# Ethidium Bromide Complexes with Self-Complementary Deoxytetranucleotides. Demonstration and Discussion of Sequence Preferences in the Intercalative Binding of Ethidium Bromide†

Rodney V. Kastrup, Michael A. Young, and Thomas R. Krugh*,‡

ABSTRACT: The binding of ethidium bromide to the self-complementary deoxytetranucleotides containing guanine and cytosine bases has been studied by circular dichroism, visible absorption, and fluorescence spectroscopies. The circular dichroism spectrum of each of the four deoxytetranucleotides was measured and compared to the spectrum calculated by a nearest-neighbor approximation method which utilized the circular dichroism spectra of the deoxydinucleotides as a basis set. Reasonable agreement was obtained between the calculated and experimental spectra for three of the four deoxytetranucleotides; however, pdC-dC-dG-dG exhibited poor agreement between the actual and nearest-neighbor calculated spectra, which suggests that pdC-dC-dG-dG may exist in an unusual conformation. A nearest-neighbor method was also used to calculate the extinction coefficients for each of the deoxytetranucleotides: these values differ substantially from values recently published by Patel and Canuel [(1977), *Proc. Natl. Acad. Sci. U.S.A. 74*, 2624–2628] in which the method used for the determination of the extinction coefficients was not stated. The magnitude of the extinction coefficients is important in determining the stoichiometry of complex formation as well as the relative sequence preferences for ethidium binding. The visible absorption titrations, the fluorescence titrations, and the circular dichroism titrations with ethidium bromide clearly show that *two* ethidiums will intercalate into a (pdC-dG-dC-dG)·(pdC-dG-dC-dG) double helix, presumably at the two (dC-dG)·(dC-dG) sequences. Results from the pdG-dC-dG-dC titrations with ethidium bromide are not as definitive. Raising the temperature from ~2 to 26 °C diminishes the strength of complex formation for the EthBr plus pdC-dG-dC-dG system and makes it more difficult to unequivocally determine the stoichiometry of the complex formation. These data thus confirm and extend our earlier observation that ethidium bromide preferentially binds to pyrimidine–purine sequences as opposed to purine–pyrimidine sequences. Experiments monitoring the binding of ethidium bromide to pdC-dC-dG-dG and pdG-dG-dC-dC indicate that ethidium will bind strongly to the (dG-dG)·(dC-dC) sequence. We conclude that the relative ordering of the sequence preferences for the binding of ethidium to the three sequences available in the tetranucleotides studied is (dC-dG)·(dC-dG) ≳ (dG-dG)·(dC-dC) ≫ (dG-dC)·(dG-dC).

A variety of spectroscopic techniques have been used to study the interaction of ethidium bromide with nucleic acids (Aktipis et al., 1975; Waring, 1965; Le Pecq and Paoletti, 1967; Angerer et al., 1974). Investigations of several model systems have been carried out in order to facilitate an understanding of ethidium bromide binding to polynucleotides (Krugh and Reinhardt, 1975; Krugh et al., 1975; Davanloo and Crothers, 1976; Patel and Canuel, 1976). In this paper we report the results from circular dichroism, fluorescence, and visible absorption measurements on the ethidium bromide complexes formed with the four self-complementary G- and C-containing deoxytetranucleotides. These four tetranucleotides are sequence isomers of one another and it is possible to deduce information on the relative preference of ethidium to bind to various nucleotide sequences which are available as intercalation sites. Ethidium demonstrated a clear preference for binding to pyrimidine–purine sites [such as (C-G)·(C-G), (T-G)·(C-A), or (T-A)·(T-A) as opposed to (G-C)·(G-C), (G-T)·(A-C), and (A-T)·(A-T), respectively] at the dinucleotide level (Krugh and Reinhardt, 1975; Krugh et al., 1975; Reinhardt and Krugh, 1978). The tetranucleotides present a new set of problems in the analysis of the binding curves because each tetranucleotide helix has three intercalation sites in addition to possible sites involving stacking of the ethidium bromide molecules on the end or, less likely, on the outside of the helix. The studies in this paper will be aimed at sorting out the binding of ethidium to the deoxytetranucleotides.

## Experimental Section

The tetranucleotides and hexanucleoside pentaphosphate, dG-dC-dG-dC-dG-dC, were purchased from Collaborative Research. The dG-dC-dG-dC-dG-dC was used without further purification, although the purity was confirmed using DEAE[1]-Sephadex chromatography. In the case of the fluorescence and visible titrations, the tetranucleotides were passed down a Chelex-100 column and lyophilized prior to use. For the CD experiments the tetranucleotides were generally used without further purification. Only the tetranucleotide pdC-dC-dG-dG, lot 587-69A, required purification prior to use. The fact that the compound was impure was established by TLC chromatography and CD spectroscopy. TLC chromatography on silica gel F plates using 1-propanol/$H_2O$/$NH_4OH$ (55:35:10, v/v/v) solvent, with both pH 2 and 7 solutions of pdC-dC-dG-dG, showed a very faint spot having an $R_f$ value of 0.5, the same as that of a dinucleotide, while a more intense spot showed an $R_f$ value of 0.4, corresponding to that of a tetranucleotide. The pH 2 solution also showed a faint spot

† From the Department of Chemistry, University of Rochester, Rochester, New York 14627. *Received January 31, 1978; revised manuscript received June 26, 1978.* This investigation was supported by Public Health Research Grant CA-17865 and Research Career Development Award K04 CA 00257 from the National Cancer Institute.

‡ Alfred P. Sloan Fellow and Research Career Development Awardee.

[1] Abbreviations used are: EthBr, ethidium bromide; CD, circular dichroism; CT DNA, calf thymus DNA; NNA, nearest-neighbor approximation; fb, fraction bound; EDTA, ethylenediaminetetraacetic acid; DEAE, diethylaminoethyl; pyr, pyrimidine; pur, purine.

0006-2960/78/0417-4855$01.00/0 

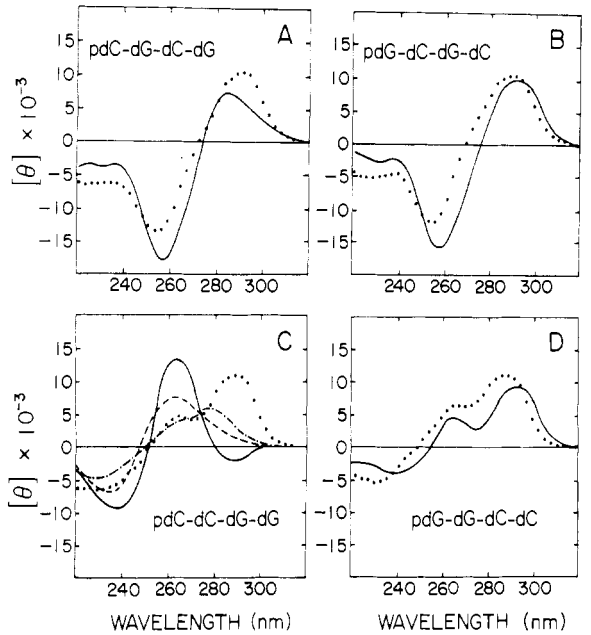



FIGURE 2: Comparison of observed and nearest-neighbor approximation CD spectra of the four self-complementary C-G tetranucleotides. All spectra were recorded in a 0.1 M NaCl, 0.1 mM NaEDTA, 5 mM phosphate solution (pH 7.0) at 0 °C, unless otherwise indicated. The dotted (···) curves represent the nearest-neighbor approximation curves for each of the deoxytetranucleotides: (A) (—) 44 μM pdC-dG-dC-dG; (B) (—) 48 μM pdG-dC-dG-dC; (C) (—) 56 μM pdC-dC-dG-dG in 5 mM phosphate (pH 7.0) at 2 °C, (- - -) 5.6 μM pdC-dC-dG-dG in 5 mM phosphate (pH 7.0) at 0 °C, (- • -) 56 μM pdC-dC-dG-dG in 5 mM phosphate (pH 7.0) at 41 °C. A spectrum of 33 μM pdC-dC-dG-dG in a 0.1 M NaCl, 0.1 mM EDTA, 5 mM phosphate solution (pH 7.0) at 0 °C had a similar shape but larger $[\theta]$ values than the spectrum shown above (—) which was recorded in a solution which did not contain the NaCl: (D) (—) 45 μM pdG-dG-dC-dC.

having an $R_f$ of 0.11. The pdC-dC-dG-dG sample was purified by a variation of the method of Bauer et al. (1975). A sample of pdC-dC-dG-dG (lot 587-69A) was desalted using a 1 × 45 cm P-2 gel column with distilled $H_2O$ as the eluant. After the material had been lyophilized, the sample (5.6 $A_{255}$ units) was dissolved in 500 μL of 0.2 M $NH_4HCO_3$ (pH 8.3) and applied to a 1 × 19 cm column of A-25 DEAE-Sephadex which had been equilibrated with 0.2 M $NH_4HCO_4$ (pH 8.3). The sample was eluted using a 400-mL gradient from 0.2 M (pH 8.3) to 0.8 M $NH_4HCO_3$ (pH 8.3) at a flow rate varying from 30 to 18 mL/h. Fractions were collected every 12 min. The elution profile (shown in Figure 1, in the supplementary material) showed closely spaced peaks, the last of which eluted at a salt concentration corresponding to that of a tetranucleotide. The fractions from this peak were pooled, lyophilized, and desalted by P2 gel chromatography, using $H_2O$ as an eluant. The CD spectrum of the purified material corresponded very well to spectra of other uncontaminated pdC-dC-dG-dG samples obtained from Collaborative Research (see Figure 2).[2]

Ethidium bromide (EthBr) was purchased from Sigma and used without further purification. Concentrations of EthBr were determined using $\epsilon_{480}$ 5860 $M^{-1}$ $cm^{-1}$ (K. R. Lee, unpublished data; Wakelin and Waring, 1974; Bresloff and Crothers, 1975). Concentrations of the tetranucleotides were determined using $\epsilon_{255}$ values that had been calculated from the di- and mononucleotide values using the nearest-neighbor approximation. $\epsilon_{255}$ values at pH 7.0 (and ~20 °C) were calculated from $\epsilon_{255}$ determined at pH 11.0 using the nearest-neighbor approximation, and the $A_{255}$ (pH 7)/$A_{255}$ (pH 11) ratio for each tetranucleotide (T. R. Krugh and C. G. Reinhardt, unpublished data). The values are: pdC-dG-dC-dG, $\epsilon_{255}$ 9440 $M^{-1}$ $cm^{-1}$; pdG-dC-dG-dC, $\epsilon_{255}$ 9460 $M^{-1}$ $cm^{-1}$; pdC-dC-dG-dG, $\epsilon_{255}$ 9800 $M^{-1}$ $cm^{-1}$; pdG-dG-dC-dC, $\epsilon_{255}$ 9800 $M^{-1}$ $cm^{-1}$. For the dG-dC-dG-dC-dG-dC the nearest-neighbor approximation yielded $\epsilon_{260}$ 8890 $M^{-1}$ $cm^{-1}$. This value compares very favorably with $\epsilon_{260}$ 8850 $M^{-1}$ $cm^{-1}$ determined by Dr. Fritz M. Pohl using enzymatic degradation (DNase I and phosphodiesterase) of dG-dC-dG-dC-dG-dC (Fritz M. Pohl, personal communication).

CD spectra were recorded on a JASCO J-40 spectrometer using a 0.5-cm jacketed cell. The absorbance was kept below 2 throughout the titrations. The cell temperature was maintained between 1 and 3 °C for the various samples and was regulated to within ±0.5 °C using a Haake FS2 circulating bath. Unless otherwise indicated, the buffer for all samples was 0.1 M NaCl, 5 or 10 mM phosphate (pH 7.0), 0.1 mM EDTA. The initial solution for all the CD titrations consisted of 525 or 550 μL of a tetranucleotide solution.

Fluorescence and visible absorption titrations were carried out simultaneously. Four-millimeter fluorescence cells were used in both spectrophotometers. Titrations were carried out using an initial solution of 225 μL of EthBr in the sample cell and 225 μL of buffer in the reference cell. Tetranucleotide solutions were added with a 10-μL Hamilton syringe to the sample cell and also to the reference cell in order to correct for minor impurities in the tetranucleotide samples which absorbed in the visible region. After each addition, the solutions were mixed with a glass stirring rod. Fluorescence spectra were obtained using a Perkin-Elmer MFP-2A fluorescence spectrophotometer operating in the direct mode. An excitation wavelength of 546 nm and an emission wavelength of 590 nm were used for all experiments. Visible absorption spectra were measured on a Cary 118 spectrophotometer. The experiment utilizing the method of continuous variation (Job, 1928) was carried out in a 1-cm cell, at 2.5 °C. The concentrations of EthBr and the dG-dC-dG-dC-dG-dC strand were varied so that the total concentration (EthBr plus hexanucleoside pentaphosphate strand) was kept constant at 150 μM. The changes in the absorbance, $\Delta A$, and the relative quantum yields, $\phi_{app}/\phi$, were calculated as previously reported (Krugh and Reinhardt, 1975). The fluorescence emission maximum of free ethidium is the same as that of ethidium bound to the oligonucleotides.

## Results

*CD Spectra of the Free Tetranucleotides.* The CD spectra of each of the four tetranucleotides, as well as spectra calculated using the nearest-neighbor approximation (NNA) (e.g., see Cech et al., 1976; Allen et al., 1972; Gray and Tinoco, 1970; Cantor and Tinoco, 1965), are shown in Figure 2. The ellipticities measured for the deoxydinucleotides were used to compute molar ellipticities, which in turn were used in the nearest-neighbor approximation. The equations used to calculate the CD spectra of the tetranucleotides are:

$$[\theta]_{\lambda,CGCG} = 0.25[4[\theta]_{\lambda,CG} + 2[\theta]_{\lambda,GC} - [\theta]_{\lambda,G} - [\theta]_{\lambda,C}]$$

$$[\theta]_{\lambda,GCGC} = 0.25[2[\theta]_{\lambda,CG} + 4[\theta]_{\lambda,GC} - [\theta]_{\lambda,G} - [\theta]_{\lambda,C}]$$

$$[\theta]_{\lambda,CCGG} = 0.25[2[\theta]_{\lambda,CC} + 2[\theta]_{\lambda,GG} + 2[\theta]_{\lambda,CG} - [\theta]_{\lambda,G} - [\theta]_{\lambda,C}]$$

$$[\theta]_{\lambda,GGCC} = 0.25[2[\theta]_{\lambda,GG} + 2[\theta]_{\lambda,CC} + 2[\theta]_{\lambda,GC} - [\theta]_{\lambda,G} - [\theta]_{\lambda,C}]$$

[2] Collaborative Research has recently changed their methods of synthesizing the deoxyoligonucleotides, which should reduce the problems with impurities.

All the $\theta$ values in these equations are molar ellipticities on a residue (nucleoside) basis.

The comparison between the measured spectra and the NNA spectra in Figure 2 shows reasonable agreement between the measured and NNA spectra for pdC-dG-dC-dG, pdG-dC-dG-dC, and pdG-dG-dC-dC. The calculated and measured spectra for pdC-dC-dG-dG, however, are very different. The observed pdC-dC-dG-dG spectrum differs from the predicted spectrum primarily in the 260–300-nm region. Since the purity of the pdC-dC-dG-dG sample was established by chromatography (see Experimental Section), it is apparent that the nearest-neighbor analysis calculation of the CD spectrum is not adequate for calculating the CD spectrum of pdC-dC-dG-dG.

It is intriguing that pdC-dC-dG-dG, with a $(pyr)_2$–$(pur)_2$ sequence, is the one deoxytetranucleotide whose NNA CD spectrum differs greatly from the observed spectrum. We note with interest that the double-stranded block oligonucleotide $(dT_{15}\text{-}dG_{15})\cdot(dC_{15}\text{-}dA_{15})$ has exhibited unusual melting properties in which there is a transmission of stability from the dG·dC region to the dT·dA region (Burd et al., 1975a,b). Also, in a recent paper, Early et al. (1977) found evidence using NMR that the junction between the dA·dT and dG·dC base-paired regions in this block oligonucleotide exhibits an unusual conformation of the double helix. Under the conditions used to calculate the NNA CD spectrum (0 °C, 0.1 $\mu$M NaCl), a fraction of the pdC-dC-dG-dG was in the double-helical conformation. However, even the spectra taken under conditions of single strandedness (i.e., high temperature or low concentration; cf. Figure 2C) do not reasonably approximate the NNA CD spectrum. It is clear that the discrepancies between the observed and the NNA CD spectra are not a result of the presence of varying amounts of double- and single-stranded tetranucleotide, which suggests either the presence of an unusual conformation that is presumably associated with the nucleotide sequence, or the failure of the nearest-neighbor approximation for calculating the CD spectrum.

*Extinction Coefficients of the Deoxyoligonucleotides.* The determination of the stoichiometry of drug–nucleic acid complexes is limited by the accuracy of the extinction coefficients of the oligonucleotides and the drugs. The extinction coefficient for EthBr is well known; however, the values used for the oligonucleotides needed to be determined. Nearest-neighbor approximation (NNA) equations similar to those used for the CD spectra were used to calculate $\epsilon_{255}$ for each of the four tetranucleotides and $\epsilon_{260}$ for the hexanucleoside pentaphosphate. Extinction coefficients listed in the P-L Biochemicals catalog 104 or 105 were used as the best available source for the extinction coefficients of the deoxydinucleosides. Using these values, we determined the extinction coefficients of the deoxydinucleotides at pH 11 and then used the NNA method to calculate the extinction coefficients of the deoxytetranucleotides at pH 11. The values obtained from the data at pH 11 were converted to pH 7 and agreed within ~5% of the values determined directly at pH 7; the extinction coefficients listed in the Experimental Section are intermediate between the two sets of data and have been weighted toward the values determined at pH 11, since stacking interactions should be less important in these solutions. In the case of dG-dC-dG-dC-dG-dC, $\epsilon_{260}$ was calculated only at pH 7. The agreement between the NNA calculated value of $\epsilon_{260}$ 8890 $M^{-1}$ $cm^{-1}$ and the experimentally determined value of $\epsilon_{259}$ 8850 $M^{-1}$ $cm^{-1}$ (by enzymatic degradation and spectrophotometric examination of the products, Fritz M. Pohl, personal communication) leads us to conclude that the NNA method is reliable for calculating extinction coefficients values. We note also that

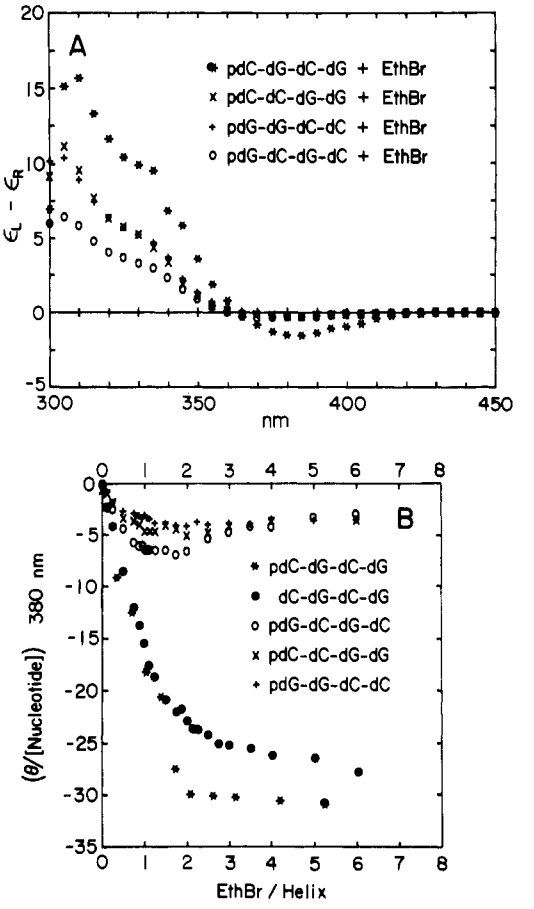



FIGURE 3: (A) Circular dichroism spectra of ethidium bromide solutions with the deoxytetranucleotides. All spectra were recorded between 0 and 3 °C in a 0.1 M NaCl, 0.1 mM EDTA, 5 mM phosphate solution at pH 7.0. The concentration of the ethidium bromide was equal to the concentration of the tetranucleotide in each solution (which is equivalent to stating that the solutions are two ethidiums per equivalent concentration of tetranucleotide helix); the concentrations of the tetranucleotides were: pdC-dG-dC-dG, 44 $\mu$M; pdC-dC-dG-dG, 84 $\mu$M; pdG-dG-dC-dC, 120 $\mu$M; pdG-dC-dG-dC, 81 $\mu$M. (B) Normalized ellipticities (i.e., ellipticities per mole of nucleotide) as a function of the concentration of ethidium bromide added (expressed in ethidium bromide/helix ratios). All titrations were measured between 0 and 2 °C in a 0.1 M NaCl, 0.1 mM EDTA, 5 mM phosphate solution (pH 7.0). The ellipticities were normalized to a constant nucleotide concentration to correct for small dilution effects during the titration. The concentrations of the nucleotides at the start of the titrations were: pdC-dG-dC-dG, 44 $\mu$M; dC-dG-dC-dG, 75 $\mu$M; pdG-dC-dG-dC, 81 $\mu$M; pdC-dC-dG-dG, 84 $\mu$M; pdG-dG-dC-dC, 120 $\mu$M.

there is a general decrease in extinction coefficients (per residue) proceeding from mononucleotide to dinucleotide, to tetranucleotide, to hexanucleotide, and to polynucleotide, as expected. This also leads us to conclude that the values we have calculated for the tetranucleotides are more reliable than the lower values (7250 to 7625 $M^{-1}$ $cm^{-1}$ per residue) reported by Patel and Canuel (1977) for three of these tetranucleotides. Differences in temperature and salt conditions cannot account for the discrepancy between our values and those of Patel and Canuel, since we find only a few percent *increase* in our values of the extinction coefficients at pH 7 when using the no salt, 70 °C conditions employed by Patel and Canuel (1977).

*CD Spectra of the Ethidium Bromide–Deoxytetranucleotide Complexes.* CD spectra of solutions containing equimolar concentrations of EthBr and each of the four deoxytetranucleotides (2 EthBr/1 helix, or alternatively stated, 2 EthBr/2 strands) are shown in Figure 3A. The spectra all show a characteristic negative band of small amplitude at 380 nm and

positive bands of larger amplitude at 330 and 307.5 nm. The wavelengths of the extrema in these spectra correspond very well to those observed for EthBr–DNA solutions. Spectra from previous experiments with EthBr and the deoxy- and ribodinucleotides also showed bands at ~307, 330, and 380 nm arising from a complex in which ethidium is intercalated between two base pairs. Differences among the CD spectra of the four drug–oligonucleotide solutions are readily apparent. The pdC-dG-dC-dG solution with EthBr exhibits the largest amplitudes at 380, 330, 307.5 nm; pdC-dC-dG-dG and pdG-dG-dC-dC exhibit intermediate values and pdG-dC-dG-dC exhibits the smallest magnitudes at 330 and 307.5 nm, although the 380-nm bands for the latter three solutions are nearly equivalent. The differences in the $\epsilon_L - \epsilon_R$ values for the ethidium solution with pdG-dC-dG-dC when compared to pdC-dG-dC-dG illustrates the effect that sequence isomers can have on the observed CD spectra. Spectra of the ethidium solutions with the other two deoxytetranucleotides, pdC-dC-dG-dG and pdG-dG-dC-dC, are quite similar to one another (Figure 3), which suggests that the EthBr environment in the two complexes is similar, even though the conformations of these two tetranucleotides in the absence of EthBr may be very dissimilar (cf. Figure 2). Titrations were carried out by monitoring the amplitude of the ellipticity at 380, 330, and 307.5 nm in an attempt to determine the stoichiometry of complex formation between EthBr and the deoxytetranucleotides. The origins of the induced optical activity observed in the 300–400-nm region of the spectrum have been discussed by a number of authors (e.g., Balcerski and Pysh, 1976; Houssier et al., 1974). It is generally believed that the 308- and 330-nm bands result from a combination of the inherent asymmetry associated with the intercalation of the phenanthridinium ring into the helix as well as exciton interactions between two intercalated ethidiums. The dye–dye interaction results in the molar ellipticity of the bound ethidium being a function of the fraction of intercalation sites occupied. On the other hand, the asymmetry of the binding site is thought to produce the optical activity at 380 nm (Houssier et al., 1974).

*$\theta^{380}$ vs. Ethidium Bromide/Helix Ratio.* The ellipticities at 380 nm (normalized to the same tetranucleotide concentration) vs. the ethidium bromide/helix ratio for each of the four tetranucleotides, as well as the tetranucleoside triphosphate dC-dG-dC-dG, are shown in Figure 3B. It should be noted that, under the present experimental conditions, the tetranucleotides exist partially in the single-stranded form and partially in the double-helical form [e.g., see Pohl (1974) for data on the double-helical formation of $(dG\text{-}dC)_n$ oligomers in *1 M NaCl* solution]. The addition of the ethidium stabilizes the tetranucleotide double helices just as it stabilizes DNA. The relatively sharp end points in the titrations with the dC-dG-dC-dG and pdC-dG-dC-dG oligonucleotides (Figure 3B) both indicate that two ethidiums can combine with two of these tetranucleotides to form a miniature double-helical complex in which two ethidiums are intercalated into each double helix. Although additional binding of ethidium to the helix may occur after this point, the behavior of the ellipticity at 380 nm indicates that two, and only two, EthBr molecules will intercalate into the double helices formed by either pdC-dG-dC-dG or dC-dG-dC-dG.

The $\theta^{380}$ titration curves for pdG-dC-dG-dC and pdC-dC-dG-dG with ethidium bromide level off at the 1 ethidium/1 helix ratio. In the pdG-dG-dC-dC titration (Figure 3B), the value of $\theta^{380}$ does not exhibit a sharp break point, but rather it exhibits a general leveling off near the 2 ethidium/1 helix ratio. It should be noted that in all but the pdC-dG-dC-dG (and dC-dG-dC-dG) titrations the small amplitude of the induced CD at 380 nm presents difficulties in interpreting the data (e.g., compare the data in Figure 3B to the spectra in Figure 3A).

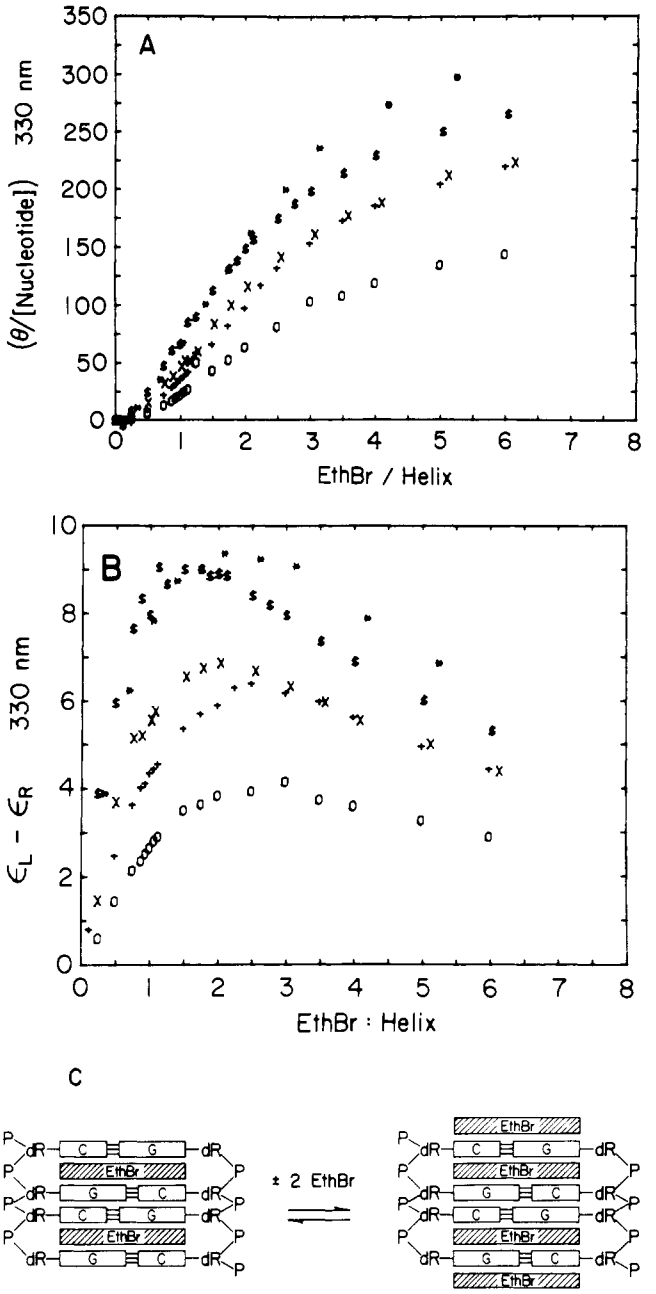



FIGURE 4: (A) Normalized ellipticities at 330 nm vs. the ethidium/helix ratios for the same titrations as shown in Figure 3. (B) Molar circular dichroism at 330 nm (based on the *ethidium bromide* concentration) for the titrations shown in Figure 3. The symbols used in A and B above are: (*) pdC-dG-dC-dG; ($) dC-dG-dC-dG; (O) pdG-dC-dG-dC; (X) pdC-dC-dG-dG; (+) pdG-dG-dC-dC. (C) Schematic illustration of a complex in which two ethidiums are intercalated into the two C-G sequences of the helix formed by two pdC-dG-dC-dG tetranucleotides (on the left above), and a schematic illustration of the possible stacking of one or two additional ethidiums to the ends of the helix when there is an excess of ethidium present in the solution.

*$\theta^{330}$ and $\epsilon_L - \epsilon_R{}^{330}$ Values as a Function of the Ethidium/Helix Ratio.* We have plotted the $\theta^{330}$ values as a function of the ethidium to helix ratio in Figure 4A. In considering these data, it is important to remember that the 330-nm bands arise, at least in part, from dye–dye interactions (e.g., see Houssier et al., 1974, and references therein) and that ethidium can stack on the ends of the oligonucleotide helices, once formed in solution (e.g., see Figure 4C). The $\theta^{330}$ do not provide

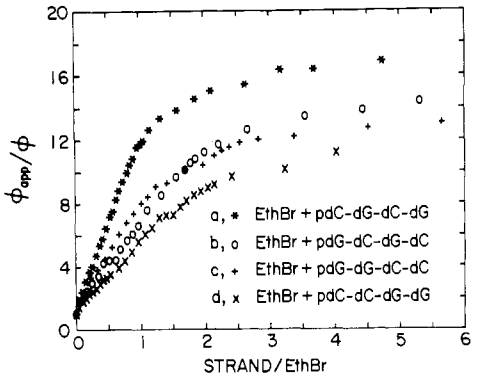



FIGURE 6: Apparent quantum yield $\phi_{app}/\phi$, as a function of the added tetranucleotide, expressed in terms of the ratio of tetranucleotide strand concentration to ethidium concentration. The concentration of ethidium at the start of each concentration was: (a) 144 $\mu$M for pdC-dG-dC-dG; (b) 142 $\mu$M for pdG-dC-dG-dC; (c) 133 $\mu$M for pdC-dC-dG-dG; (d) 124 $\mu$M for pdG-dG-dC-dC. All titrations were performed in 0.1 M NaCl, 0.1 mM EDTA, 5 mM phosphate solution (pH 7.0) between 0 and 2 °C.

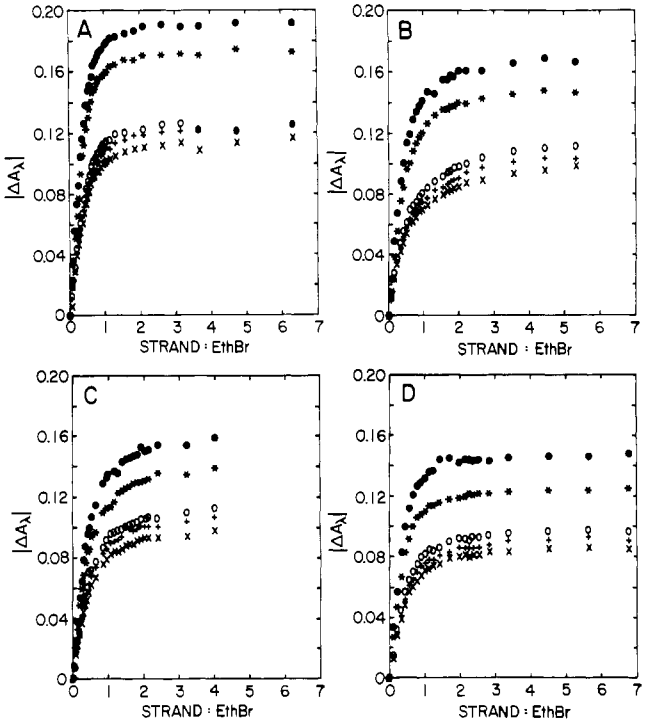



FIGURE 7: Absolute changes in the absorbance of ethidium at a given wavelength, $|\Delta A_\lambda|$, plotted as a function of the strand to ethidium ratios for the pdC-dG-dC-dG (A), pdG-dC-dG-dC (B), pdC-dC-dG-dG (C), and pdG-dG-dC-dC (D) titrations. The symbols refer to the following wavelengths: (⊗) 460 nm; (*) 480 nm; (O) 540 nm; (+) 546 nm; (X) 550 nm.

as much stoichiometric information as the $\theta^{380}$ data. The fact that the $\theta^{330}$ magnitude increases up to and beyond a ratio of 5 EthBr/1 helix for all four tetranucleotides can be accounted for by assuming that EthBr molecules which are stacked on the ends of the helices contribute to the 330-nm band, although not to the 380-nm band. Plotting the data in terms of molar circular dichroism ($\epsilon_L - \epsilon_R$) per mole of drug (Figure 4B), rather than per mole of nucleotide, shows a different trend, with increasing values up to 2 to 3 EthBr/1 helix and then a gradual decrease. We conclude that the addition of excess EthBr in the CD titrations results in the stacking of the ethidium moieties on the ends of the helical complex, as schematically illustrated in Figure 4C. Binding of ethidium to the end of the helix must be much weaker than the intercalation; otherwise, in the cases of pdC-dG-dC-dG and dC-dG-dC-dG, a sharp turnover at the 4 EthBr/1 helix ratio would be observed in the $\epsilon_L - \epsilon_R$ plot. The eventual decrease in $\epsilon_L - \epsilon_R$ for all four tetranucleotides results from the addition of (optically inactive) excess EthBr. The data for 307.5 nm show substantially the same behavior as the 330-nm data for all these systems, although the magnitudes of the observed ellipticities at 307.5 nm are larger in all cases (data not included). A small negative band in the 550-nm region was also observed in the spectra, but the magnitude of this band was even less than the 380-nm band, and thus we did not include the data in this paper. Circular dichroism spectra were also monitored in the 220–300 nm regions for titrations of pdC-dG-dC-dG, pdG-dC-dG-dC, and pdC-dC-dG-dG with ethidium bromide. These data and a brief discussion are presented in Figure 5 (in the supplementary material).

*Fluorescence Spectroscopy.* The fluorescence emission of ethidium bromide is enhanced upon intercalation into the double helix of a nucleic acid (e.g., see the review by Le Pecq, 1971; and the references cited in Reinhardt and Krugh, 1978). The ethidium bromide titrations with the tetranucleotides were monitored by fluorescence emission in order to further characterize the complex formation in these model systems. In contrast to the CD experiments, in which ethidium bromide was added to an initial solution of tetranucleotide, the fluorescence experiments were carried out by adding the tetranucleotides to a solution of ethidium bromide. The plots of $\phi_{app}/\phi$ as a function of the nucleotide/ethidium ratio are shown in Figure 6. The titration of ethidium bromide with pdC-dG-dC-dG (curve a in Figure 6) shows a linear increase in $\phi_{app}/\phi$ up to a 1 strand/1 EthBr ratio (or 2 ethidiums/helix) and then a gradual turnover. This is consistent with the results obtained for the pdC-dG-dC-dG titration using circular dichroism spectroscopy to monitor complex formation.

Somewhat different behavior is exhibited in the other three tetranucleotide titrations. The titration of ethidium bromide with pdG-dC-dG-dC shows a gradual turnover at approximately a 2 strands/1 ethidium ratio (which is the same as 1 ethidium/helix), curve b in Figure 6. The pdC-dC-dG-dG titration appears to have a "turnover" point at 2 strands/1 EthBr (i.e., 1 EthBr/1 helix). The data obtained for the EthBr plus pdG-dG-dC-dC system (curve c in Figure 6) show changes in slope at approximately one strand/1 EthBr and again at a two strands/1 EthBr ratio, i.e., at ratios of 2 EthBr/1 helix and 1 EthBr/1 helix. This suggests that two EthBr molecules can intercalate into a (pdG-dG-dC-dC)·(pdG-dG-dC-dC) helix, presumably at the (dG-dG)·(dC-dC) sites, but in the presence of an excess of tetranucleotide only one EthBr molecule intercalates per helix.

*Visible Absorption Spectroscopy.* Visible absorption titrations were carried out simultaneously with the fluorescence titrations. In addition to the change in absorbance, the change in $\lambda_{max}$ as a function of increasing nucleotide concentration was also monitored. Graphs of $\Delta A$, at five different wavelengths, are plotted vs. the strand/EthBr ratio in Figure 7. The EthBr titration with pdC-dG-dC-dG (Figure 7A) exhibits a sharp turnover in the region of 1 strand/1 EthBr, i.e., 2 EthBr/1 helix. The EthBr titration with pdG-dC-dG-dC (Figure 7B) does not exhibit a sharp turnover as observed in the pdC-dG-dC-dG titration (Figure 7A). The pdC-dC-dG-dG and the pdG-dG-dC-dC titrations with ethidium bromide exhibit a somewhat different behavior than the other two tetranucleotides (Figures 7C and 7D, respectively). The pdG-dG-dC-dC titration curve (Figure 7D) rises rapidly as tetranucleotide is added until the 1 strand/1 EthBr ratio, and

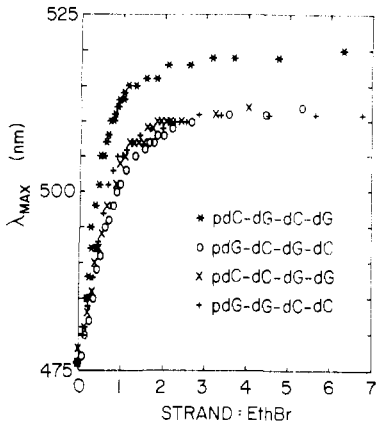



FIGURE 8: Plot of the location of the absorption maximum of ethidium bromide in the visible region, $\lambda_{max}$, as a function of added tetranucleotide during the titrations described in Figure 7.

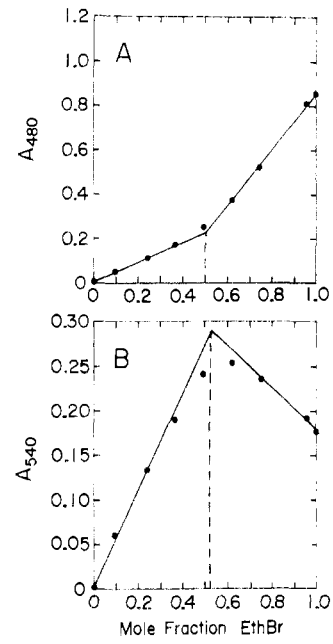



FIGURE 9: Absorbance vs. mole fraction EthBr for the continuous-variation experiment involving EthBr plus dG-dC-dG-dC-dG-dC. The concentration of EthBr plus dG-dC-dG-dC-dG-dC strand was kept at 150 $\mu$M throughout the experiment: (A) absorbance monitored at 480 nm; (B) absorbance monitored at 540 nm.

then it reaches a maximum value shortly thereafter. This strongly suggests the formation of a complex in which two ethidiums are intercalated into a (pdG-dG-dC-dC)·(pdG-dG-dC-dC) double helix.

When considering the results of the fluorescence and visible absorption titrations it is important to remember that the tetranucleotides were added to solutions of EthBr, so that at the beginning of the titrations there was a large excess of EthBr compared to the number of available binding sites. In addition to the intercalation sites, the ethidiums could also stack on the ends of helices. In previous experiments (e.g., see Reinhardt and Krugh, 1978; Krugh and Reinhardt, 1975), we have generally observed that ethidium complexes with mononucleotides or noncomplementary dinucleotides (i.e., complexes in which ethidium is stacked but not intercalated) shift the absorption maximum from 480 to ~500 nm. On the other hand, the formation of intercalated complexes is characterized by an absorption maximum in the vicinity of 510–520 nm. Consequently, we expect the stacked ethidiums to contribute more to the 460- and 480-nm absorbance changes than to the 540-, 546-, and 550-nm absorbance changes. Plots of $(\Delta A/\Delta A_{max})$ for each of the four tetranucleotides support this interpretation (data not shown). A comparison of the pdC-dG-dC-dG titration data with the pdG-dC-dG-dC titration data with ethidium bromide is informative. At the 1 strand/1 EthBr ratio, the fractional changes (in percent of $\Delta A_{max}$) are 92 (460 nm) and 87% (550 nm) for pdC-dG-dC-dG, as compared to 84 (460 nm) and 72% (550 nm) for pdG-dC-dG-dC. Thus, the fractional changes in absorbance at 550 nm lag behind those at 460 nm, especially in the pdG-dC-dG-dC titration where there is only one (dC-dG)·(dC-dG) site. The fractional changes observed for the titrations of ethidium with the sequence isomers pdG-dG-dC-dC and pdC-dC-dG-dG are more similar to each other than the other two sequence isomers discussed above. The fractional changes at 1 strand/1 EthBr are 89 (460 nm) and 84% (550 nm) for pdG-dG-dC-dC, while pdC-dC-dG-dG shows changes of 84 (460 nm) and 82% (550 nm). The pdC-dC-dG-dG tetranucleotide is unusual in that all five wavelengths show essentially the same behavior throughout the titration, which suggests that intercalation is the primary mode of binding during the titration.

*$\lambda_{max}$.* The change in $\lambda_{max}$ in the visible absorption spectrum of ethidium bromide was monitored as a function of the tetranucleotide/ethidium ratio for each visible absorption titration (Figure 8). In the pdC-dG-dC-dG titration the $\lambda_{max}$ shows sigmoidal behavior, as expected. At the 1:1 point, $\lambda_{max}$ has shifted from 476 to 515 nm, and it ultimately moves to a value between 518 and 520 nm. The values of $\lambda_{max}$ at the end of the titrations with the other three nucleotides (Figure 8) stand in contrast to the pdC-dG-dC-dG titration in that $\lambda_{max}$ shifts only to 511–512 nm. However, in one experiment (data not shown) using 23 $\mu$M EthBr and titrating to 13 strands of pdG-dC-dG-dC per EthBr, the $\lambda_{max}$ was 515 nm.

*Continuous Variation.* A further test of the relative preference of (dC-dG)·(dC-dG) sites compared to (dG-dC)·(dG-dC) sites was done using the method of continuous variation (Job, 1928) to study complex formation of dG-dC-dG-dC-dG-dC with ethidium bromide. The results for two different wavelengths are shown in Figure 9. All four wavelengths that were monitored show that the extrapolated straight lines intersect at a 0.50–0.53 mol fraction of EthBr, thus indicating the stoichiometry of complex formation to be one strand/1 drug (i.e., 2 drugs/helix). Since there are two (dC-dG)·(dC-dG) sites on this helix compared to 3 (dG-dC)·(dG-dC) sites, these experiments indicate that EthBr will form a stronger complex to a system with 2 (dC-dG)·(dC-dG) sites than to a system with 3 (dG-dC)·(dG-dC) sites. These experiments are a further confirmation that EthBr has a larger binding constant to pyr(3′–5′)pur sequences than to pur(3′–5′)pyr sequences.

*Temperature Dependence of the Ethidium Bromide plus pdC-dG-dC-dG Spectra.* In the supplementary material we have included two figures pertaining to the temperature dependence of the binding of EthBr to pdC-dG-dC-dG. Figure 10 shows CD spectra of a 2 EthBr/2 pdC-dG-dC-dG strands solution as a function of temperature. The visible absorption and fluorescence titrations of ethidium bromide with pdC-dG-dC-dG at 26 °C are shown in Figure 11. Discussion of the results is included in the figure captions.

*Stoichiometry of Complex Formation and Equilibrium Constants.* In order to analyze the titration data in a more qualitative way, one must be able to calculate the concentration of the ethidium–tetranucleotide complexes. We have chosen to use the fluorescence data in these calculations because the increase in the fluorescence intensity (under our experimental

conditions) is much larger for intercalated ethidiums than for stacked ethidiums or "outside bound" ethidiums (e.g., see Reinhardt and Krugh, 1978; Le Pecq, 1971; and the references therein). Also, an excitation wavelength of 546 nm was used in order to reflect the formation of a complex in which ethidium is intercalated, as opposed to being stacked on the end of the helices. A possible complication could arise in interpreting the fluorescence titrations, since dye–dye quenching interactions could be a problem when two ethidiums are intercalated or stacked close to each other. This would occur primarily at low strand/EthBr ratios; hence, in the region of the binding isotherm having "secondary" binding as well as intercalated EthBr, the experimentally calculated concentrations of bound EthBr may be different from the true concentrations of intercalated ethidium. The close proximity of two intercalation sites must certainly lead to a site–site interaction, but it is not easy to determine from these experiments either the sign or magnitude of $\Delta G_{\text{site-site}}$. We can conclude, however, that $\Delta G_{\text{site-site}}$ cannot be a *large positive* number for the pdC-dG-dC-dG complex formation with two ethidiums because, if it were, we would not have observed the highly cooperative binding in the titration of ethidium with this deoxytetranucleotide. We believe that these complications do not detract from the usefulness of making comparisons between model calculations and the experimental data, although the quantitative use of the calculated values (e.g., to extract thermodynamic parameters, etc.) must be done with caution.

The fraction of bound ethidium (fb) was estimated from the relative fluorescence intensity data using the equation

$$\text{fb} = \frac{(\phi_{\text{app}}/\phi) - 1}{(\phi_{\text{app}}/\phi)_{\text{max}} - 1} \tag{1}$$

where $\phi_{\text{app}}/\phi$ is the apparent quantum yield ratio, and the value of $(\phi_{\text{app}}/\phi)_{\text{max}}$ represents the relative quantum yield ratio in the presence of a large excess of nucleotide, a value which we graphically estimated from the titration data. The concentration of bound ethidium was calculated according to eq 2:

$$C_b = (C_0)(\text{fb})(V_T/V_I) \tag{2}$$

where $C_b$ is the concentration of bound EthBr; $C_0$ is the initial concentration of EthBr; fb is the EthBr fraction bound; $V_T$ is the volume of solution present at each point in the titration; and $V_I$ is the volume of solution at the initial point in the titration.

Model calculations were performed using the equilibria:

$$2[N] + [D] \overset{K_1}{\rightleftharpoons} [N_2D] \tag{3}$$

$$[N_2D] + [D] \overset{K_2}{\rightleftharpoons} [N_2D_2] \tag{4}$$

$$2[N] + 2[D] \overset{K_1K_2}{\rightleftharpoons} [N_2D_2] \tag{5}$$

where [N] is the tetranucleotide strand concentration; [D] is the free drug concentration; $[N_2D]$ is the concentration of a double helix which has one intercalated ethidium; $[N_2D_2]$ is the concentration of the double helix which has two intercalated ethidiums. Note that eq 5 is the product of equilibria 3 and 4 and thus has an equilibrium constant equal to $K_1$ times $K_2$. The concentrations of the complexes were calculated using a Fortran polynomial roots subroutine. For the case in which only $N_2D$ is formed, the program was adapted to a PDP 11/34 computer which was interfaced to a Tektronix 4051 graphics terminal and Tektronix 4662 digital plotter. For the case in which both $N_2D$ and $N_2D_2$ were present, the program was run on an IBM 360, and the results were then transferred to the Tektronix 4051. For the cases where only $N_2D$ was assumed to be formed, the concentration of the complex was calculated as a function of the nucleotide concentration for a given concentration of drug, $[D_0]$, and equilibrium constant, $K_1$. For the $N_2D_2$ model systems, the concentration of $N_2D$ was first calculated for a given set of $K_1$, $K_2$, $[D_0]$, and $[N_0]$, and then the concentration of $N_2D_2$ was determined from $K_2[D_0]$ and $[N_2D]$ (eq 4). The slopes of the curves ($\Delta[N_2D]/\Delta[N]$ and $\Delta[N_2D_2]/\Delta[N]$) were determined using spans of three points in both the theoretical and experimental curves, although it should be noted that the slopes of the theoretical curves were calculated over much smaller intervals of $\Delta[N]$ than the experimental data. We will thus expect to find only a reasonable agreement in the shape of the calculated and experimental derivative curves.

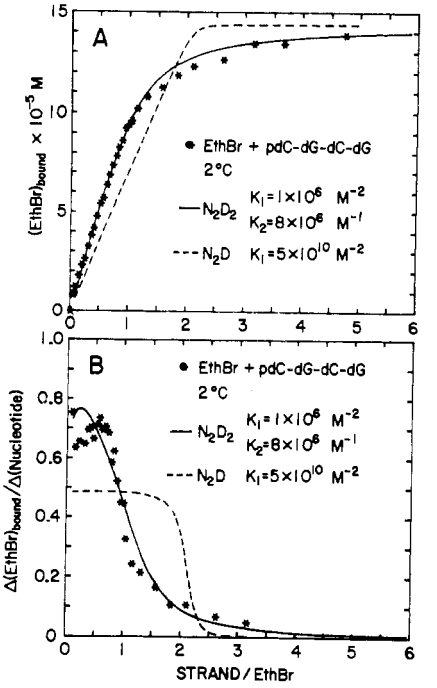



FIGURE 12: (A) Comparison of experimental (*) and calculated concentrations of bound ethidium bromide as a function of the strand/drug ratio for the pdC-dG-dC-dG titration. (B) Comparison of the experimental [(*) 3 point span] and calculated (solid and dashed lines) derivative curves for the bound ethidium bromide as a function of the strand/EthBr ratio. The solid curves in both figures were on the basis of a *bis* intercalation, while the dashed curves assume that only one ethidium intercalates per helix formed. The equilibrium constants used are given in the figures. The experimental data is from Figure 6.

The excellent agreement between the experimental and calculated concentrations of bound ethidium for the pdC-dG-dC-dG titration at 2 °C is shown in Figure 12A. The values of $K_1$ and $K_2$ listed in the figure represent our selection of the "best-fit" values in both the calculation of the bound ethidium concentration and the derivative spectra shown in Figure 12B. It should be noted that in the $N_2D_2$ model equilibrium, the individual $K_1$ and $K_2$ values are only important in determining the relative amounts of $N_2D$ and $N_2D_2$ present during the titration. The total amount of bound drug is determined by the product $K_1K_2$. The salient point of the data in Figure 12 is that the agreement between the experimental and theoretical curves unequivocally shows that a 2:1 ethidium/helix complex is formed, as schematically illustrated on the left-hand side of Figure 4C. To further document this important point, we have also included in Figure 12A a calculated curve which one would obtain if we assume that only the $N_2D$ complex is

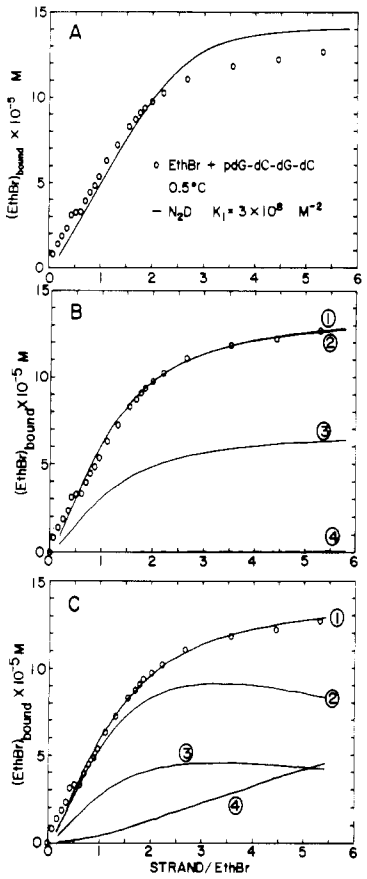



FIGURE 13: (A) Comparison of experimental (O) and calculated concentrations of EthBr for the pdG-dC-dG-dC titration assuming that only one ethidium intercalates per helix. The experimental data were calculated from the titration shown in Figure 6 using an estimated value of $(\phi_{app}/\phi)_{max}$. B and C show comparisons of experimental and calculated concentrations of ethidium bromide for the pdG-dC-dG-dC titration assuming the formation of the $N_2D_2$ complex. B shows the case for $K_1 = 1 \times 10^5$ $M^{-2}$ and $K_2 = 7 \times 10^6$ $M^{-1}$. In both B and C, curve 1 represents the total bound ethidium bromide. Curve 2 is the concentration of the ethidium in the $N_2D_2$ species, curve 3 is the concentration of the $N_2D_2$ species (curve 3 equals one-half of curve 2), and curve 4 is the concentration of the ethidium in the $N_2D$ species. C shows the case for $K_1 = 1 \times 10^7$ $M^{-2}$ and $K_2 = 7 \times 10^4$ $M^{-1}$. Note that the product $K_1K_2$ is $7 \times 10^{11}$ $M^{-3}$ in both B and c above.

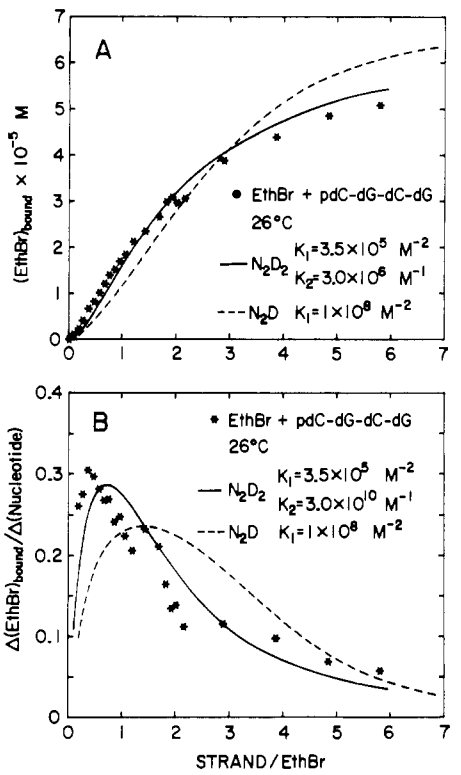



FIGURE 14: (A) Comparison of experimental (*) and calculated concentrations of bound ethidium bromide for the pdC-dG-dC-dG titration at 26 °C. The experimental data were derived from the titration shown in Figure 6. The solid curve represents the best fit for the $N_2D_2$ equilibrium, while the dashed curve represents the best fit for the $N_2D$ equilibrium with the equilibrium constants given in the figure. (B) Comparison of experimental and calculated derivative curves for the complexation of ethidium bromide with pdC-dG-dC-dG at 26 °C.

formed (eq 3). It should be obvious that no matter how large a value of $K_1$ is used it is not possible to reproduce the experimental data using only the $N_2D$ model, which further substantiates the interpretation of the pdC-dG-dC-dG + EthBr equilibrium in terms of the formation of a complex in which two ethidiums are intercalated into a miniature double helix.

The analysis of the fluorescence titration of ethidium bromide with pdG-dC-dG-dC presents quite different results than those presented above. The experimental and calculated curves are shown in Figure 13A for the $N_2D$ equilibrium, while in Figures 13B and 13C we show the results of allowing the formation of both 2:1 and 1:1 complexes (i.e., the $N_2D_2$ system). Note that the experimental data are not very well represented by the $N_2D$ equilibrium (Figure 13A) over the entire titration. The fit is much better if we allow for the formation of $N_2D_2$ complexes. The improvement in the fit in going from the $N_2D$ to the $N_2D_2$ system is anticipated simply from the introduction of one more variable to adjust ($K_2$) in order to reproduce the experimental data.

As noted previously, the total amount of drug bound in the $N_2D_2$ model depends on the product $K_1K_2$, but the individual $K_1$ and $K_2$ values determine the relative amounts of $N_2D$ and $N_2D_2$ present during the titration. An example of this is shown in Figure 13B, where $K_1 = 1 \times 10^5$ $M^{-2}$ and $K_2 = 7 \times 10^6$ $M^{-1}$, and Figure 13C, where $K_1 = 1 \times 10^7$ $M^{-2}$ and $K_2 = 7 \times 10^4$ $M^{-1}$. The product, $K_1K_2$, is $7 \times 10^{11}$ $M^{-3}$ in both cases. In the first case practically no $N_2D$ is formed throughout the titration, while in the second case the $N_2D$ predominates at ratios greater than five strands per drug.

We have also recorded the titration of ethidium bromide with pdC-dG-dC-dG at 26 °C (Figure 11, in the supplementary material). The formation of the ethidium—tetranucleotide complex is much weaker at 26 than at 2 °C, as expected. Consequently, we cannot deduce the stoichiometry of complex formation from a visual analysis of the titration data. We have compared the experimental data to the "best-fit" theoretical data for both the $N_2D$ and the $N_2D_2$ models, as shown in Figure 14A (the corresponding derivative curves are shown in Figure 14B in the supplementary material). The $N_2D_2$ model calculations reproduce the experimental data within experimental error, while the $N_2D$ model calculations exhibit a systematic deviation from the experimental data. We thus conclude that at 26 °C the predominant complex formed is one in which two ethidiums combine with two pdC-dG-dC-dG tetranucleotides to form a 2 EthBr/1 helix complex. It should be noted that at high strand to drug ratios the predominant complex may switch from a 2:1 to be a 1 EthBr/1 helix complex as a result of the excess number of intercalation sites available. Magnetic resonance experiments are required to answer this interesting question, since it reflects upon the possibility of site–site interaction.

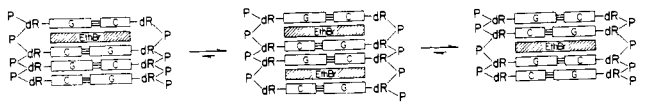


FIGURE 15: Schematic illustrations of three possible complexes in which ethidium is intercalated into the (pdG-dC-dG-dC)·(pdG-dC-dG-dC) helix: All three (and possibly other) complexes may be in equilibrium, but we conclude that under the present experimental conditions the complex on the right is the predominant complex formed.

## Discussion

All of the present data lead to the conclusion that pdC-dG-dC-dG has two strong intercalation sites, and by analogy to previous work with the deoxydinucleotides we conclude that the two ethidiums are intercalated at the two (dC-dG)·(dC-dG) sequences. On the other hand, the CD, visible, and fluorescence titrations clearly exhibit quite different behavior for the binding of ethidium to pdG-dC-dG-dC, in which there are two dG-dC sequences and only one dC-dG sequence. For example, the $\theta^{380}$ CD band (Figure 3B) levels off at approximately the 1:1 EthBr/helix ratio and remains essentially constant until more than two ethidiums have been added per helix. A decrease in $\theta^{380}$ beyond the 2 EthBr/1 helix point is probably due to the partial overlap of the positive 330-nm band with the negative 380-nm band (the crossover point moves from ~350 to ~370 nm during the titration). However, an alternative (but less likely) explanation is that the binding of the ethidiums at the end of the helices results in a change in the conformation of the miniature intercalated helix which influences the observed ellipticity of the 380-nm band. We believe that the 380-nm CD titration data of ethidium bromide with pdG-dC-dG-dC indicate that under these experimental conditions the preferred complex formed involves the intercalation of one ethidium per double helix of (pdG-dC-dG-dC)·(pdG-dC-dG-dC), presumably at the one (dC-dG)·(dC-dG) intercalation site available, as schematically illustrated in Figure 15. The relatively small magnitude of the 307- and 330-nm bands for the 2 EthBr/1 helix solutions (which suggests weak dye-dye interactions) when compared to equivalent solutions with the other deoxytetranucleotides (Figure 3A) also supports this interpretation. On the other hand, the shape of the fluorescence titration curve of ethidium with pdG-dC-dG-dC was better fit by model calculations in which there were two intercalated ethidiums per helix than one intercalated ethidium per helix. Thus, the combination of data that we have for ethidium bromide titrations with pdG-dC-dG-dC does not allow us to determine a unique stoichiometry for this system (indeed, there may not be a unique stoichiometry, since ethidium may intercalate at any sequence). It is clear, however, that the complex formed between EthBr and pdG-dC-dG-dC is either different in structure or much weaker than the complex formed between ethidium and pdC-dG-dC-dG. All of this data thus reflects that the complex formation is sequence dependent.

For the intercalation of ethidium into double-helical oligonucleotides or polynucleotides, we can separate the intercalation process into the following four steps: (1) Unwinding of the helix to create the intercalation site, which involves both a destacking of the base pairs and a change in the conformation of the sugar-phosphate backbone; (2) a change in the solvation of the nucleic acid at the intercalation site; (3) desolvation of the ethidium cation; (4) intercalation of the desolvated ethidium cation into the intercalation site created in steps 1 and 2. The thermodynamic parameters of intercalation are a sum of the $\Delta H$ and $\Delta S$ terms associated with the four hypothetical (and all inclusive, but not necessarily independent) steps outlined above. The question of preferential binding of ethidium to various intercalation sites is essentially a question of how sequence dependent are the energetics associated with the creation of the intercalation site (steps 1 and 2) and the sequence dependence of the intercalation of the desolvated ethidium into the preexisting site (step 4). A considerable body of thermodynamic data on the formation of G·C and A·U base pairs has been extracted from the study of synthetic and natural RNA oligonucleotides (e.g., Eigen and Pörschke, 1970; Pörschke, 1970; Martin et al., 1971; Uhlenbeck et al., 1971; Gralla and Crothers, 1973; Hoggett and Maass, 1971; Podder, 1972; Pörschke et al., 1973; Borer et al., 1974). However, much less detailed information is available for DNA (e.g., Pohl, 1974; Naylor and Gilham, 1966; Cassani and Bollum, 1967; Scheffler et al., 1968, 1970), but the melting of synthetic DNA polymers and oligonucleotides with simple defined sequences illustrates the effect of sequence on the stability of polynucleotides (e.g., see Wells et al., 1970; Burd et al., 1975a,b; Early et al., 1977). The important question for the present experiments is slightly different from the one addressed in the references cited above. In the present context, we are interested in determining the relative energetics associated with the creation of an appropriate intercalation site at each of the various sequences along a DNA molecule, and having created the site we are interested in determining the energetics associated with placing an ethidium (or any other planar species) into the intercalation site. There is no doubt that both of these processes will have sequence-dependent energetics: the interesting question to be answered is how different are the energies and how does nature, or how can we, exploit this phenomenon?

Both base stacking and the conformation of the sugar-phosphate backbone have been recognized as being important in determining the structure and the stability of DNA. Recent calculations and model-building studies have provided some insight into the possible sources of sequence preferences (e.g., see Alden and Arnott, 1975, 1977; Pack and Loew, 1977, 1978; Pack et al., 1977; M. Nuss and P. Kollman, personal communication; Krugh and Reinhardt, 1975; Sobell et al., 1977; Fujita and Nagata, 1976; and the references cited in these papers). However, the intercalation of the drugs involves a delicate interplay of hydrophobic, electrostatic, van der Waals, hydrogen bonding, and other forces, and thus it will be some time before a unified theory is developed. As an example of the difficulty involved, one can consider the thermodynamic data for the binding of ethidium to DNA obtained by Le Pecq and Paoletti (1967) in which they found that both $\Delta H$ and $\Delta S$ were *positive* at low ionic strengths, whereas both $\Delta H$ and $\Delta S$ were *negative* at high ionic strengths.

At the beginning of the CD experiments, the deoxytetranucleotides exist partially in the double-helical form and partially in the random-coil form: the addition of ethidium stabilizes the intercalated (helical) complex. On the other hand, in the fluorescence (and visible absorption) titrations the deoxytetranucleotides were added to an ethidium solution. During the early parts of the fluorescence titrations the tetranucleotides would be primarily in the single-stranded form if ethidium bromide were not present, and thus it is reasonable to expect the titration to be described by the equilibria given in eq 3-5. The equilibrium constants which were determined by the "best fit" of the experimental and calculated curves are not considered to be reliable enough to extract the corresponding $\Delta H$ and $\Delta S$ values. This situation will be different in future experiments with deoxyhexanucleotides where the two additional base pairs add a great deal of stability to the oligonucleotide double helix. The influence of temperature on

the stability of the complex of ethidium with pdC-dG-dC-dG is readily apparent from a comparison of either the experimental data or the equilibrium constants at 2 and 26 °C. This implies that the equilibrium given in eq 5 has a relatively large negative $\Delta H$ term. Although the values of the equilibrium constants may have a rather large uncertainty, we believe that the fit of the pdC-dG-dC-dG titration to eq 5 substantiates the experimental data which clearly shows the formation of a complex in which two ethidiums are intercalated into the (pdC-dG-dC-dG)·(pdC-dG-dC-dG) double helix [at the two (dC-dG)·(dC-dG) intercalation sites]. In the binding of ethidium to pdG-dC-dG-dC several complexes are presumably involved in the overall equilibrium (Figure 15), but we believe that the predominant complex formed in the CD experiments is the monointercalated species.

It is important to note that when ethidium intercalates into double-stranded polynucleotides it does not exhibit a requirement for any one base, nor does it exhibit any *one* sequence specificity [e.g., ethidium binds strongly to poly(dG)·poly(dC), poly(dG-dC)·poly(dG-dC), poly(dA)·poly(U), and poly(dA-dT)·poly(dA-dT)]. The polynucleotide data and our oligonucleotide data are consistent with one another, since we propose that ethidium exhibits *preferential* binding to certain sequences. The most definitive is that ethidium preferentially binds to pyr(3′-5′)pur sequences, as opposed to pur(3′-5′)pyr sequences (the present manuscript, Krugh et al., 1975; Krugh and Reinhardt, 1975; Reinhardt and Krugh, 1978). In addition, the present data indicate that the (dC-dC)·(dG-dG) sequence binds ethidium much more strongly than the (dG-dC)·(dG-dC) sequence, which also illustrates the importance of the orientation of the base pairs in the intercalation site, and leads to the ordering of the sequence preferences for ethidium binding to the deoxytetranucleotides as (dC-dG)·(dC-dG) $\gtrsim$ (dC-dC)·(dG-dG) $\gg$ (dG-dC)·(dG-dC). By inference, we believe that the same relative preferences will be observed with polynucleotides. If simple molecules such as ethidium can exhibit sequence preferences, than strategically located aromatic amino acids may play an important role in the very selective protein–nucleic acid recognition.

## Supplementary Material Available

Figures 1, 5, 10, and 11 showing elution profiles and ultraviolet and visible CD spectra (5 pages). Ordering information is given on any current masthead page.

## References


Aktipis, S., Martz, W. W., and Kindelis, A. (1975), *Biochemistry 14,* 326–331.

Alden, C. J., and Arnott, S. (1975), *Nucleic Acids Res. 2,* 1701–1717.

Alden, C. J., and Arnott, S. (1977), *Nucleic Acids Res. 4,* 3855–3861.

Allen, F. S., Gray, D. M., Roberts, G. P., and Tinoco, I., Jr. (1972), *Biopolymers 11,* 853–879.

Angerer, L. M., Georghiou, S., and Moudrianakis, E. N. (1974), *Biochemistry 13,* 1075–1083.

Balcerski, J. S., and Pysh, E. S. (1976), *Nucleic Acids Res. 3,* 2401–2409.

Bauer, S., White, M. D., and Lapidot, Y. (1975), *Nucleic Acids Res. 2,* 2355–2364.

Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974), *J. Mol. Biol. 86,* 843–853.

Bresloff, J. L., and Crothers, D. M. (1975), *J. Mol. Biol. 95,* 103–123.

Burd, J. F., Wartell, R. M., Dodgson, J. B., and Wells, R. D. (1975a), *J. Biol. Chem. 250,* 5109–5113.

Burd, J. F., Larson, J. E., and Wells, R. D. (1975b), *J. Biol. Chem. 250,* 6002–6007.

Cantor, C. R., and Tinoco, I., Jr. (1965), *J. Mol. Biol. 13,* 65–77.

Cassani, G., and Bollum, F. J. (1967), *J. Am. Chem. Soc. 89,* 4798–4799.

Cech, C. L., Hug, W., and Tinoco, I., Jr. (1976), *Biopolymers 15,* 131–152.

Davanloo, P., and Crothers, D. M. (1976), *Biochemistry 15,* 5299–5305.

Douthert, R. J., Burnett, J. P., Beasley, F. W., and Frank, B. H. (1973), *Biochemistry 12,* 214–220.

Early, T. A., Kearns, D. R., Burd, J. F., Larson, J. E., and Wells, R. D. (1977), *Biochemistry 16,* 541–551.

Eigen, M., and Pörschke, D. (1970), *J. Mol. Biol. 53,* 123–141.

Fredericq, E., and Houssier, C. (1972), *Biopolymers 11,* 2281–2308.

Fujita, H., and Nagata, C. (1976), *J. Theor. Biol. 57,* 187–202.

Gralla, J., and Crothers, D. M. (1973), *J. Mol. Biol. 78,* 301–319.

Gray, D. M., and Tinoco, I., Jr. (1970), *Biopolymers 9,* 223–244.

Hoggett, J. G., and Maass, G. (1971), *Ber. Bunsenges. Phys. Chem. 75,* 45–54.

Houssier, C., Hardy, B., and Fredericq, E. (1974), *Biopolymers 13,* 1141–1160.

Job, P. (1928), *Ann. Chim. 9,* 113.

Krugh, T. R. (1972), *Proc. Natl. Acad. Sci. 69,* 1911–1914.

Krugh, T. R., and Reinhardt, C. G. (1975), *J. Mol. Biol. 95,* 133–162.

Krugh, T. R., and Young, M. A. (1977), *Nature (London) 269,* 627–628.

Krugh, T. R. Wittlin, F. N., and Cramer, S. P. (1975), *Biopolymers 14,* 197–210.

Lee, C. H., Chang, C.-T., and Wetmur, J. G. (1973), *Biopolymers 12,* 1099–1122.

Le Pecq, J.-B. (1971), *Methods Biochem. Anal. 20,* 41–86.

Le Pecq, J.-B., and Paoletti, C. (1967), *J. Mol. Biol. 27,* 87–106.

Martin, F. H., Uhlenbeck, O. C., and Doty, P. (1971), *J. Mol. Biol. 57,* 201–215.

Naylor, R., and Gilham, P. T. (1966), *Biochemistry 5,* 2722–2728.

Pack, G. R., and Loew, G. H. (1977), *Int. J. Quantum Chem., Quantum Biol. Symp. 4,* 87–96.

Pack, G. R., and Loew, G. (1978), *Biochim. Biophys. Acta 519,* 163–172.

Pack, G. R., Muskavitch, M. A., and Loew, G. (1977), *Biochim. Biophys. Acta 478,* 9–22.

Patel, D. J., and Canuel, L. L. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 3343–3347.

Patel, D. J., and Canuel, L. L. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 2624–2628.

Podder, S. K. (1972), *Biopolymers 11,* 1395–1410.

Pörschke, D. (1971), *Biopolymers 10,* 1989–2013.

Pörschke, D., Uhlenbeck, O. C., and Martin, F. H. (1973), *Biopolymers 12,* 1313–1335.

Reinhardt, C. G., and Krugh, T. R. (1978), *Biochemistry 17* (preceding paper in this issue).

Scheffler, I. E., Elson, E. L., and Baldwin, R. L. (1968), *J. Mol. Biol. 36,* 291–304.

Scheffler, I. E., Elson, E. L., and Baldwin, R. L. (1970), *J. Mol. Biol. 48,* 145–171.

Sobell, H. M., Tsai, C.-C., Jain, S. C., and Gilbert, S. G. (1977), *J. Mol. Biol. 114,* 333-365.
Tinoco, I., Jr. (1964), *J. Am. Chem. Soc. 86,* 297-298.
Uhlenbeck, O. C. Martin, F. H., and Doty, P. (1971), *J. Mol. Biol. 57,* 217-229.
Wakelin, L. P. G., and Waring, M. J. (1974), *Mol. Pharmacol. 10,* 544-561.
Waring, M. J. (1965), *J. Mol. Biol. 13,* 269-282.
Wells, R. D., Larsen, J. E., Grant, R. C., Shortle, B. E., and Cantor, C. R. (1970), *J. Mol. Biol. 54,* 465-497.
Williams, R. D., and Seligy, V. L. (1974), *Can. J. Biochem. 52,* 281-287.

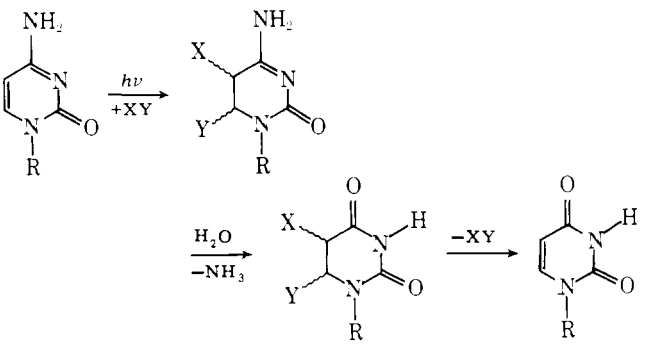
# Photochemistry of Cytosine Derivatives. 1. Photochemistry of Thymidylyl-(3′→5′)-deoxycytidine†

Fu-Tong Liu and N. C. Yang*

ABSTRACT: The photochemistry of thymidylyl-(3′→5′)-deoxycytidine (dTpdC) was studied as a model system of adjacent thymine and cytosine bases in DNA. Acetophenone-sensitized irradiation causes the cytosine moiety in dTpdC to react with the thymine moiety intramolecularly. Three unstable photoproducts are formed initially which are converted into three isomeric dinucleoside phosphates of thymine-uracil cyclobutane photodimer in a ratio of 4.2:2.2:1. Under the same irradiation condition thymidylyl-(3′→5′)-thymidine (dTpdT) yields two products in a ratio of 6:1. The structures of these products are established by chemical and spectroscopic methods. The major product in these reactions has been identified as the stereoisomer which has the same anti,anti relationship between the pyrimidine rings and the deoxyribose group as in the parent dinucleoside phosphates. The efficiency of the intramolecular dimerization of dTpdC is about one-third that of dTpdT. The results suggest that the cytosine base in DNA may be converted to a uracil base via photodimerization with an adjacent pyrimidine base, hydrolysis, and photoreactivation.

The objective of this research is to correlate the photochemistry and photophysics of nucleic acid derivatives with the photobiology of microorganisms in order to provide a molecular basis for mutagenesis induced by ultraviolet light (for a recent review of the photochemistry and photobiology of nucleic acid derivatives, see Wang, 1976). Photophysically, cytidylic acid has the lowest singlet excited state among all nucleotide units in DNA (Guéron et al., 1967; Lamola, 1973). After photoexcitation, singlet excited energy may localize at the cytosine residue in DNA. Photochemically, the reaction of cytosine and its derivatives may result in the saturation of their 5,6-double bond (Guéron et al., 1974). Consequently, the 4-amino group in the primary photoproduct is no more stabilized by the resonance existing in cytosine, and hydrolytic deamination of the primary product will lead to the formation of a dihydrouracil derivative which may undergo further chemical or photochemical transformation to give a uracil derivative (Scheme I). If such a sequence of events will occur in vivo, the formation of dU from C may lead to the miscoding of C as dU or T. Photobiologically, results from in vivo studies on T4 and S13 phages indicate that the major portion of UV-induced mutation corresponds to the C to T base transition (Drake, 1963, 1966a,b; Howard & Tessman, 1964). We have previously reported the photochemical conversion of cytosine

SCHEME I




to uracil derivatives in the presence of mercaptans including cysteine (Yang et al., 1974) and wish to explore additional pathways for such conversions.

Photodimerization of pyrimidines is the most thoroughly studied photoreaction of nucleic acid derivatives (for reviews, see Burr, 1968; Setlow, 1968; Fahr, 1969; Varghese, 1972; Fisher & Johns, 1976). The formation of pyrimidine dimers has been correlated with many of the biological effects of UV irradiation (Setlow, 1966; Meistrich, 1972; Meistrich & Drake, 1972). Thymine and its derivatives undergo photodimerization readily under a variety of experimental conditions, but the principal photochemical process of cytosine and its monomeric derivatives is the photohydration (Fisher & Johns, 1976; Yang & Liu, 1977). Although there are only a few examples of photodimerization involving cytosine reported in the literature, cytosine-containing dimers such as Thy[]Cyt and Cyt[]Cyt have been detected in UV-irradiated DNA and characterized as Thy[]Ura and Ura[]Ura, respectively (Setlow & Carrier, 1966; Setlow et al., 1965). Cytosine-containing dimers may be formed in amounts comparable to the thymine

† From the Department of Chemistry, The University of Chicago, Chicago, Illinois 60637. *Received December 7, 1977; revised manuscript received June 23, 1978.* This work was taken from the Ph.D. thesis of F.-T.L. and was presented in part at the 4th annual meeting of the American Society of Photobiology, March 1976 (Yang & Liu, 1976). We gratefully acknowledge the support of Grant CA-10220 from the National Cancer Institute of the U.S. Public Health Service. The 270-MHz NMR spectrometer used in this work was provided by grants from the National Science Foundation (GP-33116) and National Cancer Institute through the University of Chicago Cancer Research Center (CA-14599).

0006-2960/78/0417-4865$01.00/0 