© 1996 Oxford University Press Nucleic Acids Research, 1996, Vol. 24, No. 22 **4501–4505**

Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes

Naoki Sugimoto*, Shu-ich Nakano, Mari Yoneyama and Kei-ich Honda

Department of Chemistry, Faculty of Science, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658, Japan

Received July 26, 1996; Revised and Accepted September 13, 1996

ABSTRACT

To improve the previous DNA/DNA nearest-neighbor parameters, thermodynamic parameters ($\Delta H°$, $\Delta S°$ and $\Delta G°$) of 50 DNA/DNA duplexes were measured. Enthalpy change of a helix initiation factor is also considered though the parameters reported recently did not contain the factor. A helix initiation factor for DNA/DNA duplex determined here was the same as that of RNA/RNA duplex ($\Delta G°_{37}$ = 3.4 kcal/mol). The improved nearest-neighbor parameters reproduced not only these 50 experimental values used here but also 15 other experimental values obtained in different studies. Comparing $\Delta G°_{37}$ values of DNA/DNA nearest-neighbor parameters obtained here with those of RNA/RNA and RNA/DNA, RNA/RNA duplex was generally the most stable of the three kinds of duplexes with the same nearest-neighbor sequences. Which is more stable between DNA/DNA and RNA/DNA duplexes is sequence dependent.

INTRODUCTION

Thermodynamics of enthalpy change ($\Delta H°$), entropy change ($\Delta S°$) and free energy change ($\Delta G°$) for helix formation of nucleic acids and melting temperature (T_m) for helix melting can be explained by considering nearest-neighbor interactions. Nearest-neighbor parameters ($\Delta H°$, $\Delta S°$ and $\Delta G°$) of DNA/DNA and RNA/RNA double helices (1,2) were obtained on the basis of the nearest-neighbor model (3). Recently, we also reported the nearest-neighbor parameters for RNA/DNA duplexes (4). Thus, thermodynamics for double helix formation of DNA/DNA, RNA/RNA and RNA/DNA can be estimated with the nearest-neighbor parameters (1,2,4,5) and their parameters are used to predict stable secondary structures and active centers of nucleic acids (6–9).

However, it was reported that there were sometimes larger differences between measured and predicted thermodynamics of DNA/DNA double helices (10,11) when the published DNA/DNA nearest-neighbor parameters (1) were used. In order to make a much better prediction, as all DNA duplexes consist of 10 nearest-neighbor sequences of dAA/dTT, dAT/dAT, dCG/dCG, dCT/dAG, dGA/dTC, dGC/dGC, dGG/dCC, dGT/dAC, dTA/dTA and dTG/dCA, these 10 DNA nearest-neighbor parameters should be improved. Also, an initiation factor should be improved because enthalpy change of a helix initiation factor has been assumed to be 0 kcal/mol for DNA/DNA and RNA/RNA parameters (1,2). Although new nearest-neighbor parameters for predicting DNA/DNA duplex stability were reported on the basis of a small number of thermodynamic results of DNA oligomers (12), free energy changes for duplex initiation (1.82 kcal/mol for G/C pair and 2.8 kcal/mol for A/T pair) were quite different from that of RNA (3.4 kcal/mol). Such a large difference between DNA/DNA and RNA/RNA duplex initiations is amazing. The enthalpy change for DNA/DNA duplex initiation has to be considered for the improvement, because it was shown that the enthalpy contribution to the duplex initiation was more important than the entropy contribution in the case of RNA/DNA duplexes (4).

Here, we report improved DNA nearest-neighbor parameters determined by measuring thermodynamic parameters for 50 DNA duplexes in 1 M NaCl buffer. The new nearest-neighbor parameters are compared with the previous ones (1,12).

MATERIALS AND METHODS

Materials

All deoxyribonucleotides (87 nucleotides) were synthesized chemically on a solid support using phosphoramidite procedures and purified with high-performance liquid chromatography (HPLC) after deblocking operations (13). The synthesized oligonucleotides were further purified and desalted with a C-18 Sep-Pak cartridge. The final purity was confirmed to be >98% by HPLC. These oligoucleotides consist of the 10 kinds of nearest-neighbor sequences of Watson–Crick base pairs described above, and their lengths range from 5 to 14 nucleotides. The sequences were selected to have many different combinations of the nearest-neighbor base pairs.

All experiments were conducted in a buffer including 1 M NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA (pH 7.0). Each oligonucleotide concentration was determined by measuring the absorbance at 260 or 280 nm at a high temperature as described previously (14). Single strand extinction coefficients were calculated from mononucleotide and dinucleotide data by using a nearest-neighbor approximation (15), though it should be noted that the data included not a small error. The DNA strand and its complementary DNA strand were mixed with a 1:1 concentration ratio to obtain each DNA double helix.

UV measurement

Absorbance measurements in the UV region were made on Hitachi U-3200 and U-3210 spectrophotometers. Melting curves (absorbance versus temperature curves) were measured at 260 or 280 nm. Cell holders were thermostated by Hitachi SPR-7 and SPR-10 thermoprogrammers. The water condensation on the

---

*To whom correspondence should be addressed. Tel: +81 78 435 2497; Fax: +81 78 435 2539; Email: sugimoto@konan-u.ac.jp cuvette exterior at the low-temperature range was avoided by flushing with a constant stream of dry N_2 gas. The heating rate was 0.5 or 1.0°C/min.

The melting curves of all the DNA/DNA duplexes examined here showed normal melting behaviors as well as RNA/RNA and RNA/DNA duplexes (2,4). Melting data were collected and fitted with NEC PC9801 computers. T_m values of Watson–Crick base pairs were obtained with the curve fitting procedure (16). Thermodynamic parameters ($\Delta H°$, $\Delta S°$ and $\Delta G°$) for DNA/DNA double-helix formation were determined as the average values obtained by T_m^{-1} versus $\ln(C_t)$ plot and the curve fitting procedure (2,4).

Calculation of nearest-neighbor parameters

According to the nearest-neighbor model, a free energy change at 37°C ($\Delta G°_{37}$) of the helix formation consist of three terms: (i) a free energy change for helix propagation as a sum of each subsequent base pair, (ii) a free energy change for helix initiation and (iii) a free energy change of an entropy effect when the duplex is composed of self-complementary strands (1,2,4). Nearest-neighbor parameters of DNA/DNA duplexes were determined by using thermodynamic data for 50 DNA/DNA duplexes measured here and the data for 15 DNA/DNA duplexes examined previously (1,10,17) on the computer program (4). End effect by terminal dA/dT pair were not considered, because destability of the fraying was not so large and it affected the whole duplex stability by only 0.4 kcal/mol (12). Each nearest-neighbor parameter of $\Delta G°_{37}$ and $\Delta H°$ was described on the program by 0.1 kcal/mol steps so as to predict the stabilities of all the DNA/DNA duplexes applying to a two-state model with minimum sum of square-error values (4). The initiation parameter of $\Delta G°_{37}$ was once fixed at 3.4 kcal/mol, because the initiation factor is included to all 65 DNA/DNA duplexes so that the change of the initiation parameter significantly affected all base-set parameters. Once the base-set parameters were determined, all nearest-neighbor parameters including the initiation factor were calculated again. Nearest-neighbor parameters of $\Delta H°$ were obtained in the same way as for $\Delta G°_{37}$ determination. The values of $\Delta S°$ were estimated by the determined values of $\Delta G°_{37}$ and $\Delta H°$. Predicted melting temperatures at 100 µM strand concentration were also calculated.

RESULTS

Thermodynamics of DNA/DNA double helices

Measured free energy changes at 37°C ($\Delta G°_{37}$) of d(CTAGTGGA)/d(TCCACTAG), d(GCCAGTTA)/d(TAACTGGC) and d(GGTGCCAA)/d(TTGGCACC), for example, were –7.5, –8.2 and –9.0 kcal/mol, respectively. One can predict the thermodynamics of the double helix formation of these DNAs with the previous nearest-neighbor parameters (1). We have compared these measured values with predicted ones by the parameters. The predicted $\Delta G°_{37}$ values of d(GCCAGTTA)/d(TAACTGGC) and d(GGTGCCAA)/d(TTGGCACC) were –7.1 and –9.6 kcal/mol, respectively. The differences between predicted and measured values were not so small [15.5 and 6.3% for d(GCCAGTTA)/d(TAACTGGC) and d(GGTGCCAA)/d(TTGGCACC), respectively]. Moreover, the predicted $\Delta G°_{37}$ value of d(CTAGTGGA)/d(TCCACTAG) was –5.3 kcal/mol, which was 41.5% different from the measured value of –7.5 kcal/mol. It was also reported that the differences between measured and predicted thermodynamics by the parameters were sometimes larger than expected (10,11). Accordingly, we improved the nearest-neighbor parameters for DNA double helix formation.

Improved nearest-neighbor parameters for DNA/DNA duplexes

To determine new nearest-neighbor parameters, we measured the thermodynamics of duplex formation for 50 DNAs. Thermodynamic parameters of all the duplexes determined with both T_m^{-1} versus $\ln(C_t)$ plots and the curve fitting procedure (4) are provided as supplementary material (Tables S1 and S2). Melting behaviors of all 50 DNAs used here show a two-state transition, so that they are considered to show similar transition enthalpy changes as calculated from the shape of the calorimetric curves (18,19). Thermodynamic parameters for another 15 DNA duplexes (1,10,17) were also considered, so data for 65 duplexes in order to determine the nearest-neighbor parameters are listed in Table S3 as supplementary material.

DNA/DNA double helices have 10 nearest-neighbor sequences of dAA/dTT, dAT/dAT, dCG/dCG, dCT/dAG, dGA/dTC, dGC/dGC, dGG/dCC, dGT/dAC, dTA/dTA and dTG/dCA. We considered these 10 values and the enthalpy effect of a helix initiation for the prediction which had been ignored (1,2). The energies of an initiation at A/T and G/C base pairs were regarded as the same values according to the previous assumption (2,4). As the helix symmetry factor was due to an entropy effect of associating self-complementary strands, the factor was regarded as the same as that of RNA double helix (2). Thus, the negative entropy change of –1.4 cal/mol/K is used for double helix formation of self-complementary strands.

Table 1. Improved thermodynamic parameters for DNA/DNA double helix initiation and propagation in 1 M NaCl buffer[a]

Sequence	$\Delta H°$ kcal/mol	$\Delta S°$ cal/mol/K	$\Delta G°_{37}$ kcal/mol
dAA dTT	–8.0	–21.9	–1.2
dAT dTA	–5.6	–15.2	–0.9
dTA dAT	–6.6	–18.4	–0.9
dCA dGT	–8.2	–21.0	–1.7
dCT dGA	–6.6	–16.4	–1.5
dGA dCT	–8.8	–23.5	–1.5
dGT dCA	–9.4	–25.5	–1.5
dCG dGC	–11.8	–29.0	–2.8
dGC dCG	–10.5	–26.4	–2.3
dGG dCC	–10.9	–28.4	–2.1
initiation	0.6	–9.0	3.4
self-complementary	0.0	–1.4	0.4
non-self-complementary	0.0	0.0	0.0

[a] Estimated errors in $\Delta H°$, $\Delta S°$ and $\Delta G°_{37}$ are ±0.3 kcal/mol, ±1.3 cal/mol/K and ±0.1 kcal/mol, respectively.

The optimized propagation factor for 10 nearest-neighbor sequences and an initiation factor are shown in Table 1. The $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$ and T_m values of 65 DNA duplexes predicted with the improved parameters are listed as supplementary data in Table S3 together with the measured values.

DISCUSSION

Accuracy of the improved nearest-neighbor parameters

Improved nearest-neighbor parameters have the same tendency as the previous ones though the absolute values are quite different. For example, the dCG/dCG and dGC/dGC pair have a relatively large stabilization energy, while dTA/dTA and dAT/dAT have the smallest stability for a helix formation. Thermodynamics of d(CTAGTGGA)/d(TCCACTAG), d(GCCAGTTA)/d(TAACTGGC) and d(GGTGCCAA)/d(TTGGCACC) which could not be predicted with the previous parameters as described above are predicted exactly with the improved parameters. For example, the predicted $\Delta G°_{37}$ values were –7.3, –7.8 and –9.2 kcal/mol, respectively, and the differences between the measured and predicted $\Delta G°_{37}$ values were only 2.7, 5.1 and 2.2%, respectively. Recently, SantaLucia Jr *et al.* reported new nearest-neighbor parameters for DNA stability (12). Using their parameters for prediction of these duplexes, the predicted $\Delta G°_{37}$ values were –6.8, –7.5 and –8.9 kcal/mol, respectively, and the average difference was 6.9%. Thus, our improved nearest-neighbor parameters determined here are more useful to predict DNA stability than the parameters of SantaLucia Jr *et al.* including the incorrect helix initiation factor described above.

All thermodynamic values of the 65 DNAs were predicted with both improved and previous parameters, and the values of $\Delta G°_{37}$ are shown in Table 2. The average error of the predicted values by the previous parameters is 23.5%, while that by the improved parameters is only 4.8%. The improved parameters can also predict much higher accuracy of the other 15 thermodynamics (average error of only 5.6%) compared with the previous prediction (that of 27.9%). The average differences of $\Delta H°$, $\Delta S°$ and T_m of 65 DNA duplexes between measured and predicted values by the improved parameters are 6.1, 6.8 and 5.7%, respectively. These errors are much smaller than those by the previous parameters (9.3, 9.4 and 19.5%, respectively). Thus, the improved parameters can lead to much improved prediction for the stability of DNA/DNA double helices.

Table 2. Measured and predicted free energy changes for DNA duplex formations

DNA sequence[a]	$\Delta G°_{37}$ values (kcal/mol)		
	Measured	Predicted (I)[b]	Predicted (II)[c]
AGCCG	–5.6	–5.3	–5.0
ACCGCA	–6.7	–7.0	–6.6
ATGCGC	–7.3	–6.6	–6.7
CGGTGC	–7.2	–7.0	–6.6
CGTGCC	–6.9	–7.0	–6.6
TGCGCA	–6.9	–7.0	–6.9
AATACCG	–5.9	–6.0	–5.7
AGCCGTG	–8.5	–8.5	–7.9
AGCTTCA	–6.1	–6.3	–5.1
GGACTTA	–5.6	–5.3	–3.9
AAAAAAAA	–5.1	–5.0	–5.7
AAGCGTAG	–8.0	–8.3	–7.1
AATCCAGT	–6.8	–7.0	–6.1
ACATATGT	–5.7	–5.3	–3.4
ACCTAGTC	–6.5	–7.1	–4.6
ACGACCTC	–8.7	–9.0	–7.2
AGAGAGAG	–7.4	–7.1	–4.2
AGCGTAAG	–7.8	–8.3	–7.1
AGTCCTGA	–7.5	–7.9	–5.9
ATGCGCAT	–9.0	–8.8	–9.3
CACGGCTC	–10.0	–10.0	–9.3
CCATATGG	–6.8	–6.5	–6.8
CGATATCG	–7.3	–7.5	–7.0
CGCGTATA	–8.4	–8.7	–8.1
CGCTGTAA	–7.9	–8.5	–7.6
CTAGTGGA	–7.5	–7.3	–5.3
CTCACGGC	–9.8	–10.0	–9.3
CTGAGTCC	–8.0	–7.9	–5.9
GAATATTC	–4.3	–4.3	–3.8
GACTAGTC	–6.6	–6.1	–2.8
GAGTACTC	–5.8	–6.1	–2.8
GATTAATC	–4.3	–4.3	–3.8
GCATATGC	–7.8	–6.9	–6.8
GCCAGTTA	–8.2	–7.8	–7.1
GGTGCCAA	–9.0	–9.2	–9.6
GTCGAACA	–8.3	–8.3	–6.6
GTCTAGAC	–6.5	–6.1	–2.8
TAGGCCTA	–8.2	–7.5	–7.0
TATGCATA	–5.9	–5.5	–4.8
AAAAAAAAA	–6.2	–6.2	–7.4
ATAACTGGC	–9.0	–8.7	–8.3
ATCTATCCG	–8.7	–8.7	–8.2
CGCTGTTAC	–9.9	–10.0	–8.7
GCCAGTTAA	–8.8	–9.0	–8.8
AAAAAAAAAA	–6.7	–7.4	–9.1
CGGCAAGCGC	–13.3	–15.6	–17.4
TAGGTTATAA	–7.0	–7.7	–7.0
ACGTATTATGC	–10.4	–11.2	–10.6
ATTGGATACAAA	–10.3	–11.4	–12.0
ACATTATTATTACA	–11.3	–12.0	–11.7
GCGCGC[d]	–9.2	–8.7	–9.4
CAAAAAG[d]	–4.4	–4.6	–4.7
CGTCGACG[d]	–9.8	–10.6	–9.3
GAAGCTTC[d]	–6.7	–6.9	–6.0
GGAATTCC[d]	–7.4	–6.7	–7.4
GGTATACC[d]	–5.5	–6.1	–5.0
CAAAAAAAG[d]	–7.3	–7.0	–8.1
CAAACAAAG[d]	–7.9	–7.8	–7.6
CAAAGAAAG[d]	–7.3	–7.6	–7.4
CAAATAAAG[d]	–6.5	–6.4	–6.7
GCGAATTCGC[d]	–12.9	–12.7	–14.0
CGCGTACGCGTACGCG[d]	–29.1	–27.7	–28.6
CACAG[e]	–3.6	–3.0	–0.8
CAACCAACCAAC[f]	–14.1	–14.0	–14.2
CTTCCTTCCTTC[f]	–14.1	–13.4	–13.6

[a] The DNA duplex consists of the denoted DNA strand and its complementary DNA strand.
[b] The values predicted with our parameters in Table 1.
[c] The values predicted with the parameters of ref. 1.
[d] Data from Breslauer *et al.* (1).
[e] Data from Hall *et al.* (10).
[f] Data from Ratmeyer *et al.* (17).

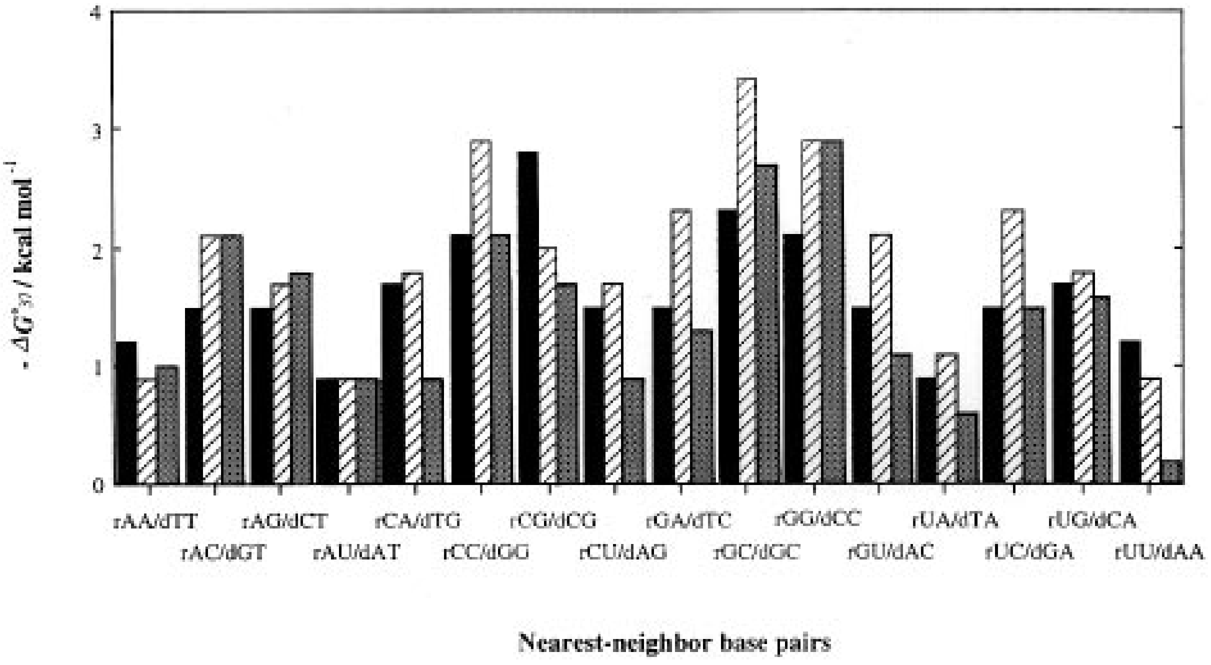

Figure 1. Comparison of ΔG°_{37} for DNA/DNA (black) (this study), RNA/RNA (stripe) (2), and RNA/DNA hybrid (gray) (4) nearest-neighbor base pairs. The denoted nearest-neighbor base pairs are for RNA/DNA hybrid formation.

Comparison of DNA/DNA nearest-neighbor parameters with RNA/RNA and RNA/DNA parameters

Free energy change of the initiation factor for DNA/DNA was determined to be 3.4 kcal/mol in this study, which was the same as that of RNA/RNA (2). As ΔG°_{37} of RNA/DNA was 3.1 kcal/mol for the duplex initiation (4), the duplex initiation processes for DNA/DNA, RNA/RNA and RNA/DNA seem to be not so different energetically. It is reasonable that these helix initiations are similar because each duplex associates two oligonucleotides in the same manner. But, though the enthalpy change of the helix formation had been ignored for the nearest-neighbor parameters of DNA/DNA (1) and RNA/RNA duplexes (2), we determined the enthalpic energy for the helix initiation process for RNA/DNA hybrids (4). The enthalpy change was 1.9 kcal/mol and the free energy change was 3.1 kcal/mol for RNA/DNA duplexes. That is, the enthalpy change seems to be dominant for RNA/DNA helix initiation (4). On the other hand, ΔH° and ΔG°_{37} of the initiation factor for DNA/DNA duplexes are 0.6 and 3.4 kcal/mol, respectively. Thus, at a helix initiation process of DNA/DNA duplexes, an entropy contribution is dominant for the free energy change.

Figure 1 shows the ΔG°_{37} values of the nearest-neighbor pairs of DNA/DNA duplexes with those of RNA/RNA and RNA/DNA ones (2,4). Generally, as shown in Figure 1, RNA/RNA double helix is the most stable with the same nearest-neighbors. Which is more stable between DNA/DNA and RNA/DNA duplexes depends on its sequence. Nearest-neighbor values of dCPu/dPyG of DNA/DNA duplexes are about 1.0 kcal/mol more stable than those of RNA/DNA hybrids. The values of dPyT/dAPu and dTPu/dPyA of DNA/DNA are also ~0.8 and 0.2 kcal/mol, respectively, more stable than those of RNA/DNA hybrid. All of these pairs of DNA/DNA duplexes are more stable than those of RNA/DNA because of their enthalpy advantages. In contrast, the values of dPuC/dGPy and dPuG/dCPy of DNA/DNA duplexes are ~0.6 kcal/mol less stable than those of RNA/DNA hybrids.

SUPPLEMENTARY MATERIAL

Three tables listing thermodynamic parameters obtained with the T_m^{-1} versus $\ln(C_t)$ plot (Table S1) and the curve fitting procedure (Table S2), and 65 measured and predicted thermodynamic data of DNA/DNA duplex formation (Table S3).

See supplementary material available in NAR Online.

ACKNOWLEDGEMENTS

We acknowledge M. Katoh, Y. Shintani, T. Ohmichi,.Y. Naraki, D. Miyoshi, I. Yasumatsu, N. Wakizaka and H. Nakamuta (Konan University) for obtaining thermodynamic data. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture, Japan to N.S.

REFERENCES

1 Breslauer, K. J., Frank, R., Blöcker, H. and Marky, L. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
2 Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
3 Tinoco, Jr. I., Uhlenbeck, O. C. and Levine, M. D. (1971) *Nature* **230**, 363–367.
4 Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M. and Sasaki, M. (1995) *Biochemistry* **34**, 11211–11216.
5 Turner, D. H., Sugimoto, N. and Freier, S. M. (1990) in Saenger, W. R. (ed.) *Landolt-Börnstein Nucleic Acids* Vol. 1c, chap. 3.6, Berlin, Germany.
6 Sugimoto, N. and Sasaki, M. (1991) *Chemistry Lett.* 345–348.
7 Sugimoto, N., Sasaki, M. and Tanaka, A. (1993) *Supramolec. Chem.* **2**, 99–102.

8 LeCuyer, K. A. and Crothers, D. M. (1993) *Biochemistry* **32**, 5301–5311.
9 Connell, G. J. and Yarus, M. (1994) *Science* **264**, 1137–1141.
10 Hall, K. B. and McLaughlin, L. W. (1991) *Biochemistry* **30**, 10606–10613.
11 Sugimoto, N., Honda, K. and Sasaki, M. (1994) *Nucleosides Nucleotides* **13**, 1311–1317.
12 SantaLucia, Jr. J., Allawi, H. T. and Seneviratne, P. A. (1996) *Biochemistry* **35**, 3555–3562.
13 Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H. and Freier, S. M. (1986) *Biochemistry* **25**, 7840–7846.
14 Sugimoto, N., Kierzek, R., Freier, S. M. and Turner, D. H. (1986) *Biochemistry* **25**, 5755–5759.
15 Richards, E. G. (1975) in Fasman, G. D., (ed.) *Handbook of Biochemistry and Molecular Biology*: *Nucleic Acids* 3rd. ed., Vol. 1, 597, CRC Press, Cleveland, OH.
16 Petersheim, M. and Turner, D. H. (1983) *Biochemistry* **22**, 256–263.
17 Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G. and Wilson, D. (1994) *Biochemistry* **33**, 5298–5304.
18 Marky, L. A. and Breslauer, K. J. (1987) *Biopolymers* **26**, 1601–1620.
19 Zieba, K., Chu, T. M., Kupke, D. W. and Marky, L. A. (1991) *Biochemistry* **30**, 8018–8026.