Nucleic Acids Research, 1996, Vol. 24, No. 11 **2095–2103**

Thermodynamic effects of formamide on DNA stability

R. D. Blake* and Scott G. Delcourt

Department of Biochemistry, Microbiology and Molecular Biology, University of Maine, Orono, ME 04469-5735, USA

Received February 5, 1996; Revised and Accepted April 18, 1996

ABSTRACT

Formamide lowers melting temperatures (T_m) of DNAs linearly by 2.4–2.9°C/mole of formamide (C_F) depending on the (G+C) composition, helix conformation and state of hydration. The inherent cooperativity of melting is unaffected by the denaturant. dT_m/dC_F for 11 plasmid domains of 0.23 < (G+C) < 0.71 generally fit to a linear dependence on (G+C)-content, which, however, is consistent with a (G+C)-independent alteration in the apparent equilibrium constant for thermally induced *helix↔coil* transitions. Results indicate that formamide has a destabilizing effect on the helical state, and that sequence-dependent variations in hydration patterns are primarily responsible for small variations in sensitivity to the denaturant. The average unit transition enthalpy $\overline{\Delta H}_m$ exhibits a biphasic dependence on formamide concentration. The initial drop of –0.8 kcal/mol bp at low formamide concentrations is attributable to a $\delta\overline{\Delta H}_m$ for exchange of solvent in the vicinity of the helix: displacement by formamide of weakly bound hydrate or counterion. The phenomenological effects are equivalent to lowering the bulk counterion concentration. Poly(dA·dT) exhibits a much lower sensitivity to formamide, due to the specific pattern of tightly bound, immobilized water bridges that buttress the helix from within the narrow minor groove. Tracts of three (A·T)-pairs behave normally, but tracts of six exhibit the same level of reduced sensitivity as the polymer, suggesting a conformational shift as tracts are elongated beyond some critical length [McCarthy,J.G. and Rich,A. (1991) *Nucleic Acids Res.* 19, 3421–3429].

INTRODUCTION

The addition of formamide ($HCONH_2$) to aqueous buffer solutions of DNAs lowers their stability, and for this reason this denaturant is a common additive in many low temperature studies. Studies by McConaughy *et al.* (1), Record (2), Casey and Davidson (3) and Hutton (4) demonstrated that melting temperatures, T_m, for thermal denaturation of DNAs decrease linearly by approximately –0.65°C/volume fraction (× 100) formamide. Earlier studies had suggested the denaturing ability of agents such as formamide is due to their lower ion-solvating power (5,6), and to their ability to increase the solubility of free bases (6). It was said these agents increase the hydrophobic character of the solvent, thereby decreasing the activity coefficients and free energies of the bases, favoring the denatured state (7–10). The results of these early studies seem to have been widely interpreted as indicating that since the denatured state was favored, the effects of formamide were largely independent of base content.

Systematic studies of the effects of formamide on DNAs of different (G+C) content have not been carried out, or at least not reported, so the relationships of (G+C) content, formamide concentration and stability are not known. The goal of this study was to examine this question, with the objective of gaining further insight into the molecular reasons for the denaturing effects of formamide. If the denaturing effects could be shown to be the consequence of differences in perturbations of ligand association constants for the binding of water and ions to helix and coil states of DNA, this agent could serve as a useful means of probing variations in local structure and hydration of DNA.

MATERIALS AND METHODS

Specimen DNAs

The principal specimens used in these studies were plasmid DNAs, primarily pN/MCS-10, -11 and -12, constructed in this laboratory by established procedures (11) to contain a repetitive sequence at a site in the plasmid with favorable energetic characteristics. A 55 bp multiple cloning sequence (MCS) with recognition sites for 10 restriction enzymes was first installed at the unique *Nru*I locus (972) of pBR322 (12). Repetitive tridecamers of sequence: $[AAGTTGAACAAAT]_n AAGTTG$, where $57 > n > 16$, were then inserted at the unique *Sma*I locus of the MCS. This locus is immediately adjacent to a unique *Kpn*I recognition sequence of the MCS, and 803 bp away from an *Eco*RV sequence, both sites used to linearize the plasmid in preparation for melting. *Escherichia coli* HB101 or SURE™ cells containing plasmid were grown in 1 l ampicillin selective LB broth, and amplified by the addition of chloramphenicol. Cells were lysed in alkaline SDS, and the plasmid precipitated from the lysate in isopropanol. Crude plasmid DNA was purified on two CsCl gradients.

In the preparation of oligomer repeats for insertion into pN/MCS, the residues of *n*-base oligonucleotides were first

* To whom correspondence should be addressed

paired with the residues of complementary overlaps of (*n*/2)-base oligonucleotides, and then ligated using T4 DNA ligase to generate tandemly repeating polynucleotide chains of 100–1000 bp. The DNA was size fractionated on low-melting agarose, and fragments >200 bp removed from the gel, and made blunt with Klenow enzyme. The resulting DNA was ligated into pN/MCS at the *Sma*I locus of the MCS (12). In preparation for insertion of the repetitive elements, pN/MCS plasmid was restricted at the unique *Sma*I site and digested with bacterial alkaline phosphatase to remove the terminal 5′-phosphate groups to prevent vector recircularization during the ligation step. The phosphatase was inactivated by the addition of SDS, followed by phenol extraction. Plasmid DNA was then precipitated and ligated to the repetitive DNA at an insert:vector mole ratio of >10:1 using T4 DNA ligase. Circular recombinant plasmids were introduced into HB101 cells and transformed cells selected on agar containing ampicillin. The repetitive inserts were checked by double-stranded sequencing over the insert region by the dideoxy chain-termination method (13). Checked in this fashion and by electrophoretic mobility, it was found that the insert length, *n*, of the repeat in pN/MCS-10 is 38 (500 bp), 57 in pN/MCS-11 (747 bp), and 16 in pN/MCS-12 (214 bp).

Equilibrium melting curves

High resolution derivative melting curves were obtained by a difference-approximation method (14,15), with a modified double-beam ratio-recording spectrophotometer (Cary). The finite difference method involves the approximation of a derivative $dA_{\lambda(nm)}/dT$ by its differential $\Delta A_\lambda/\Delta T$, with analytical reconstruction of the true derivative achieved as previously described (14). Temperatures were ramped at 6°C/h, which has been shown to provide equilibrium melting of most domains. Slower rates lead to a significant increase in thermal degradation, particularly of single-stranded coil regions (16). For this reason, specimens were never remelted; instead, replicate experiments were always carried out on fresh aliquots of the same or new preparations. The distribution of (G+C) compositions and sizes of domains responsible for subtransitions were obtained by spectral decomposition of melting curves obtained at 260, 270 and 282 nm, where the ratios of derivative extinction coefficients for A·T and G·C pairs, $d\varepsilon_{A\cdot T}/d\varepsilon_{G\cdot C}$, are 4.31, 1.00 and 0.120, respectively (15).

Commercial poly(dA·dT) (~1 µg; Sigma) of mean length >20 000 bp was added to all specimens before melting, to serve as secondary standard of solvent conditions. The sharp melting of this synthetic DNA is given by:

$$T_m^{poly(dA.dt)} = 19.07\log_{10}[Na^+] + 86.87, °C \quad \mathbf{1}$$

Solvent

The solvent consisted of the specified formamide concentration, plus 0.074 M NaCl, 0.005 M Na-cacodylate and 0.2 mM Na–EDTA (pH 6.85; $[Na^+]$ = 0.075 M).

Formamide

The formamide used in these studies was commercial Ultrapure-grade (99.9%) from USB. The A_{260} of a neat solution was <1.0, and the conductivity <100 µmho/cm. Formamide has a molecular weight of 45.04 and a density of 1.3340 g/ml at 20°C, so that while the relationship between volume fraction of formamide (V_F/V_{total}) and concentration, C_F (mol/l), is linear, that with the mole fraction is non-linear, and best represented by a third-order polynomial:

$$X_F = -2.048\times10^{-4} + 5.445\times10^{-3}[\%V_F] + 1.647\times10^{-5}[\%V_F]^2 + 2.909\times10^{-7}[\%V_F]^3 \quad \mathbf{2}$$

indicating the volume of a water molecule is less than half the volume of a formamide molecule.

van't Hoff enthalpies

Enthalpic quantities associated with conformational changes in DNA were determined by van't Hoff analysis of transition equilibria (17), where derivative curves of cooperative, two-state $h\rightarrow c$ transitions of individual domains are given by:

$$\frac{dA_{270}(T)}{dT} = \frac{\Delta A_{270}\Delta H^{total}_{m,i}/4RT^2}{\cosh^2\left[\left(\Delta H^{total}_{m,i}/2RTT_{m_i}\right)\left(T-T_{m_i}\right)\right]} \quad \mathbf{3}$$

The precise bell-shape of curves given by this equation are defined by three parameters: T_{m_i}, the melting temperature, $\Delta H^{total}_{m,i}$, the total enthalpy for dissociation of the i^{th} domain at, or close to the melting temperature, and ΔA_{270},the total integrated change in absorbance. If ΔA_{270} is normalized to unity, $dA_{270}/dT = d\theta/dT$ where θ represents the fraction of base-pairs remaining. The average enthalpy per base pair for this domain is then given by:

$$\overline{\Delta H}_{m,i} = \Delta H^{total}_{m,i}/N_i \quad \mathbf{4}$$

Values of $\Delta H^{total}_{m,i}$ were determined by non-linear regression to experimental curves. This was achieved by reading digitized experimental curves obtained at high densities into vector arrays, which were passed to a least-squares equation to minimize the weighted sum of squares, *S*, in the customary fashion:

$$S = \sum_j\left[\frac{\Phi-\hat{\Phi}}{\sigma_j}\right] \quad \mathbf{5}$$

Φ represents values for the experimental dataset, $[dA_{270}(T)/dT]_j$, and σ_j is the deviation of the j^{th} experimental point from the theoretical curve (Φ). Unless otherwise stated in the Results, ΔH^{total}_i were determined this way by extending the summation over the entire transition region. It is assumed that temperature values x_j are known exactly, and that all uncertainties occur in the y_j values, or $[dA_{270}(T)/dT]_j$.

The advantage of the van't Hoff method is its high degree of sensitivity, where excellent results can be obtained on <50 µg of precious DNA. The disadvantage is that accurate knowledge of the states of equilibria are determined indirectly, which can be problematic in studies of the equilibrium thermal denaturation of DNA. Denaturation is a complex piecemeal process, with subtransitions spread over a 15–20°C range, emanating from micro-domains of a few base pairs to macro-domains of >500. All have different stabilities and many overlap one another (18). Under such circumstances it is difficult to determine the states of equilibria for individual transitions. The problem can be circumvented, however, by constructing plasmid DNAs, such as the

pN/MCS-series described above, that harbor large, discrete domains of sharply-melting oligonucleotide repetitive elements. The equilibrium denaturation of these domains can be isolated from the background contributions of nearby transitions by subtracting denaturation profiles of plasmids without the insert from those that have it (12).

RESULTS

Thermal stability of domains

Denaturation temperatures, T_m, and fraction (G+C) contents (F_{GC}) of domains leading to discrete transitions were determined from high resolution derivative curves obtained from the loss of hypochromicity at 260 and 282 nm (*cf.* Materials and Methods). Sizes of the different domains are obtained by integration of dA_{270}/dT curves, as well as through correlations with statistical mechanical results of analyses of the sequences of specimens. Illustrative 270 nm curves of plasmid pN/MCS-12 DNA in two different solvent environments are shown in Figure 1. The uppermost curve was obtained in the standard buffer (0.075 M Na^+) without formamide, and shows 11 transitions, 10 of which originate from the plasmid. The first, very sharp transition at 65.49°C is due to the melting of poly(dA·dT), added to all DNA specimens as convenient marker and secondary standard of precise solvent conditions. The melting of poly(dA·dT) is particularly sensitive to extrinsic factors that affect the stabilities of all DNAs. The half-width of this curve is only 0.067 ± 0.007°C, while T_m of replicate experiments vary by <±0.02°C. This means T_m of plasmid transitions obtained in the identical environment as that for the poly(dA·dT) reference standard can be determined with similar precision. Sensitivities of T_m to variations in base composition and ionic strengths are therefore limited only by knowledge of the sequence of the specimen and by the accuracy with which the buffer is prepared.

The plasmid melts at higher temperatures and over a far wider, 15°C range. Thermal dissociation appears to take place by 10 separate transitions, but, in fact, it is more complicated than that (12). For example, only three of these transitions actually arise from single domains. The other seven represent superpositions of transitions associated with two or more distant domains of very similar melting temperature. With the transition for poly(dA·dT) as reference, the T_{m_i} of these transitions can be measured with a precision of ±0.03°C.

The base composition of domains contributing to transitions in curves of Figure 1 vary between 0.23 and 0.71 F_{GC}. The first transition at 76.27°C, no.1, arises from a single 214 bp domain consisting of a 13 bp repetitive insert that we designed for investigating the thermodynamic effects of formamide. The insert has the sequence: $(AAGTTGAACAAAT)_{16}AAGTTG$, and is located at the *Sma*I locus of the MCS. In this location, the insert lies between two large, relatively (G+C)-rich domains that serve as energetic barriers to further melting of the plasmid. Transitions for these neighboring barrier domains occur at much higher temperatures, corresponding to no.9 and no.10 in Figure 1. The insert therefore melts as a closed loop when the plasmid is linearized at the unique *Eco*RV sequence, 803 bp away from the insert. If linearized at the unique *Kpn*I sequence, which is immediately adjacent to the insert, it melts from the end of the DNA. The thermodynamic effects of melting the insert in these different physical environments are different. When melting as a loop there is a smaller overall change in entropy for dissociation, seen as a higher T_m The repetitive insert in the pN/MCS-12 plasmid melts from the end at 74.48°C, and as an internal loop at 76.27°C.

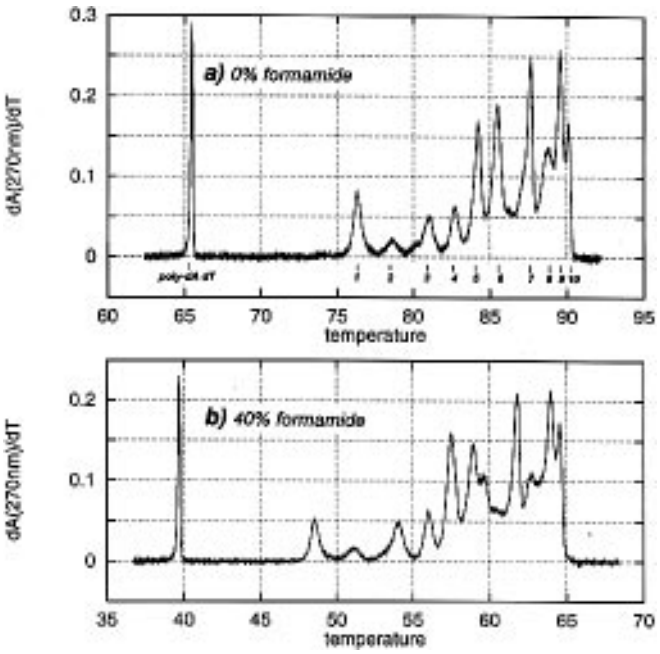

Figure 1. This figure shows derivative melting curves of the 4632 bp, linearized, pN/MCS-12 plasmid DNA in two solvent systems, one without (a) and one with (b) formamide. Both curves were obtained at 270 nm where the spectral change in absorption of A·T base pairs is equal to that for G·C pairs. Both curves were recorded digitally in continuous fashion at 6.0°C/h. Both curves were obtained with poly(dA·dT) added as an internal standard. Transitions for poly(dA·dT) are seen as sharp spikes at the low temperature ends of both curves. (**a**) Melting in the standard buffer (Na^+ = 0.075 M). The subtransition labeled no.1 represents a 214 bp repetitive sequence insert: $(AAGTTGAACAAAT)_{16}AAGTTG$, that was installed at the *Sal*I locus of a 55 bp MCS located at the unique *Nru*I locus of pBR322. (**b**) Melting of the same DNA specimen as in (a) in the same standard buffer, but with the addition of 40% formamide (v/v).

Effects of formamide on thermal stability

In Figure 1b, the solvent environment has been altered by the addition of 40% formamide (*V*/*V*), C_F = 10.1 mol/l, although the buffer concentration is identical to that in Figure 1a. The T_m of poly(dA·dT) has fallen from 65.49 to 39.65°C, with no significant reduction in sharpness. Plasmid stability has been even more affected, as can be seen from the different relationships of the transitions for poly(dA·dT) and no.1 for the repetitive insert in Figure 1a and b.

The patterns of transitions in 0 and 40% formamide are similar, but not entirely conserved. The split in transition no.6 in formamide, for example, represents an obvious differential effect of the denaturant. In the absence of formamide, three transitions are superimposed under the peak labeled no.6 in Figure 1a (12). One is less affected by formamide than the other two, and so falls behind in the pattern of Figure 1b. Also, formamide has a greater effect on domains contributing to no.7 than to no.8, so that the former separates farther from the latter in the lower profile. The temperature spread of all transitions has also increased, from 15 to 17°C, suggesting a possible dependence on base composition,

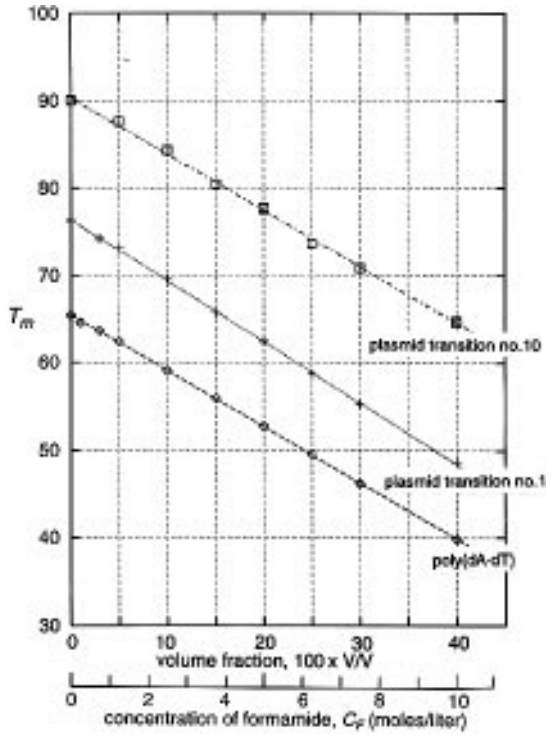

Figure 2. The dependence of melting temperature T_m on the volume fraction of formamide added for three of the transitions in Figure 1. A second x-axis for the dependence on formamide concentration is included.

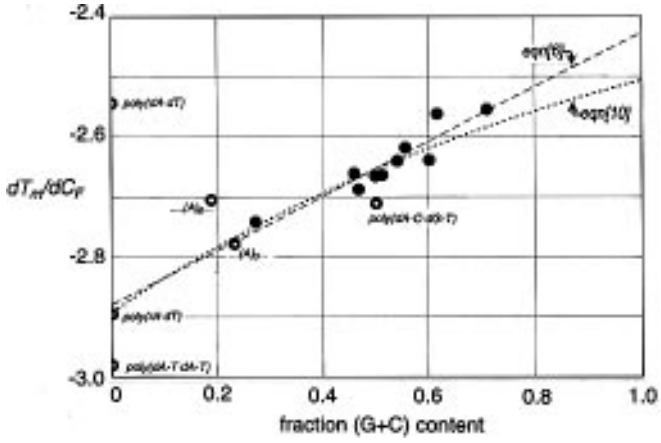

Figure 3. A plot of the dependence of the slope dT_m/dC_F on (G+C)-content, obtained from results such as represented in Figure 2. The (G+C)-contents of domains giving subtransitions were determined by spectral decomposition as described in Materials and Methods, and confirmed by reference to the denaturation map calculated by statistical mechanical analysis of melting of the plasmid sequence. The filled circles represent values of dT_m/dC_F for domains in pN/MCS-12 and pN/MCS-22, where the sizes of circles are about twice the magnitudes of the experimental errors in dT_m/dC_F. Open circles labeled …$(A)_3$… and …$(A)_6$…, represent values of dT_m/dC_F for the repetitive sequence domains containing tracts of 3 (A·T) bp (pN/MCS-12) and 6 (A·T) bp (pN/MCS-22), respectively. Other open circles represent measured dT_m/dC_F for synthetic duplexes identified in the figure by name. The dashed line represents the best-fit linear function **6** through dT_m/dC_F for domains. The dotted line represents the modified van't Hoff equation **10**, where $\Delta H_m = 7630 + 3500\cdot(G+C)$, $T_m = 337.14 + 41.97\cdot(G+C)$, and where by approximation $\partial \ln K_{app}/\partial C_F = -0.098$.

although no decrease could be detected in cooperativity of melting. The half-width of the sharp melting curve of poly(dA·dT) remained invariant (0.067 ± 0.005°C) over the full 0–40% range of formamide concentrations.

Figure 2 shows the dependence of the T_m for poly(dA·dT), and of plasmid transitions no.1 and no.10 on formamide concentration, C_F. Curves are linear with average standard errors of only ±0.15°C; that fall to only ±0.05°C if T_ms are corrected for deviations of $T_m^{poly(dA.dT)}$ from the linear function describing the behavior of that polymer. A plot of the slopes of results such as shown in Figure 2 against the fractional (G+C) composition of domains in the plasmid is shown in Figure 3, together with results for several synthetic polynucleotide duplexes. The diameters of filled circles in Figure 3 are slightly larger than estimated experimental errors for these data. Results for plasmid domains, represented by the filled points, approximate an underlying linear dependence on (G+C) content, with a slope of 0.453°C/l/mol formamide.

Small (<3%) deviations from the linear function, together with the absence of conservation in the melting patterns in Figure 1, indicate the superposition of small conformation-dependent or base sequence-dependent effects. Distinguishing between these two possibilities is better achieved with synthetic specimens consisting of biased sequences and known conformations.

Effects of formamide on the thermal stability of synthetic DNAs and (A·T)-tracts

Results for synthetic polynucleotide duplexes examined in this study deviate more widely from the expected base composition dependence. This reflects the amplification of base sequence-dependent effects that can be achieved with biased synthetic homo- and copolymers. The slope result for poly(dA·dT), –2.56 ± 0.05°C/mol/l, is 11% smaller than anticipated from the extrapolated behavior of mixed sequence domains (–2.88); whereas slopes for alternating copolymers poly(dA–T·dA–T) (–2.98) and poly(dA–C·dT–G) (–2.72) are slightly larger than expected. Poly(dA·dT) is clearly less affected by formamide than either poly(dA–T·dA–T) or poly(dA–C·dT–G). Poly(rA·dT), a hybrid duplex that adopts an A-structure, reacts differently to the presence of formamide than either poly(dA·dT) or poly(dA–T·dA–T), which adopt similar B′- and B-structures, respectively. Since effects that formamide might have on the denatured coil states of these three (A·T)-containing specimens is presumably the same, these results indicate (i) that differences between the effects of formamide on stabilities of both quasi-random and biased sequences are primarily helix conformation-dependent, and (ii) the magnitude of these differences depend in some way on the specific conformation of the helix, perhaps through differences in conformation-dependent hydration patterns.

The reduced effects of formamide on poly(dA·dT) are not restricted to the long $(A\cdot T)_\infty$-tracts of the homopolymer, but extend to smaller (A·T)-tracts that exceed some critical length as well. The very short $(A\cdot T)_3$-tract of the repeating insert sequence in pN/MCS-12: $[AAGTTGAAC(A)_3T]_{16}AAGTTG$, seems to react normally to the denaturing effects of formamide. The transition for this insert has an observed slope of –2.74 while a slope of –2.75 is expected for a domain of 0.23 (G+C)-content. However, the slope for the $(A\cdot T)_6$ tract-containing repeat domain

of pN/MCS-22: $[AAGTTGAAC(A)_6T]_{12}AAGTTG$, is only –2.73, well off the line where a value of –2.79 is expected. The difference of this insert from expected is only 32% of the observed-expected difference seen for poly(dA·dT), however the $(A{\cdot}T)_6$-tract represents only 38% of the repeat domain, and therefore should be affected by about this amount if the anomalously low reactivity of (A·T)-tracts to formamide is limited to just tract regions.

The reduced reactivity of $(A{\cdot}T)_6$-tracts does not depend on the position and physical environment of the domain during melting, since the same pN/MCS-22 plasmid linearized immediately adjacent to the repeat domain by *Kpn*I has almost the same slope value, –2.71, as it does when linearized at a remote site by *Eco*RV. The anomalous reactivity to formamide is clearly attributable to (A·T)-tracts at least six pairs in length.

Two-state behavior of the insert domain

Transition enthalpies in different formamide concentrations were determined by analysis of two-state *helix↔coil* equilibria associated with the repetitive insert domain. This domain is shown to melt in two-state fashion by several criteria when the plasmid is linearized by *Eco*RV, forcing the domain to melt as a closed loop. The experimental transition for this domain, labeled no.1 in Figure 1a, is shown in magnified form in Figure 4, represented by the heavy, somewhat noisy solid line. This domain has a (G+C) composition of only 0.23, so that its transition is almost entirely isolated from transitions at slightly higher temperatures emanating from the parent plasmid. Nevertheless, remnants of overlapping transitions that might affect the analysis were eliminated by subtracting a curve of the plasmid without the insert from the curve in Figure 1a for the plasmid with the insert, and this is the curve shown in Figure 4. The overall transition enthalpy, $\Delta H^{total}_{m_1}$, was then determined, as described in Materials and Methods, by least-squares fit of equation **3** for the two-state, van't Hoff dependence of the domain equilibrium on temperature to the experimental curve. The fitted curve is given by the smooth line in Figure 4, punctuated by diamond-shaped symbols every tenth degree. The residuals, plotted in the lower half of this figure, are small and totally without biases that would suggest the two-state expression **3** might be inappropriate.

The average enthalpy per base pair for this domain, $\overline{\Delta H}_{m_1}$, was determined from equation **4**. When the fit of equation **3** is made over the full temperature range of the transition shown in Figure 4, $\overline{\Delta H}_{m_1}$ = 7977 cal/mol-bp. Virtually the same value is obtained, 8040 cal, when the fit is limited to just the central 0.6°C temperature region of 76.0–76.6°C; which is expected of a two-state transition. $\overline{\Delta H}_{m_1}$ = 8111 ± 146 for three separate measurements on different preparations of pN/MCS-12.

Perhaps the most convincing indication of two-state behavior is that analysis of the transition for the same but much longer repetitive sequence domain in pN/MCS-10 indicates $\overline{\Delta H}_{m,1} = 7994 \pm 400$. The 500 bp tridecamer sequence in pN/MCS-10 is almost 250% longer than the same 214 bp repeat in pN/MCS-12, yet gives a unit enthalpy that is only 1.5% smaller.

A still further indication that the entire repeat transition denatures in two-state fashion is obtained from a statistical thermodynamic analysis of the pN/MCS-12 sequence. Derivative melting curves were computed with the algorithm of Poland and Scheraga (19) and Poland (20), with the Fixman and Friere (21) approximation of the loop function by a sum of exponentials. The curve for denaturation of the repetitive sequence domain in *Eco*RV-linearized pN/MCS-12, calculated with values for the various parameters from the literature (12,22), is given by the second of two smooth curves in Figure 4, punctuated by cross-shaped symbols every tenth degree. This curve was shifted –0.17°C on the temperature scale, so that it would superimpose on the experimental and van't Hoff curves, but no other adjustments were made. The agreement between observed and calculated amplitudes, breadths and areas is excellent, substantiating the two-state assumption. The denaturation map also shows the repetitive domain dissociating in two-state fashion. The small difference between observed and calculated temperature scales can probably be attributed to small errors in the ratios of $\overline{\Delta H}_{m_{LM}}$ upon $\overline{\Delta S}_{m_{LM}}$ for one or more of the 10 nearest neighbors (22).

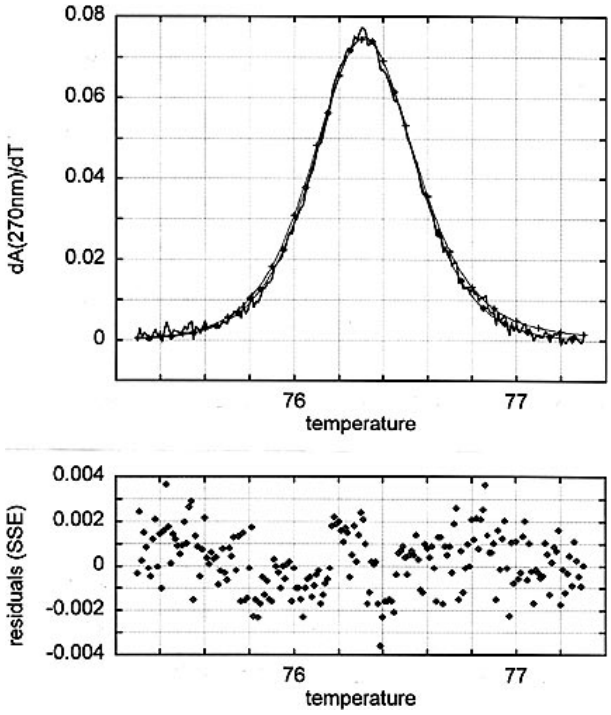

Figure 4. The transition for the repetitive sequence insert in pN/MCS-6: $[AAGTTGAAC(A)_3T]_{16}AAGTTG$, represented by the heavy, somewhat noisy line. This curve corresponds to an expanded view of the subtransition labeled no.1 in Figure 1a. The plasmid was linearized for melting at the *Eco*RV locus 803 bp away from the insert, so that the insert melts as a closed loop in two-state fashion. The curve was obtained as described in the text and in the legend to Figure 1. The curve punctuated by diamond-shaped symbols every 0.1°C was calculated from the two-state van't Hoff expression **3**, with the three adjustable parameters determined by non-linear regression to the experimental curve as described in the text. The T_m was 76.315°C, ΔA_{270} was 0.04509, and the average unit transition enthalpy for the 214 bp domain illustrated in this figure was 7977 cal/mol-bp. Residuals for the difference between experimental and van't Hoff curves are plotted below the melting curve, and indicate a totally random standard error of 3.4×10^{-4} between the two. The curve punctuated by cross-shaped symbols every 0.1°C represents a statistical mechanical melting curve, calculated from the entire sequence for the plasmid with parameters for simulating melting as described in the text. The calculated transition for the insert domain was isolated from the transition background emanating from other parts of the plasmid DNA, and superimposed on the experimental curve in this figure.

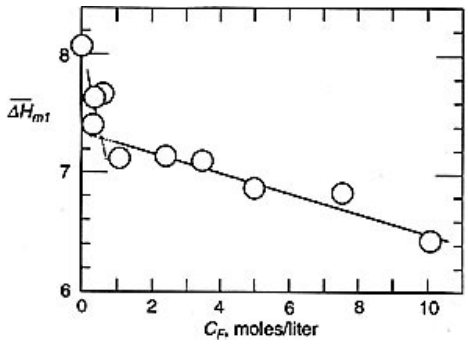

Figure 5. The dependence on formamide concentration C_F of the average unit transition enthalpy for the repetitive insert domain in pN/MCS-12. $\overline{\Delta H}_{m_1}$ were obtained as described in Figure 4 and in the text.

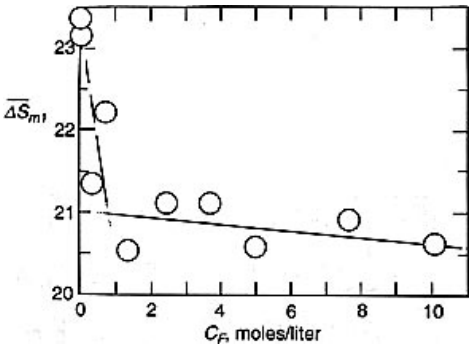

Figure 6. The dependence on formamide concentration of the average unit transition entropy for the repetitive insert domain in pN/MCS-12; obtained as described in the text.

Effects of formamide on transition enthalpies

The variation of $\overline{\Delta H}_{m_1}$ with C_F for the repetitive sequence domain is shown in Figure 5. $\overline{\Delta H}_{m_1}$ drops from 8.1 to ~7.3 kcal between 0–1.0 M formamide/l, a change of –0.8 kcal/mol-bp per 1 M formamide/l. Thereafter the enthalpy drops only 0.05 kcal/mol-bp. Since T_{m_1} drops monotonically with formamide concentration, the initial drop at low C_F is almost perfectly compensated by a similar drop in $\overline{\Delta S}_{m_1}$, as shown in Figure 6, where it is assumed that $\overline{\Delta S}_{m_1} = \overline{\Delta H}_{m_1}/T_{m_1}$. The thermodynamic effects associated with DNA melting can be apportioned to chemical and solvation effects, where the origin of the linear compensation is attributable to perturbations of the latter by formamide. The initial drop therefore suggests a small enthalpic loss associated with the exchange of formamide for bound water and possibly counterion to the helical state. The more gradual drop in $\overline{\Delta H}_{m_1}$ above 1 M formamide corresponds to an apparent heat capacity of only +18 cal/mol·deg that probably reflects a small decrease in residual single-strand stacking enthalpy with increasing T_m. Although $\overline{\Delta H}_m$ decreases over the wide range of C_F examined in this study, we did not find a corresponding decrease in the cooperativity of melting.

DISCUSSION

Results obtained in this study can be summarized as follows:

(i) Formamide lowers the melting temperature by 2.4–2.9°C per mole of formamide, depending on the (G+C) composition, helix conformation and state of hydration. The inherent cooperativity of melting, the strong dependence of the conformational state of base pairs on those of their neighbors from hydrogen bonding and stacking forces, is unaffected by the denaturant. Other results indicate that the denaturant both destabilizes the helical state and stabilizes the coil state.

(ii) The dependence on formamide concentration of the average $\overline{\Delta H}_m$ (per bp) for the repetitive insert domain of the plasmid is biphasic. Between 0 and 1.0 M formamide, the enthalpy decreases from 8.1 to 7.3 kcal/mol-bp, but thereafter the decrease is at 1/15th the rate; with compensating decreases in $\overline{\Delta S}_m$.

(iii) dT_m/dC_F for plasmid domains exhibit a dependence on (G+C) content approximated by the expression,

$$\frac{dT_m}{dC_f} = 0.453 \cdot (G + C) - 2.88 \qquad \mathbf{6}$$

The dependence is interpretable in terms of a formamide-induced shift in the phenomenological equilibrium K_{app} for $h \leftrightarrow c$ equilibria of domains that vary in base composition:

$$DNA_n \overset{K_{app}}{\leftrightarrow} DNA_c \qquad \mathbf{7}$$

(iv) While dT_m/dC_F for individual domains generally follow expression **6**, some vary slightly (±3%) from the line, reflecting local differences in sensitivity to the denaturant. We interpret these local differences as arising from sequence-dependent variations in the energy of hydration at selected sites. Measurements of such small effects is possible because of the high precision (±0.03°C) in measured T_m, achieved through the addition of the sharply melting poly(dA·dT) as an internal reference standard. Because results were obtained at high resolution, T_m of a large number of domains of different base compositions were obtained simultaneously from the same plasmid DNA specimen in a single melting experiment.

(v) Deviations from expression **6** are much greater for synthetic polynucleotide duplexes, where three different synthetic (A·T)-containing duplexes exhibit very different sensitivities toward the denaturant. Since the coil states are affected equally, the helical states must be destabilized to different extents, reflecting differences in hydration patterns and hydration energies for different helical conformations of synthetic duplexes.

(vi) Poly(dA·dT) exhibits a significantly lower sensitivity to the denaturing effects of formamide than predicted from **6** or from that of poly(dAT·dAT) or poly(rA·dT). One conformational feature that distinguishes poly(dA·dT) from the other (A·T)-containing duplexes is an exceptionally narrow minor groove, sustained by a well-ordered spine of hydration. We suggest that this tightly bound layer of hydration is not displaced by formamide, leading to a reduced sensitivity.

(vii) The degree of reduced sensitivity is not found in tracts of only two or three (A·T)-pairs, but tracts of six (A·T)-pairs exhibit the same level of reduced sensitivity as the polymer. This indicates the unusual hydration pattern and stability associated with these tracts at melting temperatures does not take place in the

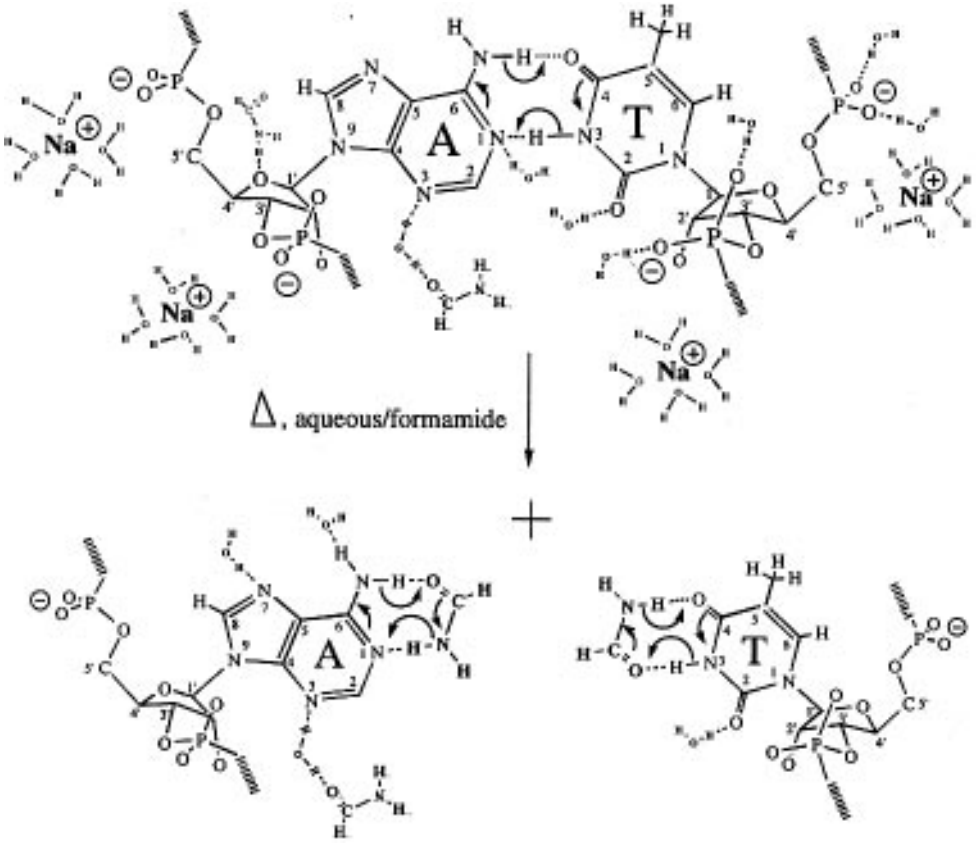

Figure 7. A proposed model for the stabilization of the coil state by 'formamidation' of the Watson–Crick binding sites of the bases. The model includes the potential enhancement of hydrogen bonding through resonance-assistance between base pairs as well as between bases and formamide.

duplex until some critical length between three and six (A·T) base pairs.

Results of this study are consistent with formamide destabilizing the helical state through displacement of loosely and uniformly bound hydrate. Formamide is a strongly associating liquid (23), capable of four hydrogen bonds, the same as water. It is a strong donor and a stronger acceptor than water (24). Calculations indicate that formamide–water hydrogen bonds are ~20% stronger than water–water bonds (25), so it seems probable that some or most of its effect as a denaturant must involve hydrogen bonding with DNA and DNA hydrate. Displacement of hydrate leads to destabilization of the helix, presumably because of the propensity of formamide to form hydrogen bonded networks and oligomeric chains in aqueous solution (24).

Destabilization of the helical state by formamide probably accounts for the small drop of –0.8 kcal/mol-bp in the transition enthalpy, by (i) displacing hydrate contributing to the support of the duplex, and/or by (ii) altering the local dielectric. Ignoring distinctions in binding energies and equilibrium exchangeable classes of hydrate, we expect δn_F formamides to displace δm_W waters and possibly δr_{Na^+} counterions as well:

$$DNA_h \cdot r_{Na^+,h} \cdot m_{h,W} + n_{h,F} \leftrightarrow DNA_h \cdot r_{Na^+,h} \cdot m_{h,W} + \delta r_{Na^+}Na^+ + \delta m_W W + \delta n_F F \quad \mathbf{8}$$

δm_W and δr_{Na^+} are >0, while δn_F <0. These changes probably lead to the small $\overline{\delta\Delta H}_{m_1}$ for **8** at low formamide concentrations.

The *helix↔coil* reaction scheme, **7**, can then be amended to take these three processes into account:

$$DNA_h \cdot r_{Na^+,h} \cdot m_{h,W} \cdot n_{h,F} \leftrightarrow DNA_c \cdot r_{Na^+,c} \cdot m_{c,W} \cdot n_{c,F} + \Delta\delta_{Na^+}Na^+ + \Delta\delta m_W W + \Delta\delta n_F F \quad \mathbf{9}$$

Displacement of hydrate is the key process. In **8**, displacement presumably involves the disproportionation of δm_W waters by δn_F formamides between helix states, whereas in **9** $\delta\Delta m_W$ represents the new difference in numbers of waters shed during melting. Also, $|\delta n_F| < |\delta m_W|$ because the volume of formamide is ~66 Å^3, more than twice the volume of water, 30 Å^3. The geometric relationships of donor and acceptor groups of these two molecules are also very different (24–27).

DNA hydrate can be categorized as either tightly or loosely bound. Altogether 22–28 waters complete the hydrodynamic mass of the generic A·T pair, whereas only 18–24 are bound to G·C pairs (28–34). One-third bind in a sequence-independent fashion with sufficient strength they become immobilized, exhibiting low levels of exchange with the surrounding solvent (32,35). Waters binding to condensed cations and anionic phosphate oxygens bind most firmly, while the hydration of sugars and backbone ester oxygens is of more moderate strength (35–39). The remaining two-thirds are found primarily at groove sites, and are more weakly bound, binding tightly only under special circumstances and only in the narrow groove (40). Presumably the weakly bound waters include a readily exchange-

able class that is displaced by formamide. If this class of hydrate is found in the grooves, and in the narrow groove in particular, displacement by formamide would alter the ionic potential, leading to changes in the numbers of sodium ions condensed to helical and coil states. Such changes are included in the first of the two schemes above, **8**, because the amount bound to the helix with its high charge density, may be affected by formamide. Thus, one explanation for the destabilization of the helical state is that formamide leads to a $\Delta\delta r_{Na+}$ between buffer systems.

There are two possible explanations for the variation of stability with (G+C)-content. In the first of these δm_W varies with (G+C)-content, and since there appears to be fewer weak groove-bound exchangeable waters at G·C than A·T sites, the amount displaced is in proportion to the amount bound. This simple explanation is contravened by other results, notably those with synthetic poly(dA·dT) and poly(dA–T·dA–T). Buoyant density measurements indicate poly(dA·dT) binds more water than poly(dA–T·dA–T) (41,42), yet the former is less affected by formamide than the copolymer.

In the second interpretation, δm_W does not change with (G+C)-content in formamide. Rather, the effect depends on displacement of a uniform class of bound water. Displacement may alter hydrate that contributes to the stability of the duplex, or it may alter the local dielectric, thereby altering counterion screening or the numbers of condensed counterion (43). In accordance with le Chatelier's principle, a change in $r_{Na+,h}$ brought on by formamide will perturb K_{app} and T_m for **9** if $\Delta\delta m_W$ or $\Delta\delta r_{Na+}$ for the different buffer systems are non-zero. According to this scheme, the phenomenological effects of formamide are equivalent to lowering the bulk concentration of counterion.

A uniform exchange process can explain the dependence on (G+C) content in the following manner. Assuming that variations in the stability of the duplex with formamide reflect variations in the equilibrium constant, $d\ln K_{app}/dC_F$, and, in accordance with multiple-binding theory (44), that these variations are constant and proportional to the displacement of essential hydrate δm_W and possibly δr_{Na+} in the helical state, the van't Hoff expression can be rewritten:

$$\frac{dT_m}{dC_F} = -\left(\frac{RT_m^2}{\Delta H_m}\right) \cdot \frac{\partial 1nK_{app}}{\partial C_F} \qquad \mathbf{10}$$

indicating the dependence of dT_m/dC_F on (G+C) content would vary directly with T_m^2, and inversely with ΔH_m. Since ΔH_m increases with (G+C) content (45), the prediction of a decrease in dT_m/dC_F with increasing (G+C) content corresponds with observation. The dependence on (G+C)-content calculated from equation **10** and given by the dotted curve in Figure 3, is in reasonably good agreement with the experimental results. Published rates of change of ΔH_m and T_m with (G+C)-content for equation **10** were obtained from the literature (22,45), while the value for $\partial \ln K_{app}/\partial C_F$ is –0.098.

The hydration-dependent thermodynamic effects of formamide are most conspicuous in the behavior of strongly biased synthetic DNAs, and in particular the (A·T)-containing specimens. Poly(dA·dT) is most inert to the effects of formamide, in the same way that (A·T)-tracts have been found to be hyporeactive to various chemical agents, whereas the reactivity of the isomeric alternating (A·T/T·A)-tracts are 'normal' (46). Tracts as short as $(A{\cdot}T)_6$ are just as inert, suggesting a conformational shift as (A·T)-tracts are elongated (47). (A·T)-tracts are known to adopt altered global structures, afforded by the opportunity for more uniform reductions in steric conflicts between neighbor pairs (48), but primarily because of the specific pattern of water bridges in the minor groove that buttresses the helix (40). Diffraction studies (49–51) indicate that (A·T)-tracts have narrow minor grooves, with an immobilized spine of hydration that buttresses an altered B′-structure (52). Proton magnetic resonance results indicate that the spine is particularly stable and inert to chemical exchange (33,53).

To account for the stabilization of the coil state and the increase in solubility of free bases found by Levine *et al* (6), it is proposed that the bases may be formamidated, as shown in Figure 7. There are obvious complementary structural features in both formamide and the bases. The C-N amide bond of formamide has substantial double-bond character, an energetic barrier to intramolecular rotation, so that the molecule is essentially planar (25,54,55). The bidentate formamide pairs with the bases in this scheme, gaining additional binding support from cooperative π-electron charge transfer interactions (26,27) that are not possible through hydration. Resonance-assisted hydrogen bonding as shown here, is said to enhance the stability of interactions by ~12% (56).

ACKNOWLEDGEMENTS

Supported in part by grants from MAES (Project No. 08402) and NIH.

REFERENCES

1 McConaughy, B.L., Laird, C.D. and McCarthy, B.I. (1969) *Biochemistry* **8**, 3289–3295.
2 Record, M.T., Jr (1967) *Biopolymers* **5**, 975–992.
3 Casey, J. and Davidson, N. (1977) *Nucleic Acids Res.* **4**, 1539–1552.
4 Hutton, J.R. (1977) *Nucleic Acids Res.* **4**, 3537–3555.
5 Helmkamp, G.K. and Ts'o, P.O.P. (1961) *J. Am. Chem. Soc.* **83**, 138–142.
6 Levine, L., Gordon, J.A. and Jencks, W. (1962) *Biochemistry* **2**, 168–175.
7 Herskovitz, T.T., Singer, S.J. and Geiduschek, E.P. (1961) *Arch. Biochem. Biophys.* **94**, 99–114.
8 Geiduschek, E.P. and Herskovitz, T.T. (1961) *Arch. Biochem. Biophys.* **95**, 114–129.
9 Hamaguchi, K. and Geiduschek, P. (1962) *J. Am. Chem. Soc.* **84**, 1329–1338.
10 Ts'o, P.O.P., Helmkamp, G.K. and Sander, C. (1962) *Proc. Natl. Acad. Sci. USA* **48**, 686–698.
11 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York.
12 Delcourt, S.G. and Blake, R.D. (1991) *J. Biol. Chem.* **266**, 15160–15169.
13 Zhang, H., Scholl, R., Browse, J. and Somerville, C. (1988) *Nucleic Acids Res.* **16**, 1220.
14 Yen, S-W.W. and Blake, R.D. (1980) *Biopolymers* **19**, 681–700.
15 Blake, R.D. and Hydorn, T.G. (1985) *J. Biochem. Biophys. Methods* **11**, 307–316.
16 Blake, R.D. and Haydock, P.V. (1979) *Biopolymers* **18**, 3089–3109.
17 Yen, S-W.W. and Blake, R.D. (1981 *Biopolymers* **20**, 1161–1181.
18 Vizard, D., White, R. And Ansevin, A.T. (1978) *Nature* **275**, 250–251.
19 Poland, D. and Scheraga, H.A. (1970) *Theory of Helix-Coil Transitions in Biopolymers*, Academic Press, New York.
20 Poland, D. (1974) *Biopolymers* **13**, 1859–1871.
21 Fixman, M. and Freire, J. (1977) *Biopolymers* **16**, 2693–2704.
22 Blake, R.D. (1996) *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol 2. (Meyers, R.A., ed.) VCH Publ, New York, 1–19.
23 Chang, Y.J. and Castner, E.W., Jr (1993) *J. Chem. Phys.* **99**, 113–125.
24 Jasien, P.G., and Stevens, W.J. (1986) *J. Chem. Phys.* **84**, 3271–3277.
25 Marchese, F.T., Mehrotra, P.K. and Beveridge, D.L. (1984) *J. Chem. Phys.* **88**, 5692–5702.

26 Cheeseman, J.R., Carroll, M.T. and Bader, R.F.W. (1988) *Chem. Phys. Lett.* **143**, 450–458.
27 Shivaglal, M.C. and Singh, S. (1992) *Int. J. Quant. Chem.* **44**, 679–690.
28 Falk, M., Hartman, K.A., Lord, J.R. and Lord, R.C. (1962) *J. Am. Chem. Soc.* **84**, 3843–3846.
29 Falk, M., Hartman, K.A., Lord, J.R. and Lord, R.C. (1963) *J. Am. Chem. Soc.* **85**, 387–394.
30 Falk, M., Poole, A.G. and Goymour, C.G. (1970) *Can. J. Chem.* **48**, 1536–1542.
31 Tunis, M.-J.B. and Hearst, J.E. (1968) *Biopolymers* **6**, 1345–1353.
32 Westhof, E. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17**, 125–144.
33 Kubinec, M.G. and Wemmer, D.E. (1992) *J. Am. Chem. Soc.* **114**, 8739–8740.
34 Schneider, B., Cohen, D.M., Schleifer, L., Srinivasan, A.R., Olsen, W.K. and Berman, H.M. (1993) *Biophysical J.* **65**, 2291–2303.
35 Westhof, E. (1987) *Int. J. Biol. Macromol.* **9**, 186–192.
36 Wing, R.M., Drew, H.R., Takano, T., Broka, C., Tanaka, S., Itikura, K. and Dickerson, R.E. (1980) *Nature* **278**, 755–757.
37 Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.* **149**, 761–786.
38 Drew, H.R. and Dickerson, R.E. (1981) *J. Mol. Biol.* **151**, 535–556.
39 Fratini, A.V., Kopke, M.L., Drew, H.R. and Dickerson, R.E. (1982) *J. Biol. Chem.* **257**, 14686–14707.
40 Chuprina, V.P., Heinemann, U., Nurislamov, A.A., Zislenkiewitz, P., Dickerson, R.E. and Saenger, W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 593–597.
41 Wells, R.D. and Wartell, R.M. (1974) in K.Burton (ed.) *Biochemistry of Nucleic Acids* Butterworths, London. pp. 41–64.
42 Marky, L.A. and Kupke, D.W. (1989) *Biochemistry* **28**, 9982–9988.
43 Manning, G.S. (1978) *Q. Rev. Biophys.* **11**, 179–246.
44 Schellman, J.A. (1978) *Biopolymers* **17**, 1305–1322.
45 Klump, H.H. (1988) In M.N.Jones (ed.) *Biochemical Thermodynamics*, 2nd Ed. Elsevier Press, Amsterdam, pp. 100–144.
46 Sullivan, J.K. and Lebowitz, J. (1991) *Biochemistry* **30**, 2664–2673.
47 McCarthy, J.G. and Rich, A. (1991) *Nucleic Acids Res.* **19**, 3421–3429.
48 Calladine, C.R. and Drew, H.R. (1986) *J. Mol. Biol.* **192**, 907–918.
49 Coll, M., Wang, C.A.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8385–8389.
50 Nelson, H.C.M., Finch, J.T., Luisi, B.F. and Klug, A. (1987) *Nature* **330**, 221–226.
51 DiGabriele, A.D., Sanderson, M.R. and Steitz, T.A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1816–1820.
52 Chuprina, V.P. (1987) *Nucleic Acids Res.* **15**, 293–311.
53 Liepnish, E., Leupin, W. And Otting, G. (1994) *Nucleic Acids Res.* **22**, 2249–2254.
54 Kurland, R.J. and Wilson, E.B. (1957) *J. Chem. Phys.* **27**, 585–590.
55 Klimkowski, V.J., Sellars, H.L. and Schafer, L. (1979) *J. Mol. Struct.* **54**, 299.
56 Jeffrey, G.A. and Saenger, W. (1991) in *Hydrogen Bonding in Biological Structures*, Springer-Verlag, NY.