# Stabilization Factors Affecting Duplex Formation of Peptide Nucleic Acid with DNA†

Naoki Sugimoto,*,‡,§ Naonori Satoh,‡ Kyohko Yasuda,‡ and Shu-ichi Nakano‡

*Department of Chemistry, Faculty of Science and Engineering, and High Technology Research Center, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan*

*Received March 9, 2001; Revised Manuscript Received May 9, 2001*

ABSTRACT: Peptide nucleic acid (PNA) is an oligonucleotide analogue in which the sugar−phosphate backbone is replaced by an *N*-(2-aminoethyl)glycine unit to which the nucleobases are attached. We investigated the thermodynamic behavior of PNA/DNA hybrid duplexes with identical nearest neighbors but with different sequences and chain lengths (5, 6, 7, 8, 10, 12, and 16 mers) to reveal whether the nearest-neighbor model is valid for the PNA/DNA duplex stability. CD spectra of 6, 7, and 8 mer PNA/DNA duplexes showed similar signal, while 10, 12, and 16 mer duplexes did not. The average difference in $\Delta G^\circ_{37}$ for short PNA/DNA duplexes with identical nearest-neighbor pairs was only 3.5%, whereas that of longer duplexes (10, 12, and 16 mers) was 16.4%. Therefore, the nearest-neighbor model seems to be useful at least for the short PNA/DNA duplexes. Thermodynamics of PNA/DNA duplexes containing 1−3 bulge residues were also studied. While the stability of the 12 mer DNA/DNA duplex decreased as the number of bulge bases increases, the number of bulge bases in PNA/DNA unchanged the duplex stability. Thus, the influence of bulge insertion in the PNA/DNA duplexes is different from that of a DNA/DNA duplex. This might be due to the different base geometry in a helix which may potentially make hydrogen bonds in a base pair and stacking interaction unfavorable compared with DNA/DNA duplexes.

Peptide nucleic acid (PNA) is an oligonucleotide analogue in which the sugar−phosphate backbone is replaced by an *N*-(2-aminoethyl)glycine unit to which the nucleobases are attached (*1*). PNA has neither charges nor chirality centers and is designed to recognize the complementary sequence of RNA and DNA through Watson−Crick base pairs (PNA/RNA or PNA/DNA hybrid duplex) (*2*). PNA is also known to associate with a complementary PNA (PNA/PNA duplex) (*3*). It is intriguing to measure thermodynamics of PNA/DNA and PNA/RNA duplexes to understand the interactions that influence the stability of PNA-containing duplexes as well as its stability. It has been reported that the stability of PNA/RNA and PNA/DNA duplexes shows a higher melting temperature ($T_m$) under a low ionic strength condition than the RNA/RNA or RNA/DNA duplex does (*2*, *4*). This observation implies that cations do not have an important role for PNA-containing duplex stability.

It is also important to establish a method that can reproduce behaviors of PNA/DNA duplexes. Two methods have been developed for the stability prediction for PNA/DNA duplexes. Griffin and Smith have reported a first approach to predict PNA/DNA duplex stability (*5*). They introduced the following free energy terms which are required for PNA/DNA duplex formation: (a) a nearest-neighbor interaction, (b) an initiation factor, (c) a dangling end, (d) a PNA/DNA stabilization per nearest-neighbor interaction, and (e) an ionic strength effect. Parameters of (a) and (b) are approximated with values determined for DNA/DNA duplexes (*6*), and parameters (c), (d), and (e) are empirically determined. Gissen and co-workers developed a method to predict $T_m$[1] for the PNA/DNA duplex (*7*). Unfortunately, both models are based on the parameters determined for DNA/DNA duplexes. Although the DNA parameters are determined in 1 M NaCl, which is supposed to be saturated, it is still unclear whether the nature of the duplex formation is the same or not.

To understand the nature of PNA-containing duplex formation, we focused on the nearest-neighbor model which is useful for DNA/DNA, RNA/RNA, and RNA/DNA duplexes. Thermodynamics of DNA/DNA, RNA/RNA, and RNA/DNA have been well investigated, and these duplex stabilities can be predicted on the basis of a nearest-neighbor model with high accuracy (*8*−*13*). The nearest-neighbor model is consistent with the fact that the major interactions in a nucleic acid duplex formation are hydrogen bonds in a base pair and stacking interaction between nearest-neighbor bases. However, there is no report whether the nearest-neighbor model is also valid for PNA-containing duplexes

† This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture, Japan, and a grant from the "Research for the Future" Program of the Japan Society for the Promotion of Science to N.S.

* To whom correspondence should be addressed at the Department of Chemistry, Faculty of Science, Konan University. Phone: +81-78-435-2497. Fax: +81-78-435-2539. E-mail: sugimoto@konan-u.ac.jp.

‡ Department of Chemistry, Faculty of Science and Engineering, Konan University.

§ High Technology Research Center, Konan University.

[1] Abbreviations: UV, ultraviolet; CD, circular dichroism; HPLC, high-performance liquid chromatography; $T_m$, melting temperature; $Na_2$-EDTA, ethylenediamine-*N*,*N*,*N*′,*N*′-tetraacetic acid disodium salt.

10.1021/bi010480m CCC: $20.00 
Published on Web 06/28/2001

which may have different interactions to be considered. If the nearest-neighbor model is also useful for PNA/DNA duplexes, it would be helpful to suppose interactions in PNA/DNA duplexes as well as to establish a new method to predict PNA/DNA duplex stability.

In this study, we investigated the applicability of the nearest-neighbor model for the stability of PNA/DNA duplexes. We measured CD spectra and thermodynamics of 5, 6, 7, 8, 10, 12, and 16 mer PNA/DNA duplexes; each pair has identical nearest-neighbors but with different sequences. Also, 12 and 16 mer PNA/DNA duplexes are designed to have 1−3 bulge residues between 8 and 4 base pair stems and between 8 base pair stems, respectively, to evaluate the number of base pairs in the stem necessary to form Watson−Crick base pairs.

## MATERIALS AND METHODS

*Materials.* PNAs were synthesized with a Boc strategy on a solid support synthesis as described (*14*). The Boc-BHA-PEG-PS resin [*tert*-butyloxycarbonylbenzhydrylamine−poly(ethylene glycol) handle−polystyrene resin, PE Biosystems] was treated with *m*-cresol/trifluoroacetic acid (TFA) (5/95 v/v) to remove the Boc group for the coupling. The Boc-PNA monomer (5 equiv) was then added with *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) (4.5 equiv) and *N*,*N*-diisopropylethylamine (DIPEA) (10 equiv) as a coupling activator. After coupling, 5% acetic anhydride and 5% piperidine were added, and then the next couplings were continued. Deprotection of benzyloxycarbonyl (Z group) at A, C, and G nucleobases was done with *m*-cresol/thioanisole/trifluoromethanesulfonic acid (TFMSA)/TFA (1/1/1/7 v/v). The final purity of the synthesized PNAs was confirmed on a YMC C18 reverse-phase column by HPLC (>92%). The mobile phase consisted of liquid A (0.1% TFA in water) and liquid B (0.08% TFA in acetonitrile). MALDI-TOF analysis was used on a Voyager-DE mass spectrometer (Perseptive Biosystems Inc.) to confirm the molecular weight of the PNA strand using a sinnapinic acid matrix. Single strand concentrations of the PNA strands were considered to be the same as those for DNA.

DNA oligonucleotides were synthesized chemically on a solid support using the phosphoramidite method on an Applied Biosystems Model 391 DNA synthesizer. The synthesized DNA oligonucleotides were purified with HPLC after the deblocking operations (*15*). These oligomers were further purified and desalted with a C-18 Sep-Pak cartridge column. The final purity of the oligomers was confirmed to be greater than 98% by HPLC. Single strand concentrations of the oligonucleotides were measured at 260 nm and high temperature (*16*), and single strand extinction coefficients were calculated with mononulceotide and dinucleotide data using a nearest-neighbor approximation (*17*). A PNA strand and its complementary DNA strand were mixed at a concentration ratio of 1:1 to obtain PNA/DNA duplexes.

*CD Measurements.* CD spectra were obtained on a JASCO J-600 spectropolarimeter equipped with a temperature controller. The experimental temperature was 5.0 °C. The cuvette-holding chamber was flushed with a constant stream of dry $N_2$ gas to avoid water condensation on the cuvette exterior. All CD spectra were measured from 320 to 200 nm in a 0.1 cm path-length cuvette.

*UV Measurements.* Absorbance was measured with Hitachi U-3200 and U-3210 spectrophotometers equipped with Hitachi SPR-7 and SPR-10 thermoprogrammers. The water condensation on the cuvette exterior at a low-temperature range was avoided by flushing with a constant stream of dry $N_2$ gas. The heating rate was 0.5 or 0.3 °C/min. UV melting curves were measured in a NaCl−phosphate buffer containing 1 M NaCl, 10 mM $Na_2HPO_4$, and 1 mM $Na_2$-EDTA adjusted to pH 7.0. Prior to the experiment, all buffers were degassed by heating to 90 °C for 10 min.

*Determination of Thermodynamic Parameters for Duplex Formations.* All melting curves were fitted with a theoretical equation to obtain the thermodynamic parameters for double helix formation ($\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$) as described elsewhere (*8*, *10*, *18*). We also evaluated these thermodynamics from $T_m^{-1}$ vs $\log(C_t/4)$ plots. From the slope and intercept of the inverse of the melting temperature, the thermodynamic parameters were deduced using the equations:

$$T_m^{-1} = 2.303R \log(C_t/4)/\Delta H° + \Delta S°/\Delta H° \quad (1)$$

$$\Delta G°_{37} = \Delta H° - 310.15\Delta S° \quad (2)$$

Here, $R$ is the gas constant and $C_t$ is the total strand concentration.

Errors for these thermodynamics ($\sigma_{\Delta H°}$, $\sigma_{\Delta S°}$, and $\sigma_{\Delta G°_{37}}$) in a curve-fitting procedure were estimated as standard deviations from melting curves measured at different $C_t$s (at least eight different concentrations in the range of 0.8−80 $\mu$M). The $\sigma_{\Delta H°}$ and $\sigma_{\Delta S°}$ values from the $T_m^{-1}$ vs $\log(C_t/4)$ method were estimated from the linearity of the plots, and $\sigma_{\Delta G°_{37}}$ was calculated by the equation (*12*, *13*, *19*):

$$(\sigma_{\Delta G°_{37}})^2 = (\sigma_{\Delta H°})^2 + 310.15^2(\sigma_{\Delta S°})^2 - 2[310.15(R_{\Delta H°,\sigma_{\Delta S°}})\sigma_{\Delta H°}\sigma_{\Delta S°}] \quad (3)$$

Here, $R_{\Delta H°,\sigma_{\Delta S°}}$ is the correlation coefficient between $\Delta H°$ and $\Delta S°$.

## RESULTS

*Structure of PNA/DNA Duplexes.* Circular dichroism (CD) spectroscopy provides information of the global structure of nucleotides. It is known that CD spectra of a PNA/DNA and a nucleic acid duplex are similar, consistent with a right-handed duplex for PNA/DNA duplexes (*2*), although the geometry of bases is supposed to be different from DNA duplexes.

The CD spectra of pna(CCGACG)/d(CGTCGG) (**6a**) and pna(CGACCG)/d(CGGTCG) (**6b**) which have identical nearest-neighbor base pairs are shown in Figure 1a. Both spectra had a similar shape with an intense positive peak around 265 nm and a negative peak around 245 nm, analogous to reported A-form spectra of DNA duplexes (*20*). Likewise, PNA/DNA duplexes pairs with identical nearest-neighbor pairs but different sequences of pna(CGTGGC)/d(GCCACG) (**6c**) and pna(CGGTGC)/d(GCACCG) (**6d**), pna(GCGACCG)/d(CGGTCGC) (**7a**) and pna(GACCGCG)/d(CGCGGTC) (**7b**), pna(GCGTTCC)/d(GGAACGC) (**7c**) and pna(GTTCGCC)/d(GGCGAAC) (**7d**), pna(CTCACGGC)/d(GCCGTGAG) (**8a**) and pna(CACGGCTC)/d(GAGCCGTG) (**8b**), and pna(ACGTACCG)/d(CGGTACGT) (**8c**) and

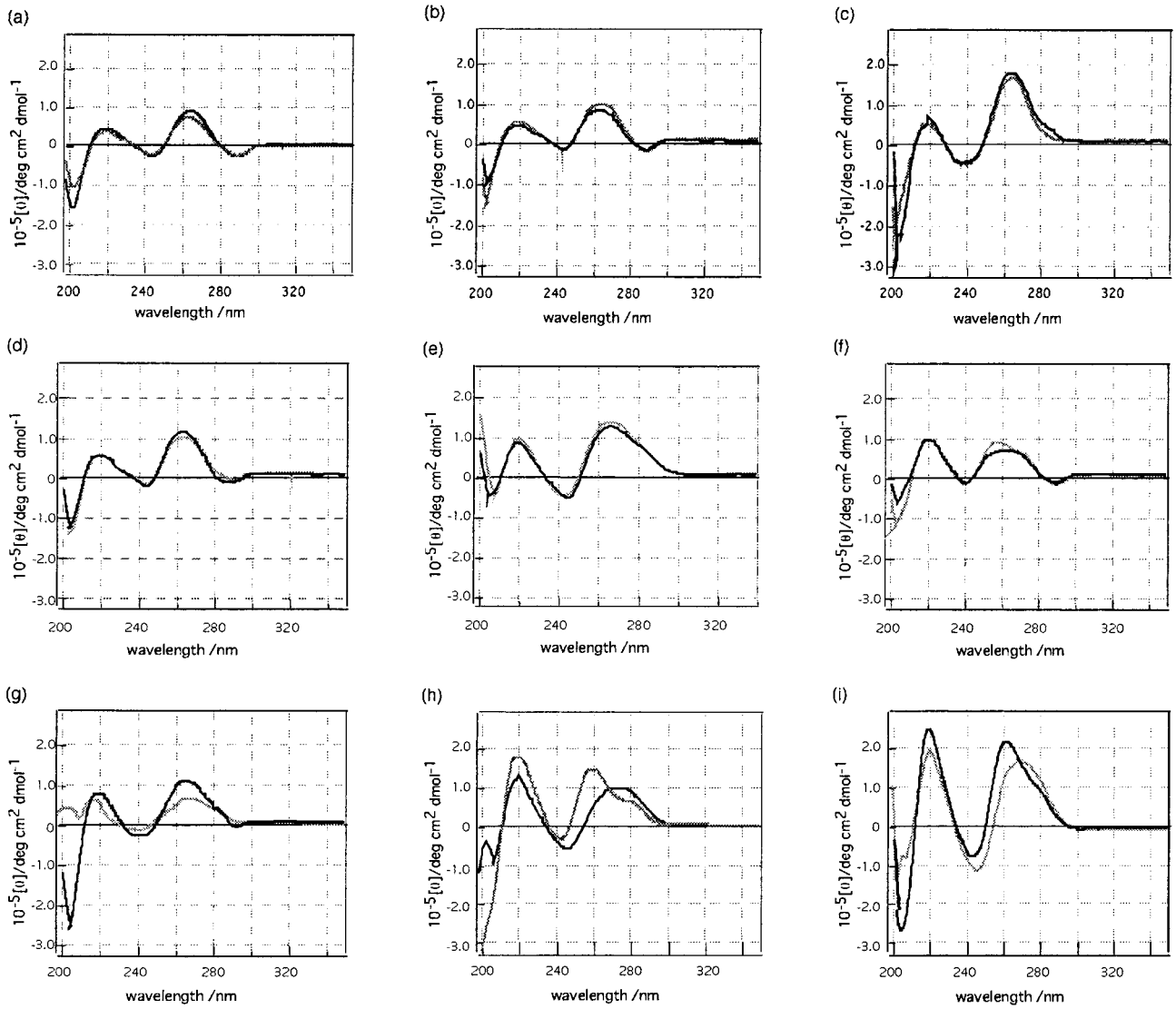


FIGURE 1: CD spectra of (a) pna(CCGACG)/d(CGTCGG) (**6a**, black line) and pna(CGACCG)/d(CGGTCG) (**6b**, gray line), (b) pna(CGTGGC)/d(GCCACG) (**6c**, black line) and pna(CGGTGC)/d(GCACCG) (**6d**, gray line), (c) pna(GCGACCG)/d(CGGTCGC) (**7a**, black line) and pna(GACCGCG)/d(CGCGGTC) (**7b**, gray line), (d) pna(GCGTTCC)/d(GGAACGC) (**7c**, black line) and pna(GTTCGCC)/d(GGCGAAC) (**7d**, gray line), (e) pna(CTCACGGC)/d(GCCGTGAG) (**8a**, black line) and pna(CACGGCTC)/d(GAGCCGTG) (**8b**, gray line), (f) pna(ACGTACCG)/d(CGGTACGT) (**8c**, gray line) and pna(ACCGTACG)/d(CGTACGGT) (**8d**, black line), (g) pna(GCTAACAGCG)/d(CGCTGTTAGC) (**10a**, black line) and pna(GCGCTACAAG)/d(CTTGTAGCGC) (**10b**, gray line), (h) pna(ATTGGATACAAA)/d(TTTGTATCCAAT) (**12a**, black line) and pna(CATTGGAATAAC)/d(GTTATTCCAATG) (**12b**, gray line), and (i) pna(ATAAATTGGATACAAA)/d(TTTGTATCCAATTTAT) (**16a**, black line) and pna(CAAATGGATTAAATAC)/d(GTATTTAATCCATTTG) (**16b**, gray line). The concentration of these samples was 70 $\mu$M, and measurements were done in 1 M NaCl−phosphate buffer (pH 7.0) at 5.0 °C.

pna(ACCGTACG)/d(CGTACGGT) (**8d**) showed similar CD spectra (Figure 1a−f). However, duplex pairs of pna(GCTAACAGCG)/d(CGCTGTTAGC) (**10a**) and pna(GCGCTACAAG)/d(CTTGTAGCGC) (**10b**), pna(ATTGGATACAAA)/d(TTTGTATCCAAT) (**12a**) and pna(CATTGGAATAAC)/d(GTTATTCCAATG) (**12b**), and pna(ATAAATTGGATACAAA)/d(TTTGTATCCAATTTAT) (**16a**) and pna(CAAATGGATTAAATAC)/d(GTATTTAATCCATTTG) (**16b**) revealed different CD spectra (Figure 1g−i). These results indicate that PNA/DNA duplexes which have identical nearest-neighbor pairs seem to have a similar duplex structure, although long duplexes do not.

*Melting Behavior of PNA/DNA Duplexes.* Figure 2a compares the melting curves of pna(CCGACG)/d(CGTCGG) (**6a**) and pna(CGACCG)/d(CGGTCG) (**6b**) at a same concentration. These melting curves were almost identical and the melting temperatures ($T_m$) were close, 30.9 and 29.3 °C for **6a** and **6b**, respectively. This result indicates these 6 mer duplexes possess similar thermostability. To evaluate the thermodynamics for these duplex formation, $T_m^{-1}$ vs log($C_t$/4) plots for **6a** and **6b** were also compared, as seen in Figure 2b. The plots for these duplexes were almost identical, and the differences of their thermodynamic values ($\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$) calculated with eqs 1 and 2 were only 2.9%, 3.9%, and 0.5%, respectively.

*Differences in Thermodynamic Values of PNA/DNA Duplexes with Identical Nearest-Neighbor Base Pairs.* We compared thermodynamics for 10 PNA/DNA duplex pairs (20 hybrid duplexs); each pair has identical nearest neighbors but with different sequences. Thermodynamic values ($\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$) and $T_m$s for these PNA/DNA duplexes are summarized in Table 1. Thermodynamics were obtained with the van't Hoff method and a curve-fitting method (*8*, *10*−*13*). Thermodynamics determined by these methods agreed

Table 1: Thermodynamic Parameters of PNA/DNA Hybrid Duplexes with Identical Nearest-Neighbor Base Pairs[a]

| | | log($C_t$) method | | | | curve fitting method | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| no. | sequence[b] | $-\Delta H°$/ kcal mol$^{-1}$ | $-\Delta S°$/ cal mol$^{-1}$ K$^{-1}$ | $-\Delta G°_{37}$/ kcal mol$^{-1}$ | $T_m$[c]/ °C | $-\Delta H°$/ kcal mol$^{-1}$ | $-\Delta S°$/ cal mol$^{-1}$ K$^{-1}$ | $-\Delta G°_{37}$/ kcal mol$^{-1}$ | $T_m$[c]/ °C | $T_m$[d]/ °C | $T_m$[e]/ °C |
| **5a** | pna(AGCGG)/d | 38.1 ± 1.7 | 102 ± 5 | 6.39 ± 0.4 | 35.9 | 35.1 ± 0.9 | 91.7 ± 4.1 | 6.65 ± 0.7 | 38.1 | 41.0 | 26.0 |
| **5b** | pna(AGGCG)/d | 35.8 ± 3.4 | 94.8 ± 9.4 | 6.41 ± 0.7 | 36.0 | 36.8 ± 2.0 | 98.5 ± 6.7 | 6.26 ± 0.1 | 34.8 | 41.0 | 26.0 |
| **6a** | pna(CCGACG)/d | 33.8 ± 1.5 | 88.0 ± 4.1 | 6.58 ± 0.3 | 37.5 | 36.4 ± 2.3 | 96.0 ± 7.8 | 6.67 ± 0.1 | 38.2 | 56.0 | 32.5 |
| **6b** | pna(CGACCG)/d | 32.8 ± 1.3 | 84.2 ± 3.4 | 6.67 ± 0.3 | 38.3 | 35.4 ± 1.6 | 92.9 ± 4.9 | 6.64 ± 0.1 | 38.0 | 56.0 | 32.5 |
| **6c** | pna(CGTGGC)/d | 52.5 ± 3.6 | 147 ± 10 | 6.87 ± 0.6 | 39.0 | 50.3 ± 4.3 | 140 ± 14 | 6.87 ± 0.2 | 39.6 | 54.9 | 33.2 |
| **6d** | pna(CGGTGC)/d | 45.4 ± 3.4 | 124 ± 10 | 7.11 ± 0.6 | 41.0 | 49.6 ± 2.6 | 134 ± 8 | 7.14 ± 0.1 | 41.8 | 54.9 | 33.2 |
| **7a** | pna(GCGACCG)/d | 46.9 ± 2.6 | 125 ± 7 | 8.19 ± 6.7 | 48.4 | 48.5 ± 4.3 | 130 ± 14 | 8.35 ± 0.2 | 49.1 | 72.9 | 47.2 |
| **7b** | pna(GACCGCG)/d | 48.3 ± 2.6 | 128 ± 7 | 8.68 ± 0.6 | 51.5 | 51.6 ± 1.7 | 138 ± 6 | 8.87 ± 0.1 | 51.7 | 72.9 | 47.2 |
| **7c** | pna(GCGTTCC)/d | 39.2 ± 1.7 | 108 ± 5 | 5.93 ± 0.3 | 32.4 | 42.0 ± 2.2 | 117 ± 8 | 5.72 ± 0.2 | 31.2 | 63.6 | 34.5 |
| **7d** | pna(GTTCGCC)/d | 38.9 ± 3.9 | 103 ± 11 | 5.07 ± 0.8 | 32.6 | 40.9 ± 4.4 | 113 ± 15 | 5.77 ± 0.3 | 31.4 | 63.6 | 34.5 |
| **8a** | pna(CTCACGGC)/d | 61.3 ± 1.8 | 168 ± 5 | 9.11 ± 0.3 | 51.2 | 64.2 ± 5.6 | 177 ± 18 | 9.16 ± 0.3 | 51.2 | 74.0 | 44.3 |
| **8b** | pna(CACGGCTC)/d | 59.6 ± 1.5 | 163 ± 4 | 9.07 ± 0.4 | 50.8 | 55.7 ± 5.3 | 150 ± 17 | 9.07 ± 0.2 | 51.8 | 74.0 | 44.3 |
| **8c** | pna(ACGTACCG)/d | 46.2 ± 5.3 | 123 ± 15 | 7.96 ± 1.3 | 46.9 | 46.3 ± 2.2 | 124 ± 7 | 8.06 ± 0.3 | 47.6 | 66.2 | 43.3 |
| **10a** | pna(GCTAACAGCG)/d | 36.3 ± 2.2 | 87.5 ± 5.5 | 9.15 ± 0.6 | 61.2 | 37.5 ± 4.0 | 91.9 ± 12 | 9.05 ± 0.4 | 59.3 | 81.2 | 56.4 |
| **10b** | pna(GCGCTACAAG)/d | 66.8 ± 2.8 | 180 ± 8 | 11.1 ± 0.5 | 59.6 | 72.6 ± 2.4 | 198 ± 7 | 11.4 ± 0.4 | 59.4 | 78.4 | 53.6 |
| **12a** | pna(ATTGGATACAAA)/d | 74.1 ± 5.0 | 201 ± 14 | 11.5 ± 0.9 | 59.2 | 72.7 ± 2.2 | 197 ± 7 | 11.5 ± 0.2 | 60.0 | 81.7 | 54.5 |
| **12b** | pna(CATTGGAATAAC)/d | 50.6 ± 4.4 | 129 ± 11 | 10.6 ± 0.8 | 63.9 | 49.0 ± 1.8 | 124 ± 7 | 10.6 ± 0.4 | 61.7 | 81.7 | 53.0 |
| **16a** | pna(ATAAATTGGATACAAA)/d | 74.5 ± 1.8 | 197 ± 5 | 13.3 ± 0.6 | 68.0 | 76.7 ± 4.1 | 204 ± 13 | 13.5 ± 0.3 | 67.3 | 92.6 | 62.9 |
| **16b** | pna(CAAATGGATTAAATAC)/d non-two-state transition | 44.6 ± 3.2 | 108 ± 8 | 11.0 ± 0.7 | 71.4 | 44.9 ± 3.8 | 109 ± 12 | 11.1 ± 0.4 | 70.7 | 92.6 | 62.2 |
| **8d** | pna(ACCGTACG)/d | 61.6 ± 2.4 | 169 ± 7 | 9.27 ± 0.5 | 51.4 | 36.9 ± 2.1 | 92.0 ± 6.7 | 8.41 ± 0.4 | 53.6 | 66.2 | 43.3 |

[a] All experiments were done in a buffer containing 1 M NaCl/10 mM $Na_2HPO_4$/1 mM $Na_2EDTA$ (pH 7.0). [b] A hybrid duplex consists of the denoted PNA strand and its complementary DNA strand. [c] Melting temperatures were calculated at the total strand concentration of 100 $\mu$M. [d] Melting temperatures were calculated by using Griffin parameters (*5*). [e] Melting temperatures were calculated by using Gissen parameters (*7*).

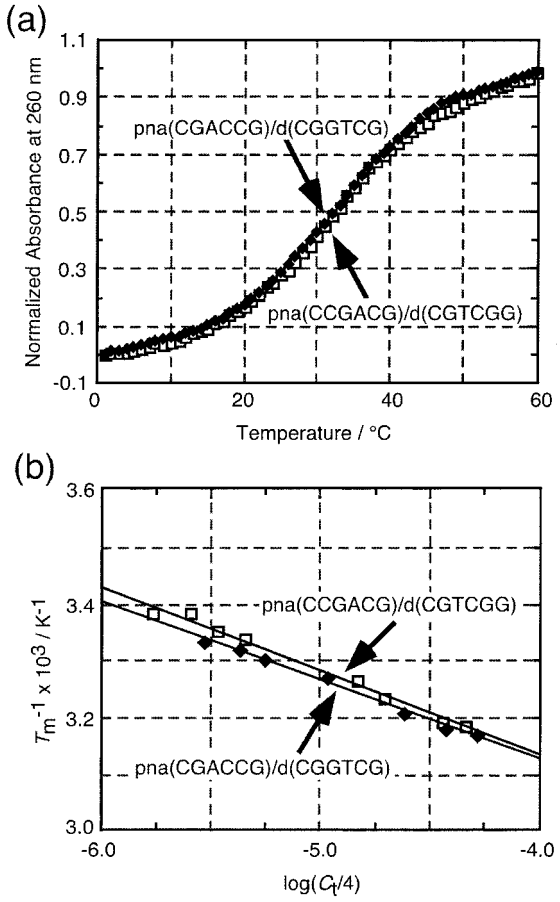


FIGURE 2: (a) Normalized melting curves and (b) $T_m^{-1}$ vs log($C_t$/4) plots of pna(CGACCG)/d(CGGTCG) (closed diamonds) and pna(CCGACG)/d(CGTCGG) (open squares) in 1 M NaCl−phosphate buffer (pH 7.0). The concentration of these samples at the above melting curves was 8 $\mu$M.

within 10%, consistent with a two-state transition from a duplex to two single strands. However, thermodynamics for **8d** obtained by the van't Hoff method and the curve-fitting method were quite different (40% for $\Delta H°$, 46% for $\Delta S°$, and 12% for $\Delta G°_{37}$ in difference). This result implies that the melting of **8d** was not a two-state transition.

Differences in $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ of PNA/DNA duplexes with identical nearest-neighbor pairs are shown in Figure 3, in which the values of **8c** and **8d** are omitted because of a non-two-state transition for **8d**. For the short duplex pairs of 5, 6, 7, and 8 mers, the differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$ were relatively small; the average differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$ were 4.5%, 5.6%, 3.5%, and 2.8 °C, respectively. These differences are similar to the values observed for RNA/DNA duplexes (3.9%, 4.5%, 4.1%, and 1.7 °C in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$, respectively) (*21*). In contrast, the average differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$ for the 10, 12, and 16 mer PNA/DNA duplexes were significant (54.5%, 64.9%, 16.4%, and 3.7 °C, respectively). Thus, the nearest-neighbor model might not be useful for the stability predictions for long PNA/DNA duplexes.

*Effect of Chain Length on PNA/DNA Duplex Formation.* We also measured 12 and 16 mer PNA/DNA duplexes containing bulge residues. The sequence and obtained thermodynamics are listed in Table 2. The 16 mer PNA/DNA duplex (**16p**) has an extra four bases at N and 3′-terminal of the 12 mer PNA/DNA duplex of **12p**. Surprisingly, thermodynamic parameters of **12p** and **16p** were not so different (2.9% for $\Delta H°$, 0.7% for $\Delta S°$, and 14.2% for $\Delta G°$ in difference), while those of DNA/DNA duplexes of **12d** and **16d** were quite different (33.8% for $\Delta H°$, 34.3% for $\Delta S°$, and 29.8% for $\Delta G°$ in difference).

In Figure 4, thermodynamic parameters of **12p**, **12p1**, **12p2**, and **12p3** are compared with the corresponding DNA/DNA duplexes of **12d**, **12d1**, **12d2**, and **12d3**. The stability of these DNA/DNA duplexes decreased with increasing number of bulge bases (Figure 4a), while the stabilities of these PNA/DNA duplexes were similar regardless of the

Table 2: Comparison with Thermodynamic Parameters between DNA/DNA and PNA/DNA Duplexes with/without T Bulge Bases[a]

| sequence[b] (3′-5′)/(5′-3′) or (C-N)/(5′-3′) | *n* | no. | log($C_t$) method $-\Delta H°$/ kcal mol$^{-1}$ | log($C_t$) method $-\Delta S°$/ cal mol$^{-1}$ K$^{-1}$ | log($C_t$) method $-\Delta G°_{37}$/ kcal mol$^{-1}$ | log($C_t$) method $T_m$[c]/ °C | curve fitting method $-\Delta H°$/ kcal mol$^{-1}$ | curve fitting method $-\Delta S°$/ cal mol$^{-1}$ K$^{-1}$ | curve fitting method $-\Delta G°_{37}$/ kcal mol$^{-1}$ | curve fitting method $T_m$[c]/ °C |
|---|---|---|---|---|---|---|---|---|---|---|
| d(AAACATAGGTTA)/ | 0 | **12d** | 71.4 ± 1.2 | 197 ± 15 | 10.3 ± 0.2 | 54.4 | 76.7 ± 5.0 | 214 ± 16 | 10.2 ± 0.3 | 54.6 |
| d(TTTGTATC**T**$_n$CAAT) | 1 | **12d1** | 60.6 ± 2.2 | 174 ± 7 | 6.51 ± 0.21 | 36.9 | 51.0 ± 2.4 | 143 ± 8 | 6.64 ± 0.15 | 40.1 |
| | 2 | **12d2** | 40.6 ± 5.7 | 111 ± 16 | 6.23 ± 0.77 | 34.7 | 43.4 ± 3.8 | 120 ± 12 | 6.13 ± 0.20 | 34.9 |
| | 3 | **12d3** | 29.3 ± 2.5 | 75.0 ± 6.6 | 6.06 ± 0.48 | 32.1 | 30.5 ± 2.1 | 78.9 ± 7.0 | 6.00 ± 0.52 | 32.0 |
| d(AAACATAGGTTAAATA)/ d(TTTGTATCCAATTTAT) | | **16d** | 94.2 ± 5.9 | 261 ± 17 | 13.0 ± 1.3 | 60.0 | 104 ± 12 | 291 ± 36 | 13.6 ± 0.7 | 60.0 |
| pna(AAACATAGGTTA)/ | 0 | **12p** | 74.1 ± 5.0 | 201 ± 14 | 11.5 ± 0.9 | 59.2 | 72.7 ± 2.2 | 197 ± 7 | 11.5 ± 0.2 | 60.0 |
| d(TTTGTATC**T**$_n$CAAT) | 1 | **12p1** | 45.1 ± 2.0 | 121 ± 66 | 7.55 ± 0.35 | 44.2 | 47.9 ± 4.0 | 130 ± 13 | 7.61 ± 0.11 | 43.5 |
| | 2 | **12p2** | 40.0 ± 2.2 | 104 ± 6 | 7.73 ± 0.45 | 44.6 | 41.3 ± 4.0 | 102 ± 16 | 7.72 ± 0.22 | 43.6 |
| | 3 | **12p3** | 36.7 ± 2.5 | 92.9 ± 6.5 | 7.86 ± 0.60 | 48.7 | 34.6 ± 1.7 | 86.2 ± 5.4 | 7.90 ± 0.89 | 46.3 |
| pna(AAACATAGGTTAAATA)/ d(TTTGTATCCAATTTAT) | | **16p** | 74.5 ± 1.8 | 197 ± 5 | 13.3 ± 0.6 | 68.0 | 76.7 ± 4.1 | 204 ± 13 | 13.5 ± 0.3 | 67.3 |

[a] All experiments were done in a buffer containing 1 M NaCl/10 mM $Na_2HPO_4$/1 mM $Na_2EDTA$ (pH 7.0). [b] The boldface character indicates the bulge nucleotide. [c] Melting temperatures were calculated at the total strand concentration of 100 $\mu$M.

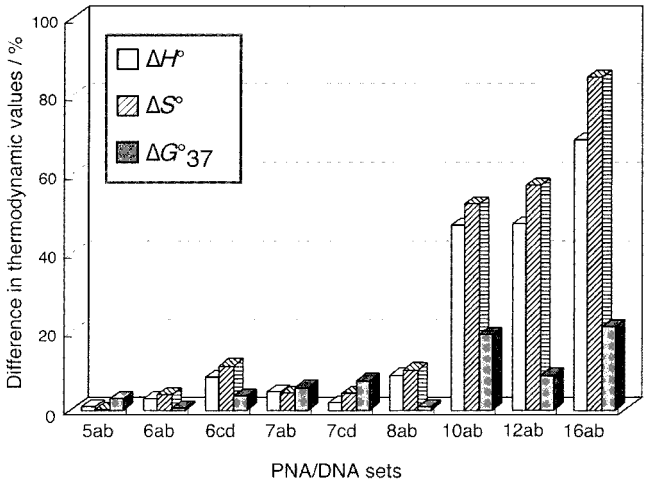


FIGURE 3: Differences in thermodynamic values ($\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$) of PNA/DNA duplexes with identical nearest-neighbor base pairs but different sequences. These sequences are indicated in Table 1.

bulge bases (Figure 4b). It is possible that since 12 mer PNA/DNA duplexes potentially have only four Watson−Crick base pairs at N and 3′-terminal followed by the bulge bases, **12p1**, **12p2**, and **12p3** do not form a bulge structure. This result is consistent with the fact that short PNA/DNA duplexes of pna(CGG)/d(CCG), pna(TAA)/d(TTA), pna(GTAA)/d(TTAC), and pna(AGTAA)/d(TTACT) did not show melting curve by temperature or CD spectra assigned to be a duplex structure (Figure S1).

## DISCUSSION

*Duplex Formation of Short PNA/DNA Duplexes.* We have investigated thermodynamics of PNA/DNA duplexes. As shown in Figure 1, CD spectra of 6, 7, and 8 mer PNA/DNA duplex pairs suggest that the global conformation of these PNA/DNA duplexes with identical nearest-neighbor base pairs was very similar. The average differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$ for the PNA/DNA duplexes of 5, 6, 7, and 8 mers were only 4.5%, 5.6%, 3.5%, and 2.8 °C, respectively. Thus, the nearest-neighbor model seems to be useful for the stability predictions of the short PNA/DNA duplexes.

The PNA/DNA duplexes of pna(CGG)/d(CCG), pna(TAA)/d(TTA), pna(GTAA)/d(TTAC), and pna(AGTAA)/d(TTACT) did not form a stable duplex structure under the experimental conditions examined here. In contrast, 5 mer PNA/DNA duplexes of **5a** and **5b** and 6 mer PNA/DNA duplexes of **6a** and **6b** showed a transition from a duplex to single strands. Thus, it is likely that a tetramer or pentamer would be required to form a stable duplex with a complementary DNA under the experimental condition, although its sequence and the experimental conditions could affect it. This idea is consistent with the result that **12p1**, **12p2**, and **12p3** did not form a bulge structure, so that the PNA/DNA duplexes with bulge nucleotide may have fraying terminal ends rather than forming a duplex containing bulge nucleotides.

*Duplex Formation of Long PNA/DNA Duplexes.* As shown in Figure 3, the differences in $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and $T_m$ for the PNA/DNA duplexes (10, 12, and 16 mers) were 54.5%, 64.9%, 16.4%, and 3.7 °C, respectively, indicating that the nearest-neighbor model may be improper for these long PNA/DNA duplexes. However, since the differences in thermodynamics and $T_m$ of **8a** and **8b** were relatively small (8.8% for $\Delta H°$, 10.2% for $\Delta S°$, 0.7% for $\Delta G°_{37}$, and 0.2 °C for $T_m$), the octamer could be applied by the stability prediction with the nearest-neighbor model, although it depends on the sequence and experimental conditions.

It is reported for a 10 mer PNA/DNA duplex that the enthalpy and free energy changes in the duplex formation obtained by DSC (differential scanning calorimetry) and ITC (isothermal titration calorimetry) were different from those determined by the UV melting method (*22*, *23*). This observation suggests that a PNA strand could easily form a self-structure in a single stand state. It has been indicated that a single stranded PNA has an intramolecular hydrogen bond at the backbone by molecular modeling (*24*, *25*), although NMR studies do not support it (*26*, *27*). If PNA strands form intramolecular hydrogen bonds in the single strand state which may be needed to be broken when it associates with a complementary strand, since the energy level of the single strand state of a PNA strand could affect the apparent thermodynamics, especially for duplex formations with long strands, so that it may be necessary for predicting the stability of the long PNA/DNA duplex to consider the single strand state of PNA. Unfortunately, since most of single stranded PNA used in Table 1 did not show two-state transition (Figure S2), we could not estimate the

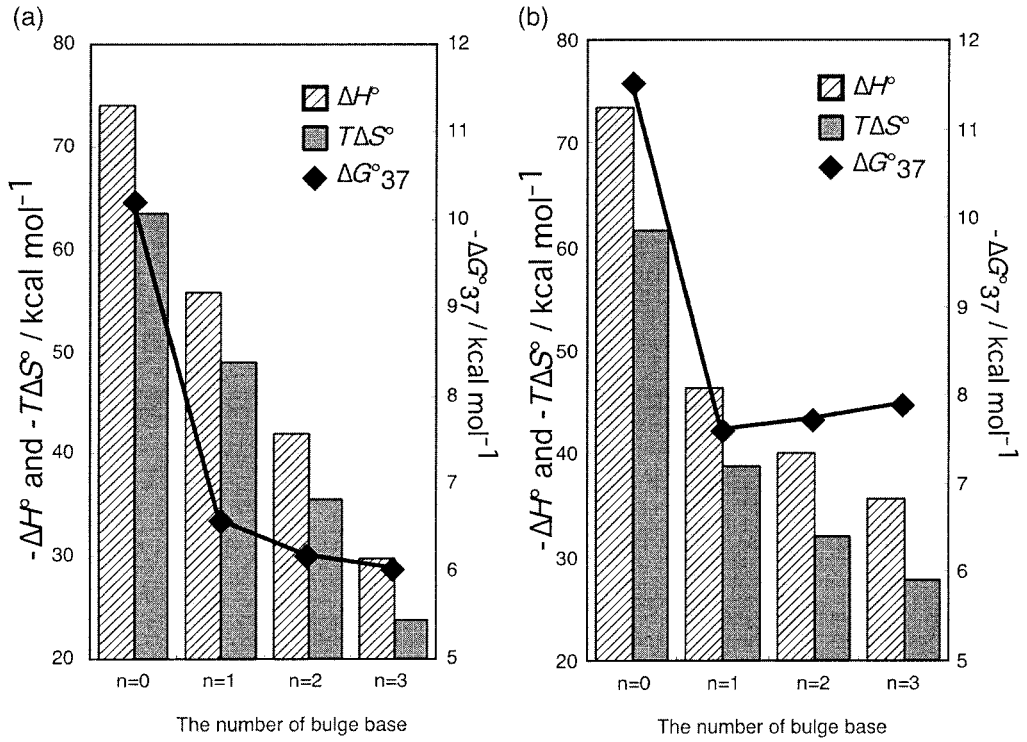


FIGURE 4: Thermodynamics for $\Delta H°$, $T\Delta S°$, and $\Delta G°_{37}$ of (a) d(AAACATAGGTTAA)/d(TAAC**T**$_n$CTATGTTT) and (b) pna-(AAACATAGGTTAA)/d(TAAC**T**$_n$CTATGTTT) in 1 M NaCl−phosphate buffer (pH 7.0). The boldface character indicates the bugle base, and *n* means the number of bulge bases.

energy level of these PNA strands. Only the single strand PNA of pna(ACCGTACG) showed a two-state melting curve (data not shown), and the melting temperature of the PNA was dependent on its concentration, indicating the PNA strand formed a self-complementary structure by itself. The self-complementary structure would influence the observed thermodynamics, showing that the melting of **8d** was not a two-state transition and showing different thermodynamics compared with those of **8c**.

*Effect of Bulge on Duplex Formation of PNA with DNA.* We also focused on the influence of bulge bases on a PNA/DNA duplex formation. A bulge loop is formed in double helical nucleic acids when the helix is interrupted by an unpaired nucleotide on only one strand. The bulge is one of the important structure elements in nucleic acids, and the features of bulge are similar in RNA and DNA. It is reported that a DNA/DNA duplex containing a single bulge base is 3.5−4.6 kcal mol$^{-1}$ less stable than a decamer of a DNA/DNA duplex of d(GCGAAAAGCG)/d(CGCTTTTCGC) regardless of the identity of the bulge nucleotide (*28*). It is also reported that an RNA/RNA duplex of (GCG**B**$_n$GCG)/(CGCCGC), where **B** is A or U, is destabilized with increment of bulge size, and this destabilization is caused by unfavorable enthalpy and favorable entropy changes (*29*). These results are consistent with the results in Figure 4a. The DNA/DNA duplex was destabilized by the single bulge insertion at about 3.6 kcal mol$^{-1}$ in $\Delta G°_{37}$. Moreover, the DNA/DNA duplex was destabilized with increment of bulge size caused by unfavorable enthalpy and favorable entropy changes. The unfavorable enthalpy change came from distortion of the continuous stacking interaction around the bulge nucleotide, and it led to increased freedom around the bulge nucleotide accompanied by favorable entropy changes. The 12 mer PNA/DNA duplex was also destabilized by 3.9 kcal mol$^{-1}$ in $\Delta G°_{37}$ when a single T bulge was inserted in the duplex. Nevertheless, the destabilization was independent of the bulge size as shown in Figure 4b, although an increment of bulge size is expected to induce unfavorable enthalpy and favorable entropy changes in PNA/DNA duplexes as well as DNA/DNA duplexes. This result indicates that **12p1**, **12p2**, and **12p3** might not form a bulge structure; these PNA/DNA may have fraying ends without bulge, which could cause the observed stability independent of the bulge size.

*Structure of PNA/DNA Duplexes.* The detailed structure of PNA/PNA and PNA/DNA duplexes has been solved by NMR and X-ray crystallography (*26*, *27*, *30*). A PNA/PNA duplex has a very wide and deep major groove and a very narrow and shallow minor groove, and the duplex has a very large pitch of 18 base pairs per turn and a large pitch height (57.6 Å) (*30*). A canonical B-form helix seen for DNA/DNA duplexes has a pitch of 10 base pairs per turn and 34 Å of pitch height (*31*), whereas a PNA/DNA duplex has a pitch of 13 base pairs per turn and 42 Å of pitch height (*27*). Because base pairs in a PNA/DNA duplex possess a different geometry compared with nucleic acid duplexes, the strength of the stacking interaction of PNA/DNA duplexes is expected to be different from that of DNA/DNA duplexes (*26*). Interestingly, CD spectra of the 10, 12, and 16 mer PNA/DNA duplexes suggest different base configuration for these duplex pairs (Figure 1). This could influence the thermodynamics observed for long PNA/DNA duplexes.

It is known for DNAs that a duplex structure that has A/T runs deviates from a canonical B-form because the hydration pattern in the minor groove at an A/T run is unusual (*32*). We have designed some duplex pairs to have the same A/T runs, such as pairs of **7c** and **7d** and **10a** and **10b** have one $A_2/T_2$ run and a pair of **16a** and **16b** has one $A_2/T_2$ and two $A_3/T_3$ runs. It is thus expected that there is the same hydration

pattern for each pair. However, what we observed was the same CD spectrum for **7c** and **7d** and different CD spectrum for longer duplex pairs. Therefore, the observation that longer PNA/DNA pairs showed different CD spectra would not be only due to the presence of A/T runs in these duplexes, although the details of the hydration depending on the A/T run sequence are remained to be addressed.

*Stability Predictions by Published Methods.* Griffin and Smith (*5*) and Gissen et al. (*7*) has developed parameters to predict the stability of PNA/DNA duplexes. They estimate the stability of PNA/DNA duplexes by using the nearest-neighbor parameters determined for DNA/DNA duplexes (*6*). Although possible nearest-neighbor pairs for the DNA/DNA duplex are only 10, 16 nearest-neighbor pairs are required to deal with PNA/DNA pairs such as RNA/DNA hybrid duplexes (*10*). Griffin and Smith assumed the free energy of the PNA/DNA duplex formation with the sum of five parameter terms: (i) a nearest-neighbor interaction term, (ii) an initiation term, (iii) a dangling end stabilization term, (iv) a PNA/DNA stabilization term, and (v) an ionic strength term (*5*). Their method can be applied to a broad salt concentration range. However, this method can not be applied for sequences containing three or more consecutive guanines or cytosines (*11*). Parameters of (iii), (iv), and (v) were determined with one sequence, and two parameters of (i) and (ii) were regarded as the same as DNA/DNA duplexes. They also used enthalpy energy determined by DNA/DNA nearest-neighbor parameters; yet the stability of the PNA/DNA duplex does not come so much from a more negative binding enthalpy but from a relatively smaller change in the binding entropy (*22*). Nielsen and co-workers calculated the thermostability using an experimental equation (*7*). Their model is based on the $T_m$ calculated for the corresponding DNA/DNA duplexes employing the nearest-neighbor approach. They also considered pyrimidine content, length of strand, and asymmetry of PNA/DNA duplexes. For model building, they studied 316 $T_m$s of PNA/DNA duplexes with 6−23 length. These duplexes were designed to form antiparallel, full-matched, and duplex structure (triplex forming homopyrimidine PNAs were excluded). Their prediction can estimate $T_m$ values, but prediction of enthalpy, entropy, and free energy has not been addressed.

$T_m$ values predicted with these methods (*5*, *7*) are listed in Table 1. The average differences of predicted $T_m$ is 19.7 °C for the Griffin and Smith model and 6.0 °C for the Gissen et al. model. The largest difference between measured and predicted $T_m$ is 31.8 °C for the Griffin and Smith model and 11.0 °C for the Gissen et al. model. We have found that the nearest-neighbor model would be useful at least for short PNA/DNA duplexes. It would be necessary to develop a new method containing the possible 16 nearest-neighbor parameters of enthalpy, entropy, and free energy for PNA/DNA duplexes. Also, it is desirable that the parameters to be used reflect interactions in PNA/DNA duplexes such as nearest-neighbor interactions for DNA/DNA and RNA/DNA duplexes (*8*, *10*). As the chain length of the PNA/DNA duplex becomes longer, base geometry in a PNA including the duplex and/or single strand state of PNA may contribute to destabilization of the duplex stability, however, these are not included in the current nearest-neighbor model used for nucleic acid duplexes. The deviation of the duplex structure, which potentially makes hydrogen bonds in a base pair and stacking interaction unfavorable, may be significant for PNA/DNA hybrid duplexes with long length.

## SUPPORTING INFORMATION AVAILABLE

Figure 1S shows normalized melting curves and CD spectra for short PNA/DNA duplexes and Figure 2S shows normalized melting curves for single-stranded PNAs. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Nielsen, P. E., Egholm, M., Berg, R. H., and Buchardt, O. (1991) *Science 254*, 1497−1500.
2. Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Nordén, B., and Nielsen, P. E. (1993) *Nature 365*, 566−568.
3. Nielsen, P. E., Buchardt, O., Egholm, M., and Nordén, B. (1994) *Nature 368*, 561−563.
4. Tomac, S., Sarkar, M., Ratilainen, T., Wittung, P., Nielsen, P. E., Nordén, B., and Gräslund, A. (1996) *J. Am. Chem. Soc. 118*, 5544−5552.
5. Griffin, T. J. and Smith, L. M. (1998) *Anal. Biochem. 260*, 56−63.
6. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746−3750.
7. Giesen, U., Kleider, W., Berding, C., Geiger, A., Ørum, H., and Nielsen, P. E. (1998) *Nucleic Acids Res. 26*, 5004−5006.
8. Sugimoto, N., Nakano, S., Yoneyama, M., and Honda, K. (1996) *Nucleic Acids Res. 24*, 4501−4505.
9. Turner, D. H., Sugimoto, N., and Freier, S. M., (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
10. Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995) *Biochemistry 34*, 11211−11216.
11. SantaLucia, J., Jr., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry 35*, 3555−3562.
12. Allawi, H. T., and SantaLucia, J., Jr. (1997) *Biochemistry 36*, 10581−10594.
13. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) *Biochemistry 37*, 14719−14735.
14. Christensen, L., Fitzpatrick, R., Gildea, B., Petersen, K. H., Hansen, H. F., Koch, T., Egholm, M., Buchardt, O., Nielsen, P. E., Coull, J., and Berg, R. H. (1995) *J. Pept. Sci. 3*, 175−183.
15. Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., and Freier, S. M. (1986) *Biochemistry 25*, 7840−7846.
16. Sugimoto, N., Kierzek, R., Freier, S. M., and Turner, D. H. (1986) *Biochemistry 25*, 5755−5759.
17. Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology*: *Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., pp 596−603, CRC Press, Cleveland, OH.
18. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 256−263.
19. Tebbutt, P. (1998) in *Basic Mathematics for Chemists*, 2nd ed., pp 88−109, John Wiley & Sons Ltd., West Sussex, U.K.
20. Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr. (2000) in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 185−199, University Science Books, Sausalito, CA.
21. Sugimoto, N., Katoh, M., Nakano, S., Ohmichi, T., and Sasaki, M. (1994) *FEBS Lett. 354*, 74−78.
22. Schwarz, F. P., Robinson, S., and Butler, J. M. (1999) *Nucleic Acids Res. 27*, 4792−4800.
23. Chakrabarti, M. C., and Schwarz, F. P. (1999) *Nucleic Acids Res. 27*, 4801−4806.
24. Almarsson, Ö., Bruice, T. C., Kerr, J., and Zuckermann, R. N. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7518−7522.

25. Almarsson, Ö., and Bruice, T. C. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9542−9546.
26. Leijon, M., Gräslund, A., Nielsen, P. E., Buchardt, O., Nordén, B., Kristensen, S. M., and Eriksson, M. (1994) *Biochemistry 33*, 9820−9825.
27. Eriksson, M., and Nielsen, P. E. (1996) *Nat. Struct. Biol 3*, 410−413.
28. LeBlanc, D. A., and Morden, K. M. (1991) *Biochemistry 30*, 4042−4047.
29. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29*, 278−285.
30. Rasmussen, R., Kastrup, J. S., Nielsen, J. N., Nielsen, J. M., and Nielsen, P. E. (1997) *Nat. Struct. Biol. 4*, 98−101.
31. Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr. (2000) in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 88−91, University Science Books, Sausalito, CA.
32. Soler-Lopez, M., Malinina, L., Liu, J., Huynh-Dinh, T., and Subirana, J. A. (1999) *J. Biol. Chem. 274*, 23683.

BI010480M