Improved Nearest-Neighbor Parameters for Predicting DNA Duplex Stability[†]

John SantaLucia, Jr.,* Hatim T. Allawi, and P. Ananda Seneviratne

Department of Chemistry, Wayne State University, Detroit, Michigan 48202

Received August 14, 1995; Revised Manuscript Received December 1, 1995[⊗]

ABSTRACT: Thermodynamic data were determined from UV absorbance vs temperature profiles of 23 oligonucleotides. These data were combined with data from the literature for 21 sequences to derive improved parameters for the 10 Watson−Crick nearest neighbors. The observed trend in nearest-neighbor stabilities at 37 °C is GC > CG > GG > GA ≈ GT ≈ CA > CT > AA > AT > TA (where only the top strand is shown for each nearest neighbor). This trend suggests that both sequence and base composition are important determinants of DNA duplex stability. On average, the improved parameters predict $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_M within 4%, 7%, 8%, and 2 °C, respectively. The parameters are optimized for the prediction of oligonucleotides dissolved in 1 M NaCl.

Accurate prediction of DNA thermal denaturation is important for several molecular biological techniques including PCR[1] (Saiki et al., 1988), sequencing by hybridization (Fodor et al., 1993), antigene targeting (Freier, 1993), and Southern blotting (Southern, 1975). In these techniques, choice of a nonoptimal sequence or temperature can lead to amplification or detection of wrong sequences (Steger, 1994). Furthermore, knowledge of the sequence dependence of DNA melting is important for understanding the details of DNA replication, mutation, repair, and transcription (Mendelman et al., 1989; Petruska et al., 1988).

One widely used method for predicting nucleic acid duplex stability, pioneered by Tinoco and co-workers (Borer et al., 1974), uses a nearest-neighbor model for helix propagation. Several nearest-neighbor parameter sets for predicting DNA duplex stability are available in the literature (Gotoh & Tagashira, 1981; Ornstein & Fresco, 1983; Vologodskii et al., 1984; Wartell & Benight, 1985; Breslauer et al., 1986; Aida, 1988; Otto, 1989; Quartin & Wetmur, 1989; Klump, 1990; Delcourt & Blake, 1991; Doktycz et al., 1992). The quantum mechanical studies (Ornstein & Fresco, 1983; Aida, 1988; Otto, 1989) were performed in the gas phase and neglected solvent and counterion interactions and, thus, do not reflect the conditions typically found *in vivo* or *in vitro*. Data for the thermal denaturation of polymers are difficult to interpret properly since their transitions are typically not two-state (i.e., many unfolding intermediates are possible) and their melting temperatures are high. Thus, polymer melting is typically performed in solutions with low salt concentration and the thermodynamic results are extrapolated to the standard state temperature (25 or 37 °C) and higher salt concentration (Breslauer et al., 1986). In addition, polymer melting does not involve a bimolecular initiation event and is dependent on only eight invariants which are linear combinations of the 10 nearest neighbor parameters required to predict oligonucleotide thermodynamics (Gray & Tinoco, 1970; Vologodskii et al., 1984; Doktycz et al., 1992). Thus, studies of polymer thermodynamics (Gotoh & Tagashira, 1981; Vologodskii et al., 1984; Wartell & Benight, 1985; Klump, 1990; Delcourt & Blake, 1991) are most applicable for the prediction of polymer behavior but do not reliably predict oligonucleotide thermodynamics (Sugimoto et al., 1994). Thus, we decided to expand the DNA oligonucleotide thermodynamic database and derive new nearest-neighbor parameters in 1 M NaCl buffer.

In this paper, thermodynamic measurements are reported for 26 oligonucleotides ranging in length from 4 to 16 base pairs. Thermodynamic data for 23 of these sequences are combined with data for 21 oligonucleotides from the literature to derive improved nearest-neighbor parameters. The parameters are able to predict the stabilities of DNA duplexes within the limits of the nearest-neighbor model.

MATERIALS AND METHODS

DNA Synthesis and Purification. Oligonucleotides were the gift of Hitachi Chemical Research and were synthesized on solid support using standard phosphoramidite chemistry (Brown & Brown, 1991). Oligomers were removed from solid support and base blocking groups were removed by treatment with concentrated ammonia at 50 °C overnight. Each sample was evaporated to dryness, dissolved in 250 μL of water and purified on a Si500F thin-layer chromatography plate (Baker) by eluting for 5 h with *n*-propanol/ammonia/water (55:35:10 by volume) (Chou et al., 1989). Bands were visualized with an ultraviolet lamp, and the least mobile band was cut out and eluted three times with 3 mL of distilled deionized water. The sample was then evaporated to dryness. Oligonucleotides were desalted and further purified with a Sep-pak C-18 cartridge (Waters). The DNA was eluted with 30% acetonitrile buffered with 10 mM ammonium bicarbonate at pH 7.0. The purity of oligonucleotides was checked by analytical C-8 HPLC (Perceptive Biosystems) and was greater than 95%.

Measurement of Melting Curves. Absorbance vs temperature profiles (melting curves) were measured with an Aviv 14DS UV−vis spectrophotometer with a five cuvette thermoelectric controller. Custom-manufactured microcuvettes (Hellma Cells) with 0.1, 0.2, 0.5, and 1.0 cm path lengths (60, 120, 300, and 600 μL volumes, respectively) were used

---

* Author to whom correspondence should be addressed.
[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1996.
[1] Abbreviations: Na₂EDTA, disodium ethylenediaminetetracetate; eu, entropy units (cal/K mol); HPLC, high-performance liquid chromatography; T_M, melting temperature; PCR, polymerase chain reaction; UV, ultraviolet.

so that melting curves could be measured with high sensitivity over a 100-fold range in oligonucleotide concentration. Aluminum adapters were used to properly position microcuvettes in the light beam and provide optimal thermal contact with the thermoelectric cuvette holder. To prevent water condensation at low temperatures, the sample compartment was purged with dry nitrogen gas.

The temperature was monitored with the temperature transducer (Analog Devices Inc.) mounted in the spindle of the Aviv thermoelectric cuvette holder. The temperature readings from the transducer were calibrated by measuring the voltage produced by a type K thermocouple inserted in a 1 cm microcuvette during a typical thermal denaturation run. We estimate the temperature measurement to be reproducible within 0.1 °C and accurate within 0.3 °C.

Oligonucleotides were dissolved in 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7, buffer. Samples were "annealed" and degassed by raising the temperature to 85 °C for 5 min and then cooling to −1.6 °C over a period of 25 min just prior to a melting experiment. While at 85 °C, the absorbances were measured at 260 nm for later calculation of oligonucleotide concentration using extinction coefficients calculated from dinucleoside monophosphates and nucleotides, as described (Richards, 1975). Oligonucleotide concentration was varied over an 80−100-fold range. Samples were then heated at a constant rate of 0.8 °C/min with data collection beginning at 0 °C and ending at 90−95 °C. Control experiments with a heating rate of 0.4 °C/min gave same results as those obtained with 0.8 °C/min indicating that thermal equilibrium was estabilished. The duplex to coil tranistion was monitored by measuring the absorbance at 280 nm. Air was used for the reference light beam. The spectral bandwidth was 1 nm. Absorbances were not corrected for thermal expansion since the correction was linear and less than 3% from 0 to 90 °C.

Determination of Thermodynamic Parameters. Thermodynamic parameters for duplex formation were obtained from melting curve data using the program MELTWIN v2.1 (Jeff McDowell and Douglas H. Turner, unpublished) as described (Petersheim & Turner, 1983). Data were truncated so that the upper and lower temperature baselines reflected the slopes in the transition region, generally using $T_M \pm 30$ °C (Petersheim & Turner, 1983). The root mean square difference between data and calculated curves is less than 0.5%, the approximate error in the absorbance reading. The enthalpy and entropy for the random coil to duplex equilibrium were obtained by two methods: (1) $\Delta H°$ and $\Delta S°$ from the fits of individual curves were averaged, and (2) plots of reciprocal melting temperature (T_M^{-1}) versus the natural logarithm of the total strand concentration (ln C_T) were fit to eq 1 (Borer et al., 1974):

$$T_M^{-1} = R/\Delta H° \ln C_T + \Delta S°/\Delta H° \quad (1)$$

The T_M is defined as the temperature at which half of the strands are in the double helical state and half are in the "random coil" state. For self-complementary oligonucleotides, the T_M for individual melting curves was calculated from the fitted parameters using

$$T_M = \Delta H°/(\Delta S° + R \ln C_T) \quad (2)$$

where R is the gas constant [1.987 cal/(K mol)], and the T_M

is given in K. For non-self-complementary molecules, C_T in eqs 1 and 2 was replaced by $C_T/4$. Both methods are essentially a van't Hoff analysis of the data, assuming the transition equilibrium involves only two states (i.e., duplex and random coil). We also assume that the difference in heat capacities (ΔC_p) of these states is zero (Petersheim & Turner, 1983; Freier et al., 1986b). These two methods depend differently on the two-state approximation. For a given oligonucleotide, agreement of parameters derived by the two methods is a necessary, but not sufficient, criterion to establish the validity of the two-state approximation (SantaLucia et al., 1990; Marky & Breslauer, 1987). The methods described above have been shown to give thermodynamic results in good agreement with those obtained by calorimetry (Albergo et al., 1981).

Choice of Sequences. Sequences were designed or selected from the literature to meet the following criteria: (1) two-state thermodynamics, (2) T_M between 20−65 °C to minimize extrapolation to 37 °C and allow the upper and lower temperature baselines to be adequately defined, (3) sequences with three or more consecutive guanine residues are not included since these sequences consistently yield lower than expected $\Delta H°$ values (Breslauer et al., 1986), (4) sequences have terminal GC base pairs to minimize helix "fraying" that could invalidate the two-state approximation, and (5) the oligonucleotides were dissolved in 1 M NaCl so that length-dependent counterion condensation effects could be neglected (Record & Lohman, 1978; Olmsted et al., 1989). Sequences measured in this study were designed to meet the above criteria and provide uniform representation of the 10 different nearest neighbors in the database. Throughout this paper nearest-neighbor base pairs are represented with a slash separating strands in antiparallel orientation (e.g., AC/TG means 5′-AC-3′ Watson−Crick base paired with 3′-TG-5′). The 10 nearest-neighbor sequences occur in this study with the following frequencies AA/TT = 43, AT/TA = 21, TA/AT =14, CA/GT = 28, GT/CA = 27, CA/GT = 23, GA/CT = 36, CG/GC = 35, GC/CG = 33, and GG/CC = 29.

Determination of Nearest-Neighbor Thermodynamic Parameters. The total difference in the free energy of the folded and unfolded states of a DNA duplex can be approximated with a nearest-neighbor model:

$$\Delta G_i(\text{total}) = \sum n_{ij}\Delta G_j + \Delta G(\text{init}) + \Delta G_i(\text{sym}) \quad (4)$$

where each different oligonucleotide duplex is given the subscript i, ΔG_j are the free energies for the 10 possible Watson−Crick nearest-neighbor stacking interactions [e.g., $\Delta G_1 = \Delta G°_{37}$ (AA/TT), $\Delta G_2 = \Delta G°_{37}$ (TA/AT), ..., etc.], n_{ij} is the number of occurrences of each nearest neighbor, j, in each sequence, i, $\Delta G(\text{init})$ is the initiation free energy, and $\Delta G_i(\text{sym})$ equals +0.4 kcal/mol if duplex i is self-complementary and zero if it is non-self-complementary (Bailey & Monahan, 1978; Cantor & Schimmel, 1980).

The thermodynamic results from 44 sequences with two-state transitions were used to construct appropriate matrices for input into the linear regression analysis. The [$\Delta G_i(\text{total})$ − $\Delta G_i(\text{sym})$] formed the column matrix $\mathbf{G_{Tot}}$ and the standard errors in the $\Delta G_i(\text{total})$, σ_i, formed the column matrix σ. The number of occurrences of each of the nearest neighbors, along with the initiation parameter formed the "stacking matrix", $\mathbf{S}$ with dimensions 44 × 11. The values of the 10 nearest neighbors and initiation, G_j, are unknown
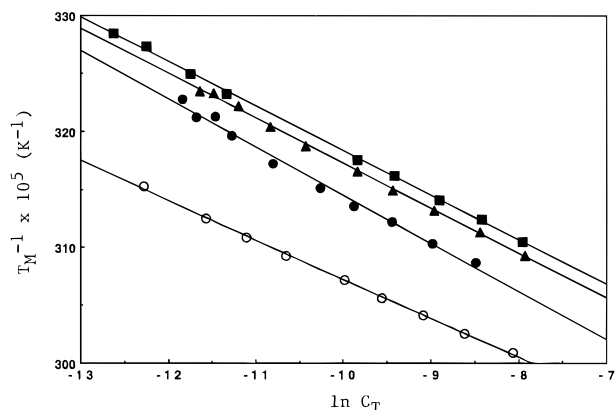
FIGURE 1: Reciprocal melting temperature vs ln C_T plots for GTTGCAAC (●), GTACGTAC (▲), GGACGTCC (○), and CGATATCG (■).

and form the column matrix, $\mathbf{G_{NN}}$. The data for all sequences is thus written:

$$\mathbf{G_{Tot}} = \mathbf{SG_{NN}} \qquad (5)$$

The solution of eq 5 for the nearest neighbors, $\mathbf{G_{NN}}$, was obtained using singular value decomposition (Press et al., 1989) which effectively minimizes the error weighted squares of the residuals (Bevington, 1969):

$$\chi^2 = \sum_{ij} |(\Delta G_i - S_{ij}\Delta G_j)/\sigma_i|^2 \qquad (6)$$

Analogous calculations were performed to obtain nearest-neighbor parameters for $\Delta H°$ and $\Delta S°$. All matrix manipulations were performed using the program MATHEMATICA (Wolfram, 1992). To verify our calculation methods, we derived the nearest-neighbor parameters for RNA and reproduced the literature values (Freier et al., 1986a).

RESULTS

Thermodynamic Data. All sequences in this study displayed monophasic melting transitions (data not shown) and showed concentration-dependent T_Ms, indicating complexes with molecularity greater than 1. Plots of T_M^{-1} versus ln C_T were linear (correlation coefficient >0.99) over the entire 80−100-fold range in concentration and are shown in Figure 1 and Supporting Information (see paragraph at the end of the paper). Thermodynamic parameters derived from the average of fits of individual melting curves and from T_M^{-1} versus ln C_T are listed in Table 1. Those sequences in which the $\Delta H°$ from the two methods agree within 20% are listed in Table 1 as two-state transitions (SantaLucia et al., 1990; Marky & Breslauer, 1987). Those with differences in $\Delta H°$ greater than 20% are listed as non-two-state transitions. Experimental heat capacity differences (Table 1), ΔC_p, were determined from the slope of $\Delta H°$ vs T_M plots (data not shown), where the $\Delta H°$s and T_Ms are from the fitted curves of each oligonucleotide at different concentrations (Petersheim & Turner, 1983; Freier et al., 1986b). For two-state transitions, parameters derived from the average of the fits and from T_M^{-1} vs ln C_T plots are equally reliable; thus the average of these parameters (Table 2) was used for the linear regression to determine nearest-neighbor increments (Table 3).

Nearest-Neighbor Parameters. Table 3 lists the nearest-neighbor thermodynamic parameters derived by multiple

linear regression from the data in Table 2. These parameters allow for self-consistent and accurate prediction of the 44 sequences in Table 2 with two-state thermodynamics. Sequences with terminal T-A base pairs were not included in the regression analysis in order to minimize systematic errors due to terminal "fraying" in these sequences. We performed a control experiment where the eight sequences in Table 2 with terminal A-T base pairs were included in the regression analysis. The results showed that a poorer fit (as judged by the χ^2 and Q parameters) was observed. This was particularly true for the $\Delta H°$ and $\Delta S°$ parameters. Inclusion of these sequences systematically made nearest neighbors with 3′-terminal T residues more stable than those with 5′-terminal T residues. For example, the $\Delta G°_{37}$s for AT/TA, TA/AT, GT/CA, TG/AC, CT/GA, and GT/CA nearest neighbors were −1.48, −0.11, −1.76, −0.99, −1.53, and −1.10 kcal/mol, respectively, while the parameters for AA/TT, CG/GC, GC/CG, GG/CC, and initiation did not change compared to the values listed in Table 3. We find that the parameters in Table 3 can predict these sequences reasonably well if a penalty of +0.4 kcal/mol is assigned (for $\Delta G°_{37}$ and $\Delta H°$) for each terminal 5′-T·A-3′ base pair. Note that sequences with terminal 5′-A·T-3′ base pairs are not assigned this penalty. Apparently, sequences with terminal 5′-T·A-3′ base pairs "fray" more than sequences with terminal 5′-A·T-3′ base pairs. The parameter for terminal 5′-T·A-3′ base pairs is included in Table 3.

The Helix Initiation Parameter. The $\Delta G°_{37}$ for helix initiation is +1.82 ± 0.24 kcal/mol (Table 3). This number applies to DNA duplexes with at least one G-C base pair and agrees reasonably well with the value of +2.3 kcal/mol determined previously (Pohl, 1974; Turner et al., 1990). Previous work indicated that initiation in sequences with only A-T base pairs is +3.4 kcal/mol (Scheffler et al., 1970; Turner et al., 1990). Thus, we have assumed the initiation at A-T pairs is +2.8 ± 1 kcal/mol (Table 3). This allows for the correct prediction of the $\Delta G°_{37}$ for A$_8$/T$_8$ (Table 2; Sugimoto et al., 1991) and TTTTATAATAAA/AAAATAT-TATTT (Bolewska et al., 1984).

The penalty for duplex initiation was assumed to be purely entropic for the reasons described previously (Freier et al., 1986a). When the initiation enthalpy was allowed to vary in the regression analysis, a favorable value with a large error was observed (−7.2 ± 3.4 kcal/mol), and the enthalpy increments for the 10 nearest-neighbors are less favorable by 1.1 kcal/mol, on average. When the initiation entropy was allowed to vary in the regression analysis, a more unfavorable value was observed (−29.1 ± 12.4 eu), and the entropy increments for the 10 nearest-neighbors are more favorable by 3.5 eu, on average.

Error Analysis. To evaluate the relative uncertainties in the nearest-neighbor parameters derived above, we determined how the experimental errors, σ_i, propagated to the errors in the nearest neighbors, σ_j, as described [Press et al. (1989) eq 14.3.19]. In the regression analysis described above, the experimental data were weighted assuming 5%, 10%, and 10% uncertainties in the $\Delta G°_{37}$, $\Delta H°$, and $\Delta S°$, respectively. The experimental uncertainties given in Table 1 were not used because they reflect the experimental reproducibility of the data (i.e., precision) not the accuracy of the data (Bevington, 1969). Data from the literature were assigned the same percent errors. Assigning the same percent error for all the data effectively weights the data for all

Table 1: Thermodynamic Parameters of Duplex Formation[a]

DNA duplex	1/T_M vs log C_T parameters				curve fit parameters			
	−Δ$G°_{37}$ (kcal/mol)	−Δ$H°$ (kcal/mol)	−Δ$S°$ (eu)	T_M (°C)[b]	−Δ$G°_{37}$ (kcal/mol)	−Δ$H°$ (kcal/mol)	−Δ$S°$ (eu)	−ΔC_p [kcal/ (K mol)][c]
			Two-State Transitions					
CCGG	3.4 ± 0.3	31.9 ± 1.1	91.9 ± 2.5	16.7	3.6 ± 0.1	29.4 ± 1.0	83.0 ± 3.5	0.2
CGCG	3.9 ± 0.4	38.5 ± 1.6	111.5 ± 3.8	23.6	4.2 ± 0.2	34.0 ± 1.7	96.2 ± 6.0	0.3
GCGC	4.3 ± 0.5	45.6 ± 2.4	133.3 ± 6.1	27.8	4.5 ± 0.1	35.8 ± 0.9	100.9 ± 2.4	
CGATCG	5.4 ± 0.4	33.1 ± 1.1	89.3 ± 2.3	34.3	5.3 ± 0.2	41.8 ± 3.8	117.9 ± 11.8	0.5
GACGTC	5.6 ± 1.3	37.2 ± 4.1	101.8 ± 9.2	36.4	5.5 ± 0.2	46.6 ± 3.8	132.6 ± 11.8	0.5
GCTAGC	5.4 ± 0.6	35.7 ± 1.8	97.7 ± 3.9	34.3	5.3 ± 0.1	42.7 ± 3.9	120.5 ± 12.2	
GGATCC	5.1 ± 0.2	30.2 ± 0.6	80.9 ± 1.3	30.8	4.9 ± 0.2	39.4 ± 3.8	111.4 ± 11.8	
CAAGCTTG	7.0 ± 0.2	54.7 ± 0.7	153.8 ± 1.8	44.6	7.0 ± 0.1	53.8 ± 0.8	150.9 ± 2.5	
CATCGATG	7.5 ± 0.6	56.3 ± 2.3	157.4 ± 5.6	47.4	7.7 ± 0.1	62.0 ± 2.4	175.1 ± 7.4	
CGATATCG	6.8 ± 0.2	51.8 ± 0.6	145.1 ± 1.4	44.0	6.9 ± 0.1	55.7 ± 3.9	157.1 ± 12.1	0.5
GAAGCTTC	6.9 ± 0.2	44.1 ± 0.8	120.3 ± 1.8	45.5	7.1 ± 0.2	54.7 ± 3.2	153.5 ± 9.8	0.5
GATCGATC	7.5 ± 0.6	53.6 ± 2.2	148.7 ± 5.2	47.7	7.7 ± 0.3	63.2 ± 5.9	178.8 ± 17.9	1.2
GATGCATC	7.2 ± 0.3	52.0 ± 1.1	144.2 ± 2.5	46.7	7.4 ± 0.2	60.3 ± 2.6	170.4 ± 7.9	
GGAATTCC	6.8 ± 0.2	46.1 ± 0.5	127.0 ± 1.2	44.5	6.8 ± 0.3	57.0 ± 2.7	161.8 ± 8.3	0.4
GGACGTCC	8.9 ± 0.2	58.6 ± 0.8	160.2 ± 1.8	55.2	9.1 ± 0.1	61.4 ± 0.8	168.7 ± 2.4	
GGAGCTCC	8.6 ± 0.3	53.4 ± 0.9	144.6 ± 2.1	54.7	8.8 ± 0.2	60.1 ± 1.8	165.4 ± 5.1	
GTACGTAC	7.0 ± 0.2	51.3 ± 0.6	142.7 ± 1.5	45.1	7.1 ± 0.2	56.4 ± 3.7	159.0 ± 11.4	0.7
GTAGCTAC	7.0 ± 0.1	51.4 ± 0.6	143.3 ± 1.3	44.9	7.0 ± 0.1	55.1 ± 2.3	155.0 ± 7.0	0.5
GTTGCAAC	7.3 ± 0.5	47.8 ± 1.6	130.5 ± 3.8	48.2	7.6 ± 0.2	59.7 ± 1.2	168.1 ± 3.7	
CCATCGCTACC/GGTAGCGATGG	13.5 ± 0.3	86.9 ± 1.2	236.8 ± 3.0	63.9	13.2 ± 0.2	82.8 ± 1.8	224.5 ± 5.2	
CCATTGCTACC/GGTAACGATGG	12.3 ± 0.3	83.9 ± 1.4	231.1 ± 3.4	59.7	12.1 ± 0.1	81.1 ± 0.6	222.4 ± 1.7	
CTGACAAGTGTC/GACTGTTCACAG	13.0 ± 0.6	88.6 ± 2.8	243.7 ± 6.9	61.6	12.1 ± 0.1	74.3 ± 2.2	200.6 ± 6.8	
CATATGGCCATATG	13.4 ± 0.5	93.7 ± 2.2	258.7 ± 5.6	65.0	12.0 ± 0.3	75.4 ± 2.4	204.3 ± 7.0	
			Non-Two-State Transitions					
GTATACCGGTATAC	13.3 ± 0.1	101.7 ± 0.7	285.3 ± 1.9	61.9	11.3 ± 0.2	74.3 ± 0.7	203.0 ± 1.9	
CATATTGGCCAATATG	14.9 ± 1.1	110.2 ± 5.8	307.4 ± 15.1	65.3	11.6 ± 0.3	67.3 ± 3.0	179.7 ± 8.9	
GTATAACCGGTTATAC	15.3 ± 0.3	113.6 ± 1.7	316.8 ± 4.3	65.9	13.2 ± 0.3	85.8 ± 1.6	234.0 ± 4.5	

[a] Listed by oligomer length and in alphabetical order. For self-complementary sequences, only the top strand is given. For non-self-complementary duplexes, both strands are given in antiparallel orientation separated by a slash. Solutions are 1 M NaCl, 10 mM sodium cacodylate, and 0.5 mM Na₂EDTA, pH 7. Errors are standard deviations from the regression analysis of the melting data. Extra significant figures are given for Δ$H°$ and Δ$S°$ to allow accurate calculation of Δ$G°_{37}$ and T_M. [b] Calculated for 10^{-4} M oligomer concentration for self-complementary sequencs and 4×10^{-4} M for non-self-complementary sequences. [c] Only those sequences with a Δ$H°$ vs T_M plot with a correlation coefficient greater than 0.8 are listed. Errors in ΔC_p are approximately 50%.

sequences equally in the regression analysis. Thus, the percent error assumed has no effect on the values of the nearest-neighbor parameters obtained, only on the propagated errors in the parameters. For example, if we assume errors of 5% for Δ$H°$, we obtain the same nearest-neighbor parameters as with 10% errors, but the error estimates for the nearest neighbors are half as large. The nearest-neighbor errors, σ_j, given in Table 3 are the standard deviations that resulted from the propagation of experimental errors, σ_i, during linear regression. The free-energy covariances (Press et al., 1989) between pairs of nearest neighbors are small (less than ±0.002 kcal²/mol² covariance) and can be neglected. The initiation parameter, however, covaries with all 10 of the nearest neighbors (−0.006 kcal²/mol² covariance, on average).

Goodness of the fits for the nearest neighbor parameters for Δ$G°_{37}$, Δ$H°$, and Δ$S°$ were evaluated from the values of χ^2 (eq 6) and the probability Q (Press et al., 1989). Q is the probability that a χ^2 larger than that observed could be obtained by chance (larger Q indicates a better fit). Q probabilities greater than 0.001 are considered statistically acceptable (Press et al., 1989). Q was calculated from the incomplete gamma function, gamma [$\nu/2$, $\chi^2/2$], where ν is the number of degrees at freedom (ν = number of independent measurements minus the number of parameters derived from the data) (Press et al., 1989). Since measurements were made on 44 different oligonucleotides and 11 parameters (10 nearest neighbors plus initiation) were determined, ν is 33 (for Δ$G°_{37}$) or 34 (for Δ$H°$ and Δ$S°$ which do not have the initiation parameter floating). For Δ$G°_{37}$, $\chi^2 = 28.4$ and $Q = 0.70$. For Δ$H°$, $\chi^2 = 33.8$ and Q = 0.47. For Δ$S°$, $\chi^2 = 50.7$ and $Q = 0.03$. These results suggest that within the estimated errors the nearest-neighbor model is a valid description of DNA thermodynamics in agreement with previous results (Sugimoto et al., 1994; Doktycz et al., 1995).

Comparison of Experimental vs Predicted Thermodynamics. Table 2 compares the experimental results for 60 oligonucleotides with those predicted using the nearest-neighbor parameters in Table 3. The 44 sequences that have two-state transitions are well predicted by the parameters in Table 3. For the T_M at 0.1 mM, the largest difference is 4.6 °C with an average deviation of 1.8 °C. The ability of the parameters in Table 3 to predict the T_M accurately is encouraging, since the parameters were not specifically optimized for the prediction of the T_M. The average deviations between experiment and prediction for Δ$G°_{37}$, Δ$H°$, and Δ$S°$ are 4%, 7%, and 8%, respectively. Previous results for RNA (Kierzek et al., 1986), DNA (Sugimoto et al., 1994), and DNA/RNA hybrid (Sugimoto et al., 1995) oligonucleotides with different sequences, but the same nearest neighbors, suggest the nearest-neighbor model should be able to predict Δ$G°_{37}$, Δ$H°$, and T_M (at 0.1 mM) with average deviations of roughly 6%, 8%, and 2 °C, respectively. Thus, the predictive capacity of the parameters in Table 3 is within the limits of the nearest-neighbor model. The χ^2 and Q parameters, described above, also suggest that the nearest-neighbor parameters given in Table 3 adequately

Table 2: Experimental and Predicted Thermodynamic Parameters of Duplex Formation[a]

		experimental				predicted			
sequence	ref[b]	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_M (°C)[c]	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	T_M (°C)[c]
		Molecules with Two-State Thermodynamics							
CCGG		3.5	30.6	87.4	16.6	3.4	23.5	64.0	12.4
CGCG		4.0	36.3	103.9	23.7	4.2	31.3	86.7	25.0
GCGC		4.4	40.7	117.1	27.5	4.4	32.3	89.6	26.2
CCGCGG	d	8.0	41.4	107.8	55.2	7.8	44.7	117.9	55.0
CGATCG		5.3	37.5	103.6	34.3	5.6	42.1	117.7	36.4
CGCGCG	e	8.3	46.4	122.8	55.7	8.6	52.5	140.6	57.3
CGGCCG	d	8.3	38.7	98.0	59.6	7.8	44.7	117.9	55.0
CGTACG	f	5.4	45.7	130.0	35.0	5.4	43.7	122.8	36.6
GACGTC		5.6	41.9	117.2	36.1	5.7	42.7	119.4	36.9
GCATGC	g	5.6	42.2	118.0	36.5	5.8	45.5	121.5	38.0
GCCGGC	d	8.5	45.2	118.3	57.7	8.0	45.7	120.8	55.4
GCGAGC/CGCTCG	h	7.7	51.4	124.0	33.2	7.5	40.7	115.1	32.0
GCGCGC	e	9.1	59.6	162.7	56.1	8.8	53.5	143.5	57.5
GCTAGC		5.3	39.2	109.1	34.3	5.3	40.7	114.8	32.6
GGATCC		5.0	34.8	96.2	30.8	5.0	35.3	97.9	30.6
GGCGCC	d	7.9	43.5	114.7	53.9	8.0	45.7	120.8	55.4
GGGACC/CCCTGG	h	6.5	32.7	84.5	44.9	6.4	36.4	96.0	45.3
GTGAAC/CACTTG	h	5.1	43.6	124.0	33.2	4.9	40.7	115.1	32.0
CAAAAAG/GTTTTTC	i	4.8	47.0	136.0	31.5	4.8	47.1	135.7	32.7
CAAAAAAG/GTTTTTTC	j	5.7	59.0	172.0	36.9	5.8	55.5	159.3	39.4
CAAGCTTG		7.0	54.2	152.3	44.6	7.2	54.9	153.7	46.0
CATCGATG		7.6	59.2	166.3	47.4	7.0	53.3	149.6	44.3
CGATATCG		6.9	53.7	151.1	44.1	6.9	54.9	155.0	43.6
CGTCGACG	k	9.8	64.1	175.1	58.3	9.8	62.9	170.4	60.2
GAAGCTTC		7.0	49.4	136.9	45.2	7.3	55.5	155.7	45.8
GATCGATC		7.6	58.4	163.7	47.6	7.2	53.9	151.6	44.1
GATGCATC		7.3	56.1	157.3	46.5	7.2	54.3	152.5	44.8
GGAATTCC		6.8	51.6	144.4	43.8	7.0	52.1	145.1	45.7
GGACGTCC		9.0	60.0	164.5	55.1	9.2	56.1	150.6	59.0
GGAGCTCC		8.7	56.8	155.0	54.4	8.8	52.1	139.7	56.6
GGTATACC	k	5.5	54.5	158.0	36.0	6.1	49.7	140.3	40.2
GTACGTAC		7.1	53.8	150.9	45.1	6.8	57.1	161.8	43.9
GTAGCTAC		7.0	53.3	149.2	44.9	6.4	53.1	150.9	40.7
GTTGCAAC		7.5	53.8	149.3	47.7	7.7	59.9	167.5	49.2
CAAAAAAAG/GTTTTTTTC	i	7.2	68.0	196.0	44.2	6.8	63.9	182.9	44.4
CAAACAAAG/GTTTGTTTC	i	7.7	64.5	183.0	47.3	7.6	63.1	178.0	48.3
CAAATAAAG/GTTTATTTC	i	6.5	58.6	168.0	41.4	6.1	59.9	173.0	40.0
CAAAGAAAG/GTTTCTTTC	i	7.3	62.8	179.0	45.2	7.4	60.9	172.1	46.7
GCGAATTCGC	k	12.9	80.0	216.3	67.9	12.2	81.8	221.7	64.8
CCATCGCTACC/GGTAGCGATGG		13.3	84.8	230.7	67.6	12.9	77.2	207.0	69.5
GCGAAAAGCG/CGCTTTTCGC	l	11.9	74.2	201.0	65.2	12.6	81.4	220.9	67.2
CCATTGCTACC/GGTAACGATGG		12.2	82.5	226.8	63.5	11.7	75.2	204.1	65.0
CTGACAAGTGTC/GACTGTTCACAG		12.6	81.5	222.1	65.7	12.9	84.0	229.2	66.2
CATATGGCCATATG		12.7	84.5	231.5	65.3	13.2	92.7	256.3	66.4
		Molecules with Terminal A-T Base Pairs[m]							
TCATGA	d	3.3	50.4	152.0	22.8	3.4	35.9	105.3	17.3
TGATCA	d	2.8	52.6	160.6	20.9	3.4	35.9	105.3	17.3
AAAAAAAA/TTTTTTTT	n	4.5	59.8	178.2	31.2	3.9	58.4	171.1	35.2
TAGATCTA	d	5.1	49.2	142.2	33.4	4.2	45.9	135.9	24.5
TCTATAGA	d	4.3	45.7	133.4	28.1	4.2	45.9	135.9	24.5
ATGAGCTCAT	o	10.0	68.0	187.0	58.1	9.5	66.5	184.7	54.4
TTTTATAATAAA/AAAATATTATTT	p	5.5	75.6	226.0	36.3	5.8	81.5	240.6	41.6
CAACTTGATATTATTA/GTTGAACTATAATAAT	q	12.4	102.0	289.0	58.8	12.7	108.8	310.2	58.1
		Molecules with Non-Two-State Thermodynamics							
CCCGGG	r	6.9	62.0	177.8	43.0	7.0	36.9	95.2	52.0
CCCAGGG/GGGTCCC	r	7.9	63.0	177.8	48.1	7.8	40.3	103.7	57.2
CGCGAATTCGCG	s	20.6	135.0	369.0	75.4	16.4	101.3	272.7	75.0
CGCATGGGTACGC/GCGTACCCATGCG	t	14.4	101.0	279.1	66.5	17.4	100.6	266.8	79.7
GTATACCGGTATAC		12.3	88.0	244.1	62.2	13.0	96.1	267.6	63.0
CATATTGGCCAATATG		13.2	88.8	243.6	65.9	15.3	109.5	303.5	67.1
GTATAACCGGTTATAC		14.3	99.7	275.4	66.3	15.0	112.9	314.8	65.8
CGCGTACGCGTACGCG	u	29.1	158.0	415.6	91.0	24.1	140.9	374.5	85.6

[a] Listed by oligomer length and in alphabetical order. For self-complementary sequences only the top strand is given. For non-self-complementary sequences both strands are given in antiparallel orientation separated by a slash. [b] Sequences without a literature reference are from Table 1 of this work. [c] Calculated for 10^{-4} M oligomer concentration for self-complementary sequences and 4×10^{-4} M for non-self-complementary sequences. [d] Sugimoto et al. (1994). [e] Senior et al. (1988). [f] Breslauer (1986). [g] Williams et al. (1989). [h] Li and Agrawal (1995). [i] Aboul-ela et al. (1985). [j] Morden et al. (1983). [k] Breslauer et al. (1986). [l] LeBlanc and Morden (1991). [m] For each 5′-terminal T-A base pair, +0.4 kcal/mol is added to both $\Delta H°$ and $\Delta G°_{37}$. [n] Sugimoto et al. (1991). [o] Li et al. (1991). [p] Bolewska et al. (1984). [q] Tibanyenda et al. (1984). [r] Arnold et al. (1987). [s] Marky et al. (1983). [t] Plum et al. (1992). [u] Raap et al. (1985).

Table 3: Thermodynamic Parameters for DNA Helix Initiation and Propagation in 1 M NaCl[a]

propagation sequence	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	$\Delta G°_{37}$ (kcal/mol)
AA/TT	-8.4 ± 0.7	-23.6 ± 1.8	-1.02 ± 0.04
AT/TA	-6.5 ± 0.8	-18.8 ± 2.3	-0.73 ± 0.05
TA/AT	-6.3 ± 1.0	-18.5 ± 2.6	-0.60 ± 0.05
CA/GT	-7.4 ± 1.1	-19.3 ± 2.9	-1.38 ± 0.06
GT/CA	-8.6 ± 0.7	-23.0 ± 2.0	-1.43 ± 0.05
CT/GA	-6.1 ± 1.2	-16.1 ± 3.3	-1.16 ± 0.07
GA/CT	-7.7 ± 0.7	-20.3 ± 1.9	-1.46 ± 0.05
CG/GC	-10.1 ± 0.9	-25.5 ± 2.3	-2.09 ± 0.07
GC/CG	-11.1 ± 1.0	-28.4 ± 2.6	-2.28 ± 0.08
GG/CC	-6.7 ± 0.6	-15.6 ± 1.5	-1.77 ± 0.06
initiation at G·C[b]	(0)	(-5.9 ± 0.8)	$+1.82 \pm 0.24$
initiation at A·T[c]	(0)	(-9.0 ± 3.2)	$(+2.8 \pm 1)$
symmetry correction[d]	0	-1.4	$+0.4$
5′-terminal T·A bp[e]	$+0.4$	0	$+0.4$

[a] Errors are standard deviations. Extra significant figures are given for $\Delta H°$ and $\Delta S°$ to allow accurate calculation of the T_M. Values in parentheses involve assumptions about the initiation process (see text). [b] Initiation parameter for duplexes that contain at least one G·C base pair. [c] Initiation parameter for duplexes that contain only A·T base pairs. [d] Symmetry correction applies only to self-complementary sequences. [e] To account for end effects, duplexes are given the penalty listed for each terminal 5′-T·A-3′ base pair. Note this penalty is not applied to sequences with terminal 5′-A·T-3′ base pairs (see text).

describe DNA thermal denaturation.

Applicability to Non-Two-State Transitions. We believe the parameters in Table 3 apply to duplexes from 4 to 20 base pairs. Beyond 20 base pairs, DNA transitions are unlikely to be two-state. Transitions that are not two-state require a statistical mechanical model for accurate predictions (Gralla & Crothers, 1973; Steger, 1994). Table 2 lists eight oligonucleotides that are not two-state. When the $\Delta G°_{37}$, $\Delta H°$, and T_M of these oligomers are predicted with the parameters in Table 3, the average deviations of measured versus predicted values are 14%, 21%, and 4.7 °C, respectively. This suggests that the two-state model can also provide reasonable approximations for oligomers that do not have strictly two-state transitions.

DISCUSSION

Application of the Nearest-Neighbor Parameters. The nearest-neighbor model asserts that the free-energy for duplex formation is the sum of three terms: (1) an unfavorable entropy associated with the loss of translational freedom upon formation of the first hydrogen bonded base pair (i.e., the initiation free energy), (2) the sum of terms for the pairwise interactions between base pairs, and (3) an entropic penalty (Bailey & Monahan, 1978; Cantor & Schimmel, 1980) for the maintenance of the $C2$ symmetry of self-complementary duplexes (eq 4). Scheme 1 illustrates the calculation of $\Delta G°_{37}$ for the sequence GCTAGC using the parameters in Table 3. Similarly, the predicted enthalpy change for GCTAGC is: $\Delta H°$(predicted) $= 2(-11.1) + 2(-6.1) + (-6.3) = -40.7$ kcal/mol. The measured value is -39.2 kcal/mol. Note that $\Delta H°$ for initiation and symmetry are zero. The predicted entropy change for GCTAGC is $\Delta S°$(predicted) $= 2(-28.4) + 2(-16.1) + (-18.5) - 5.9 - 1.4 = -114.8$ eu. The measured value is -109.1 eu. The T_M is predicted at a given oligonucleotide concentration ($C_T/4$ for non-self-complementary sequences) using eq 2 along with the predicted $\Delta H°$ and $\Delta S°$. For example, the predicted T_M for GCTAGC at 0.1 mM is T_M (predicted) $= (-40\,700$ cal/

Scheme 1. Prediction of $\Delta G°_{37}$

```
           ↓   ↓   ↓
    5′ G-C-T-A-G-C 3′
       * * * * * *
    3′ C-G-A-T-C-G 5′
           ↑   ↑
```
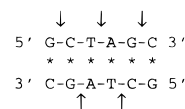
$$\Delta G°_{37}(\text{predicted}) = 2\,\Delta G°_{37}(GC/CG) + 2\,\Delta G°_{37}(CT/GA) + \Delta G°_{37}(TA/AT)$$
$$+ \Delta G°_{37}(\text{init}) + \Delta G°_{37}(\text{sym})$$
$$= 2\,(-2.28) + 2(-1.16) + (-0.60) + 1.82 + 0.4$$
$$\Delta G°_{37}(\text{predicted}) = -5.26 \text{ kcal/mol.}$$
$$\Delta G°_{37}(\text{observed}) = -5.3 \text{ kcal/mol.}$$

mol)/$(-114.8$ eu $+ 1.987$ eu $\times \ln (1 \times 10^{-4})) = 305.8$ K $= 32.6$ °C. The measured value is 34.3 °C. Note that the units for $\Delta H°$ are kcal/mol and must be multiplied by 1000 to be consistent with $\Delta S°$ and the gas constant, R, which are in eu [cal/(K mol)].

Trends in the Nearest-Neighbor Parameters. The observed trend in nearest-neighbor stabilities at 37 °C is GC > CG > GG > GA $\approx$ GT $\approx$ CA > CT > AA > AT > TA (where only the top strand is shown for each nearest neighbor). This trend suggests that both sequence and base composition are important determinants of DNA duplex stability. It has long been recognized that DNA stability depends of the percent G-C content (Marmur & Doty, 1962). The $\Delta G°_{37}$ parameters in Table 3 show that there are significant sequence dependent contributions superimposed on the general trend. On the other hand, the nearest-neighbor $\Delta H°$ parameters (Table 3) do not follow this trend. This suggests that stacking, hydrogen bonding, and other contributions to the $\Delta H°$ have a complicated sequence dependence. Perhaps, this is not surprising since it is well known that the detailed structure of DNA is profoundly dependent on sequence (Callidine & Drew, 1984; Hunter, 1993).

The average of the $\Delta S°$'s for the 10 nearest neighbor propagations is -20.9 eu. The agrees reasonably well with the sequence independent value of -24.85 ± 1.74 eu/base pair derived from polymers dissolved in 0.075 M Na$^+$ (Delcourt & Blake, 1991). Our results are also consistent with a simplistic calculation of the conformational entropy (Cantor & Schimmel, 1980):

$$\Delta S°_{\text{conf}} = 2R \ln(3 \times 7 \times 2 \times 3 \times 3 \times 2) = -26.3 \text{ eu/base pair}$$

where the numbers inside the parentheses are the assumed number of possible conformations for the α, $\beta\gamma$ (together), δ, ϵ, ζ, and χ dihedral angles. The "2" in front of the gas constant is required because two residues must be constrained to form a base pair and propagate a helix. This calculation systematically overestimates $\Delta S°_{\text{conf}}$ because many of the possible conformations would have high energies associated with them (due to steric repulsion). A more rigorous calculation would weight each of the possible conformations with a Boltzmann factor. This calculation also neglects salt effects and hydrophobic contributions to stacking (Hunter, 1993).

Comparison with Previous DNA Nearest-Neighbor Parameters. Breslauer et al. (1986) derived nearest-neighbor parameters from a data set of 19 oligonucleotides (dissolved in 1 M NaCl) and nine polymers (dissolved in low salt with

results extrapolated to high salt) and assumed a value of 5.2 kcal/mol for helix initiation (Borer et al., 1974). Our results share some similarities but also differ significantly from those of Breslauer and co-workers. Except for the GG/CC nearest neighbor, the $\Delta H°$ values of Breslauer et al. are within 2 kcal/mol of those in Table 3 (for GG/CC Breslauer reports $\Delta H° = -11.0$ kcal/mol, whereas we observe -6.7 kcal/mol). The trend for the stabilities of nearest neighbors with only AT base pairs is also similar with AA/TT > AT/TA > TA/AT. On the other hand, the parameters of Breslauer et al. (1986) have the stability order CG/GC > GC/CG ≈ GG/CC, while GC/CG > CG/GC > GG/CC in our parameter set. Breslauer et al. also have CA/GT > AA/TT > GA/CT ≈ CT/GA > GT/CA, while GA/CT ≈ GT/CA ≈ CA/GT > CT/GA > AA/TT in our parameter set. On average, Breslauer's parameters predict the $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_M of the two-state molecules given in Table 2 with average deviations of 16%, 12%, 13%, and 6.0 °C, respectively. As discussed above, the parameters in Table 3 predict the $\Delta G°_{37}$, $\Delta H°$, $\Delta S°$, and T_M of the two-state molecules given in Table 2 with average deviations of 4%, 7%, 8%, and 1.8 °C, respectively. This comparison is somewhat biased since our parameters were optimized to predict our database, while Breslauer's parameters were derived from an independent data set. Quartin and Wetmur (1989) use essentially the same data as Breslauer et al. (1986) but assume a value of $+2.2$ kcal/mol for helix initiation (Pohl, 1974). These parameters predict the $\Delta G°_{37}$ and the T_M of the two-state molecules in Table 2 with average deviations of 10% and 4.5 °C, respectively.

Comparison with RNA Nearest-Neighbor Parameters. Our DNA parameters also differ significantly from RNA parameters measured by Freier et al. (1986a). This is not surprising, however, because RNA and DNA helices are known to have different structures (i.e., A-form vs B-form). Some DNA nearest neighbors are more stable while others are less stable than the analogous RNA nearest neighbors. For example, the DNA nearest neighbors AA/TT and CG/GC are slightly more stable than the corresponding RNA nearest neighbors AA/UU and CG/GC. In all the other cases, DNA nearest neighbors are less stable than RNA. We also observe that DNA nearest neighbors with only C-G base pairs are less sequence dependent (largest difference = 0.51 kcal/mol) than the corresponding RNA nearest neighbors (largest difference 1.4 kcal/mol). The relative order for stabilities of nearest neighbors with only A-T (DNA) or A-U (RNA) base pairs are also different with AA/TT > AT/TA > TA/AT (DNA) vs UA/AU > AU/UA = AA/UU (RNA). Apparently, the relative stability of DNA and RNA duplexes depends on base sequence.

The helix initiation parameter is more favorable in DNA ($+1.82$ kcal/mol) than in RNA ($+3.4$ kcal/mol). This is somewhat puzzling as DNA and RNA duplex initiation are expected to be similar since both require two strands to associate and reduce their translational and rotational degrees of freedom by forming a hydrogen-bonded base pair. Further work is required to elucidate the origin of this effect.

Salt Dependence. The parameters in Table 3 apply to oligomers dissolved in 1 M NaCl at pH 7. To allow for approximate predictions in lower salt environments, we suggest the following preliminary equation:

$$T_M^{[Na]} = T_M^{1M\ Na} + 12.5 \log [Na^+] \qquad (7)$$

where $T_M^{1M\ Na}$ is T_M predicted from Table 3 (1 M NaCl), and $T_M^{[Na]}$ is the T_M predicted at the desired sodium concentration. This correction for the T_M is in agreement with that determined previously (Erie et al., 1987; Rentzeperis et al., 1993) for oligonucleotides but is somewhat smaller than that observed in polymers (Marmur & Doty, 1962; Schildkraut & Lifson, 1965). Between 0.1 and 1 M NaCl, this correction predicts the T_M of 26 sequences (dissolved in 0.1−0.3 M NaCl) from the literature (Aboul-ela et al., 1985; Braunlin & Bloomfield, 1991; Gaffney & Jones, 1989; Kawase et al., 1986; Williams et al., 1989; Lesnik & Freier, 1995) with an average deviation 3.5 °C (Allawi and SantaLucia, unpublished results). Below 0.1 M, this correction is not reliable. This correction assumes that trends in nearest-neighbor stability are independent of salt concentration. Counterion-condensation theory suggests this assumption is reasonable since the salt behavior depends on the spacing between phosphates which should be relatively independent of sequence (Manning, 1978). However, this theory applies strictly to polymers and salt concentration below 0.1 M, and, for short oligonucleotides, the salt behavior may depend on oligonucleotide length (Record & Lohman, 1978; Olmsted et al., 1989). Two experimental studies, however, suggest that the salt behavior of oligomers is remarkably similar to that of polymers (Williams et al., 1989; Braunlin & Bloomfield, 1991). While the above corrections were derived for sodium counterions, potassium counterions probably follow the same trend. The behavior of oligonucleotide thermodynamics in the presence of divalent cations, however, is likely to be more complicated. Previous work indicates that 1 M NaCl mimics 0.15 M NaCl/10 mM $MgCl_2$ (Williams et al., 1989)—a condition similar to those commonly used in PCR reactions. Clearly, further work on the salt dependence of oligonucleotide thermal denaturation is required (Kumar, 1995).

ACKNOWLEDGMENT

SUPPORTING INFORMATION AVAILABLE

Six figures showing $1/T_M$ vs ln C_T plots for the 20 sequences presented in Table 1 which are not shown in Figure 1 (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Aboul-ela, F., Koh, D., Tinoco, I., Jr., & Martin, F. H. (1985) *Nucleic Acids Res. 13*, 4811−4824.

Aida, M. (1988) *J. Theor. Biol. 130*, 327−335.

Albergo, D. D., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1409−1413.

Arnold, F. H., Wolk, S., Cruz, P., & Tinoco, I., Jr. (1987) *Biochemistry 26*, 4068−4075.

Bailey, W. F., & Monahan, A. S. (1978) *J. Chem. Ed. 55*, 489−493.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, pp 164−186 and 187−203, McGraw-Hill, New York.

Bolewska, K., Zielenkiewicz, A., & Wierzchowski, K. L. (1984) *Nucleic Acids Res. 12*, 3245−3256.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.

Braunlin, W. H., & Bloomfield, V. A. (1991) *Biochemistry 30*, 754−758.

Breslauer, K. J. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology* (Hinz, H., Ed.) pp 402−427, Springer-Verlag, New York.

Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746−3750.

Brown, T., & Brown, D. J. S. (1991) in *Oligonucleotides and Analogous* (Eckstein, F., Ed.) pp 1−24, IRL Press, New York.

Callidine, C. R., & Drew, H. R. (1984) *J. Mol. Biol. 178*, 773−782.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*, pp 1183−1264, W. H. Freeman, San Francisco, CA.

Chou, S.-H., Flynn, P., & Reid, B. (1989) *Biochemistry 28*, 2422−2435.

Delcourt, S. G., & Blake, R. D. (1991) *J. Biol. Chem. 266*, 15160−15169.

Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., & Benight, A. S. (1992) *Biopolymers 32*, 849−864.

Doktycz, M. J., Morris, M. D., Dormady, S. J., Beattie, K. L., & Jacobson, K. B. (1995) *J. Biol. Chem. 270*, 8439−8445.

Erie, D., Sinha, N., Olson, W., Jones, R., & Breslauer. K. (1987) *Biochemistry 26*, 7150−7159.

Fodor, S. P. A., Rava, R. P., Huang, X. C., Pease, A. C., Holmes, C. P., & Adams, C. L. (1993) *Nature 364*, 555−556.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986a) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., & Turner, D. H. (1986b) *Biochemistry 25*, 3214−3219.

Freier, S. M. (1993) in *Antisense Research and Applications* (Crooke, S. T., & Lebleu, B., Eds.) pp 67−82, CRC Press, Boca Raton, FL.

Gaffney, B. L., & Jones, R. A. (1989) *Biochemistry 28*, 5881−5889.

Gotoh, O., & Tagashira, Y. (1981) *Biopolymers 20*, 1033−1042.

Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.

Gray, D. M., & Tinoco, I., Jr. (1970) *Biopolymers 9*, 223−244.

Hunter, C. A. (1993) *J. Mol. Biol. 230*, 1025−1054.

Kawase, Y., Iwai, S., Inoue, H., Miura, K., & Ohtsuka, E. (1986) *Nucleic Acids Res. 14*, 7727−7736.

Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., & Freier, S. M. (1986) *Biochemistry 25*, 7840−7846.

Klump, H. H. (1990) in Landolt-Bornstein, New series, VII *Biophysics*, Vol. 1, *Nucleic Acids*, Subvol. c, *Spectroscopic and Kinetic Data*, Physical Data I. (Saenger, W., Ed.) pp 241−256, Springer-Verlag, Berlin.

Kumar, A. (1995) *Biochemistry 34*, 12921−12925.

LeBlanc, D. A., & Morden, K. M. (1991) *Biochemistry 30*, 4042−4047.

Lesnik, E. A., & Freier, S. M. (1995) *Biochemistry 34*, 10807−10815.

Li, Y., & Agrawal, S. (1995) *Biochemistry 34*, 10056−10062.

Li, Y., Zon, G., & Wilson, W. D. (1991) *Biochemistry 30*, 7566−7572.

Manning, G. (1978) *Q. Rev. Biophys. 11*, 179−246.

Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.

Marky, L. A., Blumenfeld, K. S., Kozlowski, S., & Breslauer, K. J. (1983) *Biopolymers 22*, 1247−1257.

Mendelman, L. V., Boosalis, M. S., Petruska, J., & Goodman, M. F. (1989) *J. Biol. Chem. 264*, 14415−14423.

Marmur, J., & Doty, P. (1962) *J. Mol. Biol. 5*, 109−118.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1983) *Biochemistry 21*, 428−436.

Olmsted, M. C., Anderson, C. F., & Record, M. T., Jr. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7766−7770.

Ornstein, R., & Fresco, J. R. (1983) *Biopolymers 22*, 1979−2000.

Otto, P. (1989) *J. Mol. Struct. 188*, 277−288.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256−263.

Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. S., Cheong, C., & Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252−6256.

Plum, G. E., Grollman, A. P., Johnson, F., & Breslauer, K. J. (1992) *Biochemistry 31*, 12096−12102.

Pohl, F. M. (1974) *Eur. J. Biochem. 42*, 495−504.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1989) *Numerical Recipes*, pp 52−64 and 498−520, Cambridge University Press, New York.

Quartin, R. S., & Wetmur, J. G. (1989) *Biochemistry 28*, 1040−1047.

Raap, J., van der Marel, G. A., van Boom, J. H., Joordens, J. J. M., & Hilbers, C. W. (1985) in *Fourth Conversation in Biomolecular Stereodynamics* (Sarma, R. H., Ed.) p 122a, Adenine Press, Guilderland, NY.

Record, M. T., Jr., & Lohman, T. M. (1978) *Biopolymers 17*, 159−166.

Rentzeperis, D., Ho, J., & Marky, L. A. (1993) *Biochemistry 32*, 2564−2572.

Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 597, CRC Press, Cleveland, OH.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S., Higuchi, R. H., Horn, G. T., Mullis, K. B., & Erlich, H. A. (1988) *Science 239*, 487−494.

SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1990) *Biochemistry 29*, 8813−8819.

Scheffler, I. E., Elson, E. L., & Baldwin, R. L. (1970) *J. Mol. Biol. 48*, 145−171.

Schildkraut, C., & Lifson, S. (1965) *Biopolymers 3*, 195−208.

Senior, M., Jones, R. A., & Breslauer, K. J. (1988) *Biochemistry 27*, 3879−3885.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503−517.

Steger, G. (1994) *Nucleic Acids Res. 22*, 2760−2768.

Sugimoto, N., Tanaka, A., Shintani, Y., & Sasaki, M. (1991) *Chem. Lett.* 9−12.

Sugimoto, N., Honda, K., & Sasaki, M. (1994) *Nucleosides Nucleotides 13*, 1311−1317.

Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yonegama, M., & Sasaki, M. (1995) *Biochemistry 34*, 1121−11216.

Tibanyenda, N., De Bruin, S. H., Haasnoot, C. A. G., van der Marel, G. A., van Boom, J. H., & Hilbers, C. W. (1984) *Eur. J. Biochem. 139*, 19−27.

Turner, D. H., Sugimoto, N., & Freier, S. M. (1990) in Landolt-Bornstein, New series, VII *Biophysics*, Vol. 1, *Nucleic Acids*, Subvol. c, *Spectroscopic and Kinetic Data*, Physical Data I. (Saenger, W., Ed.) pp 201−227, Springer-Verlag, Berlin.

Vologodskii, A. V., Amirikyan, B. R., Lyubchenko, Y. L., & Frank-Kamenetskii, M. D. (1984) *J. Biomol. Struct. Dyn. 2*, 131−148.

Wartell, R. M., & Benight, A. S. (1985) *Phys. Rep. 126*, 67−107.

Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., & Turner, D. H. (1989) *Biochemistry 28*, 4283−4291.

Wolfram, S. (1992) *MATHEMATICA* version 2.1, Wolfram Research, Inc.