Laboratory, Cold Spring Harbor, NY.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035-10038.
McIntyre, J. C., Hundley, P., & Behnke, W. D. (1987) *Biochem. J. 245*, 821-829.
Mierendorf, R. C., Percy, C., & Young, R. A. (1987) *Methods Enzymol. 152*, 458-469.
Mueckler, M., & Lodish, H. (1986) *Cell 44*, 629-637.
Rathelot, J., Canioni, P., Bosc-Bierne, I., Sarda, L., Kamoun, A., Kaptein, R., & Cozzone, P. J. (1981) *Biochim. Biophys. Acta 671*, 155-163.
Sanger, F., Miklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Sari, H., Granon, S., & Semeriva, M. (1978) *FEBS Lett. 95*, 229-234.
Scheele, G. A., Steinhilber, W., Swarovsky, B., & Kern, H. F. (1988) *J. Cell Biol. 107*, 763a.
Steinhilber, W., Poensgen, J., Rausch, U., Kern, H. F., & Scheele, G. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6597-6601.
Stenflo, J., & Fernlund, P. (1982) *J. Biol. Chem. 257*, 12180-12190.
Sternby, B., & Borgstrom, B. (1979) *Biochim. Biophys. Acta 572*, 235-243.
Sternby, B., & Erlanson-Albertsson, C. (1982) *Biochim. Biophys. Acta 711*, 193-195.
Sternby, B., & Borgstrom, B. (1984) *Biochim. Biophys. Acta 786*, 109-112.
Sternby, B., Engstrom, A., Hellman, U., Vihert, A. M., Sternby, N., & Borgstrom, B. (1984) *Biochim. Biophys. Acta 784*, 75-80.
Stratowa, C., & Rutter, W. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4292-4296.
Stratowa, C., & Rutter, W. J. (1987) *Biochemistry 26*, 1617-1625.
von Heijne, G. (1983) *Eur. J. Biochem. 133*, 17-21.
Weiloch, T., & Falk, K. (1978) *FEBS Lett. 85*, 271-274.
Weiloch, T., Borgstrom, B., Pieroni, G., Pattus, F., & Verger, R. (1981) *FEBS Lett. 128*, 217-220.

A·T and C·C$^+$ Base Pairs Can Form Simultaneously in a Novel Multistranded DNA Complex†

Eric L. Edwards,‡ Michael H. Patrick,‡,∥ Robert L. Ratliff,§ and Donald M. Gray*,‡
Program in Molecular and Cell Biology, Mail Stop FO 31, The University of Texas at Dallas, Box 830688, Richardson, Texas 75083-0688, and Genetics Group, Life Sciences Division, Mail Stop 886, Los Alamos National Laboratory, Box 1663, Los Alamos, New Mexico 87545

Received April 26, 1989; Revised Manuscript Received September 1, 1989

ABSTRACT: Previous experiments have established that in certain synthetic oligomeric DNA sequences, including mixtures of d(AACC)$_5$ with d(CCTT)$_5$, adenine–thymine (A·T) base pairs form to the exclusion of neighboring protonated cytosine–cytosine (C·C$^+$) base pairs [Edwards, E., Ratliff, R., & Gray, D. (1988) *Biochemistry 27*, 5166-5174]. In the present work, circular dichroism and other measurements were used to study DNA oligomers that represented two additional classes with respect to the formation of A·T and/or C·C$^+$ base pairs. (1) One class included two sets of repeating pentameric DNA sequences, d(CCAAT)$_{3-6}$ and d(AATCC)$_{4,5}$. For both of these sets of oligomers, an increase in the magnitude of the long-wavelength positive CD band centered at about 280 nm occurred as the pH was lowered from 7 to 5 at 0.1 and 0.5 M Na$^+$, indicating that C·C$^+$ base pairs formed. Even though it may have been possible for these oligomers to form duplexes with two antiparallel A·T base pairs per pentamer, no A·T base pairing was detected by monitoring the CD changes at 250 nm. Thus, spectral data showed that as few as 40% C·C$^+$ base pairs were stable in two sets of oligomers in which A·T base pairs did not form adjacent to, or in place of, C·C$^+$ base pairs. (2) Another class of oligomer was represented by d(C$_4$A$_4$T$_4$C$_4$), which was studied by CD, HPLC, and centrifugation experiments. We confirmed previous work that this sequence was able to form both types of base pairs as the pH and temperature were lowered [Gray, D., Cui, T., & Ratliff, R. (1984) *Nucleic Acids Res. 12*, 7565-7580]. We further established that at low pH, where both A·T and C·C$^+$ base pairs formed concurrently in d(C$_4$A$_4$T$_4$C$_4$), there was an increase in molecular weight above that expected for a duplex. Our interpretation of all these data is that those sequences that formed only A·T base pairs had an antiparallel-strand orientation while those that formed only C·C$^+$ base pairs probably had a parallel-strand orientation. We propose that d(C$_4$A$_4$T$_4$C$_4$) at low pH adopted a multistranded structure that contained both parallel-stranded C·C$^+$ base pairs and antiparallel-stranded A·T base pairs.

It is well-known that DNA can adopt a wide variety of conformations. The intensity of recent research on DNA structural polymorphism attests to the awareness that structures other than canonical B-form DNA may play important roles in transcriptional regulation (Nordheim & Rich, 1983), homologous recombination during meiosis or at immunoglobin switch regions (Sen & Gilbert, 1988), relief of torsional stress (Cantor & Efstratiadis, 1984), and telomere formation (Oka & Thomas, 1987; Henderson et al., 1987). Polypurine–polypyrimidine tracts can be located in regions of regulatory importance (Kilpatrick et al., 1986), and it is common for such sequences to have domains with high cytosine content (Evans

† This work was performed by E.L.E. in partial fulfillment of the requirements for the Ph.D. degree in the Program in Molecular and Cell Biology, The University of Texas at Dallas. Support was provided by NIH Research Grant GM 19060 and by Grant AT-503 from the Robert A. Welch Foundation. Work at the Los Alamos National Laboratory was supported in part by the U.S. Department of Energy.
‡ The University of Texas at Dallas.
∥ Current address: Box 41A, Saxeville Rd., Pine River, WI 54965.
§ Los Alamos National Laboratory.

0006-2960/90/0429-828$02.50/0

et al., 1984). Enzymatic and chemical probes indicate that some of these regions can exist in altered conformational states (Dybvig et al., 1983; Fowler & Skinner, 1986; Mirkin et al., 1987; Htun & Dahlberg, 1988; Johnson, 1988). Although it is known that these regions have noncanonical structures, it is not yet known what range of conformations is available for the cytosine-rich polypyrimidine strands to adopt. Neither is it known to what extent the local and surrounding sequences can influence the ability of a cytosine-rich region to form unusual structures.

Conformations available to several protonated cytosine-containing DNA oligomers and polymers have been characterized. CD spectroscopy is useful for studying these conformations because the formation of stacked, protonated cytosine–cytosine base pairs (C·C+ base pairs)[1] and normal Watson–Crick A·T base pairs in DNA can be readily monitored. Each type of base pair has characteristic CD spectral bands that allow its detection and quantitation. C·C+ base pairs show a greatly enhanced positive CD at wavelengths above 280 nm (Gray & Bollum, 1974; Marck et al., 1978; Gray et al., 1988), while A·T base pairs show a large increase in negative band magnitude at about 250 nm (Greve et al., 1977). The formation of C·C+ base pairs in the self-complex of poly[d(C)] results in distinctive CD spectral changes and occurs with a pK_a of 7.4 at 0.05 M Na^+ (Inman, 1964). A self-complex of poly[d(CT)] with a pK_a of 6.2 at 0.05 M Na^+ is formed upon hemiprotonation of the cytosines (Gray et al., 1980; Brown et al., 1985; Sarma et al., 1986). This self-complex consists of a central core of stacked C·C+ base pairs with the intervening thymines looped out into solution. A similar structure with looped-out thymines and paired cytosines can occur for $d(CCTT)_5$ at pH 6, 0.5 M Na^+, while its complement, $d(AACC)_5$, appears to retain its adenines intrahelical upon C·C+ base pair formation (Edwards et al., 1988).

The list of synthetic DNA sequences in which C·C+ base pairs can form has been steadily growing, but as yet it has not been possible to predict, a priori, sequence requirements that would allow both A·T and C·C+ base pairs to form adjacent to each other. We have previously found that duplexes of $d(CCTT)_5$·$d(AACC)_5$, $d(A_6C_6A_6)$·$d(T_6C_6T_6)$, or $d(A_{10}C_4T_{10})$·$d(A_{10}C_4T_{10})$ contain only A·T and not C·C+ base pairs, even at pH 5, 0.5 M Na^+, 5 °C (Edwards et al., 1988). Because the separate oligomers $d(CCTT)_5$, $d(AACC)_5$, $d(A_6C_6A_6)$, and $d(T_6C_6T_6)$ can individually form C·C+ base pairs under these same conditions, we concluded that A·T base pairs in the above duplexes prevented the formation of neighboring C·C+ base pairs. This stood in contrast to the results obtained with the acid-induced self-complex of $d(C_4A_4T_4C_4)$, in which A·T base pairs facilitated the formation of adjacent C·C+ base pairs (Gray et al., 1984).

Structural characteristics of C·C+ base pairs within oligomers that also contain adenine and thymine remain unknown. These include (1) the number of hydrogen bonds (two vs three), (2) the glycosidic linkage (syn vs anti), and (3) the strand orientation (antiparallel vs parallel). A centrosymmetric, parallel-strand pairing of cytosines with three hydrogen bonds and anti-glycosidic linkages was reported for the crystal structures of the dinucleotides r(CpA) (Westhof & Sundaralingham, 1980) and d(CpG) (Cruse et al., 1983; Coll et al., 1987). However, these crystal structures contain unusual A·A or G·G base pairs, and the structure of a C·C+ base pair may be different if it were next to other types of base pairs. A nuclear magnetic resonance study showed that a self-complex of $d(CT)_3$ at low pH has base-paired cytosines that adopted anti conformations about their glycosidic bonds (Sarma et al., 1986). On the assumption that three hydrogen bonds formed per C·C+ base pair, this would indicate a parallel-strand polarity in $d(CT)_3$. However, an antiparallel-strand orientation with anti-glycosyl linkages would be possible if a type of C·C+ base pair formed that has two hydrogen bonds (Gray et al., 1984).

In this present work we address two key questions concerning base pair formation in adenine-, thymine-, and cytosine-containing DNA sequences: (1) what is the strand orientation in those sequences that form C·C+ base pairs exclusively, and (2) why is $d(C_4A_4T_4C_4)$ an exception to other adenine-, thymine-, and cytosine-containing sequences that are able to form only one type of base pair?

Materials and Methods

Preparation of DNA Oligomers. The oligomers $d(A)_{20}$, $d(T)_{20}$, and $d(C)_{10}$ were purchased from P-L Biochemicals, Inc., and were used without further purification. Oligomers $d(A_6C_6A_6)$ and $d(T_6C_6T_6)$ were synthesized manually with a DNA synthesis kit from New England Biolabs, Inc. All other oligomers shown in Table I were synthesized by solid-phase phosphite triester methods on a Beckman System 1 DNA synthesizer and purified as previously reported (Edwards et al., 1988) except that $d(A_8T_8)$ and $d(C_4A_4T_4C_4)$ were purified on a Beckman reverse-phase C_{18} HPLC column. These two oligomers were collected as single peaks by elution with a 5–20% acetonitrile gradient in 50 mM Na^+ (phosphate), pH 7; they were then electroeluted as before and ethanol precipitated to remove any traces of acetonitrile (Edwards et al., 1988). The DNA oligomers used in the centrifugation experiments were resuspended in H_2O. Oligomers for other experiments were resuspended in 0.01 M Na^+ (phosphate) buffer at pH 7, 6, or 5. Na^+ concentrations were adjusted by the addition of 4 M NaCl. Sample pH was adjusted with HCl for oligomers used in HPLC experiments. All other oligomers were adjusted to the appropriate pH with H_3PO_4.

Extinction Coefficients. Extinction coefficients were calculated from first-neighbor equations using the known extinction coefficients of monomers and dimers (Cantor et al., 1970). For single-stranded $d(CCATT)_{3-6}$, $d(AATCC)_{4,5}$, $d(C_4A_4T_4C_4)$, and $d(A_8T_8)$, extinction coefficients at 260 nm were 8800, 9650, 8700, and 9200 $L \cdot mol^{-1} \cdot cm^{-1}$, respectively. In the calculation of the extinction coefficient for $d(A_8T_8)$, the absorbance of the $d(A)_8$ portion of the sequence was corrected by a factor of 0.8 to take into account the significant second-neighbor effects that contribute to the absorbance of poly[d(A)]. Extinction coefficients for the other oligomers have been previously reported (Edwards et al., 1988). Oligomer concentrations were determined from absorbance measurements and the calculated extinction coefficients.

Mixing Experiments. Mixtures of $d(AATCC)_5$ with $d(CCATT)_5$ were made at a molar strand ratio of 1:1. The mixtures were heated to 60 °C, and spectra were recorded at decreasing temperatures, allowing 30 min for equilibration at each temperature. No absorbance changes were noted after 30 min at a given temperature. The CD spectrum of the mixture of $d(AATCC)_5$ with $d(CCATT)_5$ at pH 6, 0.5 M Na^+, was remeasured at 5 °C after 2 weeks of storage at 5 °C. No detectable changes in the CD at 250 nm, indicative of changes in A·T base pairing, occurred during that time.

Absorption and CD Spectra. Absorption and CD spectra were measured as previously described (Edwards et al., 1988). CD values are reported as $\epsilon_L - \epsilon_R$ in units of $L \cdot mol^{-1} \cdot cm^{-1}$, per

[1] Abbreviations: C·C+, base pair of two cytosines with a shared proton; A·T, Watson–Crick base pair of adenine with thymine; CD, circular dichroism.

mole of nucleotide. During spectral measurements, samples were maintained at the specified temperature with an accuracy of ±0.5 °C. Sample loss due to evaporation was less than 0.3%.

Centrifugation of DNA Oligomers. The DNA oligomers, after purification, were resuspended in H_2O and centrifuged at 15600*g* for 15 min in an Eppendorf table-top microcentrifuge, at room temperature, to remove any particulate scattering material. The supernatant from this centrifugation was removed, NaCl was added to give 0.5 M Na^+, and the DNA solution was then adjusted to pH 7, 6, or 5 with H_3PO_4. The concentrations for all the samples at this point were between 5.98×10^{-5} and 1.46×10^{-4} M nucleotides. The absorbance was recorded at 260 nm after the samples had been held at 60 °C for 15 min to allow for strand denaturation. The DNA was then kept at 4 °C overnight and centrifuged again at 15600*g* for 30 min at 4 °C. The supernatant was removed and heated to 60 °C for 15 min, and the absorbance at 260 nm was recorded a second time. From the change in the A_{260} caused by loss of DNA from centrifugation, the percentage of precipitable DNA was calculated as [1 − (A_{260} after centrifugation)/(A_{260} before centrifugation)] × 100.

High-Pressure Liquid Chromatography of DNA Oligomers. HPLC analysis was carried out on an integrated Beckman System Gold analytical chromatographic system. Chromatography was performed on a Bio-Rad Bio-Gel TSK DEAE-5-PW anion exchange column (75 mm × 7.5 mm) run at a flow rate of 1.0 mL/min. The mean pore diameter of the beads was approximately 1000 Å. Injections were into fixed-size loops of 20, 50, or 250 μL.

Each oligomer was injected at a measured concentration and volume for each pH value tested. The DNA oligomers were injected at the same pH used for the elution gradient. All of the injected samples were at a Na^+ concentration of 70 mM. These Na^+ concentrations were less than the starting Na^+ concentration for the elution gradient (which was 0.1 M).

Elution was performed by increasing the concentration of NaCl in a phosphate buffer at the appropriate pH. Buffer A consisted of 10 mM Na^+ (phosphate) at the specified pH, and buffer B was buffer A plus 0.99 M NaCl. Buffer C was 10 mM Na^+ (phosphate), pH 12, and buffer D was buffer C plus 0.99 M NaCl. The elution gradient profile was (1) 90% buffer A plus 10% buffer B for 8.3 min and then (2) a linear gradient from 90% buffer A plus 10% buffer B to 40% buffer A plus 60% buffer B (i.e., from 0.1 to 0.6 M Na^+) during the time interval from 8.3 to 50 min. To clean the column of all DNA complexes, the column was washed with (3) 100% buffer C (high pH, low Na^+ concentration) for 10 min and (4) 100% buffer D for 10 min (high pH, high Na^+ concentration). (5) The column was finally reequilibrated for at least 15 min (six column volumes) to the starting conditions with 90% buffer A plus 10% buffer B. Before use, the buffers were filtered through Millipore GVWP Durapore 0.22-μm filters and degased.

Detection of eluted DNA oligomers was by dual-absorbance measurements at 260 and 280 nm, which allowed for the monitoring of cytosine protonation as well as the identification of adenine and thymine bases due to their distinctive A_{280}/A_{260} ratios. The Na^+ concentrations given for the elution of the various oligomers have been corrected for the void volume of the column (2.5 mL), the volume of the mixing chamber (1.2 mL), and the sample loop size (20–250 μL).

Results

C·C⁺ Base Pairs in d(CCATT)₃₋₆ and d(AATCC)₄,₅. DNA oligomers with the repeating sequence CCATT, ranging in length from 15 to 30 nucleotides, were studied under conditions of neutral and acidic pH to see if they would be able to form A·T and/or C·C⁺ base pairs. Figure 1, panel A, shows CD spectra for $d(CCATT)_6$ obtained at pH 7, 0.5 M Na^+, as the temperature was lowered from 60 to 5 °C. There was only a hint of C·C⁺ base pair formation at 5 °C for $d(CCATT)_6$ as evidenced by a slight increase in the CD at wavelengths above 290 nm. The increases in band magnitudes that were observed for this 30-mer were typical for oligomers that have increased base stacking as the temperature is lowered. There was no spectral evidence for C·C⁺ base pairing at pH 7 and 0.5 M Na^+ for the oligomers $d(CCATT)_3$, $d(CCATT)_4$, and $d(CCATT)_5$ (not shown). At the lower salt concentration of 0.1 M Na^+, there was no evidence for C·C⁺ base pairing at pH 7 for any of the oligomers in this series.

When the pH was reduced to 6 and 5, at 0.1 and 0.5 M Na^+, 5 °C, all of the $d(CCATT)_{3-6}$ oligomers showed evidence of enhanced C·C⁺ base pairing. Figure 1, panel B, shows CD

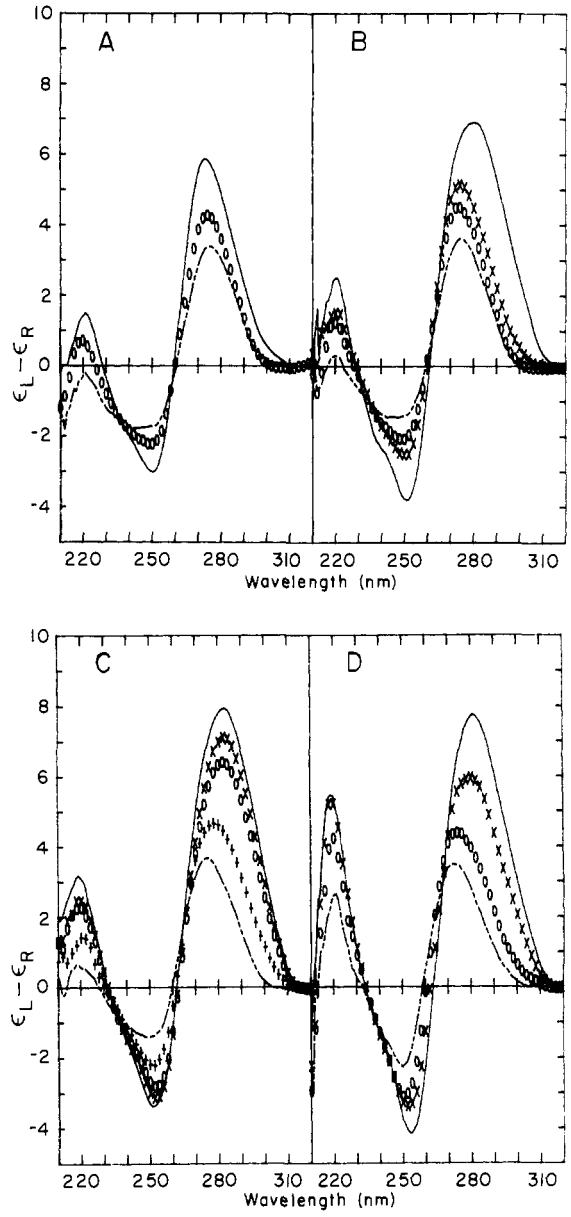

FIGURE 1: Measured CD spectra of repeating pentameric oligomers. Spectra were taken as the temperature was lowered from 60 to 5 °C at 0.5 M Na^+ and are labeled as follows: 60 °C (- - -); 40 °C (+); 30 °C (○); 20 °C (×); 5 °C (—). Spectra are of (A) $d(CCATT)_6$ at pH 7, (B) $d(CCATT)_5$ at pH 6, (C) $d(CCATT)_5$ at pH 5 and, (D) $d(AATCC)_5$ at pH 6.

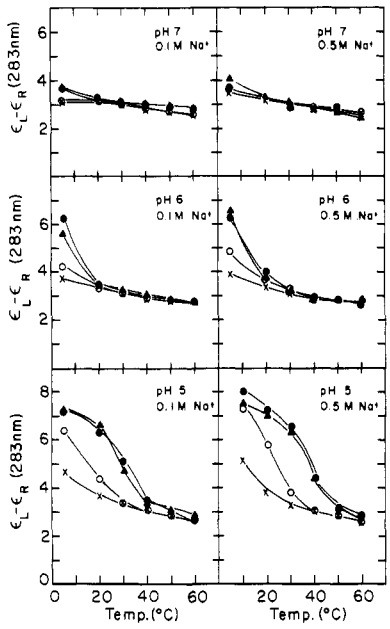

FIGURE 2: Annealing profiles for d(CCATT)$_n$, with n = 3 (×), 4 (○), 5 (●), or 6 (▲). Shown are the CD changes at 283 nm that monitored C·C$^+$ base pair formation as the temperature was lowered at the indicated pH and Na$^+$ concentration.

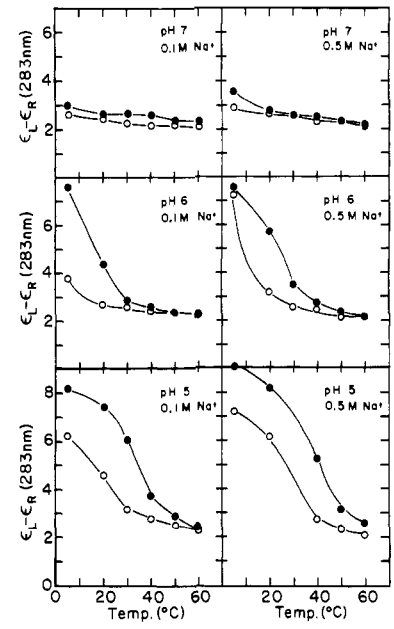

FIGURE 3: Annealing profiles for d(AATCC)$_n$, with n = 4 (○) or 5 (●). Shown are the CD changes at 283 nm that monitored C·C$^+$ base pair formation as the temperature was lowered at the indicated pH and Na$^+$ concentration.

spectra of d(CCATT)$_5$ at pH 6 as the temperature was lowered from 60 to 5 °C. These spectra exhibited a long-wavelength positive CD band that was red shifted from 274 to 280 nm and was almost doubled in magnitude. At pH 5 (Figure 1, panel C) the long-wavelength CD band was red shifted from 275 and 283 nm and more than doubled its magnitude. This showed that the formation of C·C$^+$ base pairing in d(CCATT)$_5$ commenced at higher temperatures and was more complete at 5 °C when the pH was reduced. These spectral changes of the 274-nm positive CD band were like those seen during the formation of C·C$^+$ base pairs in poly[d(C)] (Gray & Bollum, 1974). The CD magnitudes, band shapes, and crossover at wavelengths above 240 nm for the acid self-complex of d(CCATT)$_5$ were especially close to those seen during acidification of d(AACC)$_5$ (Edwards et al., 1988).

Annealing profiles for d(CCATT)$_{3-6}$ in Figure 2 show the effects of pH, Na$^+$ concentration, and oligomer length on C·C$^+$ base pair formation (monitored by an increase in the positive CD at 283 nm) as the temperature was lowered. The large CD changes over a relatively narrow temperature range at low pH indicated the cooperative nature of C·C$^+$ base pair formation during the annealing process. C·C$^+$ base pairing was obvious at pH 5 for all four of the oligomers, with the two oligomers d(CCATT)$_5$ and d(CCATT)$_6$ having the highest annealing temperatures and largest CD changes. The similarity in the annealing profiles for the two longest oligomers indicated that 10 C·C$^+$ base pairs were sufficient for maximal pairing in these sequences. Shorter oligomers showed a noticeable length dependence for C·C$^+$ base pair formation. The annealing temperature for d(CCATT)$_5$, at pH 5, 0.5 M Na$^+$, was approximately 35 °C. Under the same conditions, the annealing temperature previously measured for d(CCTT)$_5$, which presumably contained 10 C·C$^+$ base pairs, was also about 35 °C (Edwards et al., 1988).

The oligomers d(AATCC)$_{4,5}$ were also studied to see if they would form C·C$^+$ base pairs. Once the pH was lowered to 6 (Figure 1, panel D), there was significant C·C$^+$ base pairing, even at 30 °C in 0.5 M Na$^+$. CD spectra of d(AATCC)$_5$ at pH 5, 0.5 M Na$^+$ (not shown), revealed that C·C$^+$ base pairing was even more complete and was substantial at 40 °C.

CD annealing profiles of d(AATCC)$_{4,5}$ (Figure 3) were similar to the ones for d(CCATT)$_{3-6}$ (Figure 2) in that they exhibited cooperativity for the formation of C·C$^+$ base pairs. The annealing temperature for d(AATCC)$_5$ at pH 5, 0.5 M Na$^+$, was very close to the one measured for d(CCATT)$_5$ under the same conditions. Thus, d(AATCC)$_5$ and d(CCATT)$_5$ apparently could form the same number of C·C$^+$ base pairs under these conditions. Increasing the Na$^+$ concentration from 0.1 to 0.5 M Na$^+$ raised the annealing temperatures for the d(CCATT)$_{3-6}$ and d(AATCC)$_{4,5}$ series. This is also the case for d(AACC)$_5$ and d(CCTT)$_5$ (Edwards et al., 1988) but stands in contrast to the case for poly[d(C)], which is destabilized by increasing Na$^+$ concentration (Inman, 1964). Measurements of the T_m of d(AATCC)$_5$ at different oligomer concentrations showed that the annealing temperature was lowered by 5 °C as the oligomer concentration was decreased from 6.2 × 10^{-4} to 4.2 × 10^{-6} M nucleotides, which argued against the possibility of hairpin formation in these complexes (data not shown).

Mixtures of d(CCATT)$_5$ with d(AATCC)$_5$. It was previously seen that when d(CCTT)$_5$ was mixed with d(AACC)$_5$, only A·T base pairs formed (Edwards et al., 1988), even at low pH where C·C$^+$ base pairs formed in the individual strands. A similar mixing experiment was carried out with d(CCATT)$_5$ and d(AATCC)$_5$. If an antiparallel duplex could form between the different strands, it would be comprised of 60% A·T base pairs. Figure 4, panel A, shows CD spectra of the resulting mixture at pH 6, 0.5 M Na$^+$, as the temperature was lowered from 60 °C. A temperature of 60 °C was well above the annealing temperatures for cytosine pairing in the individual strands and would have allowed A·T pairs to form if these were more stable. Figure 4, panel B, shows the measured spectrum for the mixture at pH 6, 0.5 M Na$^+$, 5 °C, compared with the average of the individual strands under the same conditions. There was a close match of the two spectra, especially in the magnitudes of the bands at 250 and 280 nm, which indicated that d(CCATT)$_5$ did not interact with d(AATCC)$_5$ by pairing of adenines and thymines.

Alternative Base Pairs in d(CCATT)$_5$ and d(AATCC)$_5$. It is possible that other types of base pairs, such as A·T, A·C$^+$,

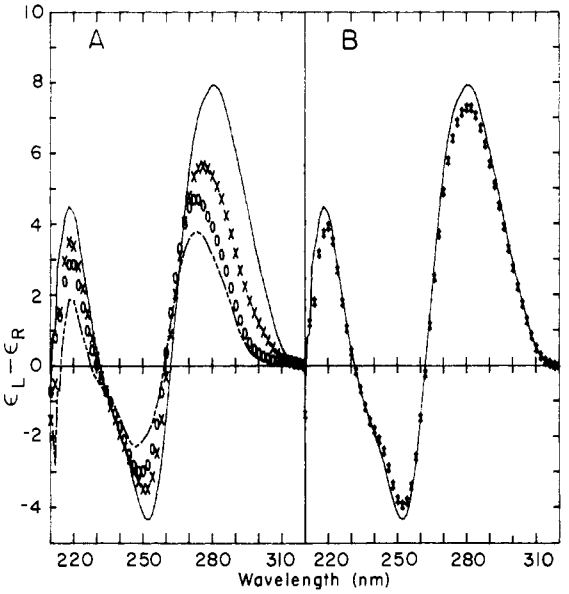

FIGURE 4: CD spectra of the 1:1 mixture of $d(AATCC)_5$ with $d(CCATT)_5$ at pH 6, 0.5 M Na^+. (A) Measured spectra are shown at 60 °C (----), 30 °C (o), 20 °C (×), and 5 °C (—). (B) The measured CD spectrum (—) of the 1:1 mix of $d(AATCC)_5$ with $d(CCATT)_5$ at pH 6, 0.5 M Na^+, 5 °C, is shown compared with the average (*) of the CD spectra of the separated strands under the same conditions. Under these conditions, the individual strands were in self-complexes.

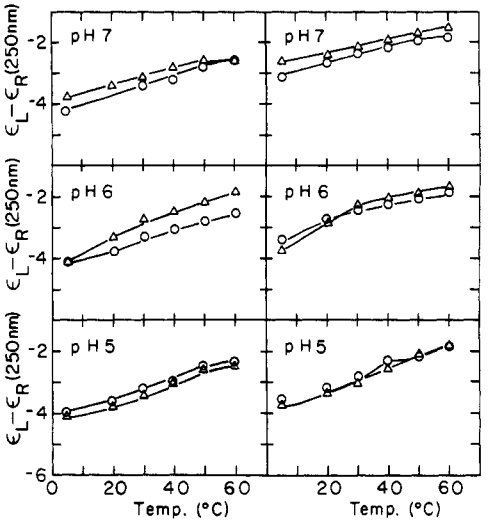

FIGURE 5: Annealing profiles for $d(CCATT)_5$ (left panels) and $d(AATCC)_5$ (right panels) at 0.1 (o) and 0.5 M (Δ) Na^+. Shown are the CD changes at 250 nm that monitored base stacking as the temperature was lowered at the indicated pH.

or $A^+ \cdot A^+$ base pairs, would have been able to form in the oligomers $d(CCATT)_5$ and $d(AATCC)_5$. A·T base pair formation can be monitored by the increase in magnitude of the negative CD band near 250 nm. The formation of A·T base pairs causes an increase by about a factor of 2 in the negative magnitude of the CD band at 250 nm for poly[d(A)]·poly[d(T)] (of −4.1 to −8.5 $L \cdot mol^{-1} \cdot cm^{-1}$) and poly[d(AT)·d(AT)] (of −3.3 to −6.2 $L \cdot mol^{-1} \cdot cm^{-1}$) (Greve et al, 1977). There were substantial increases in the CD at 250 nm for $d(CCATT)_5$ and $d(AATCC)_5$ as the temperature was lowered (Figures 1 and 4). However, increased stacking of unpaired bases during $C \cdot C^+$ base pair formation in the individual strands of $d(CCTT)_5$, $d(AACC)_5$, $d(T_6C_6T_6)$, and $d(A_6C_6A_6)$ (Edwards et al., 1988) also caused an increase in negative magnitude at 250 nm without A·T base pair formation. Thus, an important characteristic of A·T base pair formation is the cooperativity of the CD changes at 250 nm as a function of temperature. Figure 5 shows the changes in CD magnitude at 250 nm for the individual $d(CCATT)_5$ and $d(AATCC)_5$ strands at several pH and salt concentrations. The changes were essentially monotonic increases in negative-band magnitude as the temperature was lowered from 60 to 5 °C. This increase was probably caused by increased stacking of adenines and thymines as the temperature was decreased. By comparison with Figures 2 and 3, which show the cooperative nature of $C \cdot C^+$ base pair formation as monitored by the CD changes at 283 nm, there was no evidence for cooperative changes in CD magnitude at 250 nm and therefore no evidence for A·T base pair formation.

Difference spectra of the CD of the self-complexed strands at 5 °C minus the CD of the denatured single strands at 60 °C, both at pH 5, 0.5 M Na^+, are presented in Figure 6, panel A. These spectra show the changes that occurred upon base pair formation in $d(CCATT)_5$ and $d(AATCC)_5$, compared with the difference spectrum for $d(AACC)_5$. The similarity of the difference spectra, especially above 270 nm, indicated that the same type of base pairing occurred in all three oligomers as the temperature was lowered. This argued against

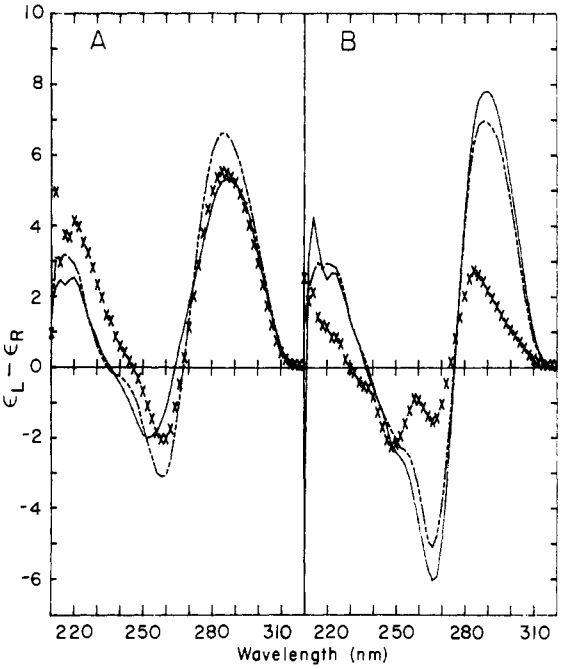

FIGURE 6: CD difference spectra (low temperature minus high temperature). (A) $d(AATCC)_5$ (----), $d(CCATT)_5$ (—), and $d(AACC)_5$ (×) all at pH 5. Spectra were calculated by subtracting the CD spectra of the single-stranded forms at pH 5, 0.5 M Na^+, 60 °C, from the CD spectra of the self-complexed forms at pH 5, 0.5 M Na^+, 5 °C. (B) $d(C_4A_4T_4C_4)$ at pH 7 (×), 6 (----), and 5 (—). These difference spectra were obtained by subtracting the CD spectra of the single-stranded forms at the specified pH, 0.5 M Na^+, 60 °C, from the CD spectra of the self-complexed forms at the same pH, 0.5 M Na^+, 5 °C.

the formation of $A \cdot C^+$ or $A^+ \cdot A^+$ base pairs, since these should have been twice as frequent in the self-complexed forms of $d(AATCC)_5$ and $d(AACC)_5$ as in that of $d(CCATT)_5$. Also, the pK_a for $A^+ \cdot A^+$ base pairs in poly[d(A)] is 4.5 (Adler et al., 1969), which is below the pH where these spectral changes were seen. The similarities in the CD difference magnitudes at 250–260 nm for the repeating pentamers compared with those for $d(AACC)_5$ were also further evidence against the presence of A·T base pairs in $d(AATCC)_5$ and $d(CCATT)_5$, since A·T base pairs could not form in $d(AACC)_5$. In summary, our results indicated that in the repeating pentameric

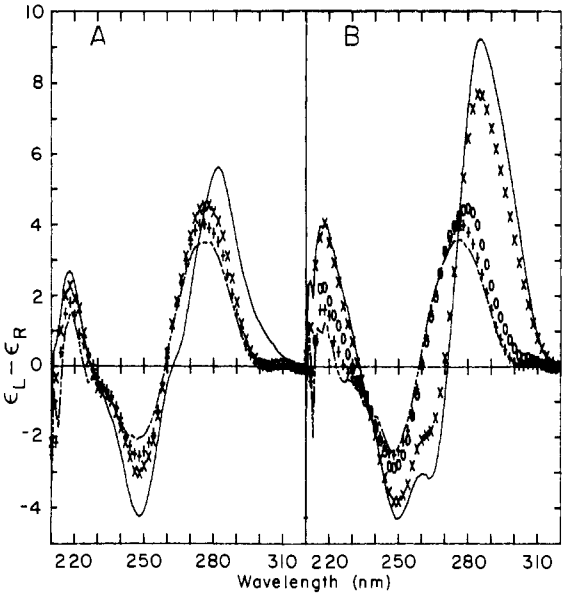

FIGURE 7: Measured CD spectra of d($C_4A_4T_4C_4$) at (A) pH 7, 0.5 M Na^+, and (B) pH 6, 0.5 M Na^+. Spectra were taken as the temperature was lowered from 60 to 5 °C and are labeled as follows: 60 °C (----); 40 °C (+); 30 °C (o); 20 °C (×); 5 °C (—).

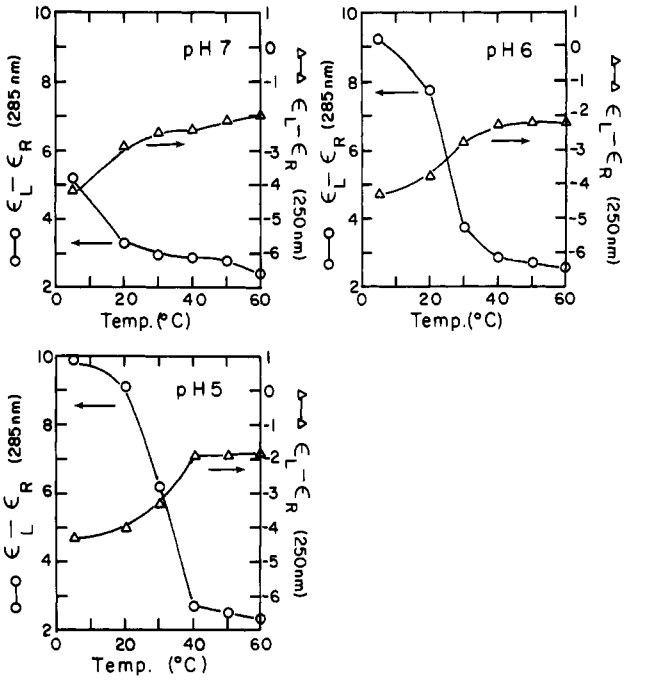

FIGURE 8: Annealing profiles for d($C_4A_4T_4C_4$) at pH 7, 6, and 5. Increases in A·T and C·C$^+$ base pair formation were followed by the CD changes at 250 (Δ) and 285 nm (o), respectively.

sequences, d(CCATT)$_5$ and d(AATCC)$_5$, the only stable base pairing, under the conditions investigated, was between cytosines.

CD Spectra of d($C_4A_4T_4C_4$). CD measurements were also used to detect and quantitate the presence (or absence) of A·T and C·C$^+$ base pairs in the oligomer d($C_4A_4T_4C_4$). Figure 7, panel A, shows CD spectra measured for d($C_4A_4T_4C_4$) at pH 7, 0.5 M Na^+, as the temperature was lowered from 60 to 5 °C. The changes between 60 and 20 °C were indicative of increased base stacking, and only at 5 °C was there any manifestation of base pairing. At 5 °C, the long-wvelength maximum was shifted from about 275 to 282 nm, providing evidence for a small amount of C·C$^+$ base pair formation. The increased magnitude of the negative band near 250 nm indicated that A·T base pairs also formed at the lowest temperature. Much larger increases in the magnitude of the long-wavelength band were seen in the spectra of this oligomer at pH 6, 0.5 M Na^+ (Figure 7, panel B), and these increases started at higher temperatures than when the oligomer was at pH 7. A negative band at 265 nm like the one seen in the acid self-complexed form of poly[d(C)] (Gray & Bollum, 1974) was also readily apparent in the CD spectra of d($C_4A_4T_4C_4$) at low temperatures at pH 6. CD spectra (not shown) of d($C_4A_4T_4C_4$) at pH 5, 0.5 M Na^+, were very similar to those seen at pH 6. The magnitudes of the bands at about 285 and 265 nm were slightly greater at pH 5 than at pH 6, and the formation of C·C$^+$ base pairs started at even higher temperatures.

The magnitude of the band near 250 nm was the same for the spectra at 5 °C at all three pH values. This is shown in the CD difference spectra in Figure 6, panel B, that detail the changes at each pH as the temperature was lowered. At all three pH values the CD difference spectra near 250 nm were equivalent, indicating equivalent amounts of A·T base pair formation. The CD differences between 260 and 270 nm and above 280 nm, indicative of C·C$^+$ base pair formation, were less at pH 7 than at pH 6 and 5. Apparently, the extent of formation of A·T base pairs was not enhanced by the lowering of the pH from 7 to 6, unlike the formation of C·C$^+$ base pairs.

Annealing profiles for d($C_4A_4T_4C_4$) at pH 7, 6, and 5, 0.5 M Na^+, are shown in Figure 8. The increase in magnitude of the 285-nm CD band was used to monitor C·C$^+$ base pair formation, and the increase in negative magnitude at 250 nm was used to monitor A·T base pair formation. Both CD bands showed cooperative changes as the temperature was lowered, unlike the cases above for d(AATCC)$_5$ and d(CCATT)$_5$ where the 250-nm band did not change cooperatively (Figure 5). While the extents of the changes at 250 nm were the same at all three pH values for d($C_4A_4T_4C_4$), it can be seen from Figure 8 that protonation of cytosine raised the temperature at which A·T base pairing occurred. The annealing temperature for both types of base pair at pH 6 was about 25 °C and was increased to about 30 °C at pH 5. The fact that the annealing temperatures were the same for both types of base pair at pH 5 demonstrated that there was simultaneous formation of A·T and C·C$^+$ base pairs in d($C_4A_4T_4C_4$) as the temperature was lowered.

The present results confirm our previous work showing that both types of base pair form simultaneously in d($C_4A_4T_4C_4$) as the temperature is lowered and that the pairing of adenines and thymines raises the pK_a of pairing of the two d(C)$_4$ tracts by about 1 pH unit from that of d(C)$_8$ (Gray et al., 1984). In addition, the data shown here emphasize that the extent of A·T base pair formation at 5 °C is independent of pH, while the extent of C·C$^+$ base pair formation is enhanced as the pH is lowered.

Centrifugation of DNA Oligomers. To determine if there was a pH-dependent formation of a large molecular weight complex in d($C_4A_4T_4C_4$), the oligomer was subjected to centrifugation at 15600*g* for 30 min, at 4 °C, in an Eppendorf table-top centrifuge. Table I shows the results for this oligomer as well as for others that were able to form only A·T or only C·C$^+$ base pairs. At pH 7, the percentage of precipitable DNA (i.e., DNA that could be centrifuged out of solution under these conditions) averaged 3 ± 2% for the oligomers in all three categories. Lowering the pH to 6 caused no change in the average precipitability of oligomers that formed only A·T or C·C$^+$ base pairs at low pH but increased by a factor of 7.5 the amount of d($C_4A_4T_4C_4$) that precipitated. For example, oligomer d(A_8T_8), also a 16-mer, showed only 2% precipitation

Table I: Types of DNA Oligomers and the Percentages That Precipitated at Low pH[a]

	pH		
	7	6	5
Oligomers That Form Only A·T Base Pairs			
$d(A_8T_8)$	1 ± 1	2 ± 1	5 ± 4
$d(AACC)_5 \cdot d(CCTT)_5$[b]	2 ± 1	5 ± 2	7 ± 3
$d(A_6C_6A_6) \cdot d(T_6C_6T_6)$[b]	4 ± 1	3 ± 1	3 ± 1
Oligomers That Form Only C·C$^+$ Base Pairs			
$d(AACC)_5$	5 ± 2	1 ± 0	4 ± 2
$d(CCTT)_5$	6 ± 2	2 ± 1	8 ± 8
$d(AATCC)_5$	1 ± 0	7 ± 2	9 ± 3
$d(CCATT)_5$	6 ± 0	5 ± 1	9 ± 1
$d(AATCC)_5$ plus $d(CCATT)_5$[c]	2 ± 1	1 ± 0	3 ± 1
Oligomer That Forms both A·T and C·C$^+$ Base Pairs			
$d(C_4A_4T_4C_4)$	2 ± 2	15 ± 5	38 ± 17

[a] Percentage of percipitable DNA oligomers as the pH was reduced from 7 to 5, at 0.5 M Na^+. The values are averages of three determinations except for those of $d(C_4A_4T_4C_4)$, for which five determinations were made. Errors are ±1 standard deviation. Centrifugations were at 15600*g* for 30 min at 5 °C. All concentrations were between 6.0×10^{-5} and 1.5×10^{-4} M nucleotides. [b] 1:1 ratio of noninteracting strands at pH 7 and interacting strands at pH 6 and 5. [c] 1:1 ratio of noninteracting strands at all pH values.

at pH 6 while $d(C_4A_4T_4C_4)$ precipitated to an extent of 15%. At pH 5, a slight increase in aggregation was noted for oligomers that form only A·T or C·C$^+$ base pairs to an average of 6 ± 3% precipitation, but the amount of $d(C_4A_4T_4C_4)$ that could be pelleted rose to nearly 40%. Thus, $d(C_4A_4T_4C_4)$ was the only oligomer tested that showed an unusually large pH dependence of precipitability, even though other oligomers were tested that also contained adenine, thymine, and cytosine. Only $d(C_4A_4T_4C_4)$ also showed spectral evidence for the simultaneous formation of C·C$^+$ and A·T base pairs.

High-Pressure Anion Exchange Liquid Chromatography. HPLC data were consistent with the centrifugation results that showed that $d(C_4A_4T_4C_4)$ formed a large molecular weight complex at pH 6 and 5. First, oligomers $d(A)_{20}$ and $d(T)_{20}$, individually, were used as standards for the elution of single-stranded DNA from a DEAE column. A mixture of the two was used to obtain the elution profile of a duplex. At pH 7, $d(A)_{20}$ eluted as a single sharp peak at a Na^+ concentration of 0.35 M and had an A_{280}/A_{260} ratio of 0.35. The elution of $d(T)_{20}$ at pH 7 occurred at a Na^+ concentration of 0.39 M, again as a single peak; this peak had a characteristic A_{280}/A_{260} ratio of 0.79. At pH 7, the mixture of $d(A)_{20}$ plus $d(T)_{20}$ eluted as a single peak at a Na^+ concentration of 0.42 M and had an intermediate A_{280}/A_{260} ratio of 0.62. No shoulders or multiple peaks were seen for $d(A)_{20} \cdot d(T)_{20}$. At pH 5, the Na^+ concentrations required for the elution of $d(A)_{20}$, $d(T)_{20}$, and $d(A)_{20} \cdot d(T)_{20}$ and the A_{280}/A_{260} ratios of these compounds were the same as those observed at pH 7. Integration of the peak areas indicated that neither $d(A)_{20}$, $d(T)_{20}$, nor $d(A)_{20} \cdot d(T)_{20}$ became irreversibly bound to the column as the pH was lowered.

The 18-mers $d(A_6C_6A_6)$ and $d(T_6C_6T_6)$ were chromatographed individually and as a mixture to examine the pH-dependent elution profiles for oligomers that contain all three bases, adenine, thymine, and cytosine, but which according to spectral evidence were able to form only A·T base pairs (Edwards et al., 1988) and which according to centrifugation data did not precipitate (Table I). As with the mixture of $d(A)_{20}$ plus $d(T)_{20}$, chromatography of $d(A_6C_6A_6)$ plus $d(T_6C_6T_6)$ gave a sharp peak at pH 5 consistent with a simple duplex form that contained only A·T base pairs.

The 16-mer $d(C_4A_4T_4C_4)$ was chromatographed under a variety of conditions where it was single stranded or contained A·T and C·C$^+$ base pairs. At pH 7, a single sharp peak eluted at a Na^+ concentration of 0.39 M and had an A_{280}/A_{260} ratio of 0.76. Lowering the pH to 6 diminished the peak eluting at 0.39 M Na^+ and caused a second peak to elute at 0.45 M Na^+ with an increased A_{280}/A_{260} ratio of ≈1.0, indicative of C·C$^+$ base pairs. Lowering the pH to 5 caused the peak eluting near 0.45 M Na^+ to become very broad, still with an A_{280}/A_{260} ratio of ≈1.0. The broad peak seen at pH 5 was indicative of the existence of a distribution of acid-induced structures. Also, as the pH was lowered from 7 to 5, more than 40% of the DNA did not elute from the column, consistent with the formation of large molecular weight complexes. This material could be eluted from the column at pH 12 (see Materials and Methods).

Discussion

$d(CCATT)_{3-6}$ and $d(AATCC)_{4,5}$ Form C·C$^+$ Base Pairs to the Exclusion of A·T Base Pairs. Slight alterations of the primary sequences of $d(AACC)_5$ and $d(CCTT)_5$ previously studied (Edwards et al., 1988) were made to allow us to investigate the effects of mismatches on C·C$^+$ base pair formation. In an effort to create DNA sequences that might form C·C$^+$ base pairs with, or without, adjacent A·T base pairs, repeating pentameric oligomers were synthesized that would contain at least one mismatch per pentamer duplex. This was done on the assumption that mismatches would permit greater helical flexibility. We studied two series of oligomers, $d(CCATT)_{3-6}$ and $d(AATCC)_{4,5}$, each of which could potentially form two A·T and two C·C$^+$ base pairs per pentamer, with either a T·T or an A·A mismatch. Increases in CD band magnitude at 280–285 nm demonstrated that C·C$^+$ base pairs were stable in the repeating pentameric sequences $d(CCATT)_{3-6}$ and $d(AATCC)_{4,5}$ at pH 6 and 5, even though all the cytosines were not contiguous and accounted for only 40% of the sequence. Previously, C·C$^+$ base pairing in sequences that contained separated stretches of cytosine had been found in poly[d(CCT)] (Gray et al., 1988), where cytosine comprised 67% of the sequence, and in poly[d(CT)] (Brown et al., 1985), $d(C_4A_4T_4C_4)$ (Gray et al., 1984), $d(CCTT)_5$, and $d(AACC)_5$ (Edwards et al., 1988), where 50% of the bases were cytosine. The $d(CCATT)_n$ series showed a length dependence for C·C$^+$ base pair formation that leveled off at $n = 5$, indicating that 10 cytosine residues served as a minimum length for stable C·C$^+$ base pair formation in these sequences.

This work defines more clearly the possibilities for C·C$^+$ base pairing in synthetic and natural DNA sequences. As described above, it is not necessary to have long, uninterrupted stretches of cytosine residues to form C·C$^+$ base pairs, since oligomers with only 40% cytosine, $d(CCATT)_5$ and $d(AATCC)_5$, could form C·C$^+$ base pairs at pH 5. On the other hand, the presence of more than 40% cytosine residues does not necessarily mean that C·C$^+$ base pairs can form. For example, in the paired duplex of $d(AACC)_5 \cdot d(CCTT)_5$, with 50% cytosine, only A·T base pairs formed, even at pH 5. The context of the cytosine sequence, and presumably the strength of different stacking arrangements, appears to be important in determining which type of base pairing occurs.

There was no evidence, such as cooperative CD changes at 250 nm, that would have demonstrated the formation of A·T base pairs in $d(CCATT)_{3-6}$ and $d(AATCC)_{4,5}$. Since the CD spectra reported for parallel-stranded A·T base pairs (van de Sande et al., 1988; Germann et al., 1989) show the same type of changes at 250 nm that have been reported before for antiparallel-stranded A·T base pairs (Greve et al., 1977), we concluded that self-complexes of the oligomers $d(CCATT)_5$

and d(AATCC)$_5$ did not contain A·T base pairs of either strand orientation.

Our present results with d(AATCC)$_5$, d(CCATT)$_5$, and d(AATCC)$_5$·d(CCATT)$_5$ and previous results with d(AACC)$_5$·d(CCTT)$_5$ (Edwards et al., 1988) raise the question of why C·C$^+$ and A·T base pairs were unable to form adjacent to each other. A possibility is that the distortion introduced into the phosphate backbone by a pyrimidine–pyrimidine base pair stacked on top of a pyrimidine–purine base pair was too extreme. However, it has been sugested that the inherent backbone flexibility is probably great enough to accommodate such fluctuations in the C1′–C1′ distances between two such neighboring base pairs (Chattopadhyaya et al., 1988). Because A·T base pairs prefer to form between antiparallel-stranded DNA and are unable to form in these repeating pentameric sequences, another possibility is that the strand orientation for C·C$^+$ base pairs has to be parallel. This would be the consequence if the paired cytosines must adopt anti-glycosidic linkages, as is usually preferred by pyrimidines, and if three hydrogen bonds are required for stable base pairing.

d(C$_4$A$_4$T$_4$C$_4$) Probably Forms A·T and C·C$^+$ Base Pairs in a Multistranded Complex. A cooperative increase in the magnitude of the long-wavelength CD band near 285 nm for d(C$_4$A$_4$T$_4$C$_4$) at pH 6 or 5, 0.5 M Na$^+$, as the temperature was decreased, showed that C·C$^+$ base pairs could form in this oligomer. Concurrently, there was a cooperative increase in the negative magnitude of the CD band centered at about 250 nm, indicative of A·T base pair formation. Both of these changes occurred at the same annealing temperature.

d(C$_4$A$_4$T$_4$C$_4$) differed from oligomers that could form *either* A·T *or* C·C$^+$ base pairs in that it was the only one that precipitated in a pH-dependent fashion. Centrifugation studies and anion exchange chromatography of the oligomer d(C$_4$A$_4$T$_4$C$_4$) established that, as the pH was lowered, the molecular weight of the resulting complex increased well beyond that of a duplex. Given the pore size of the anion exchange gel matrix and an estimated sedimentation coefficient of >100 S, the complex at low pH must have consisted of hundreds of oligomers. Comparison of d(C$_4$A$_4$T$_4$C$_4$) with other DNA sequences that formed only one of the two types of base pairs led us to conclude that d(C$_4$A$_4$T$_4$C$_4$) was unique in being able to form a multistranded complex, possibly containing parallel-stranded C·C$^+$ base pairs and antiparallel-stranded A·T base pairs.

Figure 9 shows schematic drawings of two forms of a possible multistranded structure of d(C$_4$A$_4$T$_4$C$_4$) having both strand orientations. Form I (Figure 9, panel A) of the multistranded complex of d(C$_4$A$_4$T$_4$C$_4$) consists of antiparallel duplex regions of A·T base pairs and parallel duplexes of C·C$^+$ base pairs in which the cytosines at both ends of an oligomer base pair with cytosines of a single additional oligomer. As drawn, additional oligomers could add to the A$_4$T$_4$ regions in an antiparallel orientation. Form II (Figure 9, panel B) also has different strand orientations for the A·T and C·C$^+$ base pairs, but the cytosines at the ends of one oligomer pair with the cytosines of two different oligomers. This postulated multistranded structure of d(C$_4$A$_4$T$_4$C$_4$) is not totally without precedent. The crystal structure of the tetramer d(ATAT) consists of a network in which each tetramer shares two base pairs with each of two other tetramers (Viswamitra et al., 1978).

CD data do not contain intrinsic information about strand orientation or the molecular weight of a resulting complex. Thus, in previous work on d(C$_4$A$_4$T$_4$C$_4$), Gray et al. (1984) acknowledged the possibility of the formation of a "higher

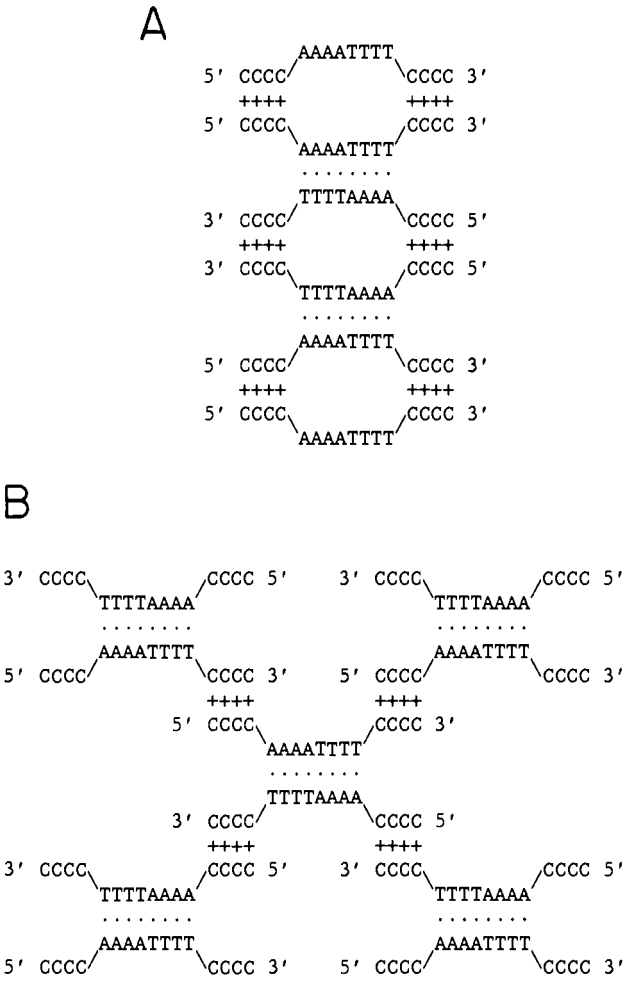

FIGURE 9: Multistranded complexes of d(C$_4$A$_4$T$_4$C$_4$) described in the text: (A) form I and (B) form II.

order structure...stabilized by interacting parallel and antiparallel duplexes". Since the formation of A·T base pairs facilitated C·C$^+$ base pair formation in d(C$_4$A$_4$T$_4$C$_4$) at a pH above that needed for C·C$^+$ base pair formation in d(C)$_8$, Gray et al. (1984) interpreted the results as indicating that the resulting complex contained both types of base pair within a d(C$_4$A$_4$T$_4$C$_4$)·d(C$_4$A$_4$T$_4$C$_4$) duplex of antiparallel strands. On the basis of the data presented in this work we now believe that that conclusion was incorrect. Instead, C·C$^+$ base pairs probably form only between parallel DNA strands, with the consequence that d(C$_4$A$_4$T$_4$C$_4$) at low pH probably exists in a novel multistranded DNA complex in which both strand orientations are present.

Recent experiments on a plasmid containing a repeating A$_2$C$_4$ motif were consistent with the pH-dependent formation of a hairpin involving C·C base pairs in an antiparallel orientation (Lyamichev et al., 1989). However, the authors acknowledge that an alternative possibility, the formation of an H-form triplex, was not excluded by the data.

Natural DNAs Have the Potential To Form C·C$^+$ Base Pairs. It is important to note that the pK_a for forming C·C$^+$ base pairs in poly[d(C)] (7.4; Inman, 1964) can be much higher than the pK_a for cytosine protonation (4.5; Saenger, 1981). The oligomers d(CCATT)$_6$, d(AATCC)$_5$, and d(C$_4$A$_4$T$_4$C$_4$) showed evidence for some C·C$^+$ base pairing at pH 7, and the pK_a for C·C$^+$ base pairing in d(C$_4$A$_4$T$_4$C$_4$) was above 6. Therefore, C·C$^+$ base pairs could potentially occur in vivo at physiological pH values (7.1–7.4; Bright et al., 1987). Because C·C$^+$ base pairing was found in sequences with

mismatches (i.e., without all opposing bases hydrogen bonded), the need for strand complementarity in vivo may not be particularly stringent.

Cytosine-rich DNA sequences have been found in immunoglobin switch regions (Shimizu & Honjo, 1984), telomeres (Oka & Thomas, 1987), and gene promoters (Evans et al., 1984; Kilpatrick et al., 1986). The immunoglobin switch regions are 5′ to the sites of recombination and are 1–10 kilobases in size, with cytosines comprising about 50% of the pyrimidine-rich strand. Commonly repeated motifs are d(ACCCC) and d(CCCCAGCTCCCC) (Sen & Gilbert, 1988). $C{\cdot}C^+$ base pairing may serve to align sequences that would bring different constant regions next to a rearranged variable region at an immunoglobin switch site. The S1 hypersensitive sites in the human β-globin gene, the rat preproinsulin gene, and the SV40 origin/enhancer region contain a high percentage of cytosine in the pyrimidine-rich strand and exhibit a nicking pattern inconsistent with B or Z DNA conformations (Evans et al., 1984). In addition to the formation of H-form triplex DNA (Mirkin et al., 1987; Htun & Dahlberg, 1988; Johnson et al., 1988), formation of parallel-stranded $C{\cdot}C^+$ base pairs may be involved in such conformational changes, and/or $C{\cdot}C^+$ base pairs might help provide the structural motif recognized by DNA-binding proteins.

Conclusions

We find that adenine-, thymine-, and cytosine-containing DNA oligomers can be grouped into three different conformational categories: (1) sequences such as $d(AACC)_5{\cdot}d(CCTT)_5$ that form duplexes having only A·T base pairs (Edwards et al., 1988), presumably with an antiparallel-strand orientation, (2) sequences such as $d(CCATT)_5$ and $d(AATCC)_5$ that form structures having only $C{\cdot}C^+$ base pairs, likely with a parallel-strand polarity, and (3) sequences that form A·T and $C{\cdot}C^+$ base pairs concomitantly such as $d(C_4A_4T_4C_4)$ (Gray et al., 1984) that may adopt a higher-order structure with both parallel- and antiparallel-stranded regions.

Registry No. $d(A_8T_8)$, 124041-81-2; $d(AACC)_5{\cdot}d(CCTT)_5$, 114956-86-4; $d(A_6C_6A_6){\cdot}d(T_6C_6T_6)$, 114978-30-2; $d(AACC)_5$, 114884-79-6; $d(CCTT)_5$, 114884-80-9; $d(AATCC)_5$, 124152-93-8; $d(CCATT)_5$, 124152-94-9; $d(C_4A_4T_4C_4)$, 93792-29-1; $d(CCATT)_6$, 124152-95-0; $d(CCATT)_3$, 124152-90-5; $d(CCATT)_4$, 124152-92-7; $d(AATCC)_4$, 124152-91-6; $d(CCATT)_5{\cdot}d(AATCC)_5$, 124152-96-1; adenine, 73-24-5; thymine, 65-71-4; cytosine, 71-30-7.

References

Adler, A. J., Grossman, C., & Fasman, G. D. (1969) *Biochemistry 8*, 3846–3858.

Bright, G. R., Fisher, G. W., Rogowska, J., & Taylor, D. L. (1987) *J. Cell Biol. 104*, 1019–1033.

Brown, D. M., Gray, D. M., Patrick, M. H., & Ratliff, R. L. (1985) *Biochemistry 24*, 1676–1683.

Cantor, C. R., & Efstratiadis, A. (1984) *Nucleic Acids Res. 12*, 8059–8072.

Cantor, C. R., Warshaw, M. M., & Shapiro, H. (1970) *Biopolymers 9*, 1059–1077.

Chattopadhyaya, R., Ikuta, S., Grzeskowiak, K., & Dickerson, R. (1988) *Nature 334*, 175–179.

Coll, M., Solans, X., Font-Altaba, M., & Subirana, J. A. (1987) *J. Biomol. Struct. Dyn. 4*, 797–811.

Cruse, W. B. T., Egert, E., Kennard, O., Sala, G. B., Salibury, S. A., & Viswamitra, M. A. (1983) *Biochemistry 22*, 1833–1839.

Dybvig, K., Clark, C. D., Aliperti, G., & Schlesinger, M. J. (1983) *Nucleic Acids Res. 11*, 8495–8508.

Edwards, E. L., Ratliff, R. L., & Gray, D. M. (1988) *Biochemistry 27*, 5166–5174.

Evans, T., Schon, E., Gora-Maslak, G., Patterson, J., & Efstratidas, A. (1984) *Nucleic Acids Res. 12*, 8043–8058.

Fowler, R. F., & Skinner, D. M. (1986) *J. Biol. Chem. 261*, 8994–9001.

Germann, M. W., Vogel, H. J., Pon, R. T., & van de Sande, J. H. (1989) *Biochemistry 28*, 6220–6228.

Gray, D. M., & Bollum, F. J. (1974) *Biopolymers 13*, 2087–2102.

Gray, D. M., Vaughan, M., Ratliff, R. L., & Hayes, F. N. (1980) *Nucleic Acids Res. 8*, 3695–3707.

Gray, D. M., Cui, T., & Ratliff, R. L. (1984) *Nucleic Acids Res. 12*, 7565–7580.

Gray, D. M., Ratliff, R. L., Antao, V. P., & Gray, C. W. (1988) in *Proceedings of the Fifth Conversation in Biomolecular Stereodynamics*, Vol. 2, DNA and Its Drug Complexes (Sarma, M. H., & Sarma, R. H., Eds.) pp 147–166, Adenine Press, Schenectady, NY.

Greve, J., Maestre, M. F., & Levin, A. (1977) *Biopolymers 16*, 1489–1504.

Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr., & Blackburn, E. H. (1987) *Cell 51*, 899–908.

Htun, H., & Dahlberg, J. (1988) *Science 241*, 1791–1796.

Inman, R. B. (1964) *J. Mol. Biol. 9*, 624–637.

Johnson, B. H. (1988) *Science 241*, 1800–1804.

Kilpatrick, M., Torri, A., Kang, D., Engler, J., & Wells, R. (1986) *J. Biol. Chem. 261*, 11350–11354.

Lyamichev, V. I., Mirkin, S. M., Danilevskaya, O. N., Voloshin, O. N., Balatskaya, S. V., Dobrynin, V. N., Filippov, S. A., & Frank-Kamenetskii, M. D. (1989) *Nature 339*, 634–637.

Marck, C., Thiele, D., Schneider, C., & Guschlbauer, W. (1978) *Nucleic Acids Res. 5*, 1979–1996.

Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495–497.

Nordheim, A., & Rich, A. (1983) *Nature 303*, 674–679.

Oka, Y., & Thomas, C. A. (1987) *Nucleic Acids Res. 15*, 8877–8898.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1986) *FEBS Lett. 205*, 223–229.

Sen, D., & Gilbert, W. (1988) *Nature 334*, 364–366.

Shimzu, A., & Honjo, T. (1984) *Cell 36*, 801–803.

Van de Sande, J., Ramsing, N., Germann, M., Elhorst, W., Kalisch, B., Kitzing, E., Pon, R., Clegg, R., & Jovin, T. (1988) *Science 241*, 551–557.

Viswamitra, M. A., Kennard, O., Jones, P. G., Sheldrick, G. M., Salisbury, S., Falvello, L., & Shakked, Z. (1978) *Nature 273*, 687–688.

Westhof, E., & Sundaralingham, M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1852–1856.