# Influence of Neighboring Base Pairs on the Stability of Single Base Bulges and Base Pairs in a DNA Fragment†

Song-Hua Ke and Roger M. Wartell*

*School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332*

*Received October 28, 1994; Revised Manuscript Received January 11, 1995*®

ABSTRACT: Temperature-gradient gel electrophoresis (TGGE) was used to determine the relative thermal stabilities of 32 DNA fragments that differ by a single unpaired base (base bulge) and 17 DNAs differing by a base pair. Homologous 373 and 372 bp DNA fragments differing by a single base pair substitution or deletion were employed. Heteroduplexes containing a single base bulge were formed by melting and reannealing pairs of 372 and 373 bp DNAs. Product DNAs were separated on the basis of their thermal stability by parallel and perpendicular TGGE. The order of stability was determined for all single unpaired bases in four different nearest neighbor environments: (GXT)•(AYC), (GXG)•(CYC), (CXA)•(TYG), and (TXT)•(AYA) with X = A, T, G, or C, and Y = no base, or visa versa. DNA fragments containing a base bulge were destabilized by 2−3.6 °C (±0.2 °C) with respect to homologous DNAs with complete Watson−Crick base pairing. Both the identity of the unpaired base and the sequence of the flanking base pairs influenced the degree of destabilization. The range of temperature shift correspond to estimated unfavorable free energies from 2.5 to 4.6 kcal/mol. Purine base bulges were generally not as destabilizing as pyrimidine base bulges. An unpaired base which was identical to one of its adjacent bases generally caused less destabilization than an unpaired base with an identity differing from its nearest neighbors. This implies that positional degeneracy of an unpaired base within a run of two or more identical bases is an important factor effecting stability. The ability of TGGE to order the stabilities of DNA fragments differing by a single base pair was used to determine the relative stabilities of base pair stacking interactions. The results determined by TGGE were consistent with the relative stabilities determined from UV melting transitions.

Extra unpaired bases or "bulges" in double-helical DNA are created when one or more consecutive bases are unpaired to bases on the opposite strand. They can arise from recombination between sequences that are imperfectly homologous or from errors during replication of DNA. A duplex with a bulged base is the proposed intermediate in the process of frame-shift mutagenesis (Streisinger et al., 1966; Ionov et al., 1993). Extra bases in duplex segments of nucleic acids are also implicated in site-specific recognition by nucleic acid binding proteins (Christiansen et al., 1985; Wu & Uhlenbeck, 1987). Understanding how different bulged bases and their sequence environment influence DNA structure and stability will help explain their role in biological processes. Information on the effect of bulged bases on DNA stability will also help optimize methods that rely on thermal stability differences to hybridize probes to specific sites.

Previous studies of unpaired bases have focused primarily on short DNA duplexes. Investigators have examined the stability and structure of single unpaired bases by NMR (Woodson & Crothers, 1987), X-ray crystallography (Miller et al., 1988; Joshua-Tor et al., 1992), UV absorbance melting studies, and gel retardation assays (Rice & Crothers, 1989; Wang & Griffith, 1991; LeBlanc & Morden, 1991). Results indicate that both the identity of the bulged base and its sequence environment influence the amount by which a duplex oligomer is destabilized. No systematic study has yet to be reported on the effects of different single-base bulges on the stability of longer DNAs. The influence of end effects on the properties of short DNA duplexes (Olmsted et al., 1991) and the desire for a more complete survey of context effects makes such a study desirable.

A vertical TGGE format (Wartell et al., 1990) was employed to determine the relative stabilities of the four possible single-base bulges on each strand in four different nearest neighbor environments within 372/373 bp DNA duplexes. This investigation utilized an approach employed in an earlier study on the influence of single base pair mismatches on DNA stability (Ke & Wartell, 1993). In TGGE DNA fragments migrate through a polyacrylamide gel with a superimposed temperature gradient. A DNA moves at a constant mobility until the least stable melting domain denatures. The branched structure of the partially melted DNA results in a large decrease in mobility. DNAs of identical lengths but differing in the stability of their first melting domain migrate to different depths before they denature. A temperature gradient parallel to the direction of electrophoresis provided the greatest resolution for detecting changes in DNA thermal stability. Experiments in which the temperature gradient was perpendicular to the direction of electrophoresis were carried out to obtain mobility transition curves.

## MATERIALS AND METHODS

*Materials.* Plasmids pUC8-31 and pUC8-36 plasmids were previously described (Tatti & Moran, 1985). The plasmids contain a 130 bp segment of the *ctc* promoter region

† This work was supported by NIH Grant GM38045.
* Author to whom correspondence should be addressed.
® Abstract published in *Advance ACS Abstracts,* March 1, 1995.

0006-2960/95/0434-4593$09.00/0 


FIGURE 1: 373 bp DNA sequence between the *Eco*RI and *Rsa*I sites from pUC8-31 plasmid. Positions −43 and −36 are indicated. The DNA fragment from the plasmid pUC8-36 has the same sequence except for a G to A substitution at position −36. The upstream primers and the downstream primer, DP15, employed in PCR are indicated. Upstream primers created base pair changes at the positions underlined. Single-base deletions in the upstream primers are indicated with a "−" at the positions.

from *Bacillus subtilis* inserted between the *Hin*dIII and *Eco*RI sites of pUC8. pUC8-31 has the wild-type *ctc* sequence, and pUC8-36 has a G•C to A•T transition at position −36. The designation "−36" denotes the base pair position located 36 base pairs upstream of the *ctc* transcription startpoint. The polymerase chain reaction (PCR) was used to amplify a 373 bp region of the plasmids containing the *ctc* promoter (Figure 1). Four sets of specific point mutations or deletions were created by PCR through the application of upstream primers with single base substitutions or deletions. The DNA oligonucleotides employed (Operon Inc., Alameda, CA) for the sixteen upstream primers and one downstream primer are shown in Figure 1. Base positions of the substitutions are underlined, and single-base deletions in the upstream primers are indicated with a "−" at the relevant positions. The downstream primer DP15 was end-labeled for some PCR amplifications with $^{32}P$ as described previously (Ke & Wartell, 1993). Thirteen 373 bp DNAs differing from each other by a single base pair were produced by PCR. Four 372 bp DNAs were created by PCR using upstream primers with one base deleted (Figure 1). The base pair changes and deletions occurred at the four positions designated −36, −38, −39, and −43.

*Taq*. DNA polymerase was obtained from Perkin Elmer Cetus Inc. or Promega Inc. The PCR protocol recommended by Perkin Elmer Cetus Inc. was employed. The 100 μL reactions contained 50 pg of plasmid DNA, 0.6 μM of each primer in a buffer of 10 mM Tris-HCl, pH 8.3, 50 mM KCl, and 2.5 mM $MgCl_2$; 200 μM of each dNTP was employed. Further details are described in Ke and Wartell (1993). Two to four microliters of each reaction was checked for size and purity on a 6.5 or 7.5% polyacrylamide gel.

Although bulged bases phased with the helical repeat of DNA and RNA induce kinking and reduce electrophoretic mobilities in gels (Tang & Draper, 1990; Wang & Griffin, 1991), no single unpaired base caused a significant perturbation of gel mobility for the DNA molecules employed. The bending induced by a single bulge near the end of the DNA is apparently not large enough to affect the mobility of the DNA. In the absence of a temperature sufficient for unwinding, all DNA fragments had the same mobility. This is best illustrated in the perpendicular gel such as Figure 6. Prior to the onset of melting the DNAs had the same mobility.

*TGGE*. TGGE was carried out using a vertical gel electrophoresis apparatus previously described (Wartell et al., 1990; Ke & Wartell, 1993). Two aluminum heating plates sandwiching the glass plates were used to establish a temperature gradient either parallel or perpendicular to the electric field. Temperature measurements in the gel (described below) confirmed that the temperature gradient was linear and uniform within the region covered by the heating plates. A 6.5% polyacrylamide gel (37.5:1, acrylamide/(bisacrylamide) was used. The gel contained 3.36 M urea and 19.2% vol/vol formamide in 0.5× TBE (0.045 M sodium borate + 0.045 M Tris + 1 mM EDTA, pH 8.1). This corresponds to 48% of what is often referred to as 100% denaturant solution (7 M urea + 40% formamide). Formamide was deionized with mixed resin AG501-X8D (Bio-Rad). The gel running buffer was 0.5× TBE. Run times

were generally from 14 to 16 h at 4.5−6 V/cm. The gels were stained with ethidium bromide and photographed.

The relative stabilities of DNA fragments were determined from experiments in which the temperature gradient was parallel to the direction of electrophoresis (e.g., Figure 2). The low and high temperatures of the gradients are described in the text. DNA mobility transitions were obtained from perpendicular temperature gradient gels (e.g., Figure 6). Approximately 1 $\mu$g of each DNA fragment was loaded into a long well along the top of the gel and electrophoresed through a gradient with the high temperature sufficient to unwind at least the first melting domain of the DNAs.

Temperatures in the gels were determined with a thermocouple probe (TMTSS-020-6, Omega, Inc.) connected to a digital thermometer (MDSD-465, Omega, Inc., accuracy estimated as ±0.1 °C). Measurements at different depths in test gels verified the linearity of the gradient. Temperature measurements were made for the perpendicular temperature gradient gels at two or more positions at the end of each transition run. The positions where the probe had been inserted were observed as dark lines in the photographs and provided a temperature scale.

Photographs of the mobility transition curves were digitized and scaled using a digitizer tablet (SummaSketch II) and input to a microcomputer. Transition curves were smoothed prior to obtaining derivative curves by the "smoothlowess" function in the Axum graphics analysis package (Trimetrix Inc., Seattle, WA). This is based on a locally weighted regression analysis (Cleveland & Devlin, 1988). The mobility transition temperature, $T_u$, was defined as the temperature at the peak of the transition's derivative curve.

## RESULTS

DNA fragments containing single base bulges were made by melting and annealing two PCR generated DNAs differing in length by a base pair. The DNAs were heated for 3 min at 97 °C, reannealed at 54 °C for at least 10 min, and allowed to slowly cool to room temperature. Two heteroduplex DNAs with a single-base bulge each and the two homoduplexes were created. The four resulting DNAs were separated by parallel TGGE. In some experiments one of the DNAs was $^{32}$P-labeled on the 5′-end of its downstream primer strand to identify separated heterodupex DNAs. The gel was stained with ethidium bromide and photographed. When $^{32}$P-labeled DNA was used the bands were located on a UV-transilluminator, excised, and their radioactivity measured by scintillation counting. DNA band identities were established from the radioactively labeled bands. Confirmation of band identities was made by switching the DNA that contained the labeled strand and/or by running one of the homoduplex DNAs at a larger concentration.

*Parallel TGGE.* Figure 2 panels A and B show parallel TGGE experiments of 372/373 bp DNAs with all possible paired, deleted, and unpaired bases at position −36. The nearest-neighbor base pairs surrounding this position are (GXT)•(AYC). The identities of the bands in Figure 2A,B are given in the figure caption and were based on the procedures described above. DNA fragments are designated by specifying the base (A, T, G, C) at position X or Y or a dash (−) for the absence of a base. Thus (GCT)•(A−C) implies the DNA with a C bulge on the top strand (position X), and (G−T)•(A−C) denotes the 372 bp homoduplex. A temperature gradient from 35.5 to 42.5 °C was used for the gel shown in Figure 2A to optimize the separation of homoduplex and heteroduplex DNAs in one gel. The uppermost band in lanes 1, 2, and 3 actually corresponds to two heteroduplex DNAs with single-base bulges. A larger separation for two heteroduplex DNAs is seen in lane 4. Figure 2B shows a gel that employed a temperature gradient from 32 to 35 °C. This shallower gradient optimized the separation of the DNAs with single-base bulges, sacrificing the ability to separate the homoduplex DNAs. The brighter bottom band in all lanes of Figure 2B contains two homoduplex DNAs. The most stable single-base bulge DNA in this series is (G−T)•(AAC) in lane 4 of Figure 2B. This band is separated by 1.6 cm from the least stable single-base bulge DNA, (GAT)•(A−C), in lane 3. We note that the A bulge in the DNA represented by (G−T)•(AAC) could also shift to the adjacent position, i.e., (GT−)•(AAC). For simplicity only one representation is used.

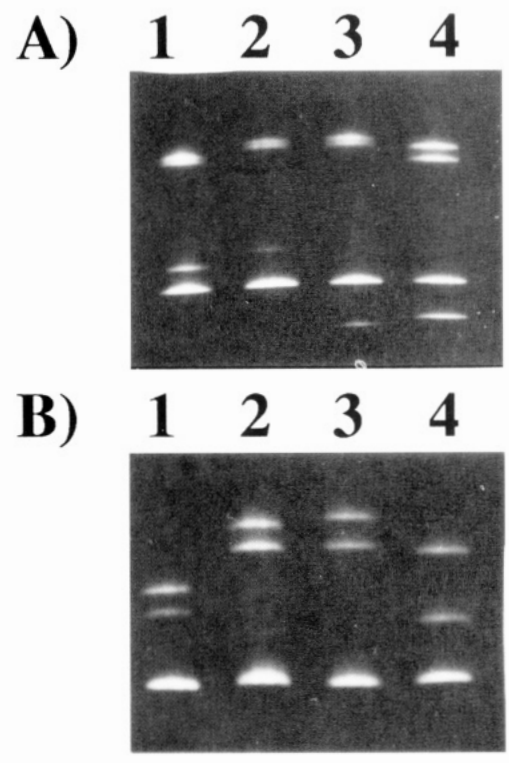


FIGURE 2: Parallel temperature gradient gel of 372/373 bp DNAs with all base pairs and single-base bulges at position −36. Samples were run for 17−20 h at 90 volts. (A) Temperature gradient was from 35.5 to 42.5 °C. From top to bottom in each lane DNA bands contain the following bases at position −36: (1)* −/C, G/−, *−/−, G•C; (2) C/−, −/G, −/−, C•G; (3) A/−, −/T, A•T, −/−; (4) T/−, −/A, T•A, −/−. (B) Temperature gradient was from 32 to 35 °C. The top two DNA bands in each lane are the same as in panel A. The lowest band in each lane contains both homoduplex DNAs, e.g., lane (1) −/C, G/−, −/−, and G•C. (*Note: −/C means no base on top strand and C on bottom strand; −/− means the base pair at the position has been deleted.)

Figures 3, 4, and 5 show parallel TGGE experiments of DNAs with all possible base pairs and deleted or unpaired bases at positions −38, −39, and −43, respectively. The base pairs surrounding these positions are (GXG)•(CYC) for −38, (CXA)•(TYG) for −39, and (TXT)•(AYA) for −43. Figures 3 and 4 used similar temperature gradients (~36 to ~42 °C) to optimize the separation of homoduplex and heteroduplex DNAs in one gel. The largest separation of

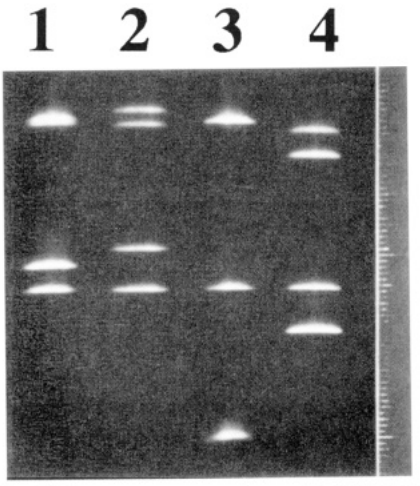


FIGURE 3: Parallel temperature gradient gel of 372/373 bp DNAs with all base pairs and single-base bulges at position −38. Temperature gradient was 36−42 °C. Samples were run for 20 h at 90 V. From top to bottom in each lane DNA bands contain the following bases at position −38: (1) −/A and T/−, T·A, −/−; (2) −/T, A/−, A·T, −/−; (3) C/− and −/G, −/−, C·G; (4) −/C, G/−, −/−, G·C.

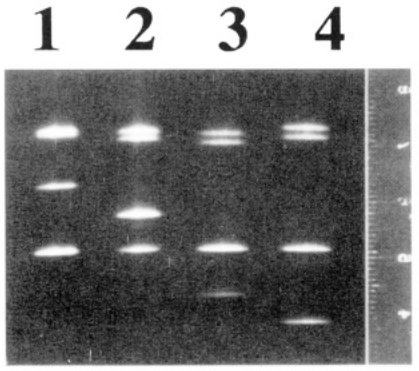


FIGURE 4: Parallel temperature gradient gel of 372/373 bp DNAs with all paired bases and single unpaired bases at position −39. Samples was run for 20 h at 90 V with a temperature gradient from 36 to 41 °C. From top to bottom in each lane DNA bands contain the following bases at position −39: (1) T/− and −/A, T·A, −/−; (2) −/T, A/−, A·T, −/−; (3) C/−, −/G, −/−, C·G; (4) −/C, G/−, −/−, G·C.

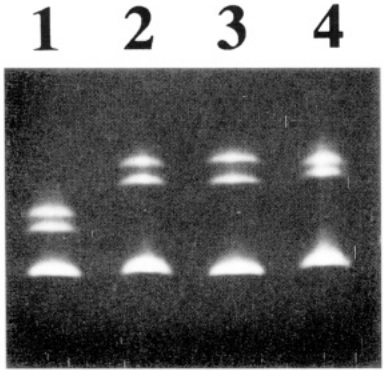


FIGURE 5: Parallel temperature gradient gel of 372/373 bp DNAs with all paired bases and single unpaired bases at position −43. Samples was run for 16.5 h at 90 V with a temperature gradient from 33 to 36 °C. From top to bottom in each lane DNA bands contain the following bases at position −43: (1) T/−, −/A, T·A, and −/−; (2) A/−, −/T, A·T, and −/−; (3) C/−, −/G, −/−, and C·G; (4) G/−, −/C, −/−, and G·C.

two homoduplex DNAs with transversion of the same base pair is observed in lanes 3 and 4 of Figure 3. The DNA fragment with the (GCG)·(CGC) sequence centered at position −38 moved 1.8 cm farther into the gel than the DNA with the (GGG)·(CCC) sequence. Figure 5 used a temperature gradient from 33 to 36 °C to optimize the separation of the DNAs with single-base bulges at position −43. Gels with a temperature gradient from 33 to 36 °C were also used to analyze the relative stabilities of DNA fragments with single-base bulges at positions −38 and −39 (not shown).

Table 1 summarizes the results from the parallel TGGE experiments. Both the identity of the unpaired base and its neighboring base pairs influence the destabilization caused by the unpaired base. A single base bulge with the same identity as a nearest neighbor base generally induces less destabilization than a base bulge with an identity different from its nearest neighbors. The data at position −36 in Table 1 illustrate this point. For each of the four unpaired bases the most stable situation occurs when the bulged base is identical to a neighboring base, e.g., (GGT)·(A−C) > (G−T)·(AGC). A similar situation occurs for the other positions as well.

At position −43, except for the two most stable situations, (T−T)·(AAA) and (TTT)·(A−A), the base bulge located within the stretch of A's is more stable than the base bulge located in the strand with the stretch of T's. A similar observation was made in a study of single-base bulges in a DNA oligomer (LeBlanc & Morden, 1991). Table 1 also shows that at each position examined the most stable single-base bulge is a purine and the least stable base bulge is a pyrimidine. The greater stability of purine bulges is consistent with NMR evidence indicating that unpaired purines are generally stacked into the helix (Woodson & Crothers, 1988, 1989) while unpaired pyrimidines may be extrahelical or intrahelical depending on the sequence context and temperature (Morden et al., 1990; van der Hoogen et al., 1988; Kalnik et al., 1989).

*Mobility Transition Curves.* Perpendicular TGGE experiments display the temperature-induced mobility transition of intact double-stranded DNA to the denatured state. Figure 6 shows transition curves of the first melting domain of five of the DNAs examined. The upper temperature limit employed in Figure 6 was selected to induce unwinding of only the first melting domain. Several steps are observed in this DNA's complete mobility transition (Wartell et al., 1990, Figure 7). The initial quasilinear increase in mobility with increasing temperature appears to be due to the effect of temperature on gel viscosity and/or pore size. The larger sigmoidal decrease in mobility is due to the unwinding of DNA strands.

The two leftmost transitions in Figure 6 correspond to DNAs altered at position −36 to have bulges (GTT)·(A−C) and (G−T)·(AAC). The middle transition corresponds to the DNA with the T·A base pair at position −36. The two transitions on the right correspond to DNAs with the base pair deletion and the G·C base pair at position −36. The transition curves of the DNAs with bulged bases are noticably broader than the other DNAs. This may indicate that internal melting from the bulged base plays a role in the melting process. It may also reflect an influence of the extra base on the configuration of partially melted DNA. Broader transitions were also observed for the 373 bp DNAs containing certain mismatched base pairs (Ke & Wartell, 1993).

The transition temperatures of the DNA fragments' first melting domain, $T_u$, are given in Table 2. The $T_u$ data were

Table 1: Comparison of Bulged Base Stabilities in 373 bp DNA Fragment

| bulge position | 5′ flanking base pair[a] | 3′ flanking base pair[a] | ranking of single-base bulge stability |
|---|---|---|---|
| −36 | G·C | T·A | G/− > −/A > T/− > −/C ≥ −/T ≥ −/G > A/− > C/− |
| −38 | G·C | G·C | G/− > −/C ≥ A/− > T/− ≥ −/G > C/− ≥ −/A > −/T |
| −39 | C·G | A·T | −/G > A/− ≥ G/− > C/− ≥ −/A > −/T > T/− > −/C |
| −43 | T·A | T·A | −/A > T/− > −/G = −/T ≥ −/C > G/− = A/− ≥ C/− |

[a] E.g., the single-base bulge designated A/− at position −36 represents the following sequence:

5′-GAT-3′
3′-C−A-5′

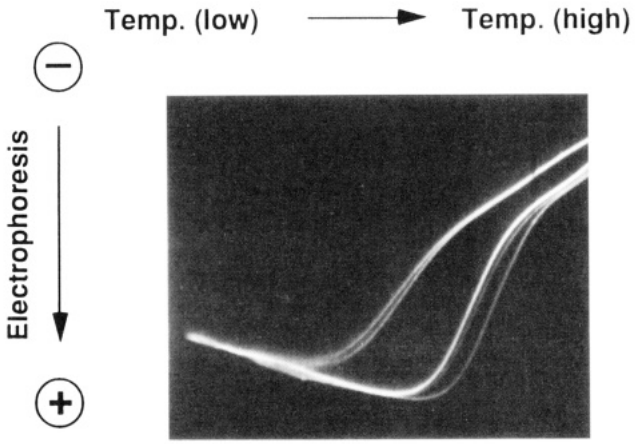


FIGURE 6: Typical perpendicular temperature gradient gel of 373 bp PCR fragments. The transition curves only show the first melting domain of the DNA fragments. Electrophoresis was conducted for 14.5 h at 90 V. The temperature gradient was 26−44 °C from left to right. The sample contained melted and reannealed DNAs with −/− and T·A at position −36 and the native DNA with G·C at the same position. The transition from left to right correspond to DNAs with the following bases at position −36: T/−, −/A, T·A, −/−, and G·C.

Table 2: Perpendicular TGGE $T_u$[a] Measurements of First Melting Domain of 372/373 bps DNA

| − 36 set (°C) | −38 set (°C) | −39 set (°C) | −43 set (°C) |
|---|---|---|---|
| C·G: 37.1 | C·G: 38.8 | G·C:[b] 37.1 | G·C: 37.7 |
| G·C:[b] 37.1 | G·C: 38.0 | C·G: 36.8 | C·G: 37.3 |
| −/−:[c] 36.5 | −/−:[c] 37.6 | −/−:[c] 36.2 | −/−:[c] 37.3 |
| T·A: 36.2 | T·A: 37.4 | A·T: 35.8 | T·A:[b] 37.1 |
| A·T: 36.1 | A·T:[b] 37.1 | T·A: 35.3 | A·T: 36.5 |
| G/−: 34.1 | G/−: 35.6 | −/G: 34.0 | −/A: 35.1 |
| −/A: 33.9 | −/C: 34.7 | A/−: 33.8 | T/−: 34.9 |
| T/−: 33.7 | A/−: 34.7 | G/−: 33.7 | −/G: 34.6 |
| −/C: 33.4 | T/−: 34.5 | C/−: 33.5 | −/T: 34.6 |
| −/T: 33.4 | −/G: 34.3 | −/A: 33.5 | −/C: 34.6 |
| −/G: 33.4 | C/−: 34.3 | −/T: 33.4 | G/−: 34.4 |
| A/−: 33.2 | −/A: 34.2 | T/−: 33.4 | A/−: 34.4 |
| C/−: 33.2 | −/T: 34.0 | −/C: 33.2 | C/−: 34.4 |

[a] $T_u$ was defined as the temperature at the peak of the derivative curve calculated from the smoothed DNA mobility transition profile. Estimated precision in $T_u$ relative to the standard DNA fragment is ±0.2 °C averaged over all DNAs. Each experiment contained the pUC8-31 DNA fragment as an internal standard. The mean $T_u$ of this DNA was 37.1 °C (±0.4 °C) based on 15 repeated experiments. [b] This is the same pUC8-31 DNA fragment. [c] −/− corresponds to the 372 bp duplex DNAs where the base pair at the designated position has been deleted.

consistent with the results from the parallel TGGE (Table 1). When $T_u$ values appear to be the same for two DNAs, the order of stability listed in Table 2 was determined from Table 1 since resolution of the parallel gradient gels was better than the perpendicular temperature gradient gels. $T_u$ values of the DNAs with a single-base bulge were lower than the values for corresponding 372 bp homoduplex DNAs by 2−3.6 °C (±0.2). Mobility transitions of the first domains of homoduplex DNAs in gels with 48% and 58% denaturant differed by 5.3 °C but gave essentially identical slopes and the same order of homoduplex stabilities (not shown). This is consistent with previous melting studies indicating that urea and formamide lower the thermal stability of G·C and A·T base pairs by roughly equivalent amounts (Hutton, 1977; Klump & Burkart, 1977; Casey & Davidson, 1977).

Table 3: Comparison of Base Pair Stacking Parameters

| DNA fragments[a] (position) | $\Delta T_u$ (°C) | $\{[\delta G_{JK} + \delta G_{KL}] - [\delta G_{PQ} + \delta G_{QR}]\}$[b] (cal/mol) |
|---|---|---|
| [TTT − TAT] (−43) | 0.6 | −81 |
| [TGT − TCT] (−43) | 0.5 | −186 |
| [GCG − GGG] (−38) | 0.8 | −615 |
| [GTG − GAG] (−38) | 0.3 | −186 |
| [GTT − GAT] (−36) | 0.1 | −240 |
| [GCT − GGT] (−36) | 0.05 | −179 |
| [CAA − CTA] (−39) | 0.5 | −27 |
| [CGA − CCA] (−39) | 0.2 | −250 |

[a] Each DNA is designated by the three top-strand bases that include the position of the base pair transversion in the center. Thus for the second row TGT is the DNA with a G·C base pair at position −43. [b] $\delta G_{JK}$ corresponds to the stacking energy parameter for the base pair stack (JpK)·(K′pJ′). Values from Delcourt and Blake (1991) were used.

*Nearest Neighbor Base Pair Stacking.* The $T_u$ data of the homoduplex DNAs in Table 2 provides information on the relative stability of nearest neighbor stacking for each pair of DNAs with an A·T to T·A or G·C to C·G transversion at a specific position. Although $T_u$ is not defined as the temperature where 50% of the domain's base pairs are melted—the $T_m$—we consider $\Delta T_u = \Delta T_m$ for the homoduplex DNAs. This assumption is supported by an analysis of the full DNA mobility curves which indicate $T_u \approx T_m$ for the first melting domain (see below). Since the slopes of the mobility transitions of the first domains of the homoduplex DNAs are the same (e.g., Figure 6) and the temperature differences are small, the $\Delta T_u$ values may be related to stacking free energy differences.

Table 3 lists the $\Delta T_u$ value for pairs of DNAs with A·T to T·A or G·C to C·G changes at specific positions. Within the framework of the nearest neighbor model of DNA stability, the DNA with the higher $T_u$ value has the more stable pair of stacking interactions at the specified position. The adjacent column lists the difference in stacking energy terms for the pairs of DNAs based on stacking parameters evaluated by Delcourt and Blake (1991). For all eight cases examined, the TGGE data and the parameters of Delcourt and Blake give the same relative ranking of stacking interactions, i.e., the DNA with the higher $T_u$ has the lower

net stacking energy. This was also observed using stacking parameters evaluated from other studies (Wartell & Benight, 1985; Breslauer et al., 1986; Doktycz et al., 1992). A quantitative correlation between $\Delta T_u$ and the net stacking energy is not observed for any of the stacking parameter sets. This lack of quantitative agreement is not unexpected given the difference in solvents between the TGGE experiments and that used to obtain the melting parameters.

A comparison was made between the $\Delta T_u$ values of DNAs differing by one base pair at a specific position with the $\Delta T_m$ values predicted from calculated melting curves. The calculations employed the Delcourt and Blake (1991) stacking parameters and other parameters corresponding to a solvent of 0.1 M $Na^+$ (McCampbell et al., 1989). The $\Delta T_m$'s predict the same order of stability for the DNAs as the measured $\Delta T_u$'s although quantitative agreement is again not observed. One interesting observation correctly predicted by the theory is that deleting a base pair at each of the four positions always produced a DNA with a stability higher than the DNA with an A•T (or T•A) pair and lower than the DNA with a G•C (or C•G) pair.

The predictions of DNA melting theory were also compared with the full experimental mobility curve of the wt 373 bp DNA. Lerman et al. (1984) introduced an equation to describe the electrophoretic retardation of DNA accompanying partial melting in a polyacrylamide gel. The relative mobility of a long duplex DNA with a short melted domain may be written as

$$\mu(T) = \mu_o \exp[-p(T)/L_r] \quad (1)$$

where $\mu_o$ is the mobility of the completely duplex DNA, $L_r$ is the length of the flexible unit of a DNA strand in number of bases, and $p(T)$ is the length of the melted region at temperature $T$ measured in the number of melted base pairs. $p(T)$ is calculated from melting theory and is the sum of probabilities of each base pair being melted (Lerman et al., 1984). This is equivalent to the number of opened base pairs if strand dissociation is neglected. Since eq 1 is applicable only when the melted domain is a small fraction of the entire DNA, neglecting strand dissociation is a reasonable assumption. Equation 1 is not expected to accurately predict the mobility transition curve beyond the initial stages of melting.

The mobility curve of the wild-type 373 bp DNA was obtained from a perpendicular temperature gradient gel. A 48% denaturant 6.5% polyacrylamide gel was used with a temperature gradient from 30 to 55 °C. The mobility curve was first normalized by dividing the distance to which DNA migrated at each temperature by the distance migrated at 30 °C. Then the small pretransition linear increase in DNA mobility with temperature (see, e.g., Figure 6) was subtracted from the entire transition. This subtraction corrects for the effect of temperature on gel pore size and/or viscosity, and its influence on the mobility of duplex DNA. The solid line in Figure 7A shows the adjusted experimental mobility transition. A theoretical mobility curve was calculated from eq 1. The melting theory parameters employed were from solution studies in ~0.1 M $Na^+$ (McCampbell et al., 1989; Delcourt & Blake, 1991) except for the melting temperatures of the average AT and GC base pairs, $T_{AT}$ and $T_{GC}$. A uniform decrease in these parameters was needed to compensate for the lower ionic strength of the gel buffer (0.045 M Tris borate) and the urea–formamide denaturant. The

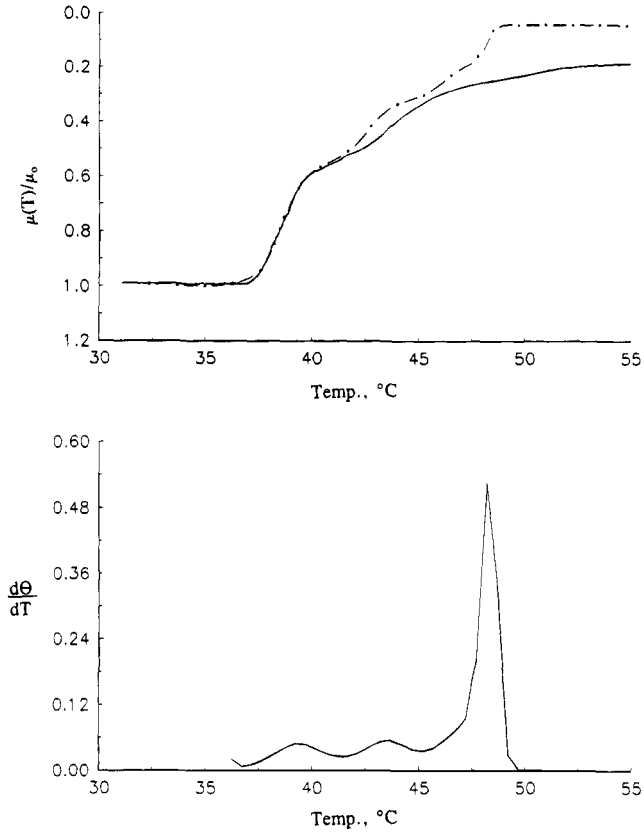


FIGURE 7: (A, top) Normalized experimental mobility transition curve of the 373 bp DNA with wild-type sequence (—) is compared to the predicted mobility curve (— • —) calculated using eq 1. In the experiment the DNA was run for 14 h at 90 V in a perpendicular temperature gradient gel containing 48% denaturant and 6.5% polyacrylamide. Normalization procedures for the experimental curve and the theoretical parameters for the calculated curve are described in the text. (B, bottom) Derivative melting curve of the 373 bp DNA calculated using the parameters described in the text for the mobility curve of panel A.

dash-dot line in Figure 7A shows the predicted mobility transition with values of $T_{AT}$ = 24.1 °C and $T_{GC}$ = 68.0 °C and $L_r$ = 100. The $T_{AT}$ and $T_{GC}$ values are about 40 °C lower than values appropriate for 0.1 M $Na^+$.

Figure 7A shows that eq 1 using DNA melting theory and its parameters provides an excellent fit to the first phase of the mobility melting curve. A comparison of the calculated number of opened base pairs for the first phase of the experimental mobility curve shows a very close correspondence (<0.1 °C) between the calculated $T_m$ and experimental $T_u$. The good agreement between the calculated and experimental curves supports the application of the theory for interpreting the mobility curve data. Figure 7B shows the calculated derivative melting curve for the 373 bp DNA. The first melting domain is well separated from unwinding of the rest of the DNA. The area under the first melting domain corresponds to the unwinding of 52 base pairs. Analysis of the DNA's melting map (not shown) shows this region is from the *Eco*RI end of the sequence where the single base bulges occur.

## DISCUSSION

The degree of destabilization due to a single-base bulge was influenced by both the base type and its neighboring base sequence. The results from this study provide information that may be related to mutational events and illustrate

how TGGE can be applied to determine the stabilities of specific non-Watson−Crick structures in long nucleic acid duplexes.

Frame-shift mutations are known to occur predominantly in simple repeated sequences (Ionov et al., 1993; Calos & Miller, 1981). According to a model of frame-shift mutation proposed by Streisinger et al. (1966), slippage of one strand of DNA causes one or more base to loop out. If the stability of a base bulge is the dominant factor determining the likelihood of a frame-shift mutation, one may expect the mutation frequency at a site to be correlated with the stability of the base bulge. NMR studies (Woodson & Crothers, 1987) on (GATG$_3$CAG)•d(CTGC$_4$ATC) and d(GATG$_3$-CAG)•d(CTGAC$_3$ATC) indicated that a bulged C base, when identical to neighboring bases, is more stable than a bulged A base which differs from neighboring bases.

We have observed as a general phenomena that a single base bulge with the same identity as one of its adjacent bases causes less destabilization than the bulged base when it has different nearest neighbors. This can be rationalized on the basis of the positional degeneracy of an unpaired base within a homogeneous tract which increases the entropy and therefore the stability of the DNA. If this phenomena minimizes error correction of random base insertions, the results suggest that homogenous tracts would tend to persist or perhaps enlarge over evolutionary time.

The analysis of the full mobility transition of the 373 bp DNA showed that the unwinding of the first melting domain is well separated from the remainder of the DNA molecule (Figure 7). Since the single base bulges are in this domain and its unwinding can be treated as a separate transition, the shift in the domain's melting temperature due to a bulged base can be related to a free energy change. The destabilizing free energy of a bulge can be expressed as (Woodson & Crothers, 1987)

$$\delta(\Delta G) = -(\Delta S)\delta T_m$$

in which $\Delta S$ is the entropy change for melting the fully base paired domain, and $\delta T_m$ is the shift in $T_m$ due to a bulge. We assume $\delta T_u \approx \delta T_m$ and $\Delta S$ is independent of temperature. $\Delta S$ can be estimated by multiplying the number of base pairs in the first melting domain, $N_1$, times $\Delta S°$, the average entropy change/base pair. Analysis of the 373 bp DNA melting curve using the parameters employed for Figure 7 indicates that $N_1 \approx 52$ bp. This value was determined from the relative area under the first domain of the derivative melting curve, 14%, and confirmed by a melting map analysis of the DNA [e.g., Wartell et al., (1990)]. The latter analysis also shows that the first melting domain is an end-melting domain encompassing the bulge sites. An estimate of error for $N_1$ was made by calculating and analyzing the melting curves of the 373 bp DNA using several different sets of stacking parameters in addition to those of Delcourt and Blake (1991) (McCampbell et al., 1989; Gotoh & Tagashira, 1981; Doktycz et al., 1992). We found $N_1 = 52.5 \pm 3.5$ bp.

Utilizing an average value per base pair of $\Delta S° = -24.8$ cal K$^{-1}$ (mol bp)$^{-1}$ (Delcourt & Blake, 1991), the destabilizing free energy for temperature shifts of 2−3.6 °C ranges from $2.5 < \delta(\Delta G) < 4.6$ kcal/mol. These values are in the same range determined for single-base bulges in DNA oligomers (Morden et al., 1983; Woodson & Crothers, 1987; LeBlanc & Morden, 1991). Given the errors in the temperature shifts, $N_1$, and $\Delta S°$, the value of $\delta(\Delta G)$ for a specific bulge is approximately ±10%.

The relative stabilities of the single base bulges determined by TGGE are in good agreement with data where available from short DNA oligomers. NMR studies were carried out on the four self-complimentary duplexes each with a single-base bulge (underlined): d(CCGGAATTCACGG), d(CCGAGAATTCCGG), d(CCGTGAATTCCGG), and d(CCGGAATTCTCGG) (Kalnik et al., 1989b, 1990). The results demonstrated that the 13-mer duplex with (GAG)•(C−C) is more stable than the duplex with (G−G)•(CAC), and the 13-mer duplex with (GTG)•(C−C) is more stable than the duplex with (G−G)•(CTC). Comparing the $T_m$ values of the bulged duplexes in these two studies, the following order of stabilities is obtained: (GAG)•(C−C) > (GTG)•(C−C) > (G−G)•(CAC) > (G−G)•(CTC). We observed the same ranking in the equivalent nearest neighbor environment at position −38 (Table 1). UV absorbance melting curves (LeBlanc & Morden, 1991) were obtained from oligomers d(GCGA$_2$XA$_2$GCG)•d(CGCT$_4$CGC) and d(GCGA$_4$GCG)•d(CGCT$_2$YT$_2$CGC) with X = C, T, or G in A-strand and Y = C, A, or G in T-strand. Comparison of DNA oligomer $T_m$ values yielded the following relative order of stability of single-base bulges: bulges in A-strand > bulges in T-strand; T bulge > C bulge > G bulge in A-strand; and C bulge > G bulge > A bulge in T-strand. In our experiments, a similar ranking was observed in the −43 set where the nearest-neighbor sequence is the same as the above oligomers. The only differences were the relative order of duplexes which have very similar thermal stabilities. This may be due to the difference between the two experimental conditions or the oligomer vs polymer contexts.

The mobility transition curves provide a novel approach for directly evaluating the heirarchy of stacking interactions in a long DNA. It differs from the least-squares analysis fitting the 10 stacking parameters to melting data from a number of DNAs [e.g., Delcourt & Blake (1991) and Breslauer et al. (1986)]. The consistency between the results obtained by TGGE and the stacking interactions obtained by least-squares analysis lends support to the determined heirarchies. A more meaningful and quantitative test of stacking interaction values will require a quantitative comparison using data obtained under similar solvent conditions.

## ACKNOWLEDGMENT

We acknowledge Thomas Maier for technical assistance and equipment maintenance.

## REFERENCES

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

Calos, M. P., & Miller, J. H. (1981) *J. Mol. Biol. 153*, 39−66.

Casey, J., & Davidson, N. (1977) *Nucleic Acids Res. 4*, 1539−1552.

Christiansen, J., Douthwaite, S. R., Christensen, A., & Garrett, R. (1985) *EMBO J. 4*, 1019−1024.

Cleveland, W. S., & Devlin, S. J. (1988) *J. Am. Stat. Assoc. 83*, 596−610.

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.

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

Hutton, J. R. (1977) *Nucleic Acids Res. 4*, 3537−3555.
Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D., & Perucho, M. (1993) *Nature 363,* 558−561.
Joshua-Tor, L., Frolow, F., Applella, E., Hope, H., Rabinovich, D., & Sussman, J. L. (1992) *J. Mol. Biol. 225*, 397−431.
Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F., & Patel, D. J. (1989a) *Biochemistry 28*, 294−303.
Kalnik, M. W., Norman, D. G., Swann, P. F., & Patel, D. J. (1989b) *J. Biol. Chem. 264,* 3702−3715.
Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F., & Patel, D. J. (1990) *J. Biol. Chem. 265*, 636−647.
Ke, S.-H., & Wartell, R. M. (1993) *Nucleic Acids Res. 21*, 5137−5143.
Klump, H., & Burkart, W. (1977) *Biochim. Biophys. Acta 475*, 601−604.
LeBlanc, D. A., & Morden, K. M. (1991) *Biochemistry 30*, 4042−4047.
Lerman, L. S., Fischer, S. G., Hurley, I., Silverstein, K., & Lumelsky, N. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 399−424.
McCampbell, C. R.,Wartell, R. M., & Plaskon, R. R. (1989) *Biopolymers 28*, 1745−1758.
Miller, M., Harrison, R. W., Wlodawer, A., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 85−86.
Morden, K. M., Gunn, B. M., & Maskos, K. (1990) *Biochemistry 29*, 8835−8845.
Olmsted, M. C., Anderson, C. F., & Record, M. T., Jr. (1991) *Biopolymers 31,* 1593−1604.
Rice, J. A., & Crothers, D. M. (1989) *Biochemistry 28*, 4512−4516.
Streisinger, G., Okada, Y., Emrich, J. N., Tsugita, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77−84.
Tang, R. S., & Draper, D. E. (1990) *Biochemistry 29*, 5232−5237.
Tatti, K. M., & Moran, C. P., Jr. (1985) *Nature 314*, 190−192.
van den Hoogen, Y. Th., van Beuzekom, A. A., van den Elst, H., van der Marel, G. A., van Boom, J. H., & Altona, C. (1988) *Nucleic Acids Res. 16*, 2971−2986.
Wang, Y.-H., & Griffith, J. (1991) *Biochemistry 30*, 1358−1363.
Wartell, R. M., & Benight, A. S. (1985) *Phys. Rep. 126*, 67−107.
Wartell, R. M., Hosseini, S. H., & Moran, J. D. (1990) *Nucleic Acids Res. 18*, 2699−2750.
Woodson, S. A., & Crothers, D. M. (1987) *Biochemistry 26*, 904−912.
Woodson, S. A., & Crothers, D. M. (1988) *Biochemistry 27*, 3130−3141.
Woodson, S. A., & Crothers, D. M. (1989) *Biopolymers 28*, 1149−1177.
Wu, H.-N., & Uhlenbeck, O. C. (1987) *Biochemistry 26*, 8221−8227.

BI9425213