Volume 5 Number 10 October 1978 Nucleic Acids Research

A comparison of the circular dichroism spectra of synthetic DNA sequences of the homopurine · homopyrimidine and mixed purine- pyrimidine types

Donald M.Gray*, A.Richard Morgan+ and Robert L.Ratliff†

*Program in Mol. Biol., University of Texas at Dallas, Box 688, Richardson, TX 75080, USA, +Dep. Biochem., University of Alberta, Edmonton, Alberta, T6G 2H7, Canada, and † Cell. and Mol. Biol. Group, Los Alamos Scientific Laboratory, University of California, Los Alamos, NM 87545, USA

Received 1 September 1978

ABSTRACT

We have obtained the ultraviolet circular dichroism spectra of two repeating trinucleotide DNAs, poly[d(A-G-G)·d(C-C-T)] and poly[d(A-A-G)·d(C-T-T)], that have all purines on one strand and all pyrimidines on the other. These spectra, together with spectra of other synthetic polymers, can be combined to give 3 first-neighbor calculations of the spectrum of poly[d(A)·d(T)] and 2 first-neighbor calculations of the spectrum of poly-[d(G)·d(C)]. The results show (1) that first-neighbor calculations utilizing only spectra of homopurine·homopyrimidine DNA sequences are no more accurate than are similar calculations that involve spectra of mixed purine-pyrimidine sequences, demonstrating that double-stranded homopurine·homopyrimidine sequences do not obviously belong to a special class of secondary conformations, and (2) that the wavelength region above 250 nm in the CD spectra of synthetic DNAs is least predictable from first-neighbor equations, probably because this region is especially sensitive to sequence-dependent conformational differences.

INTRODUCTION

The first-neighbor equations described by Gray and Tinoco[1] provide a means of combining the measured CD spectra of a limited number of (independent) nucleic acid sequences to obtain estimates of the CD spectra of other (dependent) sequences. In three previous instances, such first-neighbor estimations of CD spectra of synthetic double-stranded DNA sequences chosen to be dependent differed significantly from the measured spectra.[2,3] The discrepancies between the calculated and measured DNA spectra could be due to (a) differences in the geometry and, hence, in the CD contributions of given first-neighbor base pairs within different DNA sequences and/or (b) CD contributions from bases further removed than the first-neighbors. It is known that CD measurements are sensitive to differences in the solution conformation of DNA caused by ethanol dehydration,[4,5] changes in salt concentration,[5] and the presence of superhelical turns in a closed, circular DNA.[6,7] The extent to which non-nearest neighbors contribute to the CD spectrum of a nucleic acid polymer is unknown. First-neighbor contributions

alone can account for polymer CD spectra, as shown by our previous results with the synthetic RNA, poly[r(A-A-U)·r(A-U-U)].[3] The spectrum of this polymer was found to be quite well approximated by the first-neighbor combination of the spectra of poly[r(A-U)·r(A-U)] and poly[r(A)·r(U)]. This result is important since theoretical calculations by Johnson and Tinoco[8] and, more recently, by Cech and Tinoco[9] predict that first-neighbor estimations should be less accurate for sequences in the RNA A-conformation than for sequences in a DNA B-conformation. For both types of geometry, Cech and Tinoco[9] calculated that the exciton-type coupling of A_{260} and U_{262} oscillators in runs of A·U base pairs (regardless of strand orientation) should lead to especially poor first-neighbor approximations. They calculated that runs of G·C base pairs (all with the same strand orientation) should also lead to coupling of oscillators on nonadjacent bases. Only for sequences in which A·U and G·C base pairs alternate, or appear only in doublets, did they predict that first-neighbor estimations of CD spectra should be good. Our previous finding that a first-neighbor estimation agrees with the measured spectrum of poly[r(A-A-U)·r(A-U-U)], contrary to theoretical predictions, suggests that first-neighbor interactions may be relatively stronger than calculated from theory.

A further indication that first-neighbor CD interactions are important in the CD spectra of DNAs is that successful spectral analyses can be performed to estimate the first-neighbor frequency distributions of natural DNAs. Gray et al.[10] showed that first-neighbor frequency distributions of 11 natural DNAs can be estimated from their CD spectra with good accuracy using a limited basis set of 8 synthetic DNA spectra. Arnotts proposal[11] that the CD spectra of three D. virilis satellite DNAs contain significant CD contributions from all of the second neighbors was not confirmed, since optimal first-neighbor analyses for these DNAs did not require that spectra of polymers with the same triplet sequences be included in the basis set. However, it did appear that some triplet sequences, like d(A-A-A)·d(T-T-T) and perhaps d(T-A-G)·d(C-T-A), may affect the CD spectrum of a DNA sequence to an unusual extent.[10,12] Independently, Marck and Guschlbauer[13] have concluded from their analyses of DNA CD spectra that the first-neighbor hypothesis is substantially correct.

If the CD spectra of polynucleotides are largely determined by the CD contributions of the first-neighbors, then discrepancies between first-neighbor CD calculations and measured CD spectra may be due to sequence-dependent

conformational differences that affect the CD contributions of given first-neighbor base pairs within different DNA sequences. From preliminary X-ray diffraction evidence, Langridge[14] specifically suggested that sequences with all purines on one strand and all pyrimidines on the other (homopurine·homopyrimidine sequences) may have conformational differences from sequences having mixed purines and pyrimidines on each strand. In subsequent work, differences between first-neighbor calculations and measured CD spectra of DNAs have been explained as being partially due to geometrical differences between these two types of sequence.[1,2,3,15] Support for this explanation came from observations by Wells et al.[2] in the comparison of properties of five pairs of synthetic DNA sequence isomers. They found that the homopurine·homopyrimidine member of each pair had a lower UV extinction coefficient and were in 4 out of 5 cases less thermostable than the mixed purine-pyrimidine member. Although it is now known that the conformations of nucleic acids depend on sequence in a more complex way than suggested by Langridge, at least in fibers at reduced relative humidity,[16] it is still possible that the class of homopurine·homopyrimidine polymers might be conformationally distinct enough that their CD spectra would be more accurately interrelated to each other by first-neighbor equations than to spectra of mixed purine-pyrimidine sequences.

We have measured for the first time the CD spectra of two repeating trinucleotide DNAs, poly[d(A-G-G)·d(C-C-T)] and poly[d(A-A-G)·d(C-T-T)], which are of the homopurine·homopyrimidine type and which allow a test of the above possibility. In addition, we have obtained the CD spectrum of the new mixed purine-pyrimidine DNA poly[d(A-C-C)·d(G-G-T)], and we have remeasured the spectrum of poly[d(A-A-C)·d(G-T-T)], previously reported by Wells et al.[2] These mixed purine-pyrimidine repeating trinucleotide sequences have the same base compositions as the above homopurine·homopyrimidine DNAs and also include either (G-G)·(C-C) or (A-A)·(T-T) base pairs. In this paper, we report the results of similar first-neighbor calculations using the CD spectra of the two types of DNA sequence. We also report our characterization of the CD spectra of different forms of the polymers in neutral and acidic solutions.

MATERIALS AND METHODS

The repeating dinucleotide and trinucleotide DNAs were prepared as previously described.[3,4,17,18] For spectra at neutral pH, all the DNAs were dialyzed against several changes of a series of 3 solutions: (a) 0.5 M NaCl,

0.01 M EDTA, and 0.005 M Na_2HPO_4 adjusted to pH 7.0 with H_3PO_4; (b) 0.01 M NaCl and 0.005 M Na_2HPO_4, pH 7.0; (c) 0.005 M Na_2HPO_4 (i.e. 0.01 M Na^+), pH 7.0. For spectra at low pH, either direct additions of diluted H_3PO_4 were made to the solutions after dialysis as above, or else the samples were dialyzed into a pH 5.0 buffer. Samples dialyzed to pH 5.0 were dialyzed against several changes of 0.5 M NaCl, 0.01 mM di-Na EDTA, and 0.05 M sodium acetate, adjusted to pH 5 with acetic acid.

Techniques for the spectral measurements have been described.[3] CD data are presented as the CD per mole of monomer, $\epsilon_L - \epsilon_R$, in units of liter/(mol·cm). Sample concentrations at neutral pH were obtained from measured optical densities and the following extinction coefficients at 260 nm: 6700 liter/(mol·cm) for poly[d(A-G-G)·d(C-C-T)] (by phosphate analysis; A. R. Morgan, unpublished data), 5150 for poly[d(A-A-G)·d(C-T-T)],[2] 7600 for poly[d(A-A-C)·d(G-T-T)],[2] 6150 for poly[d(A-C-C)·d(G-G-T)] (by phosphate analysis; R. L. Ratliff, unpublished data), 6500 for poly[d(A-C)·d(G-T)],[2] and 5700 for poly[d(A-G)·d(C-T)].[2] For samples dialyzed to pH 5.0, if the CD spectra were different from the neutral pH spectra, sample concentrations were estimated from the optical densities prior to dialysis and measurements of volume changes during dialysis; these could be in error up to ± 20%.

CD and absorption spectra at elevated temperatures were corrected for volume expansion of the sample solutions, which was less than 4%. Melting profiles were obtained by manually increasing the temperature setting (in increments of 0.5 to 1.0°C through the transition) and reading the absorbance after it had reached a steady value at that temperature. Accuracy of the reported temperature is ± 0.5°C, in the sample cell.

CD spectra for poly[d(A)·d(T)]and poly[d(G)·d(C)]were taken from our previous work.[15,19] Digitized values for these two spectra are given in Ref. 10.

SPECTRA OF THE INDIVIDUAL POLYMERS

The polymer samples were characterized by their melting profiles and CD spectra under various solution conditions. Figure 1A shows the melting profiles at 260 nm of the six polymers at neutral pH, 0.01 M Na^+. Each of the profiles showed one fairly sharp transition, except for poly[d(A-A-C)·d(G-T-T)], which showed a minor second transition at 35-40°C amounting to 4% of the total hyperchromicity and likely due to a small amount of poly[d(A-T)·d(A-T)] in the preparation. The single major transition shown by each sample is evidence that the samples consist mainly of duplexes. We further show

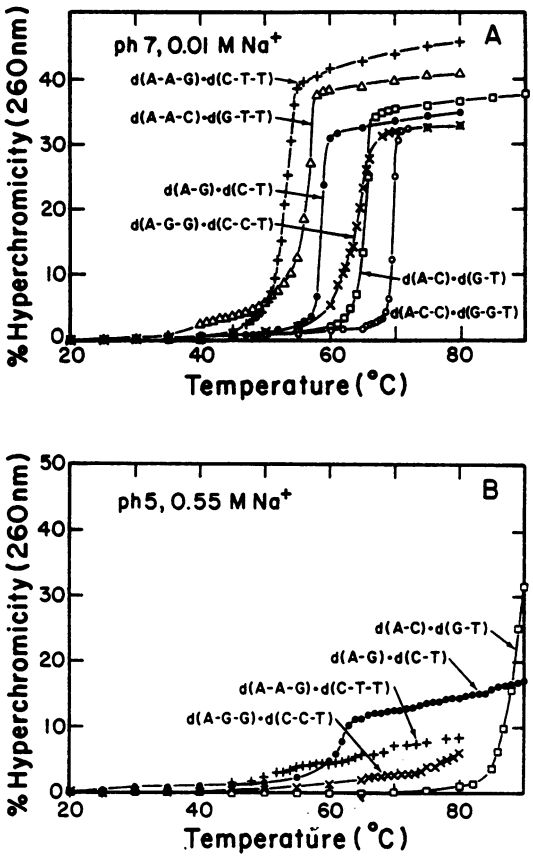

Figure 1. (A) Melting profiles at neutral pH of two repeating dinucleotide and four repeating trinucleotide DNAs. (B) Melting profiles at pH 5 of the two repeating dinucleotide and the two homopurine . homopyrimidine repeating trinucleotide DNAs.

below that the CD spectra attributed to the duplex forms are distinct from the spectra of multiple-stranded forms that may arise from exposure to acid conditions.

The two (A+T)-rich polymers, poly[d(A-A-G)·d(C-T-T)] and poly[d(A-A-C)·d(G-T-T)], had the lowest melting temperatures, 53.5 and 56.5°C, and the highest total hyperchromicities at 80°C, 46 and 41%, respectively. (Since poly[d(A-T)·d(A-T)] has a hyperchromicity of 42%,[3] its presence would not

significantly alter the value measured for the poly[d(A-A-C)·d(G-T-T)] sample.) The two (G+C)-rich polymers, poly[d(A-G-G)·d(C-C-T)]and poly[d(A-C-C)·d(G-G-T)], both had low hyperchromicities of 33% and melted at 64.5 and 70°C, respectively. Poly[d(A-G)·d(C-T)] and poly[d(A-C)·d(G-T)]had intermediate hyperchromicities of 35 and 36.5% and had melting temperatures of 58.5 and 65.5°C, respectively. Melting temperatures determined on several of these polymers at both Dallas and Los Alamos are in agreement. The Tm value for poly[d(A-C)·d(G-T)] is consistent with the value previously reported of 71.5°C in 0.02 M Na^+ (phosphate buffer, pH 7.0). We note that the above values obtained in 0.01 M Na^+ (phosphate buffer, pH 7.0) are 6-10°C lower than melting temperatures published [2,18] for 5 of these polymers (excluding poly[d(A-C-C)·d(G-G-T)]) at the same Na^+ concentration but in a poorly buffered medium (i.e. 0.01 M NaCl plus 0.0001 M sodium phosphate, pH 7.4).

CD spectra of the repeating trinucleotide DNAs poly[d(A-G-G)·d(C-C-T)] and poly[d(A-C-C)·d(G-G-T)]are shown in Figure 2. The spectra of both polymers are altered at low pH, with the CD changes depending upon the procedure used. If H_3PO_4 is added, without increasing the Na^+ concentration, the spectra of both polymers show an increase at long wavelengths (spectra shown as open circles). Such an increase is a characteristic of the formation of $C{\cdot}C^+$ base pairs in poly[d(C)] [19] and could be caused by the formation of intrastrand $C{\cdot}C^+$ base pairs in these polymers. The CD change is completely reversed, at least for poly[d(A-G-G)·d(C-C-T)], by neutralization and heating to 50°C.

If the polymers are dialyzed into a low pH buffer with increased salt, only the homopurine·homopyrimidine polymer shows a significant CD change (spectra shown as closed circles). This low pH form of poly[d(A-G-G)·d(C-C-T)] does not have increased CD at long wavelengths. This self-complex does not melt below 80°C (Figure 1B). It is interesting that the spectrum of this acid form of poly[d(A-G-G)·d(C-C-T)] has an increased negative band at low wavelengths (≤220 nm), a characteristic usually associated with the DNA A-conformation.[4] Arnott and Selsing[20] found by X-ray diffraction that the strands of triple-stranded poly[d(T)·d(A)·d(T)] in fibers are of the A-type. Morgan and Wells[21] showed in previous studies that a triple-stranded complex could be formed between homopurine·homopyrimidine poly[d(A-G)·d(C-T)] and the homopyrimidine RNA poly[r(C-U)] under acid conditions to give poly-[r(C^+-U)·d(A-G)·d(C-T)]. More recently, Johnson and Morgan[22] have presented data to show that poly[d(A-G)·d(C-T)] may undergo a transition to a tetraplex

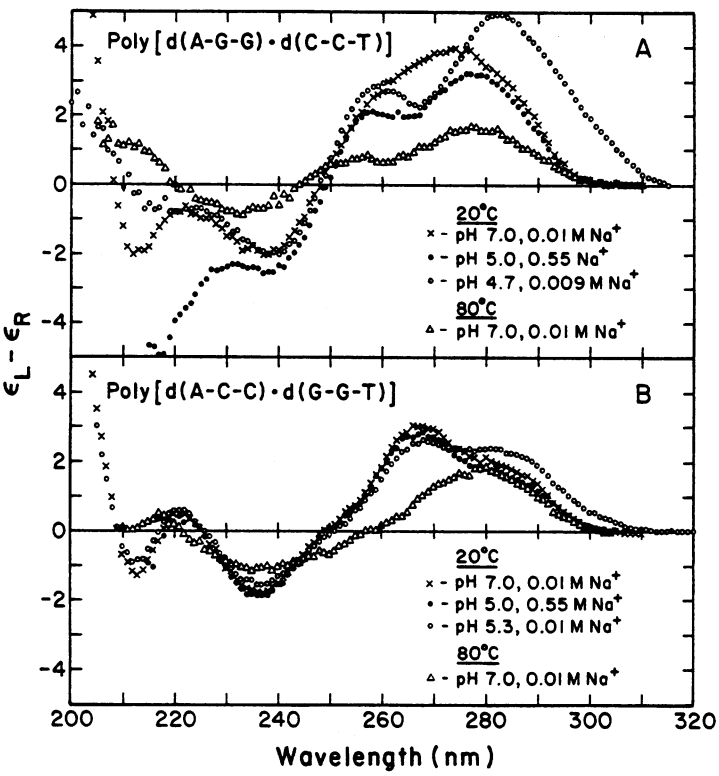

Figure 2. (A) CD spectra of the native homopurine . homopyrimidine repeating trinucleotide DNA poly [d(A-G-G) . d(C-C-T)] at neutral pH (XXX), after direct addition of acid (ooo), or after dialysis to low pH with increased salt (●●●). Also shown is a spectrum of the heat-denatured polymer (ΔΔΔ). (B) CD spectra of the mixed purine-pyrimidine repeating trinucleotide DNA poly [d(A-C-C) . d(G-G-T)] under similar conditions.

form, poly[d(T-C)·d(G-A)·d(T-C^+)·d(G-A)], on lowering the pH at moderately high ionic strengths. Similar tetraplex formation is possible between poly-[d(A-G-G)] and poly[d(C-C-T)] strands. Therefore, it seems likely that the pH 5.0 form we find for poly[(A-G-G)·d(C-C-T)] at 0.55 M Na^+ consists partly of a triple- or multiple-stranded complex that could form only among homopurine and homopyrimidine strands and that involves protonated C^+ residues, but not C·C^+ base pairs.

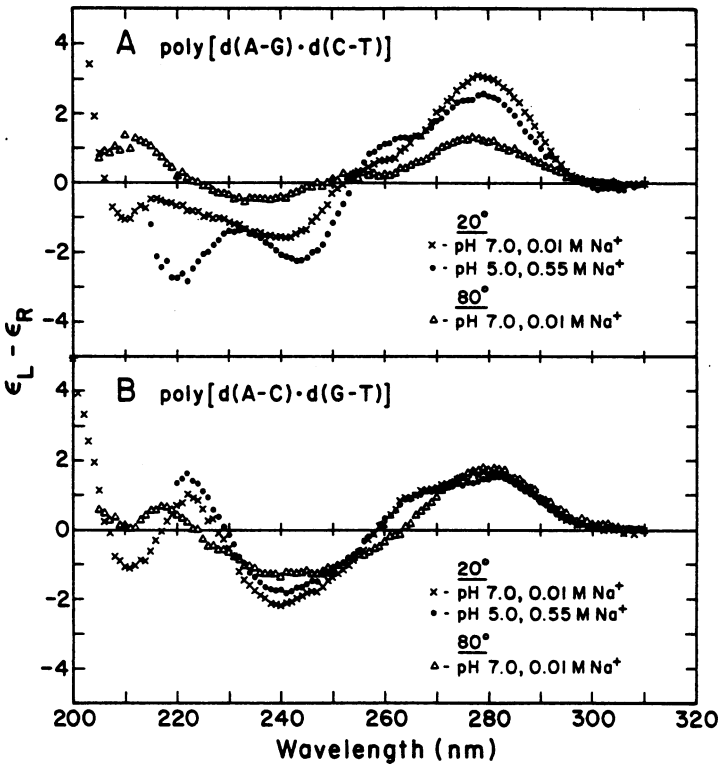

Figure 3. (A) CD spectra of the native homopurine . homopyrimidine repeating dinucleotide DNA poly [d(A-G) . d(C-T)] at neutral pH (XXX), and after dialysis to low pH (●●●). Also shown is a spectrum of the heat-denatured polymer (Δ Δ Δ). (B) CD spectra of the mixed purine-pyrimidine DNA poly [d(A-C) . d(G-T)] under similar conditions.

Similar results were found for the repeating dinucleotide DNAs, poly-[d(A-G)·d(C-T)] and poly[d(A-C)·d(G-T)] upon dialysis to pH 5.0 at 0.55 M Na^+ concentration. Figure 3 shows that the spectrum of the homopurine·homo-pyrimidine DNA was altered at low pH, with the addition of a negative band near 220 nm, while the spectrum of the mixed purine-pyrimidine DNA was not greatly affected. The CD spectrum of poly[d(A-G)·d(C-T)] at low pH under these conditions is like that obtained by Johnson and Morgan[22] for a possible tetraplex form. The low pH spectrum of poly[d(A-G)·d(C-T)] does not show an increase in the 280-310 nm region which might be caused by $C \cdot C^+$ base pairs,

as do the low pH, low salt, spectra of poly[d(A-G-G)·d(C-C-T)] and poly-[d(A-C-C)·d(G-G-T)] shown in Figure 2. Thiele et al.[23] have previously reported CD changes of poly[d(A-G)·d(C-T)] upon acid titration to a much lower pH (2.4-3.0) in 0.15 M NaCl, during which increases do occur in the 280-310 nm region. These CD changes are not fully reversible upon neutralization, but are upon neutralization and heating to 60°C. Whatever unusual polymer structure(s) can form in poly[d(A-G)·d(C-T)] under the conditions described by Thiele et al.[23], it does not exist in our neutral pH sample, as shown both by the absence of an enhanced CD in the 280-310 nm region and by the absence of a second melting transition that would presumably occur well below 60°C in 0.01 M Na^+ (Figure 1A).

The pH 5 form of poly[d(A-G)·d(C-T)] showed a distinct transition at about 62°C in its melting profile, but did not reach its full hyperchromicity (Figure 1B). This transition is at a lower temperature than expected for denaturation of the double helix, since poly[d(A-C)·d(G-T)] melts at 88-89°C as might be expected for this polymer at a salt concentration of 0.55 M Na^+, as seen in Figure 1B. The CD spectra of the poly[d(A-G)·d(C-T)] at pH 5.0 at temperatures above (ΔΔΔ) and below (●●●) the transition are compared in Figure 4. Both spectra contain the 220 nm negative band, unlike the spectrum of the native DNA. Differences between these two spectra are remarkably like those found by Gray and Ratliff[24] for disruption of the self-complex of poly-[d(G-T)]. Future experiments will be necessary to clarify the relation of these observations to the existence of tetraplex or other forms of poly-[d(A-G)·d(C-T)] at low pH.

Except for poly[d(A-C)·d(G-T)], the polymers reanneal only very slowly after heat denaturation in 0.01 M Na^+ at pH 7.0. The homopurine·homopyrimidine polymers are particularly reluctant to resume their native double-stranded forms. For example, when a sample of heat-denatured poly[d(A-G)·d(C-T)] is placed in a 20°C cell holder to cool, it loses 90% of its hyperchromicity in 15 minutes, but it does not regain its original spectrum. The spectrum of the cooled sample instead contains a 220 nm negative band and resembles the low pH form at 70°C; see Figure 4.

The homopurine·homopyrimidine polymer poly[d(A-A-G)·d(C-T-T)] is not as greatly affected by dialysis into the pH 5.0 medium as the other two homopurine·homopyrimidine polymers, as shown in Figure 5A. However, the CD does become negative at low wavelengths as for the other polymers. The melting profile for this polymer at pH 5.0 shows a continuous increase in

absorption but no discrete transitions below 80°C.

We should note that our main purpose in studying the CD spectra of the above polymers was to be assured that the spectra utilized in first-neighbor equations were reasonably those of the double-stranded forms of the polymers, since other multiple-stranded complexes may form with the homopurine·homopyrimidine polymers.

The spectra we have presented for the double-stranded forms of poly-[d(A-G-G)·d(C-C-T)], poly[d(A-A-G)·d(C-T-T)], and poly[d(A-C-C)·d(G-G-T)] are new. Spectra at neutral pH of poly[d(A-C)·d(G-T)] and poly[d(A-G)·d(C-T)] are in good agreement with those previously reported for these polymers.[2,4] There is disagreement between our spectrum of poly[d(A-A-C)·d(G-T-T)] (Figure 5B) and that previously published[2] in the magnitudes of bands below 260 nm and in crossover wavelengths, although not in general shape; these differences are larger than may be explained by the presence of a small amount of poly[d(A-T)·d(A-T)] in our sample.

RESULTS AND DISCUSSION

The measured CD spectra of the above 6 polymers may be used in 4 simple first-neighbor equations to provide 2 calculations each of the spectra of poly[d(A)·d(T)] and poly[d(G)·d(C)]:

$$CD_{calc}(\text{poly[d(A)·d(T)]}) = 3\cdot CD_{meas}(\text{poly[d(A-A-C)·d(G-T-T)]}) - 2\cdot CD_{meas}(\text{poly[d(A-C)·d(G-T)]}), \quad [I]$$

$$CD_{calc}(\text{poly[d(A)·d(T)]}) = 3\cdot CD_{meas}(\text{poly[d(A-A-G)·d(C-T-T)]}) - 2\cdot CD_{meas}(\text{poly[d(A-G)·d(C-T)]}), \quad [II]$$

$$CD_{calc}(\text{poly[d(G)·d(C)]}) = 3\cdot CD_{meas}(\text{poly[d(A-C-C)·d(G-G-T)]}) - 2\cdot CD_{meas}(\text{poly[d(A-C)·d(G-T)]}), \quad [III]$$

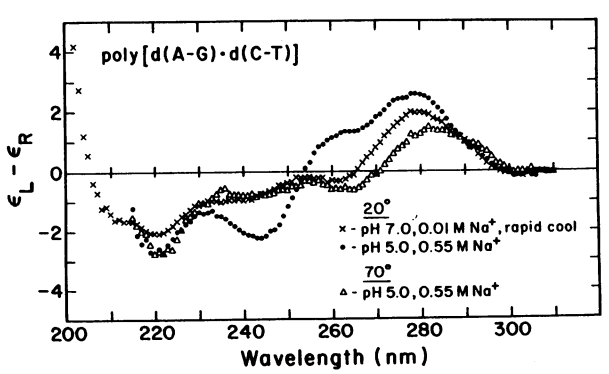

Figure 4. A comparison of the CD spectrum of poly [d(A-G) . d(C-T)] at neutral pH and 20°C after heat-denaturation (XXX) with spectra obtained at low pH and 20°C (●●●) or 70°C (ΔΔΔ).

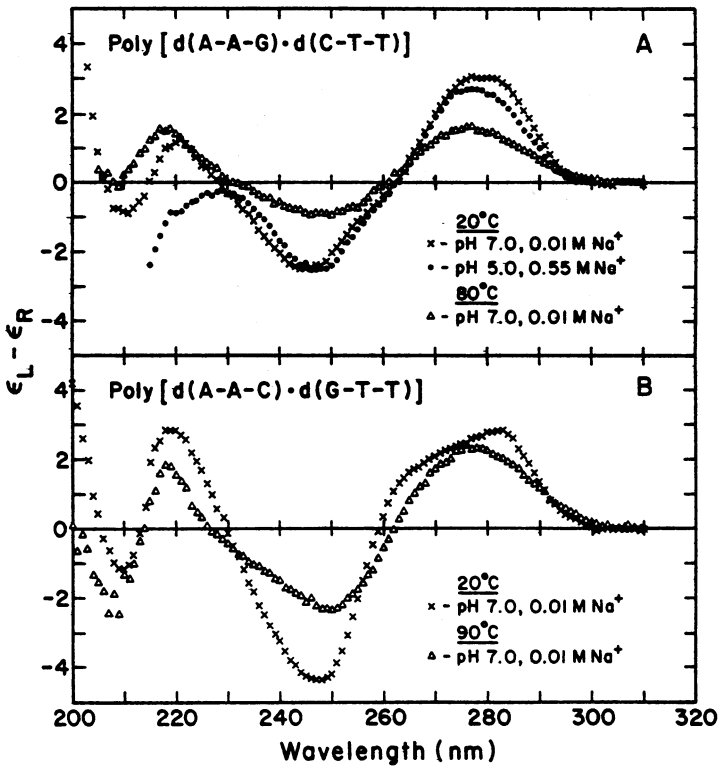

Figure 5. (A) CD spectra of the native homopurine . homopyrimidine repeating trinucleotide DNA poly [d(A-A-G) . d(C-T-T)] at neutral pH (XXX) and after dialysis to low pH (●●●). Also shown is a spectum of the heat-denatured polymer (△△△). (B) CD spectra of the mixed purine-pyrimidine poly [d(A-A-C) . d(G-T-T)] at neutral pH; native polymer (XXX) and heat-denatured polymer (△△△).

$$\text{and} \quad CD_{calc}(\text{poly[d(G)·d(C)]}) = 3 \cdot CD_{meas}(\text{poly[d(A-G-G)·d(C-C-T)]}) - 2 \cdot CD_{meas}(\text{poly[d(A-G)·d(C-T)]}). \quad [IV]$$

The spectra calculated by these equations for poly[d(A)·d(T)] and poly[d(G)·d(C)] may be compared with measured spectra for these polymers.[15,19,25] (Alternatively, of course, the spectra of the repeating dinucleotide or trinucleotide DNAs could be calculated by making use of the measured poly[d(A)·d(T)] and poly[d(G)·d(C)] spectra.) Equations [II] and [IV] are combinations of spectra of homopurine·homopyrimidine polymers to predict the spectra of other homopurine·homopyrimidine polymers; Equations [I] and [III] are combinations of spectra of mixed purine-pyrimidine DNA polymers. One additional simple combination of spectra of mixed purine-pyrimidine DNA polymers may be made from published spectra:[3]

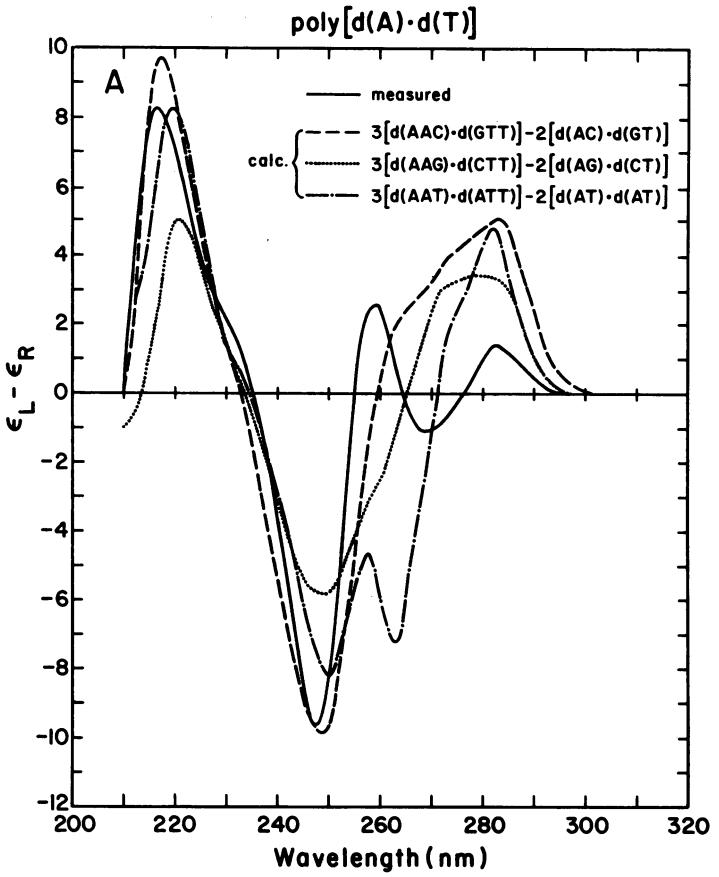

FIGURE 6A

Measured CD spectrum of poly[d(A)·d(T)] compared with spectra calculated from first-neighbor Equations [I], [II], and [V].

Extinction coefficients calculated by equations analogous to Equations [I] - [V] deviate from the measured extinction coefficients for poly[d(A)·d(T)] and poly[d(G)·d(C)] by at least 23%. Extinction coefficients for the calculated spectra from Equations [I] - [V] are 9800, 4050, 5450, 8700, and 7500 liter/(mol·cm), respectively, at 260 nm. Measured extinction coefficients for poly[d(A)·d(T)] and poly[d(G)·d(C)] are 6000[2] and 7060[31], respectively. If we assume that the absorption need not obey first-neighbor equations and that the extinction coefficients of the individual polymers are correct, then the CD spectra may be compared with the magnitudes as shown in Figures 6A and 6B.

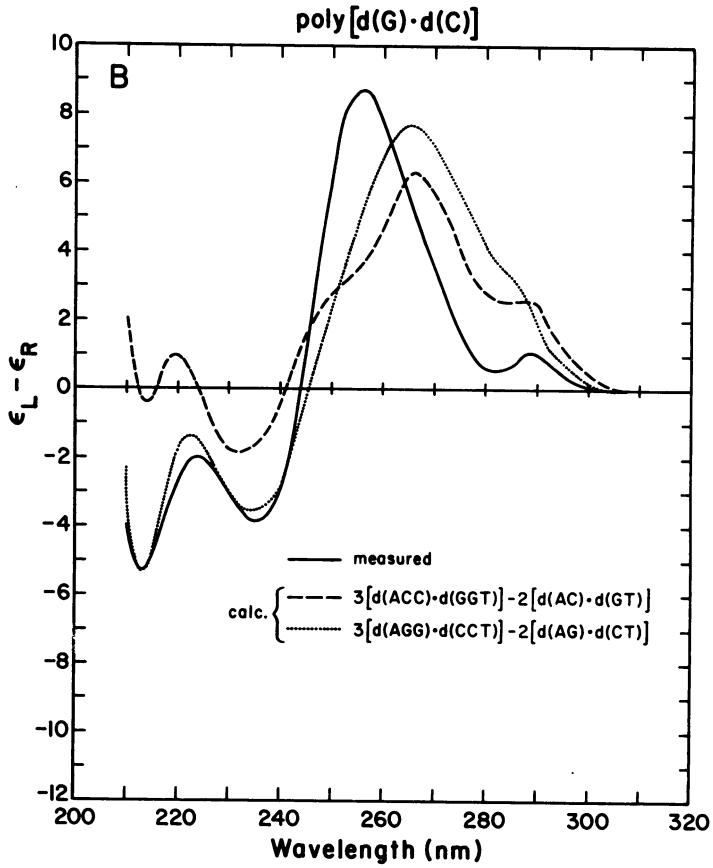

FIGURE 6B

Measured CD spectrum of poly[d(G)·d(C)] compared with spectra calculated from first-neighbor Equations [III] and [IV].

$$CD_{calc}(\text{poly[d(A)·d(T)]}) = 3 \cdot CD_{meas}(\text{poly[d(A-A-T)·d(T-T-A)]}) - 2 \cdot CD_{meas}(\text{poly[d(A-T)·d(A-T)]}). \quad [V]$$

These five equations represent all of the simplest first-neighbor combinations possible from published DNA polymer CD spectra. (One additional first-neighbor combination that involves five DNA polymer spectra is possible,[2] but is not considered here.)

In each of these equations, the CD contributions of the 2 first-neighbors of the repeating dinucleotide polymer are subtracted from the 3 first-neighbor CD contributions of the repeating trinucleotide polymer, leav-

ing the CD contributions of only the d(A-A)·d(T-T) or d(G-G)·d(C-C) first-neighbor base pairs. To the extent that the measured poly[d(A)·d(T)] and poly[d(G)·d(C)] spectra are determined by their first-neighbors, these measured spectra directly show the CD contributions of the d(A-A)·d(T-T) or d(G-G)·d(C-C) first-neighbors in the homopolymer pairs.

Figure 6A shows that the three spectra from Equations [I], [II], and [V] all resemble the measured poly[d(A)·d(T)] spectrum in having a strong positive band at low wavelengths near 220 nm, a shoulder at 230-235 nm and a strong negative band near 250 nm. Above 250 nm, however, none of the calculated spectra compare favorably with the measured poly[d(A)·d(T)] spectrum. Comparisons of spectra by Equation [I] and a rearrangement of Equation [V] have been reported before,[2,3] and the lack of agreement was partially attributed to a difference in the conformation of the mixed purine-pyrimidine DNAs used in the calculations and that of the homopurine·homopyrimidine poly-[d(A)·d(T)]. Figure 6A shows for the first time that a calculation involving only homopurine·homopyrimidine polymers does not agree any better with the measured poly[d(A)·d(T)] spectrum. This means that the CD contribution of the d(A-A)·d(T-T) first-neighbor base pair in the homopurine·homopyrimidine repeating trinucleotide poly[d(A-A-G)·d(C-T-T)] is not more closely related to the CD contribution of this first-neighbor in poly[d(A)·d(T)] than are the CD contributions extracted from the mixed purine-pyrimidine polymers. Moreover, it is a striking result that the discrepancies between the calculated and measured spectra are largely restricted to the wavelength region above 250 nm. Although this could be seen in each of the previous two comparisons,[2,3] it was not commented upon. We will return to this point below.

Figure 6B shows two new comparisons of the measured poly[d(G)·d(C)] spectrum with calculated spectra from Equations [III] and [IV]. Or, in other words, the CD contribution of the d(G-G)·d(C-C) first-neighbor base pair in poly[d(G)·d(C)] is compared with the CD contribution of this first-neighbor in poly[d(A-C-C)·d(G-G-T)] and in poly[d(A-G-G)·d(C-C-T)]. The spectrum calculated from only homopurine·homopyrimidine polymers is closer to the measured poly[d(G)·d(C)] spectrum at wavelengths below 250 nm in having similar magnitudes of negative bands at 213 and 235 nm. These bands are present in the spectrum calculated from the mixed purine-pyrimidine polymers, but they are smaller in magnitude. Above 250 nm, however, neither calculation agrees very closely with the measured spectrum in that the largest positive band appears at about 265 nm instead of at 256 nm. Also, the long-wavelength

positive band near 290 nm is relatively larger in the calculated spectra than in the measured spectrum. (The presence of this band in the calculated spectra does confirm, however, that the measured poly[d(G)·d(C)] spectrum, which has been corrected for the presence of poly[d(C)·d(C^+)] , does indeed have a band in this region which is not due to residual poly[d(C)·d(C^+)]; see Ref. 19 and 25.) Although the spectrum obtained from homopurine·homopyrimidine polymers is closer to the measured poly[d(G)·d(C)] spectrum at low wavelengths, it would be difficult to conclude that the homopurine·homopyrimidine polymer spectra were in better first-neighbor agreement. Rather, the situation may be seen as being parallel to that for the poly[d(A)·d(T)] spectrum in that both calculations correctly acknowledge the measured band positions below 250 nm in the poly[d(G)·d(C)] spectrum, and one calculation is close to the measured spectrum in magnitude as well. Above 250 nm, neither calculation is very close to the measured spectrum in the position of the major positive band.

We conclude that first-neighbor equations do not obviously relate the CD spectra of homopurine·homopyrimidine polymers more accurately to each other than to the spectra of mixed purine-pyrimidine polymers. An additional observation that the CD contributions of d(G-G)·d(C-C) extracted from poly[d(A-G-G)·d(C-C-T)] and poly[d(A-C-C)·d(G-G-T)] agree with each other fairly well above 250 nm (as do the CD contributions of d(A-A)·d(T-T) extracted from poly[d(A-A-G)·d(C-T-T)] and poly[d(A-A-C)·d(G-T-T)] to some extent) also argues that any conformational differences influencing the CD of the first-neighbors are not simply assigned to two classes of polymer sequence.

Why do the extracted CD contributions of d(A-A)·d(T-T) and d(G-G)·d(C-C) not agree better with the measured spectra of poly[d(A)·d(T)] and poly[d(G)·d(C)]? One possibility is that the first-neighbor approximiation is especially poor, as calculated by Cech and Tinoco,[9] for runs of A·T or G·C base pairs. However, it may be seen from their chain length calculations that the effect of non-nearest neighbors for homopolymer sequences does not influence just one region of the CD spectrum. Moreover, for poly[d(G)·d(C)] the effect of increasing chain length is predicted to shift the major positive band to <u>longer</u> wavelengths. We have found that the major positive band of poly[d(G)·d(C)] is at <u>shorter</u> wavelengths than the CD contributions assigned to d(G-G)·d(C-C) first-neighbors in two trinucleotide polymers. This fact, and the observation that, especially for the d(A-A)·d(T-T) CD contributions, deviations of the calculated from the measured spectra are largely restricted to the long wavelength portion of the spectrum, supports the arguments in the

Introduction that the CD spectra of polynucleotides are not generally influenced by large CD contributions from second- and further-neighbors.

On the other hand, it is well-known that it is the long-wavelength CD bands of DNA spectra that are especially sensitive to changes in solution conditions.[5,7,26-28] Therefore, it seems reasonable to tentatively assign the CD differences at long wavelengths seen in Figure 6 to conformational differences among the polymers, differences that are not restricted to either homopurine·homopyrimidine or mixed purine-pyrimidine sequences. This conclusion is in agreement with one of the "rules" for interpreting CD spectra recently proposed by Greve et al.[29] These authors suggested that the long-wavelength portion of CD spectra can be used to assess whether or not a single-stranded polynucleotide has the same conformation as one strand of a double-stranded polynucleotide. If the above interpretation is correct, the comparisons in Figure 6 show that the conformations of many polymer sequences may be different under the same solution conditions. It is possible that conformational differences among the polymers would be reduced under other solution conditions. In fact, the CD spectrum of poly[d(A)·d(T)] shows an enhancement of the long-wavelength positive band near 280 nm and a reduction of the positive band near 260 nm during premelting changes[27] and in solutions of high salt (e.g. 5 M NaCl),[28] changes that bring the measured spectrum of the polymer into better first-neighbor agreement with 2 of the d(A-A)·d(T-T) CD contributions shown in Figure 6A.

The type of DNA conformational change that may be correlated with changes in the long-wavelength portion of the CD spectrum is not known with certainty. Changes in the secondary conformation between B- and C-conformations[30] or changes in the tertiary conformation that affect the nearest-neighbor CD interactions in bent or flexed regions[7] could be involved.

The result that discrepancies in first-neighbor spectral comparisons of synthetic polymers may be largely restricted to a portion of the UV CD spectrum may explain why a spectral analysis of complex DNAs for first-neighbor frequencies can proceed successfully using the spectra of synthetic polymers.[10,13] Our results also show that the accuracy of CD analyses of DNA CD spectra should be greatly enhanced once the factors that limit spectral comparisons of simple sequences are more thoroughly understood.

ACKNOWLEDGEMENTS

We are indebted to Marilyn Vaughan for expert technical assistance. This work was supported by PHS Research Grant No. 19060 from the National Institute of General Medical Sciences and Grant AT-503 from the Robert A.

Welch Foundation. Work at the University of Alberta was supported by the MRC of Canada. Work at the Los Alamos Laboratory was supported by the U. S. Energy Research and Development Administration.

REFERENCES

1. Gray, D.M., and Tinoco, I., Jr. (1970) Biopolymers 9, 223-244.
2. Wells, R.D., Larson, J.E., Grant, R.C., Shortle, B.E., and Cantor, C.R. (1970) J. Mol. Biol. 54, 465-497.
3. Gray, D.M., Ratliff, R.L., and Williams, D.L. (1973) Biopolymers 12, 1233-1245.
4. Gray, D.M., and Ratliff, R.L. (1975) Biopolymers 14, 487-498.
5. Ivanov, V.I., Minchenkova, L.E., Schyolkina, A.K., and Poletayev, A.I. (1973) Biopolymers 12, 89-110.
6. Maestre, M.F., and Wang, J.C. (1971) Biopolymers 10, 1021-1030.
7. Gray, D.M., Taylor, T.N., and Lang, D. (1978) Biopolymers 17, 145-157.
8. Johnson, W.C., Jr., and Tinoco, I., Jr. (1969) Biopolymers 7, 727-749.
9. Cech, C.L., and Tinoco, I., Jr. (1977) Biopolymers 16, 43-65.
10. Gray, D.M., Hamilton, F.D., and Vaughan, M.R. (1978) Biopolymers 17, 85-106.
11. Arnott, S. (1975) Nucleic Acids Res. 2, 1493-1502.
12. Gray, D.M., Lee, C.S., and Skinner, D.M. (1978) Biopolymers 17, 107-114.
13. Marck, C., and Guschlbauer, W. (1978) Nucleic Acids Res. 5, 2013-2031.
14. Langridge, R. (1969) J. Cell Physiol. 74 (suppl. 1), 1-20.
15. Allen, F.S., Gray, D.M., Roberts, G.P., and Tinoco, I., Jr. (1972) Biopolymers 11, 853-879.
16. Arnott, S. (1976) in DNA Secondary Structures in Organization and Expression of Chromosomes. Allfrey, V.G., Bautz, E.K.F., McCarthy, B.J., Schimke, R.T., and Tissieres, A., Eds., Dahlem Konferenzen, Berlin.
17. Ratliff, R.L., Hoard, D.E., Hayes, F.N., Smith, D.A., and Gray, D.M. (1976) Biochemistry 15, 168-176.
18. Morgan, A.R., Coulter, M.B., Flintoff, W.F., and Paetkau, V.H. (1974) Biochemistry 13, 1596-1603.
19. Gray, D.M., and Bollum, F.J. (1974) Biopolymers 13, 2087-2102.
20. Arnott, S., and Selsing, E. (1974) J. Mol. Biol. 88, 509-521.
21. Morgan, A.R., and Wells, R.D. (1968) J. Mol. Biol. 37, 63-80.
22. Johnson, D., and Morgan, A.R. (1978) Proc. Nat. Acad. Sci., U.S. 75, 1637-1641.
23. Thiele, D., Sarocchi, M.-T., Guschlbauer, W., Lezius, A., and Marck, C. (1973) Mol. Biol. Rep. 1, 155-160.
24. Gray, D.M., and Ratliff, R.L. (1977) Biopolymers 16, 1331-1342.
25. Marck, C. and Thiele, D. (1978) Nucleic Acids Res. 5, 1017-1028.
26. Ivanov, V.I., Minchenkova, L.E., Minyat, E.E., Frank-Kamenetskii, M.D., and Schyolkina, A.K. (1974) J. Mol. Biol. 87, 817-833.
27. Greve, J., Maestre, M.F., and Levin, A. (1977) Biopolymers 16, 1489-1504.
28. Zimmer, C., and Luck, G. (1974) Biochim. Biophys. Acta 361, 11-32.
29. Greve, J., Maestre, M.F., Moise, H., and Hosoda, J. (1978) Biochemistry 17, 887-893.
30. Tunis-Schneider, M.J.B., and Maestre, M.F. (1970)J. Mol. Biol. 52, 521-541.
31. Inman, R.B., and Baldwin, R.L. (1964) J. Mol. Biol. 8, 452-469.