Thermodynamic Characterization of Deoxyribooligonucleotide Duplexes Containing Bulges†

Darryl A. LeBlanc and Kathleen M. Morden*
Department of Biochemistry, Louisiana State University, Baton Rouge, Louisiana 70803
Received October 22, 1990; Revised Manuscript Received January 16, 1991

ABSTRACT: Ultraviolet absorption techniques were used to study the thermodynamics of duplex formation for a DNA decamer, d(GCGAAAAGCG)·d(CGCTTTTCGC), and a series of related duplexes, each of which contains a bulged base centered in the A·T tract. Thermodynamic parameters were obtained from nonlinear least-squares fits of the melting curves and the concentration dependences of the melting temperatures. Duplexes containing a localized single-base bulge were found to be 3.5–4.6 kcal/mol less stable than the decamer at 37 °C. These results indicate that both the identity of the bulged base and the strand in which it is located may influence the amount by which the duplex is destabilized. Bulged bases located in the T-strand, d(CGCTTYTTCGC), in position Y, were observed to be slightly more destabilizing than those located in the A-strand, d(GCGAAXAAGCG), in position X. Bulged purines may be more destabilizing than bulged pyrimidines.

The availability of synthetic oligonucleotides and the advancement of techniques applicable to the study of nucleic acid structure have led to a wealth of detailed structural information on the various families of nucleic acid helices. With this foundation, we can begin to explore the effects of various helical perturbations. Of particular interest are perturbations in nucleic acid helices where only one of the two strands contains an unpaired base or bases. This perturbation is referred to as a bulge, and the unpaired base is commonly known as a bulged base. A clear understanding of the factors which modulate the stability of helices with bulged bases would greatly facilitate research in several areas. Such knowledge is vital to testing Streisinger's model for the mechanism of frameshift mutation (Streisinger et al., 1966). It will also aid in the design of oligonucleotide probes by allowing the identification of alternative binding sites generated through the incorporation of one or more bulged bases. Many investigations aimed at bulges in DNA have utilized nuclear magnetic resonance (NMR) (Haasnoot et al., 1980; Patel et al., 1982; Morden et al., 1983, 1990; Hare et al., 1986; Roy et al., 1987; Woodson & Crothers, 1987, 1988a,b, 1989; van den Hoogen et al., 1988; Kalnik et al., 1989a,b, 1990; Nikonowicz et al., 1989) or X-ray diffraction (Joshua-Tor et al., 1988; Miller et al., 1988) techniques to obtain highly detailed structural information about the helical deformations introduced by the presence of the bulged base. However, a complete characterization of the effects that a bulged base has on duplex stability also requires an investigation of the global thermodynamic properties of duplex formation.

Ultraviolet absorption techniques can be utilized to obtain thermodynamic parameters for the process of duplex formation (Marky & Breslauer, 1987). The stability of DNA duplexes has been investigated in this manner with the goal of predicting duplex stability based upon nucleotide sequence (Breslauer et al., 1986). UV absorption techniques have also been used to estimate the destabilizing free energies of bulges in both DNA (Morden et al., 1983; Woodson & Crothers, 1987) and RNA (Fink & Crothers, 1972; Longfellow et al., 1990). We describe below the thermodynamic characterization of a duplex DNA decamer and a series of related oligomers, each of which contains a bulged base centered in an A·T tract as shown in Scheme I. The melting characteristics of two individual single strands are also briefly discussed.

Scheme I

5'GCGAAAAGCG3' 3'CGCTTTTCGC5'	decamer
5'GCGAAXAAGCG3' 3'CGCTT-TTCGC5'	A-strand bulges, X = C, T or G
5'GCGAA-AAGCG3' 3'CGCTTYTTCGC5'	T-strand bulges, Y = C, A or G

Experimental Procedures

Materials

Synthetic DNA oligomers purified by semipreparative anion-exchange HPLC were purchased from the Midland Certified Reagent Co. (Midland, TX) as lyophilized single strands. Purity was confirmed in our laboratory by analytical strong anion-exchange HPLC. Oligomers were suspended in deionized, glass-distilled H_2O and used without additional purification.

Methods

Melting Curves. UV absorbance at 260 nm was monitored as a function of temperature from 5 to 95 °C for the thermal denaturation of the DNA oligomers. Experiments were performed using a Gilford Response II spectrophotometer. Absorbance readings were taken every 0.5 °C. The temperature of the cuvettes was controlled electronically by dual Peltier devices within the cuvette holder. The spectral bandwidth was 0.5 nm. All experiments were performed in 10 mM phosphate buffer, pH 7, containing 1.0 M NaCl and 0.1 mM ethylenediaminetetraacetic acid (EDTA).[1] Appropriate individual strands in distilled H_2O at concentrations of 5–10 μM were mixed to form a dilute solution of the desired duplex, which

† This work was supported by National Institutes of Health Grant GM 38137 and Louisiana Education Quality Support Fund LEQSF(86-89)-RD-A-12. D.A.L. was supported by a fellowship from the U.S. Department of Agriculture (Grant 87-GRAD-9-0091).

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; NLS, nonlinear least squares.

0006-2960/91/0430-4042$02.50/0

was then lyophilized and resuspended to generate a concentrated DNA stock solution. Samples for each melt were prepared by using the DNA stock, a concentrated buffer stock, and degassed, deionized, glass-distilled H_2O. Each duplex was studied at seven different concentrations ranging from 10 μM to 1 mM single strand. Each single strand was also studied independently, in the absence of the complementary strand, at several concentrations ranging from 5 to 500 μM. Measurements were made by using cuvettes with path lengths ranging from 0.01 to 1.0 cm, so that measured absorbances were between 0.6 and 1.8. The experiments were conducted according to the procedures described by Nelson et al. (1981). Absorbances obtained at 5 and 25 °C before and after each experiment were compared; samples which showed a greater than 2% increase in absorption (due to evaporation) were rejected and repeated. Extinction coefficients at 25 °C were calculated by nearest-neighbor approximation (Fasman, 1975). Extinction coefficients in units of mM^{-1} cm^{-1} at 25 °C are as follows: dGCGAAAAGCG, 104; dGCGAACAAGCG, 112; dGCGAATAAGCG, 114; dGCGAAGAAGCG, 116; dCGCTTTTCGC, 80.4; dCGCTTATTCGC, 94.4; dCGCTTGTTCGC, 91.1; dCGCTTCTTCGC, 87.6. Concentrations of mixtures of complementary strands were calculated from absorbances measured at 85 °C, where all strands exist only in the single-stranded state. Extinction coefficients at 85 °C were calculated by ϵ(85 °C) = ϵ(25 °C)[A(85 °C)/A(25 °C)] (Albergo et al., 1981).

Data Analysis. The absorbance versus temperature plots were normalized to an absorbance of 1.00 at 85 °C and were fit by a nonlinear least-squares (NLS) routine using the method of Marquardt (Bevington, 1969) on a MicroVAX II computer. The extinction coefficients for the double- and single-stranded states (ϵ_{ds} and ϵ_{ss}, respectively) are described by the equations:

$$\epsilon_{ds}(T) = b_{ds} + m_{ds}T \quad (1)$$

$$\epsilon_{ss}(T) = b_{ss} + m_{ss}T + C_1T^2 + C_2T^3 \quad (2)$$

where T is the temperature in degrees kelvin and b and m represent the intercepts and slopes of the lines describing the temperature dependencies of ϵ_{ds} and ϵ_{ss} (the lower and upper base lines, respectively). The constants C_1 and C_2 are introduced to account for the nonlinear nature of the temperature dependence of the single-strand extinction coefficient, ϵ_{ss}. Equation 2, in effect, describes the temperature dependence of the combined extinction coefficients of the two single strands in solution *in the absence of duplex formation.* Because our oligomers are non-self-complementary, the temperature dependence of the extinction coefficient for each strand can be determined experimentally. The upper base line required for the NLS fits of the melting curves for a particular duplex is obtained by summing the melting curves of the appropriate two individual strands and then normalizing to an absorbance of 1.00 at 85 °C. Upper base lines for all duplexes were derived in this manner. However, two of the single strands (dGCGAAGAAGCG and dGCGAATAAGCG) exhibit concentration-dependent hypochromism. The lowest concentrations were used to derive the single-strand (upper) base lines, and hypochromism was observed only below 35 °C. A similar oligomer, dGCGAACAAGCG, does not display hypochromism under these conditions. Therefore, the behavior of dGCGAAGAAGCG and dGCGAATAAGCG at temperatures below 35 °C in the absence of hypochromism was approximated based on the melting curve of dGCGAACAAGCG. Each of the corrected curves was combined with the melting curve of dCGCTTTTCGC to produce the upper base lines for the NLS fits of the melting curves of the A-strand G-bulge and A-strand T-bulge. The approximation resulted in only slight changes (~1%) in the thermodynamic parameters obtained, probably because the melting temperatures of the heteroduplexes are well above 35 °C.

The extinction coefficient (ϵ) and absorbance (A) of a solution at any temperature T can be calculated as

$$\epsilon(T) = A(T)/LC_T = \epsilon_{ss}(T) - \alpha[\epsilon_{ss}(T) - \epsilon_{ds}(T)] \quad (3)$$

where L is the path length of the cell, C_T is the total single-strand concentration, α is the fraction of strands in the duplex state, and ϵ_{ss} and ϵ_{ds} are given in eq 1 and 2. The term α is related to the equilibrium constant K for the association of two non-self-complementary strands to form a dimer by

$$K = \frac{\alpha}{(C_T/2)(1-\alpha)^2} = \exp(-\Delta H^\circ/RT + \Delta S^\circ/R) \quad (4)$$

where, again, C_T is the total single-strand concentration with each strand present in an equal concentration of $C_T/2$, and α is the fraction of single strands in the duplex state. The parameters ΔH° and ΔS° are the changes in enthalpy and entropy, respectively, for the process of duplex formation (Marky & Breslauer, 1987). The six adjustable parameters for the nonlinear-least squares fits to eq 1–4 were ΔH°, ΔS°, m_{ss} b_{ss}, m_{ds}, and b_{ds}. Although m_{ss} and b_{ss} were variable parameters, significant variation from the experimental values was never observed, probably because the curvature of the single-strand base lines was maintained by the constants C_1 and C_2 (eq 2). The thermodynamic parameters ΔH° and ΔS° for a particular duplex were calculated by averaging the values obtained from the fits of the seven experimental melting curves. The free energies of duplex formation (ΔG°) were calculated by using the standard thermodynamic relationship described by eq 5. In cases such as this, when a variable is calculated

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (5)$$

from two highly correlated parameters, it can be determined with a high degree of accuracy, even if errors in the correlated parameters are large (Bevington, 1969). Errors in ΔH° and ΔS° from the NLS fits can be as large as 15%, but because variations in these parameters are highly correlated, ΔG° can be calculated quite accurately near the T_m, in most cases to within less than ±5%.

An alternative method for determining ΔH° and ΔS° is from the concentration dependence of the melting temperature, T_m which is defined as the point at which half of the strands exist in the duplex state. A plot of T_m^{-1} vs ln ($C_T/4$) is used to determine the thermodynamic parameters according to the equation:

$$T_m^{-1} = (R/\Delta H^\circ) \ln (C_T/4) + \Delta S^\circ/\Delta H^\circ \quad (6)$$

where R is the ideal gas constant and the other variables are as described above for eq 4. Melting temperatures were obtained from the NLS fits. In this case, errors in ΔH° and ΔS° reflect errors in the slope and intercept and are less than or equal to 5%. The free energy of duplex formation at a specific T_m can also be obtained (indirectly) from eq 6, by using the slope, intercept, and the chosen temperature to calculate the term ln ($C_T/4$). At the T_m, the equilibrium constant K is equal to $4/C_T$ (Marky & Breslauer, 1987), and ΔG° can be obtained from

$$\Delta G^\circ(T_m) = -RT_m \ln K \quad (7)$$

Results

Melting Curves and Thermodynamics. Experimental melting curves for the A-strand C-bulge at seven different

Table I: Thermodynamic Parameters for Helix Formation

	$-\Delta H°$ (kcal/mol)		$-\Delta S°$ (eu)		$-\Delta G°$(37 °C) (kcal/mol)		T_m (°C) at
	NLS[a]	T_m^{-1} vs ln (C_T/4)[b]	NLS	T_m^{-1} vs ln (C_T/4)	NLS	T_m^{-1} vs ln (C_T/4)	C_T = 100 μM
perfect duplex	76 ± 5	74 ± 2	207 ± 16	201 ± 7	11.9 ± 0.3	11.7 ± 0.4	60.7 ± 0.4
calculated[c]	86.7		215		15.1		
A-strand							
T-bulge	63 ± 7	57 ± 2	176 ± 23	157 ± 6	8.4 ± 0.5	8.1 ± 0.4	46.3 ± 0.4
C-bulge	64 ± 9	55 ± 1	179 ± 28	151 ± 3	8.4 ± 0.6	8.0 ± 0.2	45.7 ± 0.2
G-bulge	58 ● 6	54 ± 2	162 ± 17	150 ± 4	7.9 ± 0.6	7.8 ± 0.2	44.5 ± 0.3
T-strand							
C-bulge	56 ± 7	56 ± 1	155 ± 23	154 ± 2	8.0 ± 0.2	7.9 ± 0.1	44.6 ± 0.1
G-bulge	52 ± 7	48 ± 1	145 ± 22	132 ± 3	7.5 ± 0.1	7.3 ± 0.2	42.6 ± 0.3
A-bulge	54 ± 6	48 ± 1	149 ± 19	131 ± 3	7.3 ± 0.4	7.2 ± 0.2	41.2 ± 0.2

[a] Nonlinear least-squares fits. Standard errors from the NLS fits are ≤15% for $\Delta H°$ and $\Delta S°$ and <8% for $\Delta G°$. [b] Plots of T_m^{-1} vs ln (C_T/4). Standard errors from these plots are ≤5% for $\Delta H°$, $\Delta S°$, and $\Delta G°$. [c] Calculated for the process of helix formation for the decamer by the method of Breslauer et al. (1986).

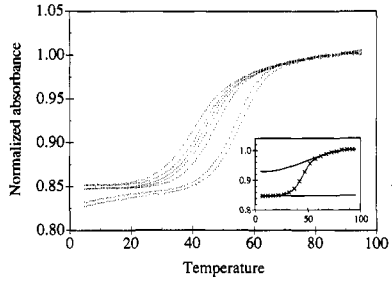

FIGURE 1: Melting curves for the A-strand C-bulge at seven different concentrations. Concentrations in units of micromolar duplex are, from left to right, 5.38, 11.7, 16.3, 23.9, 47.2, 303, and 542. Inset: One melting curve for the A-strand C-bulge (sigmoidal curve) including the upper and lower base lines used in the NLS fitting routine and the fit to the experimental data [every tenth point is represented by an (×)]. Axes are the same for the figure and the inset.

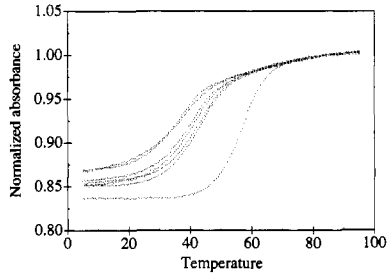

FIGURE 2: Melting curves for all seven duplexes at similar concentrations. Concentrations in micromolar duplex are, from left to right, T-strand G-bulge (21.4), T-strand A-bulge (19.4), A-strand G-bulge (18.4), T-strand C-bulge (22.0), A-strand T-bulge (21.0), A-strand C-bulge (23.9), and decamer (22.2).

concentrations are shown in Figure 1. Inset in Figure 1 are data for one concentration, along with the base lines and melting curve calculated from the NLS fit. Agreement between the experimental data and the calculated curve is excellent. For the A-strand C-bulge, the upper base line was obtained by summing the experimental melting curves for dGCGAACAAGCG and dCGCTTTTCGC. Melting curves for all seven duplexes at similar concentrations are shown in Figure 2. Figure 3 contains a plot of T_m^{-1} vs ln (C_T/4) for all seven duplexes. Thermodynamic parameters obtained from both methods are listed in Table I, along with calculated melting temperatures for all of the duplexes at 10^{-4} M single strand. Both methods of analysis assume that the process of duplex formation occurs in an all-or-none fashion, such that partially base-paired states are never significantly populated. We have also assumed that $\Delta H°$ and $\Delta S°$ are independent of temperature, i.e., that the specific heat capacities of the single- and double-stranded states are the same. If the above assumptions are valid, parameters obtained from the different

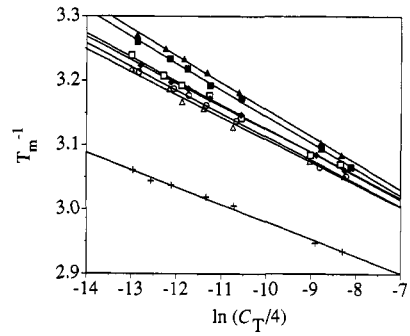

FIGURE 3: Plot of T_m^{-1} (×1000) vs ln (C_T/4) for all seven duplexes. Filled symbols represent T-strand bulges, and open symbols represent A-strand bulges: (▲) A-bulge; (■) G-bulge; (♦) C-bulge; (□) G-bulge; (○) C-bulge; (△) T-bulge; (+) decamer. Melting temperatures (T_m) are in degrees kelvin and concentrations (C_T) in moles of single strand per liter.

Table II: Change in $\Delta G°$(37 °C)[a]

	$\Delta\Delta G°$(37 °C) (kcal/mol)	
	NLS fits	$1/T_m$ vs ln(C_T/4)
A-Strand Bulges		
T-bulge	3.5	3.6
C-bulge	3.5	3.7
G-bulge	4.0	3.9
T-Strand Bulges		
C-bulge	3.9	3.8
G-bulge	4.4	4.4
A-bulge	4.6	4.5

[a] $\Delta\Delta G° = \Delta G°$(bulge) − $\Delta G°$(decamer).

analyses should be similar. Agreement of parameters from the two methods to within 15% is considered sufficient (Freier et al., 1986). In only one case, the A-strand T-bulge, is there greater than 15% disagreement. For the other six duplexes, values for $\Delta H°$ agree to within an average of 8% and values for $\Delta S°$ to within an average of 9%. In all cases, values for $\Delta G°$ agree to within 5%. However, we do observe a slight concentration dependence for the thermodynamic parameters derived from the NLS fits, which indicates that the system may not be absolutely two-state in nature. The small amount of non-two-state behavior that was observed may be due to aggregation or the existence of a slight difference in the specific heat capacities of the single- and double-stranded states. An absolute determination of the extent to which the system obeys the two-state model would require the use of a technique such as calorimetry to monitor an observable parameter other than UV absorbance. The amount of non-two-state behavior observed was small, indicating that such experiments are unwarranted. Also, optical studies are believed to provide accurate values for $\Delta G°$ for both two-state and non-two-state

Scheme II

```
5' GCGAAGAAGCG 3'   A-strand G-bulge
3'GCGAAGAAGCG  5'   possible homoduplex
```

transitions, even though the accuracies of $\Delta H°$ and $\Delta S°$ are somewhat lower (Longfellow et al., 1990).

The free energy differences for each of the bulges relative to the decamer [$\Delta\Delta G°$(37 °C)] are listed in Table II. All of the bulge-containing helices are destabilized relative to the decamer. Bulges located on the T-strand are slightly less stable than those located on the A-strand. Values for $\Delta\Delta G°$(37 °C) range from 3.5 to 4.6 kcal/mol. As has been observed for RNA (Groebe & Uhlenbeck, 1989), purine bulges may be marginally more destabilizing than pyrimidine bulges.

Homoduplex Formation. Two of the single strands, dGCGAAGAAGCG and dGCGAATAAGCG, exhibit significant hypochromism and sigmoidal melting behavior. These phenomena are attributed to homoduplex formation. The possibility of hairpin formation is ruled out by an observed concentration dependence. The homoduplex which may be formed by dGCGAAGAAGCG is shown in Scheme II and contains four G·A and two A·A mismatches. NMR evidence (unpublished results) indicates that the homoduplex has the 3′ overhang shown and not the (anticipated) 5′ overhang. The analogous homoduplex formed by dGCGAATAAGCG would contain two A·T base pairs in place of two of the G·A mismatches. Homoduplexes involving G·A mismatches have been observed by others (Li et al., 1991; Longfellow et al., 1990; Wilson et al., 1988). Longfellow et al. (1990) found that the presence of the complementary strand greatly shifts the equilibrium toward the formation of the heteroduplex containing a bulge, rather than the formation of a homoduplex, even when the melting temperatures are similar; thus, formation of the homoduplex appears to be a complicating factor only when its melting temperature is significantly higher than that of the heteroduplex. Melting temperatures for both $(dGCGAAGAAGCG)_2$ and $(dGCGAATAAGCG)_2$ are at least 15 °C below those of the respective heteroduplexes formed with dCGCTTTTCGC. At the conditions under which the single strands were mixed to form the heteroduplexes (see Melting Curves under *Methods*), the T_m of the homoduplex is almost certainly well below room temperature. Heating to temperatures above the T_m of the heteroduplex followed by slow reannealing prior to performing the melting experiments did not alter the thermodynamic parameters obtained. This result is consistent with the absence of significant homoduplex formation.

Discussion

We have characterized the thermodynamic parameters for the formation of a decanucleotide duplex and a series of related duplexes, each of which contains a bulged base in one of the two strands and centered within an A·T tract. The information obtained is pertinent to several important areas of research. Streisinger et al. (1966) have proposed a model for the mechanism of frameshift mutation in which these mutations arise in DNA through strand slippage in areas of repeating base pairs to produce a bulged base. The model explains the mutagenicity of intercalative molecules by suggesting that they stabilize the bulged bases. Thus, prior to testing this theory, it is necessary to understand the ramifications of introducing a bulged base and how these effects are modulated by the identity of the bulged base and surrounding sequence. Results obtained in this and similar investigations can also be used in the design of oligonucleotide probes and should allow the identification of possible alternative hybridization sites generated through the incorporation of one or more bulged bases.

The stability of nucleic acid helices is determined by a number of different factors that can be categorized as either enthalpic or entropic. The formation of a duplex in solution is a bimolecular association reaction which increases the order of the bases and is accompanied by an uptake of counterions (Cantor & Schimmel, 1980). All of these effects are entropically unfavorable. The increase in electrostricted water and the decrease in hydrophobically bound water that accompany duplex formation are approximately compensatory (Noguchi et al., 1971). Thus, the process of duplex formation is opposed by an overall negative (unfavorable) change in entropy. At physiological temperatures, the $T\Delta S$ term of eq 5 is overcome by large favorable enthalpic forces, resulting in an overall negative (favorable) change in free energy for the process of duplex formation. The enthalpic forces that drive duplex formation under these conditions arise mainly from the favorable phenomenon of base stacking (Saenger, 1983). Duplexes are also stabilized by the association of counterions with the negatively charged phosphate groups and hydration of the phosphates (both of which reduce electrostatic repulsions). Hydration of polar groups in the major and minor grooves of the helix contributes to duplex stabilization as well (Kopka et al., 1983).

Because the bulges that we have studied are located in the center of an A·T tract, we must also consider specific properties known to be associated with runs of several A·T base pairs. It has long been known that poly(dA)·poly(dT) is more stable than poly[d(AT)]·poly[d(AT)] (Inman & Baldwin, 1962; Riley et al., 1966). Recently, tracts of four or more A·T base pairs have been shown to exhibit several properties which may account for this increased stability (Coll et al., 1987; Nelson et al., 1987). These include a narrow minor groove containing an ordered spine of hydration and a high degree of propeller twisting which maximizes base stacking interactions. The large amount of propeller twisting may be stabilized by the formation of a network of bifurcated hydrogen bonds between an adenine N6 amine and O4 of the thymine on the opposite strand in the adjacent base pair (Coll et al., 1987; Nelson et al., 1987; Yoon et al., 1988). NMR results obtained in our laboratory show that the decamer, and to a lesser extent the A-strand T-bulge, exhibits spectroscopic properties characteristic of A·T tracts (Morden et al., 1990).

In all cases, the presence of a bulged base significantly destabilizes the duplex. The difference between the free energies of duplex formation for the decamer and for a duplex containing a bulged base ($\Delta\Delta G°$) is a measure of this destabilization. Values for $\Delta\Delta G°$(37 °C) range from 3.5 to 4.6 kcal/mol (Table II), depending on the identity of the bulged base and in which strand the bulge is located. Morden et al. (1983) reported a value of $\Delta\Delta G°$(25 °C) = 2.9 kcal/mol for the extrahelical bulged cytosines in the duplex d(CAAACAAAG)·d(CTTTTTTG). Our value of 3.5 kcal/mol for $\Delta\Delta G°$(37 °C) for the A-strand C-bulge is in fair agreement. The melting temperatures of the duplexes currently under investigation are significantly higher than those of d(CAAACAAAG)·d(CTTTTTTG) and other short duplexes studied in the past. Direct comparison is therefore difficult, as errors in thermodynamic parameters obtained by extrapolation beyond the range of experimentally observed melting temperatures can be quite large.

The change in free energy of helix formation caused by the introduction of a bulged base reflects changes in both the $\Delta S°$ and the $\Delta H°$ of helix formation. The formation of a duplex that contains a bulge is entropically more favorable and en-

Scheme III

```
A|A        A|G|A
T|T        T   T
```

thalpically less favorable than the formation of the decamer (Table I). These results do not have unique origins. For both thermodynamic parameters ($\Delta S°$ and $\Delta H°$), the presence of a bulged base introduces both favorable and unfavorable effects which interact to yield the above results. Some of these effects are described below.

Each duplex that contains a bulge also contains an extra phosphate group, which will bind both counterion and water molecules, a negative entropic effect. This negative entropy contribution cannot be as important as might be expected, since the presence of a bulge is entropically favorable. Experiments that follow the stability of the duplexes as a function of salt concentration should aid in determining the importance of this effect and are currently underway. One source of the overall positive entropy change that is observed upon introducing a bulged base may be increased conformational freedom. Bulged bases can exist in two different conformations, intrahelical (stacked within the helix) and extrahelical. It should be kept in mind that references to an intrahelical or extrahelical bulged base in solution usually refer to equilibria in which the indicated conformation is not present exclusively but merely more heavily populated. One might anticipate that extrahelical bulged bases should experience greater conformational freedom than intrahelical bulged bases. However, molecular mechanics calculations performed in our laboratory (unpublished results) indicate that extrahelical bulged bases may be accommodated in the grooves of the helix in order to avoid highly unfavorable interactions between the nonpolar base and the solvent molecules. Therefore, the two conformations may exhibit similar degrees of conformational freedom. Another important factor that may serve to make the formation of a duplex containing a bulged base entropically more favorable than the formation of the decamer is the disruption of the spine of hydration in the minor groove. On the basis of studies in which netropsin was bound to the minor groove of poly(dA)·poly(dT) and poly[d(AT)]·poly[d(AT)], Marky and Kupke (1989) have shown that poly(dA)·poly(dT) is much more hydrated than poly[d(AT)]·poly[d(AT)]. Most of the large favorable (positive) entropy change that is associated with the binding of netropsin to poly(dA)·poly(dT) is attributed to the release of hydrating waters. Both extrahelical and intrahelical bulged bases could disrupt the minor groove spine of hydration to some extent.

The enthalpy of formation for a duplex that contains a bulged base is less favorable (more positive) than the enthalpy of formation for the decamer (Table I). Insertion of an intrahelical bulged base results in the loss of stacking interactions between the flanking bases on both strands. Base stacking interactions between the bulged base and the flanking bases on the same strand are gained, compensating for a portion of the stacking interactions lost. Base stacking interactions (represented by a vertical line) in the absence and presence of an intrahelical A-strand G-bulge are depicted in Scheme III. An extrahelical bulged base may influence base stacking interactions via distortions translated through the backbone of the duplex. Other factors which may serve to make the enthalpy of formation for a duplex containing a bulged base less favorable than that of the decamer include the displacement of hydrating water molecules and the interruption of the network of bifurcated hydrogen bonds associated with A·T tracts.

Because we have systematically varied the identity of the bulged base and the strand in which it is located, we can now begin to discuss the influence of the identity of the bulged base and the surrounding sequence on the relative stabilities of different single-base bulges. Bulges located in the A-strand, between two purines, were observed to be slightly more stable than those in the T-strand, between two pyrimidines. Papanicolaou et al. (1984) previously suggested that such a phenomenon should occur. In investigations of tRNA and 5S RNA secondary structure, their algorithm found the correct structure more often when a stabilization factor of 1.6 kcal/mol was added for single-base bulges located between two purines versus those located between two pyrimidines. The rationalization for this phenomenon was based on an intrahelical model for the bulged base. It was theorized that a bulged base stacked within the helix would be easier to accommodate if it required only the disruption of pyrimidine–pyrimidine stacking interactions on the opposite strand, rather than purine–purine stacking interactions. We have observed a smaller effect of 0.1–0.5 kcal/mol (Table II). Longfellow et al. (1990) did not observe this effect in their investigations of bulges located in duplex RNAs.

Bulges consisting of a single purine appear to have a slightly greater destabilizing effect relative to the decamer than bulges that consist of a single pyrimidine. This trend is observed for bulges in both the A- and T-strands (Table II). Groebe and Uhlenbeck (1989) have reported that this phenomenon is also seen for bulges located within the stem of an RNA hairpin. Thus, the effect is observed for two very different forms of nucleic acid helices. As mentioned above, the bulged bases may be accommodated in the grooves of the helix, and a portion of the destabilizing effects attributable to the bulged base may be due to the displacement of hydrating water molecules in the major and/or minor grooves. The slightly greater destabilizing effect of purines may be due to the displacement of a larger number of hydrating waters.

NMR experiments performed in our laboratory indicate that the A-strand G-bulge is the only one of the bulges under examination that exists primarily in an intrahelical conformation, which it adopts at both 25 and 37 °C under NMR conditions of 2 mM duplex and 0.1 M NaCl (unpublished results). All of the other bulges discussed here are believed to exist primarily in an extrahelical conformation (Morden et al., 1990; unpublished results). It might be anticipated that bulged bases having different conformations (intrahelical versus extrahelical) would be thermodynamically distinct from one another. However, the A-strand G-bulge, which is conformationally different from the other bulges, follows the same thermodynamic trends as the other bulge-containing duplexes. It is clear that compensating factors do not allow intrahelical and extrahelical bulged bases in this system to be distinguished from one another based on thermodynamic data.

Conclusions and Future Work

Based on the change in the free energy of duplex formation that accompanies the introduction of a bulged base, we have shown that the identity of the bulged base and the strand in which the bulge is located may influence the amount by which the duplex is destabilized. Our results are comparable to those obtained in a previous study of a bulged cytosine located in an A·T tract (Morden et al., 1983). Bulges located in the A-strand were found to be slightly less destabilizing than those located in the T-strand. Duplexes with pyrimidine bulges may be marginally more stable than those containing purine bulges. We are currently beginning an investigation of the effects that covalently bound and intercalating drugs have on the stability of duplexes containing bulges.

ACKNOWLEDGMENTS

We thank Dr. Jeffrey W. Nelson for useful discussions and for generously providing the nonlinear least-squares fitting routine and other Fortran programs. We also thank Stella Kim for performing some of the UV absorption experiments.

REFERENCES

Albergo, D. D., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1409–1418.

Bevington, P. R. (1969) in *Data Reduction and Error Analysis of the Physical Sciences*, pp 204–246, McGraw-Hill, New York.

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

Cantor, C. R., & Schimmel, P. R. (1980) in *Biophysical Chemistry, Part III, The behavior of biological macromolecules*, W. H. Freeman, San Francisco.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Fasman, G. D., Ed. (1975) *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed., Vol. I, Chemical Rubber Company, Cleveland, OH.

Fink, T. R., & Crothers, D. M. (1972) *J. Mol. Biol. 66*, 1–12.

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

Groebe, D. R., & Ulenbeck, O. C. (1989) *Biochemistry 28*, 742–747.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., van Boom, J. H., & Altona, C. (1980) *Nucleic Acids Res. 8*, 169–181.

Hare, D., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7456–7464.

Inman, R. B., & Baldwin, R. L. (1962) *J. Mol. Biol. 5*, 172–184.

Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 82–84.

Kalnik, M. W., Norman, D. G., Swann, P. F., & Patel, D. J. (1989a) *J. Biol. Chem. 264*, 3702–3712.

Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F., & Patel, D. J. (1989b) *Biochemistry 28*, 294–303.

Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F., & Patel, D. J. (1990) *J. Biol. Chem. 265*, 636–647.

Kopka, M. L., Fratini, A. V., Drew, H. R., & Dickerson, R. E. (1983) *J. Mol. Biol. 163*, 129–146.

Li, Y., Zon, G., & Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26–30.

Longfellow, C. E., Kierzek, R., & Turner, D. H. (1990) *Biochemistry 29*, 278–285.

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

Marky, L. A., & Kupke, D. W. (1989) *Biochemistry 28*, 9982–9988.

Miller, M., Harrison, R. W., Wlodawer, A., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 85–86.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1983) *Biochemistry 22*, 5557–5563.

Morden, K. M., Gunn, B., & Maskos, K. (1990) *Biochemistry 29*, 8835–8845.

Nelson, H. M. C., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature (London) 330*, 221–226.

Nelson, J. W., Martin, F. H., & Tinoco, I., Jr. (1981) *Biopolymers 20*, 2509–2531.

Nikonowicz, E., Roongta, V., Jones, C. R., & Gorenstein, D. G. (1989) *Biochemistry 28*, 8714–8725.

Noguchi, H., Arya, S. K., & Yang, J. T. (1971) *Biopolymers 10*, 2491–2498.

Papanicolaou, C., Gouy, M., & Ninio, J. (1984) *Nucleic Acids Res. 12*, 31–45.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 445–451.

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

Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359–389.

Roy, S., Sklenar, V., Appella, E., & Cohen, J. S. (1987) *Biopolymers 26*, 2041–2052.

Saenger, W. (1983) in *Principles of Nucleic Acid Structure*, pp 116–158, Springer-Verlag, New York.

Streisinger, G., Okada, Y., Emrich, J. N., Tsugita, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.

van den Hoogen, Y. Th., van Beuzekrom, A. A., van den Elst, H., van der Marel, G. A., van Boom, J. H., & Altona, C. (1988) *Nucleic Acids Res. 16*, 2971–2986.

Wilson, W. D., Dotong, M.-H., Zou, E. T., & Zon, G. (1988) *Nucleic Acids Res. 16*, 5137–5151.

Woodson, S. A., & Crothers, D. M. (1987) *Biochemistry 26*, 904–912.

Woodson, S. A., & Crothers, D. M. (1988a) *Biochemistry 27*, 436–445.

Woodson, S. A., & Crothers, D. M. (1988b) *Biochemistry 27*, 3130–3141.

Woodson, S. A., & Crothers, D. M. (1989) *Biopolymers 28*, 1149–1177.

Yoon, C., Privé, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332–6336.