Biochemistry

 Volume 30, Number 25 June 25, 1991

Accelerated Publications

Structure, Stability, and Thermodynamics of a Short Intermolecular Purine–Purine–Pyrimidine Triple Helix†

Daniel S. Pilch,‡ Corey Levenson,§ and Richard H. Shafer*,‡,∥

Graduate Group in Biophysics and Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143, and Cetus Corporation, 1400 43rd Street, Emeryville, California 94608

Received March 22, 1991; Revised Manuscript Received April 18, 1991

ABSTRACT: We have investigated the structure and physical chemistry of the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triple helix by polyacrylamide gel electrophoresis (PAGE), ^1H NMR, and ultraviolet (UV) absorption spectroscopy. The triplex was stabilized with $MgCl_2$ at neutral pH. PAGE studies verify the stoichiometry of the strands comprising the triplex and indicate that the orientation of the third strand in purine–purine–pyrimidine (pur-pur-pyr) triplexes is antiparallel with respect to the purine strand of the underlying duplex. Imino proton NMR spectra provide evidence for the existence of new purine–purine (pur·pur) hydrogen bonds, in addition to those of the Watson–Crick (W-C) base pairs, in the triplex structure. These new hydrogen bonds are likely to correspond to the interaction between third-strand guanine NH1 imino protons and the N7 atoms of guanine residues on the purine strand of the underlying duplex. Thermal denaturation of the triplex proceeds to single strands in one step, under the conditions used in this study. Binding of the third strand appears to enhance the thermal stability of the duplex by 1–3 °C, depending on the DNA concentration. The free energy of triplex formation (−26.0 ± 0.5 kcal/mol) is approximately twice that of duplex formation (−12.6 ± 0.7 kcal/mol), suggesting that the overall stability of the pur·pur base pairs is similar to that of the W-C base pairs. In addition, under identical solution conditions with the exception of pH (7.3 vs 5.5), the stability of the third strand in the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex is approximately twice that of the third strand in the corresponding $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex (−6.4 ± 0.5 kcal/mol) [Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990) *Nucleic Acids Res. 18*, 5743–5750]. This marked enhancement in stability, coupled with the lack of an acidic pH requirement, suggests that pur-pur-pyr triplexes are appealing choices for use in applications involving oligonucleotide targeting of duplex DNA in vitro and in vivo.

The structural diversity of nucleic acids has become increasingly evident in recent years. The triple helix is one form among a variety of unusual structures (Wells et al., 1988) that has generated a substantial degree of interest. Much of this interest was generated when intramolecular triple helices, termed H-DNA, were observed in supercoiled plasmids at acidic pH (Mirkin et al., 1987; Voloshin et al., 1988; Htun & Dahlberg, 1988; Hanvey et al., 1988). Subsequently, the ability of a third strand to bind to double-helical DNA via triple-helix formation has led to possible applications of triplexes as sequence-specific artificial nucleases (Moser & Dervan, 1987; Strobel et al., 1988; Sun et al., 1989; Hélène et al., 1989; Perrouault et al., 1990), modulators of DNA-binding proteins (Maher et al., 1989), and regulators of gene expression (Minton, 1985; Caddle et al., 1990; Glaser et al., 1990; Hélène & Toulmé, 1990). The last of these applications highlights the potential for using triplex-forming nucleic acids

† This research was supported by U.S. Public Health Service Grant CA27343, awarded by the National Cancer Institute, and the University of California Cancer Research Coordinating Committee. D.S.P. is supported, in part, by Training Grant GM08284, awarded by the National Institute for General Medical Sciences, Department of Health and Human Services.

* Author to whom correspondence should be addressed.
‡ Graduate Group in Biophysics, University of California, San Francisco.
§ Cetus Corp.
∥ Department of Pharmaceutical Chemistry, University of California, San Francisco.

0006-2960/91/0430-6081$02.50/0

as chemotherapeutic agents (Cooney et al., 1988).

Early studies showed that triplex structures consisting of two pyrimidine strands and one purine strand (pyr-pur-pyr triplexes) could be readily formed with synthetic polynucleotides under appropriate solution conditions (Felsenfeld et al., 1957; Riley et al., 1966; Thrierr & Leng, 1972). Triple-helical structures were also observed containing two strands of poly(dG) and one of poly(dC) (Marck & Thiele, 1978; Marck et al., 1978). Only recently, however, has it been possible to produce a triple helix containing two strands of poly(A) and one strand of poly(U) (Broitman et al., 1987). In this case, it was observed that the length of the poly(A) strands must be within a certain range in order to allow formation of a pur-pur-pyr triple helix. Thus, it appears that the stability of pur-pur-pyr triplexes is dependent on base composition to a greater extent than pyr-pur-pyr triplexes.

Recently, there have been structural and thermodynamic studies on a variety of short pyr-pur-pyr triple helices, made up of T·A·T and C^+·G·C base triplets (Rajagopal & Feigon, 1989; Santos et al., 1989; Pilch et al., 1990a,b; Umemoto et al., 1990; Manzini et al., 1990; Plum et al., 1990). Since the latter triplet involves protonation of one cytosine at its N3 position, formation of GC-containing pyr-pur-pyr triplexes requires either acidic pH or methylation of cytosine at its C5 position (Lee et al., 1984). Cooney et al. (1988) described formation of a triple helix in which the third strand, consisting of 27 purine-rich bases, was synthesized parallel to the purine strand of a sequence in the promoter region of the *c-myc* gene. The resulting triplex was shown to inhibit *c-myc* expression in vitro. More recently, Durland et al. (1990) presented evidence for triplex formation at neutral pH where the third strand, which contained only G and T bases, was antiparallel to the purine strand of the target duplex. Maher et al. (1989) were able to form pyr-pur-pyr triple helices by addition of third strands targeted to several DNA sequences within or flanking protein-binding sites and thereby inhibit the binding of the proteins to their designated sites. However, upon synthesizing a third strand, consisting of only purine bases, in a parallel orientation with respect to the purine strand of the target duplex, they were unable to form the corresponding pur-pur-pyr triplex. Bernuês et al. (1989, 1990) have recently demonstrated evidence for intramolecular pur-pur-pyr triplex formation at neutral pH in plasmids containing $(GA)_n$·$(TC)_n$ inserts. To date, no detailed structural or physicochemical studies on pur-pur-pyr triplexes have been reported.

Recently, we presented a thermodynamic analysis (Pilch et al., 1990b) of the pyr-pur-pyr triple-helical system, d-$(C^+_3T_4C^+_3)$·d$(G_3A_4G_3)$·d$(C_3T_4C_3)$, stabilized by acidic pH and either Na^+ or Mg^{2+} cations. We now describe experiments probing the structure, stability, and thermodynamics of the corresponding pur-pur-pyr triple helix, d$(C_3T_4C_3)$·2[d-$(G_3A_4G_3)$], at neutral pH in the presence of Mg^{2+} cations. Using PAGE techniques, we establish the stoichiometry of this triplex and determine the general orientation of the third strand in pur-pur-pyr triplexes. ^{1}H NMR spectra of the imino proton region provide evidence for the presence of new hydrogen bonds in addition to those of the underlying duplex. Thermal denaturation curves are analyzed to provide thermodynamic parameters for the dissociation of the triplex, which proceeds from triplex to single strands in one step. The free energy of triplex formation is approximately twice that of duplex formation under identical solution conditions, suggesting that the collective stability of the pur·pur base pairs is similar to that of the W-C base pairs. These experiments represent the first structural and physicochemical analysis of a short intermolecular pur-pur-pyr triple helix.

Materials and Methods

Chemicals and Oligodeoxynucleotides. Synthesis and purification of the d$(G_3A_4G_3)$ and d$(C_3T_4C_3)$ oligomers were carried out as previously described (Pilch et al., 1990b). The oligomers d(TTCCTCTCCC) (**Pyr**), d(GGGAGAGGAA) (**Pur1**), and d(AAGGAGAGGG) (**Pur2**) were synthesized as previously described (Pilch et al., 1990a) and purified by standard PAGE techniques. The concentrations of all oligomers were determined spectrophotometrically. The previously reported extinction coefficients (Pilch et al., 1990b) were used for d$(G_3A_4G_3)$ [ϵ_{255} = 11 500 cm^{-1} (mol of base)$^{-1}$ L] and d$(C_3T_4C_3)$ [ϵ_{271} = 8300 cm^{-1} (mol of base)$^{-1}$ L]. The extinction coefficients for **Pyr**, **Pur1**, and **Pur2** were determined as described by Griswold et al. (1951). All oligomer concentrations, except where noted otherwise, are reported on a strand basis. The buffer reagents were obtained from Sigma Chemical Co.

PAGE Retardation Assay. Nondenaturing polyacrylamide gel electrophoresis was carried out at 4 °C. The 15% polyacrylamide gels [29:1 acrylamide:bis(acrylamide)] contained 90 mM Tris–borate buffer (TB) (pH 8.0) and 50 mM $MgCl_2$, which comprised the tank buffer as well. Gels were run at 3 V/cm (~10 mA) for 16–18 h. All DNA samples were preheated at 80 °C for 5 min, slowly cooled, and loaded in 90 mM TB (pH 7.8), 5% Ficoll (type 400), and 50 mM $MgCl_2$. For the triplex based on the d$(C_3T_4C_3)$ strand, DNA samples contained either 0 or 100 μM d$(G_3A_4G_3)$ and concentrations of d$(C_3T_4C_3)$ ranging from 0 to 350 μM. For the triplex based on the asymmetric **Pyr** strand, DNA samples contained 19 μM **Pyr**, either 19 or 38 μM **Pur1**, and either 0 or 19 μM **Pur2**. Gels were initially visualized by UV shadowing and then photographed. Gels were subsequently stained for 5 min in a 5 μg/mL solution of ethidium bromide (EtdBr), briefly destained in distilled water, visualized with a UV transilluminator, and then photographed.

Imino Proton NMR. NMR experiments were conducted on a 500-MHz General Electric GN-500 spectrometer, equipped with an Oxford Instruments magnet and a Nicolet 1280 computer. NMR measurements were made on 400 μL of solution containing 380 μM d$(C_3T_4C_3)$ and 760 μM d-$(G_3A_4G_3)$ in aqueous phosphate buffer [15 mM sodium phosphate (pH 7.2), 10% D_2O/90% H_2O]. Aliquots of 0.5 M $MgCl_2$ were added to a final concentration of 14 mM, and the imino proton spectrum was acquired after each addition. The $1\bar{3}3\bar{1}$ solvent suppression pulse sequence (Hore, 1983) was used to acquire the spectra, with the carrier frequency set at the H_2O resonance. The pulse repetition time was 3 s and the interval delay was 96 μs. The temperature was fixed at 35 °C.

Helix–Coil Transitions. The UV absorbance of solutions containing 10 mM sodium cacodylate (pH 7.3), 50 mM $MgCl_2$, 0.1 mM EDTA, and either 1–25 μM duplex or 1–100 μM triplex was measured once per minute at 260 nm (duplex) or 280 nm (triplex) on a Gilford 2600 spectrophotometer, interfaced to an Optima Systems AT-10 computer for data collection and analysis. All DNA solutions were preheated at 80 °C for 5 min and slowly cooled prior to UV analysis. The temperature was increased at a rate of 0.25 °C/min with a Gilford 2527 thermoprogrammer, and the total run time was 340 min. The cell path length was 1 cm for duplex or triplex concentrations of up to 7.5 μM and 0.1 cm for solutions containing higher concentrations. First derivatives were calculated over a window of ±2.5 °C.

Thermodynamic Analysis. The monophasic helix–coil

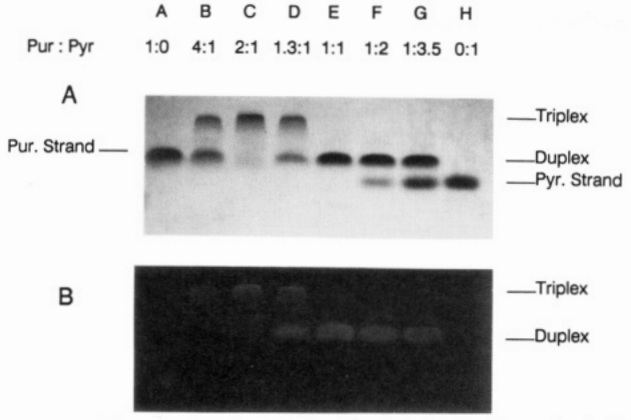

FIGURE 1: Photographs of a UV-shadowed (A) or EtdBr-stained (B) 15% polyacrylamide gel containing 90 mM TB (pH 8.0), 50 mM $MgCl_2$, and differing stoichiometric ratios of $d(G_3A_4G_3)$(Pur) to $d(C_3T_4C_3)$ (Pyr). Lanes A–G contain 100 μM $d(G_3A_4G_3)$ and 0, 25, 50, 75, 100, 200, and 350 μM $d(C_3T_4C_3)$, respectively. Lane H contains 250 μM $d(C_3T_4C_3)$ alone. DNA samples were loaded in 90 mM TB (pH 7.8), 50 mM $MgCl_2$, and 5% Ficoll and electrophoresed at 4 °C. Triplex refers to $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ and duplex refers to $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$.

transition of the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex corresponds to the equilibrium reaction:

$$d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)] \leftrightarrow d(C_3T_4C_3) + 2[d(G_3A_4G_3)]$$

The equilibrium constant for this reaction may be written as

$$K = \frac{\theta}{4C_{pyr}^2(1-\theta)^3}$$

where θ is the fraction of $d(C_3T_4C_3)$ strands in the triplex form and C_{pyr} is the concentration of $d(C_3T_4C_3)$ strands. Using a derivation similar to that reported by Gralla and Crothers (1973), we determined the value for θ at which the $d(A)/d(1/T)$ vs T curves reach their maxima ($T = T_{max}$) to be 0.366.

According to the method we have previously reported (Pilch et al., 1990b), the variation of T_{max} with pyrimidine strand concentration can be written as

$$\frac{1}{T_{max}} = \frac{2.3R}{\Delta H^\circ} \log C_{pyr}^2 + \frac{\Delta S^\circ + 2.03}{\Delta H^\circ}$$

If the changes in the standard enthalpy (ΔH°) and standard entropy (ΔS°) are assumed to be temperature-independent, they can be determined from plots of $1/T_{max}$ vs $\log C_{pyr}^2$ (Albergo et al., 1981; Pilch et al., 1990b). All $1/T_{max}$ vs log (concentration) data were analyzed by linear regression. The standard free energy changes (ΔG°) of triplex formation, as well as all the thermodynamic parameters for duplex formation, were determined as previously reported (Pilch et al., 1990b).

Results and Discussion

PAGE Retardation Assay. The appropriate stoichiometries associated with the interactions of single-stranded nucleic acids are often determined from UV mixing curves (Felsenfeld et al., 1957; Riley et al., 1966; Broitman et al., 1987; Pilch et al., 1990a,b). A second technique by which these stoichiometries, as well as the number of existing structural species, can be determined is a PAGE retardation assay, in which mixtures containing different stoichiometric ratios of strands are electrophoretically separated. The electrophoretic migration patterns of solutions containing differing stoichiometric ratios of $d(G_3A_4G_3)$ to $d(C_3T_4C_3)$ in a Mg^{2+}-containing gel are shown in Figure 1. At $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratios of 4:1, 2:1, and 1.3:1 (Figure 1A, lanes B, C, and D, respectively), a band is present that migrates more slowly than any other on the gel. This band corresponds to the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex. The other bands in lanes B and D reflect single-stranded $d(G_3A_4G_3)$ and the $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$ duplex, respectively. Only a single band, corresponding to the duplex, is present at a 1:1 $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratio (Figure 1A, lane E), while two bands, reflecting the duplex and single-stranded $d(C_3T_4C_3)$, are present at $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratios of 1:2 and 1:3.5 (Figure 1A, lanes F and G, respectively). Hence, the number and types of species present in solution over the entire range of $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratios may be summarized as follows: First, mixtures of triplex and single-stranded $d(G_3A_4G_3)$ are present at ratios in excess of 2:1. Second, mixtures of triplex and duplex are present between ratios of 2:1 and 1:1. Third, mixtures of duplex and single-stranded $d(C_3T_4C_3)$ are present at ratios in which the $d(C_3T_4C_3)$ concentration exceeds that of $d(G_3A_4G_3)$. UV mixing curves, measured at 225 nm, show slope changes at $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratios of 2:1 and 1:1 (data not shown), thus corroborating the stoichiometry of strands in the triplex observed in the gel retardation analysis.

The retarded gel mobility of the triplex relative to the duplex or either single strand is consistent with the triplex having a higher molecular weight yet similar charge density than these other species and has been observed by others (Cooney et al., 1988; Shea et al., 1990). By similar reasoning, one would expect the gel mobility of the duplex to be retarded relative to either single strand. However, only the $d(C_3T_4C_3)$ strand migrates more rapidly than the duplex, while the $d(G_3A_4G_3)$ strand comigrates with the duplex (Figure 1A, compare lanes A, E, and H). At strand concentrations of up to 100 μM, control thermal denaturation and UV absorption spectroscopy experiments provide little or no evidence that $d(G_3A_4G_3)$ self-associates or aggregates in the presence of 50 mM $MgCl_2$. The anomalously slow migration of the $d(G_3A_4G_3)$ strand is therefore not likely to be a result of either of these two processes. A variation in the coiled structure induced by the Mg^{2+} cations or a reduced negative charge density, resulting from $d(G_3A_4G_3)$ binding more Mg^{2+} ions than $d(C_3T_4C_3)$, may account, in part, for this unusually slow migration. At a 1:2 $d(G_3A_4G_3)$:$d(C_3T_4C_3)$ ratio, there is no evidence of a band that migrates more slowly than that corresponding to the duplex (Figure 1A, compare lanes E and F). The $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex is therefore not formed under the gel conditions used in this study (50 mM $MgCl_2$, pH 8.0). This result is not surprising, since formation of this triplex would require protonation of the third-strand cytosine residues and thus an acidic pH.

The gel pictured in Figure 1A is also shown in Figure 1B after being stained with EtdBr. Only the bands corresponding to the duplex and triplex are stained by the EtdBr, suggesting that ethidium bromide binds to both the duplex and triplex forms. Scaria and Shafer (1991) have recently demonstrated that EtdBr binds to the poly(dA)·2[poly(dT)] triplex through intercalation. The results from Figure 1B indicate that the binding of EtdBr to triple-helical DNA may extend to short oligonucleotides of differing sequence and third strands consisting of homopurine residues. Since EtdBr does not stain single-stranded DNA, this technique provides additional support as to which bands in Figure 1A reflect single-stranded species.

Imino Proton NMR. Recent NMR studies (Rajagopal & Feigon, 1989; Santos et al., 1989; Pilch et al., 1990a; Umemoto

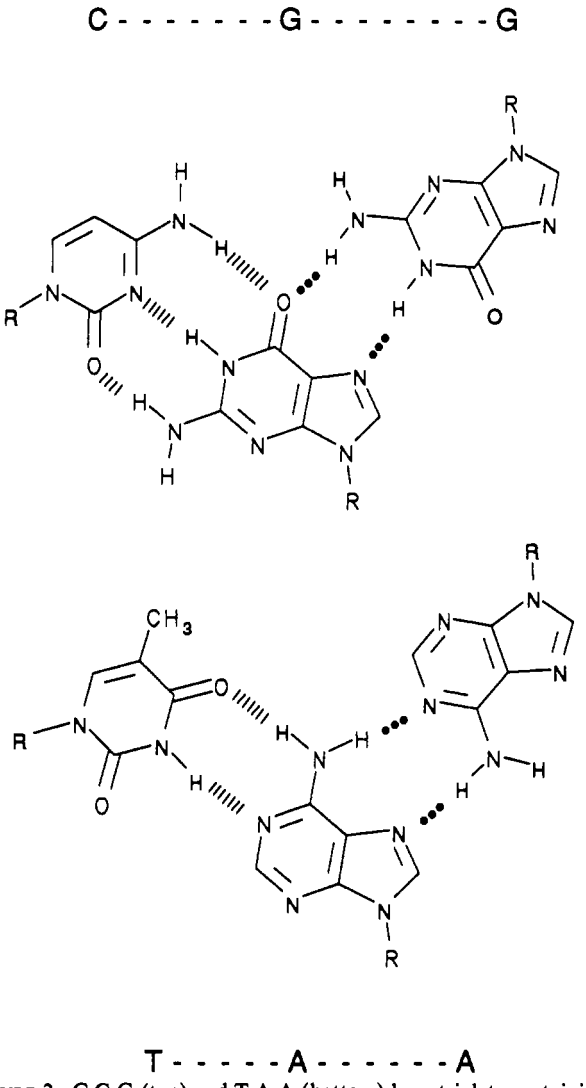

FIGURE 2: C·G·G (top) and T·A·A (bottom) base triplets containing both Watson–Crick- (hatched lines) and purine–purine-type (filled circles) hydrogen bonds as discussed in the text.

et al., 1990) have demonstrated the presence and structural characteristics of T·A·T and C^+·G·C base triplets in pyr-pur-pyr triple-helical systems. Figure 2 shows a possible model for the C·G·G and T·A·A base triplets of the $d(C_3T_4C_3)$·2-$[d(G_3A_4G_3)]$ triplex. The model for the C·G·G triplet postulates that one of the two hydrogen bonds in the G·G base pair is between the NH1 imino proton of the third-strand guanine residue and the N7 atom of the guanine residue on the purine strand of the underlying duplex. Upon formation of this hydrogen bond, the exchange of the imino proton with the solvent should sufficiently decrease so as to allow an observable resonance line to arise. Figure 3 shows the imino proton NMR spectra of a solution containing a 2:1 stoichiometric ratio of $d(G_3A_4G_3)$ to $d(C_3T_4C_3)$ in the absence (A) and presence (B) of 14 mM $MgCl_2$ at pH 7.2 and a temperature of 35 °C. In the absence of any $MgCl_2$, the imino proton spectrum (Figure 3A) is identical with that for the $d(C_3T_4C_3)$·$d(G_3A_4G_3)$ duplex under identical conditions (data not shown). The second equivalent of the $d(G_3A_4G_3)$ strand is therefore present in an unassociated single-stranded form under these conditions. Upon addition of $MgCl_2$ to a final concentration of 14 mM, new imino proton resonances arise (Figure 3B), which likely correspond to formation of the G·G hydrogen bonds discussed above and suggest stabilization of the triplex structure. Since adenine bases do not have imino protons at neutral pH, none of the new imino proton resonances

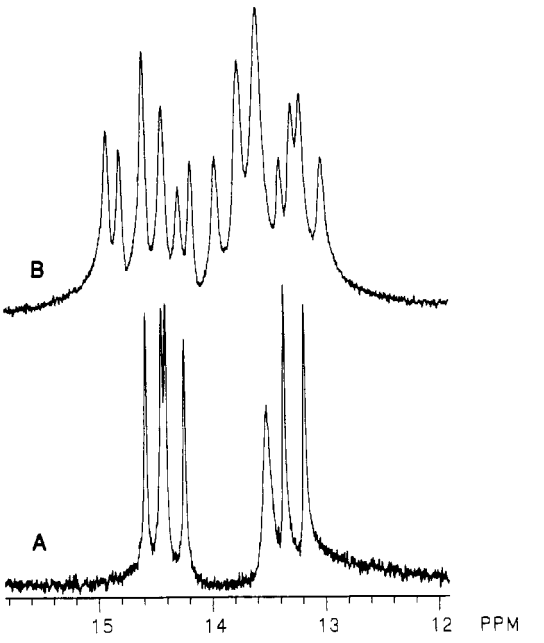

FIGURE 3: Imino proton NMR spectra of a solution containing a 2:1 stoichiometric ratio of $d(G_3A_4G_3)$ to $d(C_3T_4C_3)$ and either 0 (A) or 14 mM (B) $MgCl_2$. Experiments were carried out in 15 mM sodium phosphate (pH 7.2) plus 10% D_2O/90% H_2O at 35 °C.

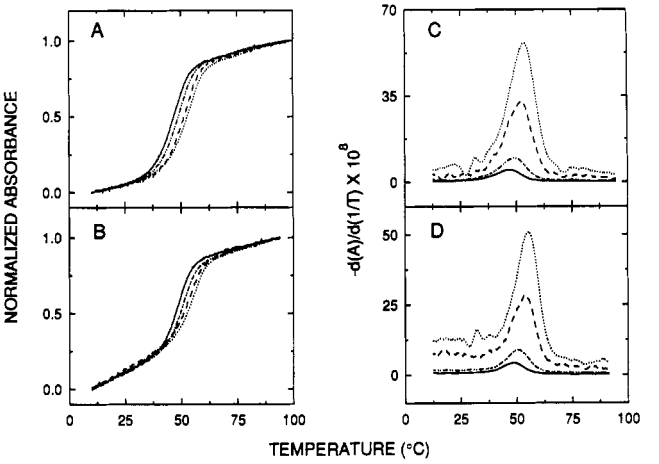

FIGURE 4: Helix–coil transitions and first-derivative plots for four different concentrations of the $d(C_3T_4C_3)$·$d(G_3A_4G_3)$ duplex (panels A and C, respectively) and the $d(C_3T_4C_3)$·2[$d(G_3A_4G_3)$] triplex (panels B and D, respectively). The helix (either duplex or triplex) concentrations are 2.5 (solid line), 5.0 (dashed–double dot line), 15 (dashed line), and 25 μM (dotted line). Raw absorbances were normalized by the following relation: normalized absorbance = $(A - A_{initial})/(A_{final} - A_{initial})$, where A = absorbance. All thermal denaturation experiments were conducted in 10 mM sodium cacodylate (pH 7.3), 50 mM $MgCl_2$, and 0.1 mM EDTA.

reflect A·A interactions. However, changes that arise in the amino proton region (6–10 ppm) of the NMR spectrum upon addition of $MgCl_2$ (data not shown) may correspond to formation of the A·A base pairs, as modeled in Figure 2.

Helix–Coil Transitions. Thermal denaturation profiles of both the $d(C_3T_4C_3)$·$d(G_3A_4G_3)$ duplex and the $d(C_3T_4C_3)$·2-$[(G_3A_4G_3)]$ triplex were determined spectrophotometrically at 260 and 280 nm, respectively, in 50 mM $MgCl_2$, pH 7.3. The helix–coil transitions at differing DNA concentrations and their first derivatives with respect to reciprocal temperature are shown in Figure 4, panels A and C, for the duplex and in Figure 4, panels B and D, for the triplex. For both structures, the helix–coil transitions are monophasic, reflecting duplex ↔ coil (Figure 4A) and triplex ↔ coil (Figure 4B) equilibria. The monophasic nature of the thermal denaturation profile

Table I: Thermodynamic Parameters for Formation of the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ Triplex and the $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$ Duplex[a]

structure	T_{max} (°C)[d]	$\Delta H°$ (kcal/mol)[e,f]	$\Delta S°$ (eu)[f]	$\Delta G°_{25}$[g] (kcal/mol)[e,f]
duplex[b]	52.0 ± 1.0	−71.6 ± 9.7	−198 ± 30	−12.6 ± 0.7
triplex[c]	54.0 ± 1.0	−152.5 ± 6.5	−424 ± 20	−26.0 ± 0.5

[a] Data were obtained in 10 mM sodium cacodylate, 50 mM $MgCl_2$, and 0.1 mM EDTA at pH 7.3. [b] The reaction is $d(C_3T_4C_3)$ + $d(G_3A_4G_3) \rightarrow d(C_3T_4C_3)\cdot d(G_3A_4G_3)$. [c] The reaction is $d(C_3T_4C_3)$ + $2[d(G_3A_4G_3)] \rightarrow d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$. [d] T_{max} at a pyrimidine strand concentration of 15 μM. [e] mol refers to mole of structure formed (duplex or triplex). [f] Errors reflect the magnitudes of the 95% confidence limits from linear regression analyses. [g] $\Delta G°_{25}$ refers to the standard free energy change at 25 °C.

observed for the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex is in marked contrast to that previously determined for the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex under identical conditions with the exception of pH (Pilch et al., 1990b). The melting profile of this latter triplex was biphasic in nature, with the first (lower temperature) phase corresponding to the dissociation of the third strand and the second (higher temperature) phase reflecting the denaturation of the remaining duplex.

In the first-derivative plots of the melting profiles for both the duplex (Figure 4C) and the triplex (Figure 4D), T_{max} increases with increasing DNA concentration, indicating that these processes are at least bimolecular in nature. In addition, depending on the DNA concentration, the T_{max} values for the triplex–coil transitions are 1–3 °C higher than those for the duplex–coil transitions under identical conditions (compare panels C and D of Figure 4 and see Table I). Hence, binding of the third strand appears to enhance the thermal stability of the duplex. Fiber diffraction (Arnott & Selsing, 1974; Arnott et al., 1974, 1976) and, more recently, NMR (Rajagopal & Feigon, 1989; Santos et al., 1989) studies have shown that the duplex undergoes significant structural changes when bound by the third strand. These changes may contribute to the observed increase in duplex stability.

An important difference between the thermal denaturation data for the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex presented in this study and those previously determined for the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex (Pilch et al., 1990b) can be found in the stability of the third strand. Under conditions differing only with respect to pH (7.3 in this study vs 5.5) and at equivalent DNA concentrations, the T_{max} values of the triplex–coil transitions shown in Figure 4D are approximately 30 °C higher than those of the lower temperature phases of the biphasic melting profiles for the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex. The $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex is therefore remarkably more stable than the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex. Furthermore, unlike the requirements for stabilization of the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex, formation of the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex does not require an acidic pH and will thus readily occur at physiologically relevant values of pH. These two key dissimilarities suggest that pur-pur-pyr triplexes may be better choices than the corresponding pyr-pur-pyr triplexes for use as sequence-specific artificial nucleases and regulators of gene expression in vitro and in vivo. A caveat to this approach lies in the limited number sequences capable of forming pur-pur-pyr triplexes. We have never been able to find solution conditions under which formation of the $d(T)_{10}\cdot 2[d(A)_{10}]$ triplex would occur (data not shown). The findings of Broitman et al. (1987), which demonstrate that only poly(A) strands between 28 and 150 bases in length will form the poly(A·A·U) triplex, would predict this negative result. Hence, the presence of guanine residues appears to be crucial for stabilization of short pur-pur-pyr triplexes. The dependence of the stability of short pur-pur-pyr triplexes on overall composition and sequence with respect to guanine residues remains to be determined.

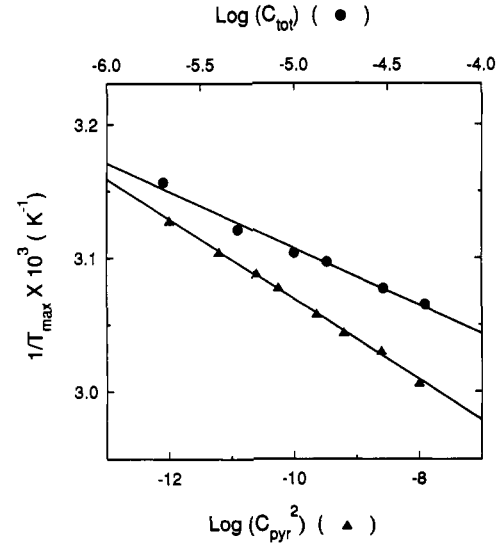

FIGURE 5: Plots of $1/T_{max}$ vs log C_{tot} (upper abscissa) and $1/T_{max}$ vs log C_{pyr}^2 (lower abscissa) for the $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$ duplex (●) and the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex (▲), respectively. C_{tot} is the total concentration of all strands and C_{pyr} is the concentration of the $d(C_3T_4C_3)$ strands. K refers to degrees kelvin. See Materials and Methods for the definition of T_{max}.

Thermodynamic Analysis. Plots of $1/T_{max}$ vs log C_{tot} and $1/T_{max}$ vs log C_{pyr}^2 are shown in Figure 5 for the $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$ duplex–coil and the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex–coil transitions, respectively. All the experiments were conducted under identical solution conditions. The helix–coil transitions of both the duplex and the triplex have $1/T_{max}$ values that decrease linearly with increasing log C_{tot} and log C_{pyr}^2, respectively. In addition to confirming that these transitions are at least bimolecular in nature, this linear correlation is consistent with duplex and triplex dissociation occurring as all-or-none processes. Table I summarizes the thermodynamic parameters derived from the plots in Figure 5 as described under Materials and Methods. In this table, the values of $\Delta G°$, $\Delta H°$, and $\Delta S°$ refer to formation of a duplex or triplex from separated single strands.

The enthalpy, entropy, and free energy changes for formation of the $d(C_3T_4C_3)\cdot d(G_3A_4G_3)$ duplex shown in Table I are very close to those previously determined for this duplex under identical solution conditions except for a pH of 5.5 (Pilch et al., 1990b). Therefore, a pH change from 5.5 to 7.3 has little or no effect on the stability of the duplex. Since the thermodynamics of $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex formation in 50 mM $MgCl_2$ at pH 5.5 has not as yet been evaluated, the effect, if any, of such a pH change on the stability of this triplex is unknown. However, PAGE analysis in 50 mM $MgCl_2$ at pH 5.5 demonstrates formation of this structure under these conditions (data not shown).

The enthalpy, entropy, and free energy changes for formation of the $d(C_3T_4C_3)\cdot 2[d(G_3A_4G_3)]$ triplex are approximately twice those for formation of the corresponding duplex. An estimate of the free energy change for formation of the pur·pur base pairs alone may be obtained from

$$\Delta\Delta G_{25} = [\Delta G_{25}(\text{triplex}) - \Delta G_{25}(\text{duplex})]$$

The free energy change derived in this way can only be con-

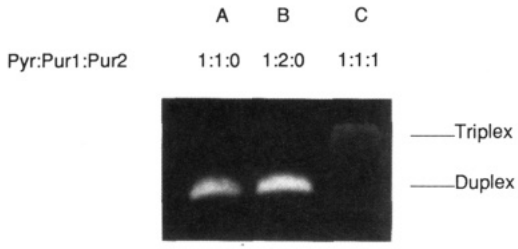

FIGURE 6: Photograph of an EtdBr-stained polyacrylamide gel containing molar ratios of either 1**Pyr**:1**Pur1** (lane A), 1**Pyr**:2**Pur1** (lane B), or 1**Pyr**:1**Pur1**:1**Pur2** (lane C). Lanes A, B, and C contain 19, 38, and 19 μM **Pur1**, respectively, and 19 μM **Pyr**. Lane C also contains 19 μM **Pur2**. Experimental conditions were identical with those described in the legend of Figure 1. Duplex refers to **Pyr·Pur1** and triplex refers to **Pyr·Pur1·Pur2**. See text for the exact sequences of the oligomers.

sidered an estimate, since the structure and stability of the duplex appear to differ when it is free or bound by a third strand. The difference between the $\Delta\Delta G_{25}$ for formation of the pur·pur base pairs (-13.4 ± 1.2 kcal/mol of triplex) and the ΔG_{25} for formation of the W-C base pairs (-12.6 ± 0.7 kcal/mol of duplex) is within the experimental error. Hence, under these conditions, the collective stabilities of the pur·pur base pairs in this sequence are very similar to those of the W-C base pairs. This similarity is in marked contrast to the results we previously reported for the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex in 50 mM $MgCl_2$ at pH 5.5, where the Hoogsteen base pairs had only half the stability (-6.4 kcal/mol of triplex) of their W-C counterparts (-12.8 kcal/mol of duplex). Although the pH values differed in the two studies, this result strongly suggests that, under identical solution conditions, pur-pur-pyr triplexes are significantly more stable structures than pyr-pur-pyr triplexes identical in sequence with respect to their W-C base-paired portions. The greater stability of pur-pur-pyr relative to pyr-pur-pyr triplexes may prove advantageous in oligonucleotide targeting of duplex DNA. It should be noted, however, that different concentrations and/or other types of cations, such as Na^+ and spermine^{4+}, can greatly enhance the stability of pyr-pur-pyr triplexes, so as to yield monophasic triplex-to-coil melting behavior as well (data not shown).

Orientation of the Third Strand in a Pur-Pur-Pyr Triplex. Hogan and co-workers have recently reported on two different triplexes in which the third strands consisted of predominantly purine bases (Cooney et al., 1988; Durland et al., 1990). In one case the third strand was synthesized parallel to the purine strand of the duplex, while in the other case it was designed to be antiparallel. In an effort to resolve this issue of orientation, we synthesized an asymmetric homopyrimidine decamer (**Pyr**), its W-C homopurine complement (**Pur1**), and a second homopurine strand antiparallel to the first one (**Pur2**). The electrophoretic migration patterns of mixtures containing stoichiometric ratios of either 1**Pyr**:1**Pur1** (lane A), 1**Pyr**:2**Pur1** (lane B), or 1**Pyr**:1**Pur1**:1**Pur2** (lane C) are shown in Figure 6. This gel was stained with EtdBr; thus bands corresponding to single-stranded species do not appear. At a ratio of either 1**Pyr**:1**Pur1** or 1**Pyr**:2**Pur1** (lanes A and B, respectively), only a single band is present, which corresponds to the **Pyr·Pur1** duplex. Therefore, triplex formation does not occur when a 1:2 ratio of **Pyr** to **Pur1** is present. In contrast, at a 1**Pyr**:1**Pur1**:1**Pur2** ratio (lane C), a single more slowly migrating band is present, which corresponds to the **Pyr·Pur1·Pur2** triplex. The orientations of all the strands in this triplex may therefore be schematically represented as

5′ - AAGGAGAGGG - 3′ (**Pur2**)
| | | | | | | | | | |
3′ - AAGGAGAGGG - 5′ (**Pur1**)
| | | | | | | | | | |
5′ - TTC CTC TCCC - 3′ (**Pyr**)

In general, this result indicates that the third strands of pur-pur-pyr triplexes, including the symmetric $d(C_3T_4C_3)\cdot 2\text{-}[d(G_3A_4G_3)]$ triplex discussed above, are antiparallel to the purine strands of the underlying duplexes. It is interesting to note that although equivalent with respect to **Pyr** concentration, the triplex in lane C was stained by EtdBr to a much lesser degree than the duplex (lanes A and B). EtdBr may therefore have a lower affinity for this particular triplex relative to the corresponding duplex.

SUMMARY

In this work we have demonstrated that the $d(C_3T_4C_3)\cdot 2\text{-}[d(G_3A_4G_3)]$ triplex is stabilized by Mg^{2+} ions at neutral pH. The third strand [or second $d(G_3A_4G_3)$ strand] interacts with the $d(G_3A_4G_3)$ strand of the underlying duplex via novel hydrogen bonds, of which those likely involving guanine imino protons can be observed by ^1H NMR techniques. For pur-pur-pyr triplexes in general, the orientation of the third strand is antiparallel to the purine strand of the duplex and can be easily verified by PAGE methodologies. Under the conditions used in this study, the triplex dissociates to single strands in one step when heated. The overall stability of the pur·pur base pairs is approximately equal to that of the W-C base pairs of the duplex. Furthermore, the stability of the third strand is roughly twice that of the third strand in the $d(C^+_3T_4C^+_3)\cdot d(G_3A_4G_3)\cdot d(C_3T_4C_3)$ triplex (Pilch et al., 1990b), under solution conditions differing only with respect to pH (7.3 vs 5.5). The enhanced stability relative to pyr-pur-pyr triplexes, coupled with the lack of a requirement for acidic pH, make pur-pur-pyr triplexes appealing choices for use as sequence-specific artificial nucleases and/or regulators of gene expression in vitro and in vivo.

ADDED IN PROOF

Recent reports on formation of an intermolecular pur-pur-pyr triplex (Beal & Dervan, 1991) and a pur-pur-pyr triplex from a single strand (Chen, 1991) have appeared. Both reports indicate a third-strand orientation, which is consistent with the result presented here.

ACKNOWLEDGMENTS

We thank Ms. Laurie Goda of the Cetus Corp., Emeryville, CA, for her help during the initial stages of the PAGE work and Dr. Ching C. Wang of the University of California, San Francisco, for use of his gel photography equipment.

REFERENCES

Albergo, D. D., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1409–1413.
Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521.
Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C., & Watts, L. (1974) *J. Mol. Biol. 88*, 523–533.
Arnott, S., Bond, P. J., Selsing, E., & Smith, P. J. C. (1976) *Nucleic Acids Res. 3*, 2459–2470.
Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360–1363.
Bernués, J., Beltrán, R., Casasnovas, J. M., & Azorín, F. (1989) *EMBO J. 8*, 2087–2094.
Bernués, J., Beltrán, R., Casasnovas, J. M., & Azorín, F. (1990) *Nucleic Acids Res. 18*, 4067–4073.

Broitman, S. L., Im, D. D., & Fresco, J. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5120–5124.
Caddle, M. S., Lussier, R. H., & Heintz, N. H. (1990) *J. Mol. Biol. 211*, 19–33.
Chen, F.-M. (1991) *Biophys. J. 59*, 494a.
Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science 241*, 456–459.
Durland, R. H., Kessler, D. J., Duvic, M., & Hogan, M. (1990) in *Molecular Basis of Specificity in Nucleic Acid–Drug Interactions* (Pullman, B., & Jortner, J., Eds.) pp 565–578, Kluwer Academic Publishers, Boston.
Felsenfeld, G., Davies, D. R., & Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023–2024.
Glaser, R. L., Thomas, G. H., Siegfried, E., Elgin, S. C. R., & Lis, J. T. (1990) *J. Mol. Biol. 211*, 751–761.
Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301–319.
Griswold, B. L., Humoller, F. L., & McIntyre, A. R. (1951) *Anal. Chem. 23*, 192–194.
Hanvey, J. C., Shimizu, M., & Wells, R. D. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6292–6296.
Hélène, C., & Toulmé, J. J. (1990) *Biochim. Biophys. Acta 1049*, 99–125.
Hélène, C., Thuong, N. T., Saison-Behmoaras, T., & François, J. C. (1989) *Trends Biotechnol. 7*, 310–315.
Hore, P. J. (1983) *J. Magn. Reson. 54*, 539–542.
Htun, H., & Dahlberg, J. E. (1988) *Science 241*, 1791–1796.
Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603–6614.
Maher, L. J., III, Wold, B., & Dervan, P. B. (1989) *Science 245*, 725–730.
Manzini, G., Xodo, L. E., Gasparotto, D., Quadrifoglio, F., Van Der Marel, G. A., & Van Boom, J. H. (1990) *J. Mol. Biol. 213*, 833–843.
Marck, C., & Thiele, D. (1978) *Nucleic Acids Res. 5*, 1017–1028.
Marck, C., Thiele, D., Schneider, C., & Guschlbauer, W. (1978) *Nucleic Acids Res. 5*, 1979–1996.
Minton, K. W. (1985) *J. Exp. Pathol. 2*, 135–148.
Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filipov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495–497.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Perrouault, L., Asseline, U., Rivalle, C., Thuong, N. T., Bisagni, E., Giovannangeli, C., Le Doan, T., & Hélène, C. (1990) *Nature 344*, 358–360.
Pilch, D. S., Levenson, C., & Shafer, R. H. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1942–1946.
Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990b) *Nucleic Acids Res. 18*, 5743–5750.
Plum, G. E., Park, Y., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436–9440.
Rajagopal, P., & Feigon, J. (1989) *Biochemistry 28*, 7859–7870.
Riley, M., Maling, B., & Chamberlin, M. J. (1966) *J. Mol. Biol. 20*, 359–389.
Santos, C. D. L., Rosen, M., & Patel, D. (1989) *Biochemistry 28*, 7282–7289.
Scaria, P. V., & Shafer, R. H. (1991) *J. Biol. Chem. 266*, 5417–5423.
Shea, R. G., Ng, P., & Bischofberger, N. (1990) *Nucleic Acids Res. 18*, 4859–4866.
Strobel, S. A., Moser, H. E., & Dervan, P. B. (1988) *J. Am. Chem. Soc. 110*, 7927–7929.
Sun, J. S., François, J. C., Montenay-Garestier, T., Saison-Behmoaras, T., Roig, V., Thuong, N. T., & Hélène, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9198–9202.
Thrierr, J. C., & Leng, M. (1972) *Biochim. Biophys. Acta 272*, 238–251.
Umemoto, K., Sarma, M. H., Gupta, G., Luo, J., & Sarma, R. H. (1990) *J. Am. Chem. Soc. 112*, 4539–4545.
Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P., & Frank-Kamenetskii, M. D. (1988) *Nature 333*, 475–476.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939–2949.