetics

As suggested by the comparison of the T_m values in Figure 3, the thermodynamic parameters in Table 1 indicate that the ΔG^0_{37} value of the A-series duplexes is in the range from -7.4 to -12.6 kcal mol^{-1} in ΔG^0_{37}, which is a substantially larger range than those of the X-series and the Z-series duplexes (-8.8 to -10.7 kcal mol^{-1} and -10.0 to -12.8 kcal mol^{-1}, respectively). In the A-series duplexes, the duplex stability containing the A/G mismatch is greater than those containing the A/C and A/A mismatches, in agreement with the nearest-neighbor parameters for the single mismatch formations determined previously (25–27). Although incorporation of a single abasic site in the duplex has a strong destabilization effect (5.2 kcal mol^{-1}), the stability is partially restored by the displacement of A by **X** (3.0 kcal mol^{-1}) or **Z** (4.1 kcal mol^{-1}) opposite the abasic site. Because the duplex containing F provides a space unoccupied by a nucleotide base, the aromatic hydrocarbon group of **X** and **Z** opposite the abasic site can intercalate between the adjacent base pairs, by adopting the amide linker geometry suitable for the aromatic hydrocarbon group stacking in the helix.

In Table 1, the ΔG^0_{37} values of the duplexes containing the **X**/C and **X**/T mismatches (-10.7 and -10.5 kcal mol^{-1}) are similar to that containing the **X**/F pair (-10.4 kcal mol^{-1}). Furthermore, the ΔH^0 and ΔS^0 values of these duplexes are identical within experimental uncertainty (-72.8 to -73.4 kcal mol^{-1} and -200 to -203 cal mol^{-1}K^{-1}, respectively), strongly suggesting that the interactions in the duplexes containing the **X**/pyrimidine pair are similar to those in the duplex containing **X**/F. This observation also implies that the pyrimidine base unstacking involves lower energy costs and that the rigidity of the duplexes is similar. On the other hand, the thermodynamic parameters of the duplexes containing the **X**/purine pair are similar (-57.5 and -56.6 kcal mol^{-1} in ΔH^0, -157 and -153 cal mol^{-1} in ΔS^0 and -8.8 and -9.2 kcal mol^{-1} in ΔG^0_{37}). The same tendencies are also found in the Z-series duplexes. It is also demonstrated that the stability of the duplexes containing **X** or **Z** paired with a pyrimidine base was greater than those paired with a purine base and the greater stabilization observed for the duplexes containing the deoxyadenosine/pyrimidine pair results from the ΔH^0 term. This observation suggests the existence of a greater penalty for the unstacking of a purine base than a pyrimidine base in duplex formation, probably because the purine bases can stack with adjacent nucleotide bases more efficiently than the pyrimidine bases.

The penalty for the nucleotide base unstacking is also implicated in Table 2. The $\Delta\Delta G^0_{37}$ value in Table 2 that reflects the stabilization energy by the displacement of A at W_1 by **X** is context-dependent, and greater stabilization is observed for A**X**A/TAT and G**X**G/CAC (-1.7 and -1.5 kcal mol^{-1}, respectively), while those for C**X**C/GAG and T**X**T/AAA are relatively small (-0.5 and -1.1 kcal mol^{-1}, respectively).

Table 3. Thermodynamic parameters for the trinucleotide formations[a]

Sequence	ΔH^0 (kcal mol^{-1})	ΔS^0 (cal mol^{-1}K^{-1})	ΔG^0_{37} (kcal mol^{-1})
5′-A**X**A-3′ 3′-TAT-5′	3.9	15	−0.5
5′-G**X**G-3′ 3′-CAC-5′	1.0	9	−1.4
5′-C**X**C-3′ 3′-GAG-5′	10.7	34	0.2
5′-T**X**T-3′ 3′-AAA-5′	5.9	20	0.4

[a]The parameters were calculated on the basis of the nearest-neighbor model using the thermodynamic parameters of the Watson–Crick base pairs (19,20,29).

These results indicate that the adenine base unstacking next to two adjacent purine (Pu) bases (Py**X**Py/PuAPu) may require greater energy than that next to two adjacent pyrimidine (Py) bases (Pu**X**Pu/PyAPy). Table 3 showing the thermodynamic parameters for the trinucleotide formation also indicates lower stabilization energy for the adenine next to two adjacent purine bases, accompanied by larger positive ΔH^0 and ΔS^0 values. Strong stacking of a purine base next to two adjacent purine bases and less stacking of a purine base next to two adjacent pyrimidine bases were also demonstrated in the fluorescence emission spectra in Figure 5.

Energetics for the duplex formations

Figure 6 compares the ΔG^0_{37} values for the A-series, X-series and Z-series duplex formations. By the displacement of A at W_1 by **X**, the stability of the duplexes with A, C and F at W_2 increased, while those with G and T decreased. Stabilization of the duplexes containing A and C at W_2 results from the ΔS^0 term, consistent with the intercalation of **X** leading to bury its aromatic hydrocarbon group accompanied by dehydration and/or base unstacking leading to a less constrained conformation. The ΔS^0-driven stabilization was also observed for the DNA duplexes containing **X** and **Z** at the 5′ dangling ends (17). In contrast, stabilization of the duplex containing the abasic site analogue is achieved by the ΔH^0 term, which is in agreement with the occupancy of the abasic site by efficient stacking of the phenyl group of **X**. On the other hand, destabilization of the duplexes with G and T at W_2 results from the loss of enthalpy energy, partly from the loss of the base pair hydrogen bonds by the replacement of A by **X** at W_1. The displacement of **X** by **Z** stabilized all the duplexes, and the stabilization originates from the ΔH^0 term, probably owing to greater stacking interaction by the naphthyl group. It is probable that **Z** is more efficient in the base flipping because of incorporation of the larger aromatic group into the duplex, consistent with more similarity in the CD spectra of the Z-series duplexes (Figure 2C) than those of the X-series duplexes (Figure 2B). In addition, the stacking of the aromatic group may also be partly driven by the release of water molecules leading to favorable ΔS^0 contributions, as suggested for the adenosine derivatives as a 5′ dangling end (17) and this could be more significant for the larger aromatic group.

Since it is concluded that the duplexes containing **X**/T and **Z**/T adopt a conformation similar to that of other types of duplexes, it is probable that the deoxyadenosine derivatives

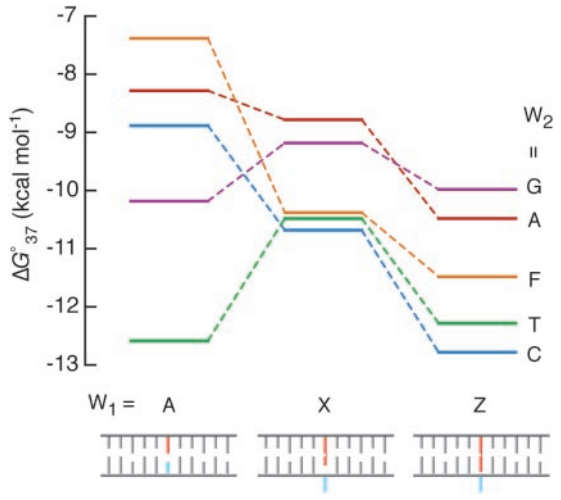


Figure 6. Comparison of the ΔG^0_{37} values for the DNA duplex formations. A model of the base flipping is also indicated, containing the W_1 (red) and the W_2 (blue) nucleotides.

do not form a base pair with thymine. We have also previously found for the RNA/DNA duplexes containing **X** that the phosphorodiester bond of an RNA strand on the 3′-side of the ribonucleotides (U as well as A, G, C) opposite **X** in the DNA strand was hydrolysed with similar efficiency (33). This finding implies that the stacking of the aromatic hydrocarbon group of **X** and **Z** is sufficiently strong to disrupt the base pair formation. The ability of the deoxyadenosine derivatives to stack in a duplex is surprising because only the stacking interaction can compensate for the loss of the energy for nucleotide base stacking and interstrand hydrogen bonds. However, it has been reported that a porphyrin with a copper atom binds the DNA duplex of $(CGATCG)_2$ by intercalation between cytosine and guanosine accompanied by extrusion of the terminal cytosine from the helix (34), and that quinoline at the 5′ end of the duplex of $(TGCGCA)_2$ disrupts the terminal T/A base pair formation and stacks on the G/C base pair next to the terminal A/T pair (35). These observations indicate the disruption of the Watson–Crick base pair by interacting with aromatic ring molecules that can efficiently stack with the nucleotide bases. Furthermore, a biascridine macrocyle (BisA) is able to recognize DNA-base mismatches and induce base flipping by the insertion of BisA into the DNA helix displacing the mispaired thymine into an extrahelical position (36). Polycyclic aromatic DNA-base surrogates, such as 1-pyrenyl, 1-napthyl, 2-naphthyl, acenaphthyl and 4-biphenyl residues, have been synthesized and found to disrupt the base stacking of the opposite strand (37,38). It has also been reported that the pyrene nucleotide in a DNA strand opposite uracil acts as a wedge to push the uracil from the base stack and restore the catalytic activity of mutant uracil-DNA glycosylases (39–41). Like the polycyclic aromatic nucleosides, the non-hydrogen-bonding Watson–Crick base pair analogues of **X** and **Z**, although these nucleosides do not have a large contiguous π-system but have an amide linker that covalently connects deoxyadenosine with a phenyl group or a naphthyl group, would be useful to flip the target base into an extrahelical position, and the lack of the pairing selectivity would have an advantage for the site-selective base flipping in a target sequence. The base flipping can be used for molecular biology and biotechnology, such as a selective cleavage of a target nucleotide and protein associations that recognize an unpaired nucleotide.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

This work was supported in part by Grants-in-Aid for Scientific Research and 'Academic Frontier' Project (2004–2009) from MEXT, Japan. Funding to pay the Open Access publication charges for this article was provided by MEXT, Japan.

Conflict of interest statement. None declared.

REFERENCES

1. Turner,D.H., Sugimoto,N., Kierzek,R. and Dreiker,S.D. (1987) Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.*, **109**, 3783–3784.
2. Bommarito,S.V., Peyret,N. and SantaLucia,J.Jr (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.*, **28**, 1929–1934.
3. Moody,E.M. and Bevilacqua,P.C. (2003) Folding of a stable DNA motif involves a highly cooperative network of interactions. *J. Am. Chem. Soc.*, **125**, 16285–16293.
4. Morales,J.C. and Kool,E.T. (2000) Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. *Biochemistry*, **39**, 12979–12988.
5. Kool,E.T., Morales,J.C. and Guckian,K.M. (2000) Mimicking the structure and function of DNA: insights into DNA stability and replication. *Angew. Chem. Int. Ed. Engl.*, **39**, 990–1099.
6. Hendrickson,C.L., Devine,K.G. and Benner,S.A. (2004) Probing minor groove recognition contacts by DNA polymerases and reverse transcriptases using 3-deaza-2′-deoxyadenosine. *Nucleic Acids Res.*, **32**, 2241–2250.
7. Berger,M., Luzzi,A., Henry,A.A. and Romesberg,F.E. (2002) Stability and selectivity of unnatural DNA with five-membered-ring nucleobase analogues. *J. Am. Chem. Soc.*, **124**, 1222–1226.
8. Liu,H., Gao,J., Maynard,L., Saito,D. and Kool,E.T. (2004) Toward a new genetic system with expanded dimensions: size-expanded analogues of deoxyadenosine and thymidine. *J. Am. Chem. Soc.*, **126**, 1102–1009.
9. Arkin,M.R., Stemp,E.D., Pulver,S.C. and Barton,J.K. (1997) Long-range oxidation of guanine by Ru(III) in duplex DNA. *Chem. Biol.*, **4**, 389–400.
10. Kuzuya,A., Mizoguchi,R., Morisawa,F., Machida,K. and Komiyama,M. (2002) Metal ion-induced site-selective RNA hydrolysis by use of acridine-bearing oligonucleotide as cofactor. *J. Am. Chem. Soc.*, **124**, 6887–6894.
11. Lermer,L., Roupioz,Y., Ting,R. and Perrin,D.M. (2002) Toward an RNaseA mimic: a DNAzyme with imidazoles and cationic amines. *J. Am. Chem. Soc.*, **124**, 9960–9961.
12. Lewis,F.D., Liu,J., Weigel,W., Rettig,W., Kumikov,I.V. and Beratan,D.N. (2002) Donor-bridge-acceptor energetics determine the distance dependence of electron tunneling in DNA. *Proc. Natl Acad. Sci. USA*, **99**, 12536–12541.
13. Tanaka,K., Tengeiji,A., Kato,T., Toyama,N. and Shionoya,M. (2003) A discrete self-assembled metal array in artificial DNA. *Science*, **299**, 1212–1213.
14. Liu,H., Gao,J., Lynch,S.R., Saito,Y.D., Maynard,L. and Kool,E.T. (2003) A four-base paired genetic helix with expanded size. *Science*, **302**, 868–871.
15. Amann,N., Huber,R. and Wagenknecht,H.A. (2004) Phenanthridinium as an artificial base and charge donor in DNA. *Angew. Chem. Int. Ed. Engl.*, **43**, 1845–1847.

16. Okamoto,A., Ichiba,T. and Saito,I. (2004) Pyrene-labeled oligodeoxynucleotide probe for detecting base insertion by excimer fluorescence emission. *J. Am. Chem. Soc.*, **126**, 8364–8365.
17. Nakano,S., Uotani,Y., Nakashima,S., Anno,Y., Fujii,M. and Sugimoto,N. (2003) Large stabilization of a DNA duplex by the deoxyadenosine derivatives tethering an aromatic hydrocarbon group. *J. Am. Chem. Soc.*, **125**, 8086–8087.
18. Kierzek,R., Caruthers,M.H., Longfellow,C.E., Swinton,D., Turner,D.H. and Freier,S.M. (1986) Polymer-supported RNA synthesis and its application to test the nearest-neighbor model for duplex stability. *Biochemistry*, **25**, 7840–7846.
19. Sugimoto,N., Nakano,S., Yoneyama,M. and Honda,K. (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.*, **24**, 4501–4505.
20. Nakano,S., Kanzaki,T. and Sugimoto,N. (2004) Influences of ribonucleotide on a duplex conformation and its thermal stability: study with the chimeric RNA–DNA strands. *J. Am. Chem. Soc.*, **126**, 1088–1095.
21. Richards,E.G. (1975) Use of tables in calculation of absorption, optical rotatory dispersion and circular dichroism of polyribonucleotides. In Fasman,G.D. (ed.), *Handbook of Biochemistry and Molecular Biology, 3rd edn*. CRC Press, Cleveland, OH, Vol. 1, pp. 596–603.
22. Marky,L.A. and Breslauer,K.J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, **26**, 1601–1620.
23. Puglisi,J.D. and Tinoco,I.,Jr (1989) Absorbance melting curves of RNA. *Methods Enzymol.*, **180**, 304–325.
24. Maskos,K., Gunn,B.M., LeBlanc,D.A. and Morden,K.M. (1993) NMR study of G.A and A.A pairing in $(dGCGAATAAGCG)_2$. *Biochemistry*, **32**, 3583–3595.
25. Allawi,H.T. and SantaLucia,J.,Jr (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry*, **37**, 2170–2179.
26. Allawi,H.T. and SantaLucia,J.,Jr (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry*, **37**, 9435–9444.
27. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J.,Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry*, **38**, 3468–3477.
28. Barsky,D., Foloppe,N., Ahmadia,S., Wilson,D.M.III. and MacKerell,A.D.,Jr (2000) New insights into the structure of abasic DNA from molecular dynamics simulations. *Nucleic Acids Res.*, **28**, 2613–2626.
29. Sugimoto,N., Nakano,M. and Nakano,S. (2000) Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. *Biochemistry*, **39**, 11270–11281.
30. Gelfand,C.A., Plum,G.E., Grollman,A.P., Johnson,F. and Breslauer,K.J. (1998) Thermodynamic consequences of an abasic lesion in duplex DNA are strongly dependent on base sequence. *Biochemistry*, **37**, 7321–7327.
31. Sági,J., Guliaev,A.B. and Singer,B. (2001) 15-mer DNA duplexes containing an abasic site are thermodynamically more stable with adjacent purines than with pyrimidines. *Biochemistry*, **40**, 3859–3868.
32. Pompizi,I., Häberli,A. and Leumann,C.J. (2000) Oligodeoxynucleotides containing conformationally constrained abasic sites: a UV and fluorescence spectroscopic investigation on duplex stability and structure. *Nucleic Acids Res.*, **28**, 2702–2708.
33. Nakano,S., Uotani,Y., Uenishi,K., Fujii,M. and Sugimoto,N. (2005) Site-selective RNA cleavage by DNA bearing a base pair-mimic nucleoside. *J. Am. Chem. Soc.*, **127**, 518–519.
34. Lipscomb,L.A., Zhou,F.X., Presnell,S.R., Woo,R.J., Peek,M.E., Plaskon,R.P. and Williams,L.D. (1996) Structure of DNA-porphyrin complex. *Biochemistry*, **35**, 2818–2823.
35. Tuma,J., Connors,W.H., Stitelman,D.H. and Richert,C. (2002) On the effect of covalently appended quinolones on termini of DNA duplexes. *J. Am. Chem. Soc.*, **124**, 4236–4246.
36. David,A., Bleimling,N., Beuck,C., Lehn,J.-M., Weinhold,E. and Teulade-Fichou,M.-P. (2003) DNA mismatch-specific base flipping by a bisacridine macrocycle. *Chembiochem.*, **4**, 1326–1331.
37. Singh,I., Hecker,W., Prasad,A.K., Parmar,V.S. and Seitz,O. (2002) Local disruption of DNA-base stacking by bulky base surrogates. *Chem. Commun.*, **7**, 500–501.
38. Beuck,C., Singh,I., Bhattacharya,A., Hecker,W., Parmar,V.S., Seitz,O. and Weinhold,E. (2003) Polycyclic aromatic DNA-base surrogates: high-affinity binding to an adenine-specific base-flipping DNA methyltransferase. *Angew. Chem. Int. Ed. Engl.*, **42**, 3958–3960.
39. Jiang,Y.L., Kwon,K. and Stivers,J.T. (2001) Turning On uracil-DNA glycosylase using a pyrene nucleotide switch. *J. Biol. Chem.*, **276**, 42347–42354.
40. Jiang,Y.L., Stivers,J.T. and Song,F. (2002) Base-flipping mutations of uracil DNA glycosylase: substrate rescue using a pyrene nucleotide wedge. *Biochemistry*, **41**, 11248–11254.
41. Kwon,K., Jiang,Y.L. and Stivers,J.T. (2003) Rational engineering of a DNA glycosylase specific for an unnatural cytosine:pyrene base pair. *Chem. Biol.*, **10**, 351–359.