Binding Strength and Specificity in DNA Interactions: The Design of A·T Specific Intercalators

W. David Wilson,* Yueh-Hwa Wang, S. Kusuma, S. Chandrasekaran, N. C. Yang,† and David W. Boykin*

Contribution from the Department of Chemistry and Laboratory for Microbial and Biochemical Sciences, Georgia State University, Atlanta, Georgia 30303-3083. *Received November 26, 1984*

Abstract: There are hundreds of compounds which bind to DNA by intercalation with G·C base pair specificity. Very few intercalators have been convincingly demonstrated to have A·T binding specificity, and one view of these results is that intercalation and A·T specificity are intrinsically incompatible. With the goals of designing compounds of high A·T binding preference and of understanding the mechanism for this preference, we have synthesized three derivatives, naphthothiophene (**1**), phenanthrene (**2**), and anthracene (**3**) derivatives, with side chains containing an hydroxy group on the α position and a cationic amino group at the β carbon. All compounds are shown to bind to DNA by intercalation based on increases in viscosity of linear DNA, unwinding of closed circular superhelical DNA, and downfield ^{31}P and upfield ^{1}H (base pair imino proton) NMR chemical shifts of DNA on titration with **1-3**. Binding strength and specificity have been evaluated by using NMR methods with calf thymus DNA and spectrophotometric titrations with three natural DNA samples (*C. perfringens*, calf thymus, *M. lysodeikticus*), poly[d(A-T)·d(A-T)], and poly[d(G-C)·d(G-C)]. Differential shifts of the A·T relative to the G·C imino proton in ^{1}H NMR experiments as well as binding constant differences establish unequivocally that **1-3** bind to DNA with significant A·T specificity. For the phenanthrene derivative **2**, the ratio of binding constants for the A·T and G·C polymers is greater than 45, indicating that this compound is the most A·T specific intercalator found to date. We propose that the mechanism for this A·T specificity involves either a direct hydrogen bond between the α-hydroxy group of the intercalators and the C-2 carbonyl of the thymine ring in the DNA minor groove or a water mediated hydrogen bonding interaction between these same two groups. The solvent specific interaction is intriguing because it can explain A·T specificity for intercalators such as tilorone which do not contain α-hydroxy groups and it can also explain the usually observed G·C specificity of intercalators as resulting from their inability to interact with the solvent lattice in the DNA minor groove.

One of the most interesting and challenging problems in research on nucleic acid interactions is the delineation of the molecular mechanisms involved in specific base pair and sequence recognition. One method of approach to this question is to design series of small probe molecules which vary in their base pair binding specificity and to investigate in detail the factors which give rise to this specificity. Since intercalation binding of small molecules with DNA has been extensively investigated at the molecular level and since intercalators, are important as drugs against several diseases,[1] we have used intercalators as our first class of base pair probe molecules. We have chosen an operational definition of an A·T specific compound as one which has an equilibrium constant for binding to A·T base pairs which is at least twice the equilibrium constant for binding to G·C base pairs (e.g., the binding constant of a compound for poly[d(A-T)·d(A-T)] is at least twice its binding constant for poly[d(G-C)·d(G-C)]). Using the factor of two eliminates from consideration compounds which might be nonspecific but which show slight preference for A·T binding due to small random or systematic errors in binding measurements.

With the goal of understanding the pronounced G·C base pair specificity of actinomycin,[2] Müller, Crothers, and co-workers[3,4] have addressed the specificity question in a systematic manner and have published an extensive study on intercalating heterocycles. All of the compounds which they investigated, which included such common intercalators as ethidium, quinacrine, and proflavine, varied from slightly to very G·C specific in their binding.[2-4] Howe-Grant and Lippard[5] analyzed the binding specificity of a series of metallointercalators and found that the intercalating compounds were G·C specific. We have found that intercalators such as naphthalene diimides and anthraquinone analogues of mitoxantrone show significant G·C binding specificity.[6] Cain, Denny, and co-workers[7] have analyzed binding specificity (using the ethidium displacement fluorescence assay) of analogues of the acridine anticancer drug amsacrine. In a study of 64 derivatives substituted on the 9-anilino ring,[7a] they have found 61 not to have statistically significant base pair specificity, two to have slight G·C specificity, and one to have slight A·T specificity ($K_{AT}/K_{GC} \sim 2$). In a companion study Denny et al.[7b] looked at the binding specificity of 22 amsacrine analogues substituted at the 3 and 5 positions of the acridine ring. Several compounds exhibited pronounced G·C specificity, but none had significant A·T specificity. Atwell et al.[7c] have prepared a related series of 23 model compounds substituted with a cationic group at the 4 position of the acridine ring. These compounds vary from having no significant specificity to being quite G·C specific. Pasternack et al.,[8] using spectral and kinetic experiments, have proposed that planar prophyrin compounds bind to DNA by intercalation only at G·C base pairs and bind to DNA in some external type complex at A·T base pairs.

Compounds which bind in the grooves of the double helix are not as well characterized as intercalators. Those which have been extensively investigated such as netropsin,[9] distamycin,[10] hydroxystilbamidine,[11] berenil,[3] and a series of similar compounds reported by Müller and Gautier[12] have all been found to be highly A·T specific. In fact, it has been observed that if proflavine is substituted with *tert*-butyl groups so as to sterically block inter-

† Department of Chemistry, Agnes Scott College, Decatur, Georgia.

(1) Wilson, W. D.; Jones, R. L. *Adv. Pharmacol. Chemother.* **1981**, *18*, 177-222.
(2) Müller, W.; Crothers, D. M. *J. Mol. Biol.* **1968**, *35*, 251-290.
(3) Müller, W.; Crothers, D. M. *Eur. J. Biochem.* **1975**, *54*, 267-277.
(4) Müller, W.; Bünemann, H.; Dattagupta, N. *Eur. J. Biochem.* **1975**, *54*, 279-291.
(5) Howe-Grant, M; Lippard, S. J. *Biochemistry* **1979**, *18*, 5762-5769.
(6) Lown, J. W., Morgan, A. R.; Yen, S.-F.; Wang, Y.-H.; Wilson, W. D. *Biochemistry*, in press.
(7) (a) Baguley, B. C.; Denny, W. A.; Atwell, G. J.; Cain, B. F. *J. Med. Chem.* **1981**, *24*, 170-171. (b) Denny, W. A.; Atwell, G. J.; Baguley, B. C. *J. Med. Chem.* **1983**, *26*, 1619-1625. (c) Atwell, G. J.; Cain, B. F.; Baguley, B. C.; Finley, G. J.; Denny, W. A. *J. Med. Chem.* **1984**, *27*, 1481-1485.
(8) (a) Pasternack, R. F.; Gibbs, E. J.; Villafranca, J. J. *Biochemistry* **1983**, *22*, 2406-2414. (b) Pasternack, R. F.; Gibbs, E. J.; Villafranca, J. J. *Biochemistry* **1983**, *22*, 5409-5417.
(9) Wartell, R. M.; Larson, J. E.; Wells, R. D. *J. Biol. Chem.* **1974**, 6719-6731.
(10) Luck, G.; Triebel, H.; Waring, M.; Zimmer, C. *Nucleic* Acids *Res.* **1974**, *1*, 503-530.
(11) Festy, B.; Strum, J.; Daune, M. *Biochim. Biophys. Acta* **1975**, *407*, 24-42.
(12) Müller, W.; Gautier, F. *Eur. J. Biochem.* **1975**, *54*, 385-394.

0002-7863/85/1507-4989$01.50/0

Table I. Spectrophotometric Binding Results[a]

compd	poly[d(A-T)·d(A-T)] $K \times 10^{-3}$	poly[d(A-T)·d(A-T)] n	*C. perfringens* $K \times 10^{-3}$	*C. perfringens* n	calf thymus $K \times 10^{-3}$	calf thymus n	*M. lysodeikticus* $K \times 10^{-3}$	*M. lysodeikticus* n	poly[d(G-C)·d(G-C)] $K \times 10^{-3}$	poly[d(G-C)·d(G-C)] n	K_{AT}/K_{GC}[b]
1	34.7	2.6	17.6	3.4	21.4	2.9	11.3	3.0	8.2	2.8	4.2
2	1050	2.5	370	3.1	129	2.6	63.2	3.1	23.0	2.3	46
3	216	2.7	57.9	3.1	57.2	2.8	34.6	2.8	34.6	2.6	6.4

[a] Best fit K and n values, using eq 1, are given for binding of **1–3** to the DNA samples indicated. [b] The ratio of binding constants for poly[d(A-T)·d(A-T)] and poly[d(G-C)·d(G-C)].

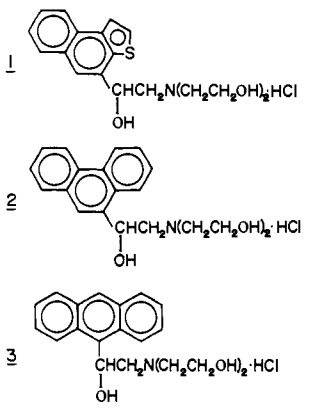


Figure 1. Structures for compounds **1–3**.

calation, it is converted to an outside binding compound and its binding specificity changes from a G·C to an A·T preference.[13] These findings led Müller and Guatier[12] to hypothesize that "a specific interaction with A·T base pairs might be basically different in type from the intercalation reaction". While the number of compounds in each category (A·T specific outside binders and G·C specific intercalators) makes this a quite compelling hypothesis, studies with a series of naphthothiophene derivatives[14,15] and with tilorone[16,17] have suggested that A·T specificity and intercalation binding are compatible.

The original naphthothiophene derivatives with di-*n*-butylamino side chains that were synthesized by Boykin and co-workers as anti-malarials[18,19] were quite difficult to study because of their low water solubility.[14,15] As part of a long-range study to investigate group free energy contributions in DNA interactions and mechanisms of specific base pair and sequence recognition, we have synthesized a naphthothiophene derivative with a diethanol amino side chain which has improved water solubility (Figure 1). In an attempt to design and synthesize intercalators with very high A·T binding specificity, we have also synthesized related phenanthrene and anthracene derivatives (Figure 1). We show here, for the first time, that intercalators can have very high specificity for A·T base pairs ($K_{AT}/K_{GC} > 45$) and propose a mechanism to account for this specific recognition.

Results

Structures of the three compounds used in these experiments are shown in Figure 1. CPK model building studies with these compounds indicated that steric interactions were minimized with the α-hydroxy groups on one side of the planar aromatic ring system and the alkyl amino on the other side of the ring system. Free side to side torsional motions are possible in the phenanthrene and naphthothiophene derivatives; but there is a significant barrier to rotation of the hydroxy and amino groups from one side of the ring to the other, and the barrier appears larger with the phenanthrene. The anthracene derivative is much more sterically restricted, and, in fact, slight bond angle and length distortions are required to build its model. The α-hydroxy and amino groups are locked essentially in positions perpendicular to the aromatic anthracene ring due to steric crowding by the two peri protons adjacent to the substituent. The barrier to rotation of these groups to opposite sides of the anthracene ring is apparently quite high.

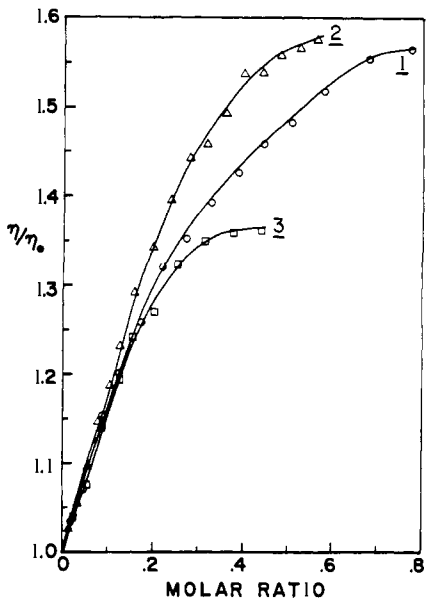


Figure 2. Viscometric titrations of sonicated calf thymus DNA with **1–3**. The reduced specific viscosity ratio (η/η_0) is plotted as a function of the molar ratio of intercalator to DNA base pairs. The titrations were conducted at 28 °C in PIPES buffer.

Viscometric Titrations. The viscosity of sonicated calf thymus DNA is increased by **1–3** as can be seen in Figure 2. The increases in viscosity produced by **1** and **2** are quite similar while the increase on titration of DNA by **3** is smaller.

We have also analyzed the interaction of **1–3** with closed circular supercoiled DNA, and the results are shown in Figure 3. All of the compounds cause the increase, maximum, and following decrease in a viscometric titration expected for an intercalating compound as it first removes the DNA supercoils to an open circle and then causes reverse supercoiling of the DNA.[1,20] Compounds **2** and **3** peak at approximately the same molar ratio of intercalator to DNA base pairs in Figure 3 while **1**, which binds more weakly to DNA (results presented later), requires higher concentrations of compound to unwind the DNA. Using the binding constants from Table I (discussed later) to correct the titrations for the actual amount of bound intercalator[20] indicates that each compound is bound at a ratio of 0.10–0.11 mol per mol of DNA base pairs at the maxima in the titrations in Figure 3. Or, in other words, the compounds have very similar intercalation unwinding angles.[1,20]

NMR Experiments. In addition to hydrodynamic experiments,

(13) Müller, W.; Crothers, D. M.; Waring, M. *Eur. J. Biochem.* **1973**, *39*, 223–234.
(14) Panter, J. W.; Boykin, D. W.; Wilson, W. D. *J. Med. Chem.* **1973**, *16*, 1366–1369.
(15) Olmstead, E. J.; Panter, J. W.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **1975**, *14*, 521–526.
(16) Strum, J.; Schreiber, L.; Daune, M. *Biopolymers* **1981**, *20*, 765–785.
(17) Strum, J. *Biopolymers* **1982**, *21*, 1189–1206.
(18) Das, B. P.; Campbell, J. A.; Samples, F. B.; Wallace, R. A.; Whisenant, L. K.; Woodard, R. W.; Boykin, D. W. *J. Med. Chem.* **1972**, *15*, 370–374.
(19) Das, B. P.; Nuss, M. E.; Boykin, D. W. *J. Med. Chem.* **1974**, *17*, 516–519.
(20) Jones, R. L.; Lanier, A. C.; Keel, R. A.; Wilson, W. D. *Nucleic Acids Res.* **1980**, *8*, 1613–1624.

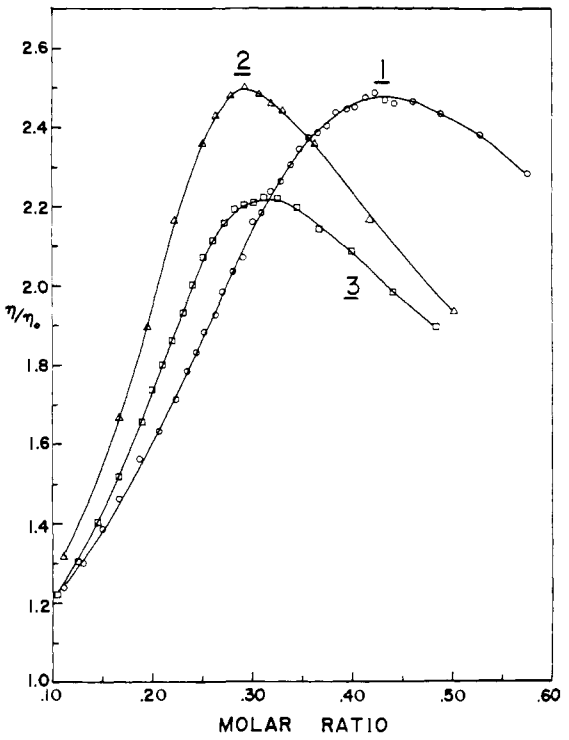


Figure 3. Viscometric titrations of closed circular supercoiled DNA with **1–3**. The plot and experimental conditions are the same as in Figure 2.

NMR spectral shifts, particularly ^{31}P and ^{1}H imino proton results, are becoming specific diagnostic tests for binding by intercalation.[21-27] We have shown that the unwinding of the DNA backbone in the intercalation reaction is related to a downfield shift in the DNA ^{31}P NMR signal.[22] This is expected for a g,g to g,t conformational transition on intercalation in agreement with the shift predictions of Gorenstein and co-workers.[24] DNA ^{31}P NMR spectra as a function of added ratio of **2** are shown in Figure 4A. The DNA spectrum broadens and shifts downfield as an exchange averaged peak on titration with **2** as has been seen with other intercalators.[21,22] The chemical shift varies from −4.3 for pure DNA to approximately −3.8 for DNA saturated with **2**. This is less than the chemical shift change obtained with ethidium, more than for daunorubicin, and similar to the change obtained for quinacrine. Compounds **1–3** give very similar line broadening and limiting shifts in the DNA ^{31}P NMR spectrum as shown in Figure 4B. The limiting shifts for **2** and **3** are the same while the limiting shift for **1** is approximately 0.1 ppm less than for **2** and **3**. This suggests that the unwinding of DNA on intercalation of **1–3** is essentially the same. These results agree quite well with the unwinding results predicted from the viscometric titrations of superhelical DNA shown in Figure 3.

Imino proton spectra of DNA have peaks near 13.7 and 12.7 ppm at 30 °C which are characteristic of A·T and G·C base pair imino protons, respectively. It has been found that intercalators typically cause ring current induced upfield shifts of these protons.[25-27] The A·T specific outside binding compound, netropsin, causes downfield, not upfield, shifts in the A·T imino band.[27] Similar proton spectra for DNA and its complexes with **1–3** at several ratios are shown in Figure 5A–C. As can be seen from spectra in Figure 5B at ratios of **2** to DNA base pairs from 0 to 0.28, the A·T resonance shifts upfield considerably more than the G·C resonance. In fact, at the 0.28 ratio the G·C resonance is visible only as a shoulder on the A·T peak. As saturation of the A·T sites is approached, the G·C peak begins to shift upfield and when the neighbor exclusion binding limit is reached, it has shifted equally as far as the A·T peak. The limiting shifts at high ratio of **2** to base pairs are 12.5 and 11.5 ppm for the A·T and G·C exchange averaged peaks, respectively. This behavior indicates that **2** intercalates at both A·T and G·C base pairs, but has a strong A·T base pair binding preference. Increasing the temperature of the sample at a saturating ratio (bottom of Figure 5B) at 60 °C causes no significant changes in the spectrum. A similar titration of **1** is shown in Figure 5A. Once again the shifts of the A·T peak are greater than for the G·C peak up to a ratio of approximately 0.28, but the difference is not as great as seen with **2** in Figure 5B. From the 0.28 ratio to saturation of DNA binding sites with **1**, the G·C resonance gives larger shifts. The limiting shifts are 12.5 and 11.6 ppm for the A·T and G·C base pairs, respectively. Heating the sample at the 0.52 ratio to 50 °C results in slight downfield shifts for both A·T and G·C exchange averaged peaks, suggesting that some dissociation may be occurring as the temperature is increased. These results indicate that **1** can intercalate at both A·T and G·C base pairs with some A·T base pair binding specificity. The specificity does not appear to be as great as for **2**, and the temperature results qualitatively suggest that the binding of **1** to DNA is weaker than for **2** binding to DNA. The results for a similar titration of DNA by **3** are shown in Figure 5C. The relative shifts of A·T and G·C base pairs during the titration are intermediate between those of **1** and **2**. The A·T resonance shifts upfield in the titration and by the 0.28 ratio the G·C peak is only a shoulder on the A·T peak. During the remainder of the titration, more G·C sites are available and binding of **3** results in an upfield shift of the G·C resonance. The limiting shifts are 12.5 and 11.5 ppm for the A·T and G·C peaks, respectively. Heating the saturation complex of **3** with DNA from30 to 60 °C results in a sharpening but no significant shift of the imino proton signals. This suggests that like **2**, **3** binds more strongly to DNA than **1** and has A·T base pair binding specificity greater than **1** but less than **2**.

Spectrophotometric Titrations. Spectral shifts for the addition of DNA to **3** are shown, for example, in Figure 6. A decrease in the extinction coefficient, shifts of the wavelength of maximum absorbance to the lower energy range, and isosbestic behavior are observed. Similar results are obtained with **1** and **2**. Similar spectral shifts and values for bound extinction coefficients are obtained for **1–3** on titrating with DNA samples of widely differing base pair composition. When these spectral changes are used for the determination of binding constants, the broadest range in a binding isotherm can be covered by placing DNA in a cuvette and titrating with the compound of interest with use of predetermined free and bound extinction coefficients at an optimum wavelength (Experimental Section).[28] With use of this method, Scatchard binding isotherms were generated for the binding of **1–3** to natural DNA samples and synthetic DNA polymers. Example plots for the binding of **3** to several DNA samples are shown in Figure 7. The points in the figure are experimental and the solid lines are determined by a nonlinear least-squares computer program using the site exclusion equation of McGhee and von Hippel:[29]

$$\nu/c = K[1 - n\nu][(1 - n\nu)/(1 - (n - 1)\nu)]^{n-1} \qquad (1)$$

where ν is the moles of compound bound per DNA base pair, c is the free compound concentration, K is the binding equilibrium

(21) Wilson, W. D.; Keel, R. A.; Mariam, Y. H. *J. Am. Chem. Soc.* **1981**, *103*, 6267–6269.
(22) Wilson, W. D.; Jones, R. L. *Nucleic Acids Res.* **1982**, *10*, 1399–1410.
(23) Levy, G. C.; Ejchart, A.; Marchetti, P. S. *J. Magn. Reson.* **1984**, *57*, 130–135.
(24) (a) Goldfield, E. M.; Luxon, B. A.; Broie, V.; Gorenstein, D. G. *Biochemistry* **1983**, *22*, 3336–3344. (b) Gorenstein, *Annu. Rev. Biophys. Bioeng.* **1981**, *10*, 355–386.
(25) Feigon, J.; Denny, W. A.; Leupin, W.; Kearns, D. R. *J. Med. Chem.* **1984**, *27*, 450–465.
(26) (a) Chandrasekaran, S.; Krishnamoorthy, C. R.; Jones, R. L.; Smith, J. C.; Wilson, W. D. *Biochem. Biophys. Res. Commun.* **1984**, *122*, 804–809. (b) Chandrasekaran, S.; Jones, R. L.; Wilson, W. D. *Biopolymers*, in press.
(27) Patel, D. J.; Kozlowski, S. A.; Rice, J. A.; Broka, C.; Itakura, K. *Proc. Nat. Acad. Sci. U.S.A.* **1981**, *78*, 7281–7284.
(28) (a) Wilson, W. D.; Lopp, I. G. *Biopolymers* **1979**, *18*, 3025–3041. (b) Wilson, W. D.; Krishnamoorthy, C. R.; Wang, Y. H.; Smith, J. C., *Biopolymers*, in press.
(29) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469–489.

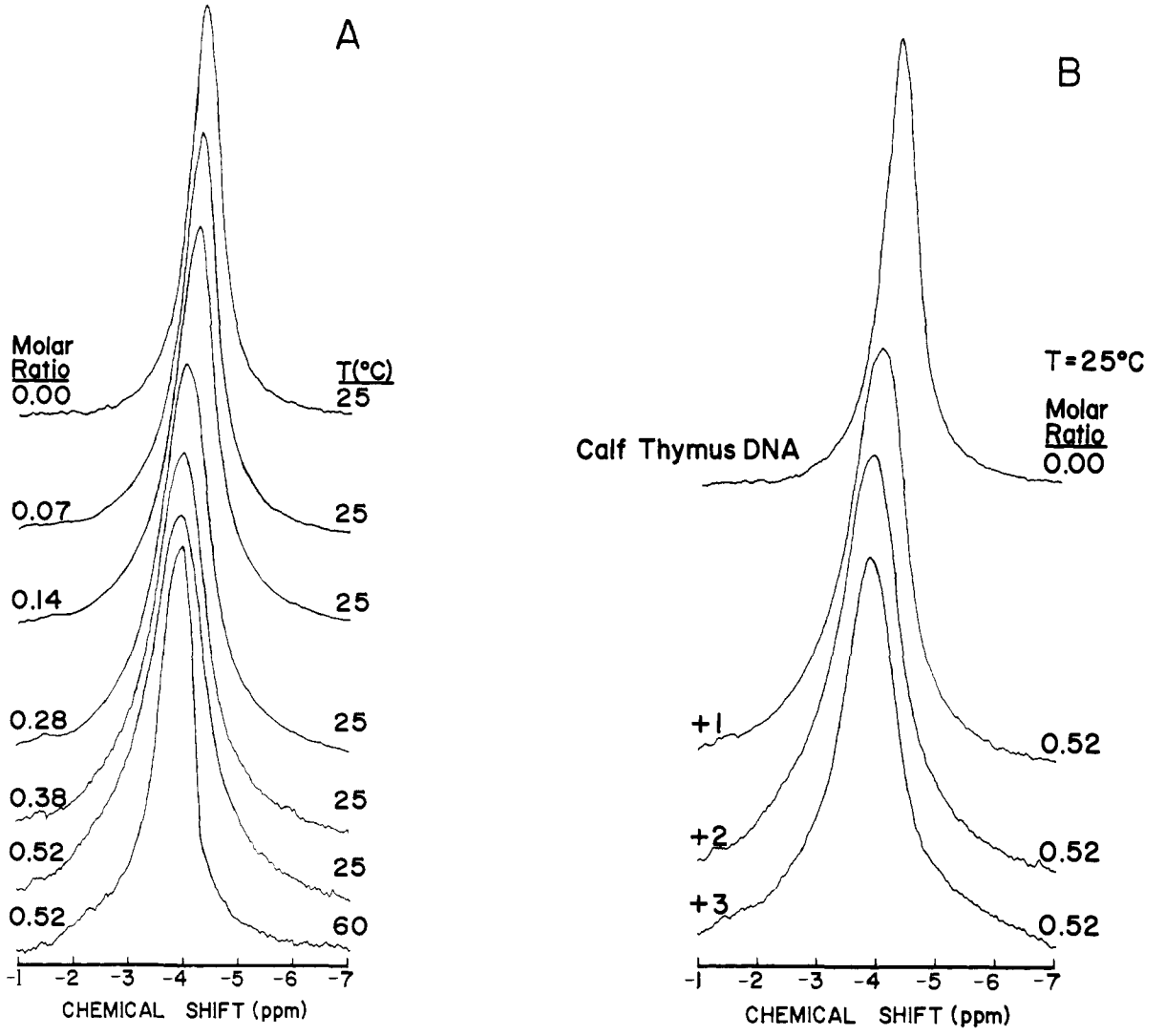


Figure 4. (A) ^{31}P NMR spectra of intercalation complexes of **2** with calf thymus DNA as a function of ratio (**2** to base pairs) and temperature. The samples (1.5 mL in volume) in 10-mm tubes were 20 mM in DNA phosphate molarity in 9% H_2O–PIPES and contained trimethyl phosphate as internal reference. The number of scans was 3000 (2 h of total accumulation time) for all the spectra. (B) ^{31}P NMR spectra of intercalation complexes of **1**, **2**, and **3** with calf thymus DNA (molar ratio = 0.52) at 25 °C. The spectral conditions were similar to those for Figure 4A.

constant, and n is the number of base pairs per binding site.[28] K and n values for all compounds and DNA samples, determined in this manner, are collected in Table I. For all compounds the highest binding constants are seen with A·T polymer and the lowest binding constants with the G·C polymer. The high G·C percentage *M. lysodiekticus* DNA has the next highest binding constant above the G·C polymer. The binding constants for **1** and **3** are approximately the same with *C. perfringens* and calf thymus DNA samples. The binding constant of **2** with *C. perfringens* DNA is significantly higher than its binding constant with calf thymus DNA. The preference of **2** for A·T base pairs is much greater than the preference for **1** and **3** as can be seen from the K_{AT}/K_{GC} ratios in Table I. Compound **1** binds more weakly to all DNA samples than **2** and **3**. Compound **2** binds most strongly to all DNA samples except poly[d(G–C)·d(G–C)] which has a higher binding constant for **3**. The binding constants for **1**–**3** are not directly proportional to the A·T base pair percentage, suggesting that in addition to a general A·T preference there may also be some sequence selectivity in binding. The number of base pairs per binding site, in general, is in the range from 2.6 to 3.1 with *C. perfringens* DNA showing the largest number of base pairs per binding site.

The isosbestic behavior in DNA titrations with **1**–**3** indicates that secondary binding modes with DNA and self association of the compounds through stacking of their aromatic rings are not significant under these conditions. We obtained Beer's law behavior for the three compounds over the concentration ranges used in the binding experiments. Spectral analysis of **3** at higher concentration indicated that slight spectral shifts and loss of Beer's law behavior occurred above the 10^{-4} M concentration range. With use of the spectral changes of **3** at high concentration, a dimerization constant of 400 M^{-1} was estimated[30] in the same buffer system used for the DNA binding determination. This dimerization constant is far too low to have any effect on the binding measurements which are generally conducted in long cuvettes at low concentrations (e.g., 10^{-5} M and below) of the intercalators.

Table II. The Effect of **1**–**3** on DNA Thermal Melting[a]

		Tm (°C)		
compd	poly[d(A-T)·d(A-T)]	C. perfringens	calf thymus	M. Lysodeikticus
none	47.2	61.4	66.0	82.5
1	52.5	63.5	68.5	82.7
2	56.5	68.5	70.0	83.4
3	53.8	63.6	69.1	83.2

[a] All experiments were conducted in PIPES buffer at a ratio of 0.4 mol of intercalator per mol of DNA base pairs.

Thermal Denaturation. The effects of **1**–**3** on the midpoint of the thermal denaturaion curve (Tm) for poly[d(A–T)·d(A–T)]

(30) Gough, A. N.; Jones, R. L.; Wilson, W. D. *J. Med. Chem.* **1979**, *22*, 1551–1554.

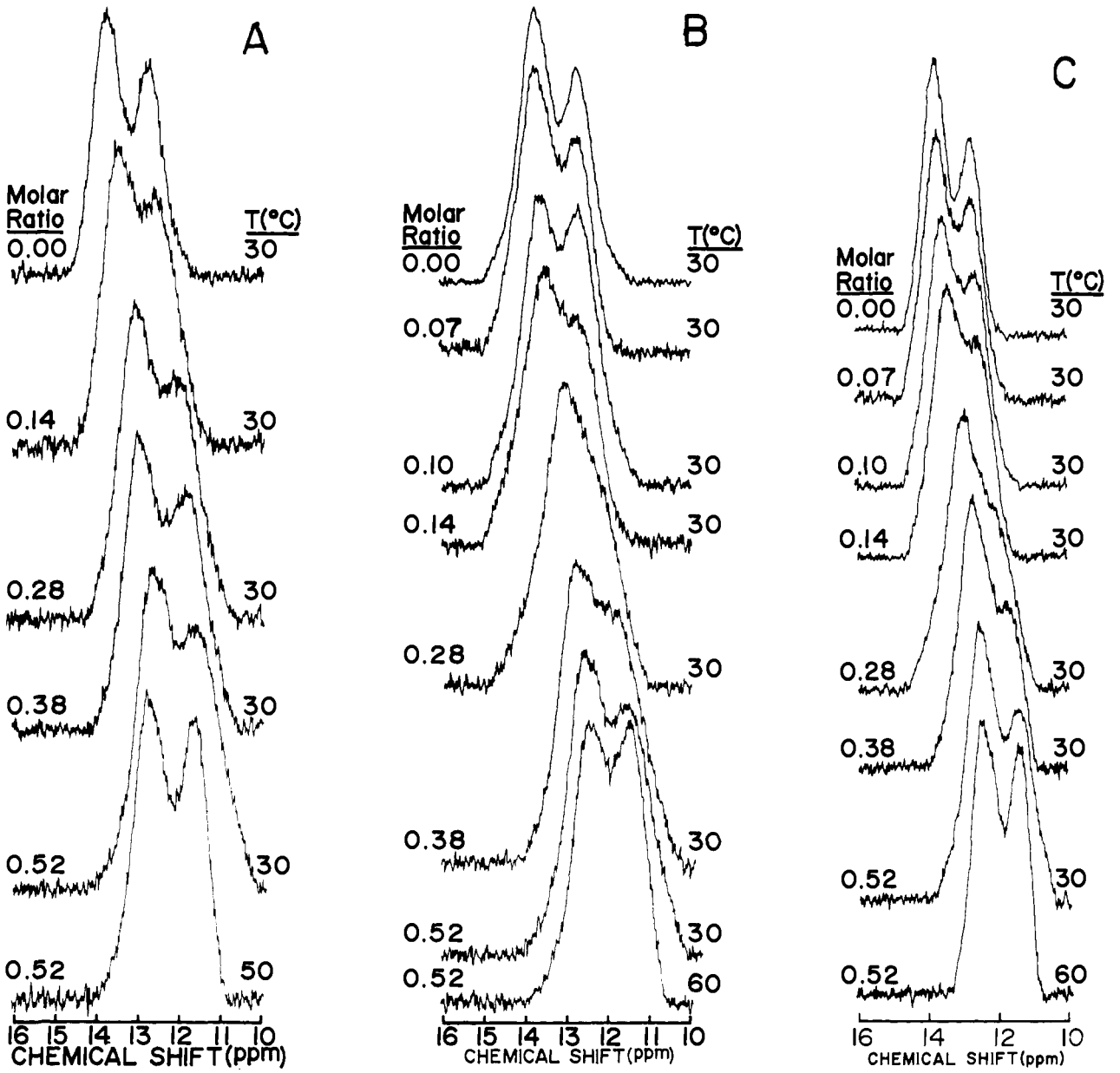


Figure 5. Imino proton magnetic resonance spectra of intercalation complexes of (A) **1**, (B) **2**, and (C) **3** with calf thymus DNA as a function of ratio (**1**, **2**, or **3** to base pairs) and temperature. The samples (0.825 mL in volume) in 5-mm NMR tubes were 20 mM in DNA phosphate molarity in 9% D_2O in H_2O–PIPES buffer and contained TSP as reference. The number of scans was typically 15000 (2 h of total accumulation time) for all the spectra.

and natural DNA samples of varying base pair percentage were also determined. Poly[d(G–C)·d(G–C)] melts too high under these conditions to be analyzed by this method. The Tm values for the four DNA samples in the presence of **1–3** at a ratio of 0.4 mol of intercalator per base pairs are listed in Table II. Less than a one degree increase in Tm is found for the three compounds with the high G·C content *M. lysodeidikus* DNA. Larger increases are obtained with calf thymus DNA with the compounds increasing the Tm in the order **2** > **3** > **1**. With the higher A·T content *C. perfringens* DNA, **2** gives a larger Tm increase than with calf thymus DNA. The increases with **1** and **3** are similar for these two DNA samples. All three compounds give the maximum increase in Tm with poly[d(A–T)·d(A–T)] and the compounds increase the Tm in the order **2** > **3** > **1**. All of these results are in excellent agreement with the results obtained from spectrophotometric binding experiments. As with the binding experiments, the lack of a direct linear correlation between the Tm increase obtained for the compounds and the A·T percentage for the DNA samples suggests that there is a general A·T preference and some sequence specificity in binding.

Discussion

A question of critical importance for understanding DNA interactions which is addressed by this work is whether intercalation binding modes can exhibit significant A·T base pair binding specificity. The answer is a very definite yes. Our earlier, tentative proposal[14,15] that naphthothiophene derivatives intercalate and can exhibit pronounced A·T specificity, although hampered in a quantitative sense by the low solubility of those initial compounds, is correct. The results from sonicated DNA viscosity increases (Figure 2), unwinding of supercoiled DNA (Figure 3), downfield shifts in ^{31}P NMR spectra (Figure 4), and upfield shifts for the base pair imino protons (Figure 5) establish unequivocally that **1–3** bind to DNA by intercalation. In addition, the differential shifts of A·T and G·C imino protons (Figure 5) and the larger binding constants (Table I) and Tm increases (Table II) for DNA samples with high A·T content establish unequivocally that these compounds show significant A·T base pair binding specificity.

These results then lead to the intriguing question of what is the mechanism of this observed A·T binding preference. One possibility is that the naphthothiophene compound **1** exhibits A·T

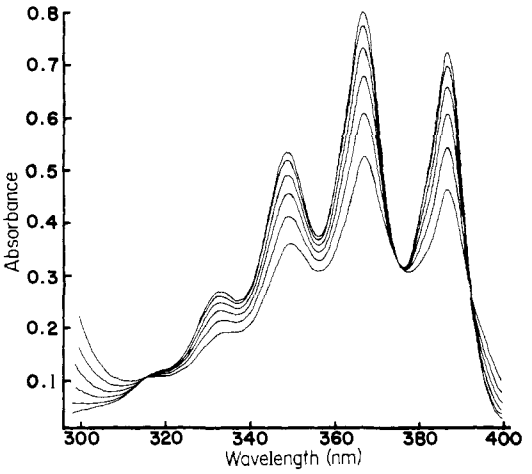


Figure 6. Spectra for **3** in the presence of varying amounts of calf thymus DNA. Scans were taken on a Cary 219 spectrophotometer at 27 °C in 10-cm cells in PIPES buffer. The concentration of **3** in each scan was 1×10^{-5} M. From top to bottom the DNA concentrations (base pair molarity) were 0, 2.21×10^{-6}, 5.90×10^{-6}, 1.11×10^{-5}, 1.85×10^{-5}, 2.96×10^{-5} M.

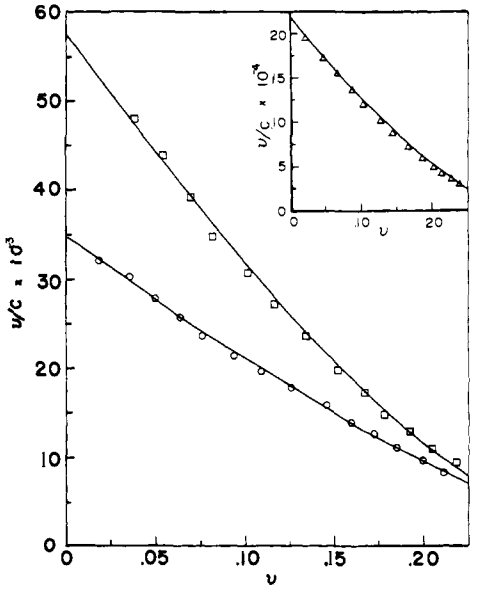


Figure 7. Scatchard plots for the binding of **3** to poly[d(A–T)·d(A–T)], △; calf thymus DNA, □; and poly[d(G–C)·d(G–C)], ○. Several titrations were conducted for each DNA sample at concentrations of DNA through the range 1.5 to 3.5×10^{-5} M base pairs and the concentration of **3** varied in the titrations. All titrations were used to generate the solid lines in the figure which were determined by using eq 1 and a computer nonlinear least-squares fit to the data points.

specificity due to a hydrogen bond between the α-hydroxy group and the thymine carbonyl oxygen at position 2 of the pyrimidine ring.[14,15] CPK model building studies with this compound suggest that a linear hydrogen bond of reasonable bond length can be formed between these two oxygen atoms with the cationic nitrogen located quite near the anionic phosphate groups of the DNA double helix backbone. The naphthothiophene ring is then located between adjacent base pairs so that stacking with the base pairs can be obtained.

We observed from these model building studies that the sulfur atom in the thiophene ring was shielded from the solvent (dehydrated) and, due to its size, caused some distortion of stacking interactions of **1** with the base pairs of DNA on either side of the intercalation site. Because of steric effects with the thiophene ring, it seemed that the optimum orientation of the α-hydroxy and cationic amino groups with respect to interaction with DNA might not be obtained. Model building suggested that replacing the naphthothiophene ring with a phenanthrene ring could overcome some of these problems while maintaining A·T specificity, and we, therefore, synthesized **2**. As can be seen from the figures and from Tables I and II, the results are indeed striking. Phenanthrene **2** binds to poly[d(G–C)·d(G–C)] approximately three times better than **1**, but it binds to poly[d(A–T)·d(A–T)] 30 times better than **1**. The difference in binding with the G·C polymer is probably due to a less optimized stacking with the base pairs of the naphthothiophene relative to the phenanthrene ring system. Considering the A·T polymer, however, the binding constant of **2** increases by approximately a factor of 10 more than the binding constant of **1**. We feel that this is due to the fact that the bulky sulfur atom in **1** near the α-hydroxy and amino groups prevents and optimum orientation of the hydrogen bond to the thymine carbonyl and interaction of the amino group with the DNA phosphate groups. In **2** both of these interactions as well as ring stacking can be optimized.

The anthracene compound **3** was of interest for several reasons. First, the aromaticity and polarizability are quite different from **1** and **2** as demonstrated by the fact that its absorption maximum wavelength is approximately 80 to 90 nm longer than those for **1** and **2**, respectively. Second, as stated earlier, the substituent group in **3** is also much more rigidly held than in **1** and **2**. Third, it has also been proposed that the anthracene ring system should have particularly good stacking interactions with the DNA base pairs.[31] The viscosity increase with sonicated DNA on titration with **3** is less than for **1** and **2**. This could be due to a slight bending of DNA on intercalation of the substituted anthracene. The substituents of **1** and **2** are freer to rotate so as not to sterically inhibit interaction of the ring systems with the DNA base pairs, but in their more rigid configuration in **3**, they could interact with the base pairs and cause some bending as has been seen with ethidium bromide[32] and asymmetricallyl substituted phenanthrolines.[33] It is interesting to note, however, that **3** has the highest binding constant with poly[d(G–C)·d(G–C)], suggesting that even if some bending is occurring the anthracene ring still has a quite favorable interaction with the DNA base pairs. CPK model building studies, however, indicate that the linear anthracene ring system prevents the optimum stacking with base pairs, thymine carbonyl hydrogen bond, and amino cation interaction with the DNA phosphate groups seen with the phenanthrene derivative **2**. The anthracene compound then has a lower A·T preference than **2** but a greater preference and generally stronger binding than **1** (Tables I and II). Unlike G·C binding specificity, polarizability seems to play no significant role in A·T binding specificity.

Another explanation for the A·T specificity of **1–3**, still involving hydrogen bonding, is that there is a water mediated hydrogen bond between the thymine C-2 carbonyl oxygen and the α-hydroxy oxygen of **1–3**. There are two favorable aspects to this proposal. One is that a similar mechanism, with a bridging hydrogen bonded water molecule between the thymine carbonyl and the intercalator carbonyl group, could also account for the A·T specificity of tilorone. The second is that the single crystal X-ray studies of Dickerson and co-workers[34] have suggested that there is a lattice of water molecules hydrogen bonded to A·T base pairs in the minor groove of A·T rich regions of DNA. This water lattice is disrupted by G·C base pairs. Perhaps intercalators, in general, also disrupt this water lattice in A·T regions and this may contribute to the usually observed G·C binding preference of intercalators. With **1–3** and tilorone, on the other hand, the water lattice may be enhanced leading to their A·T binding preference. A compound which has no significant base pair specificity may maintain the water lattice but not enhance it, so that no significant free energy difference is seen between binding to A·T and G·C base pairs.

The question of enantiomeric selectivity in the binding of **1–3** with DNA has not been directly addressed. Visible spectral

(31) Miller, K. J.; Newlin, D. D. *Biopolymers* **1982**, *21*, 633–652.
(32) Jain, S. C.; Sobell, H. M. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1161–1177 and 1179–1194.
(33) Gabbay, E. J.; Scofield, R. E.; Baxter, C. S. *J. Am. Chem. Soc.* **1973**, *95*, 7850–7857.
(34) Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535–556.

changes and NMR chemical shifts indicate (work in progress) that both enantiomers must bind to DNA. Our binding studies with homopolymers (Table I) do not detect any significant biphasic binding which suggests that both enantiomers must have approximately the same binding constant with DNA. Model building studies also indicate that both enantiomers could form a hydrogen bond in the minor groove with the thymine C-2 carbonyl oxygen or with an associated water molecule. Preliminary studies with achiral carboxamide and ester analogues of **1-3** indicate binding interactions with DNA similar to **1-3**, and this also suggests that enantiomeric recognition is not an important factor in the binding of the hydroxy compounds with DNA.

Experimental Section

Materials. DNA samples were sonicated, filtered, phenol and ether extracted, ethanol precipitated, dialyzed into PIPES buffer (0.01 M PIPES, 10^{-3} M EDTA, pH 7.0), and characterized as previously described.[26] Samples for binding and viscosity studies were sonicated for shorter periods and had an average length of 500-600 base pairs. Samples for NMR experiments were sonicated for longer times and had an average length of 150-200 base pairs. All sonications were done with application of pulse power at near 0 °C in PIPES buffer with 0.5 M NaCl added. Closed circular supercoiled Col E_1 DNA was prepared as previously described.[20] The compounds **1-3** were synthesized with use of methods previously worked out;[18,19,35] the details of their synthesis will be published elsewhile.[36]

Spectrophotometric Studies. Absorbance measurements in the UV-visible region were made on a Cary 219 spectrophotometer interfaced to an Apple IIe microcomputer through a bidirectional digital communications port. Cell holders were thermostated with use of a Haake circulating water bath. Wavelength scans and extinction coefficient measurements were made in cells from 1 to 10 cm pathlength at the wavelength range appropriate for the compound being investigated. Extinction coefficients of compounds bound to DNA were determined at the same wavelength as the extinction coefficient measurements of the free compound, but a larger molar excess of DNA was present ([DNA-P]/[compound] > 100). The Cary 219 was also used in spectrophotometric binding studies. To remove some of the random error, for each absorbance measurement in the absence or presence of DNA, the microcomputer calculated the average of 100 acquired absorbance readings at the preselected wavelength for the compound under study. These averaged absorbance values were converted by the microcomputer to ν (moles of compound bound/mole of DNA base pairs) and free ligand concentrations using the free and bound extinction coefficients for the compound. At the end of a titration, the computer plotted the digitized data which was in the fraction bound range 0.2 to 0.8. Any binding results outside of this range are subject to large systematic errors as a result of experimental errors in extinction coefficients. The computer then calculated nonlinear least-squares best fit K and n values from the site exclusion method of McGhee and von Hippel[29] as defined in eq 1.

Thermal Melting. Denaturation experiments were also conducted on the Cary 219 spectrophotometer. The data were directly plotted on the Cary 219 chart paper as absorbance vs. temperature. The data were also sent to the Apple IIe computer, absorbance averaged as described above, and stored on disk for more accurate Tm analysis if desired. Melting experiments were conducted in 1-cm cells in the five chamber temperature control unit of the Cary 219. Temperature was increased at 0.5 °C per min during the melting experiment by a Haake A81 programmable temperature bath and Haake PG20 programmer unit.

NMR. Proton spectra were obtained at 270.05 MHz on a JEOL GX-270 spectrometer by using the Redfield 21412 pulse sequence under the following conditions: typically 15000 scans; 0.5-s pulse repetition rate; 4-Hz line broadening; carrier frequency at 13.5 ppm; 8K data points; TSP reference; 10000 Hz spectral width; 9% D_2O in H_2O-PIPES 00 buffer; 20 mM DNA phosphate molarity; and 0.825-mL sample volume in a 5-mm NMR tube.

Phosphorus spectra were obtained at 109.25 MHz with a JEOL GX-270 spectrometer under the following conditions: typically 3000 scans; 45° pulse width with a pulse repetition time of 2.5 s; 4 Hz line broadening; broad band bilevel decoupling; 2K data points zero filled to 8K data points; trimethyl phosphate reference; 2000 Hz spectral width; 9% D_2O in H_2O-PIPES 00 buffer; 20 mM DNA phosphate molarity; and 1.50-mL sample volume in a 10-mm NMR tube. Temperature in ^{31}P NMR experiments was monitored by using the ^{31}P "thermometer" method described by Gorenstein and co-workers.[37]

Viscometric Measurements. Viscometric titrations of both sonicated calf thymus and closed circular superhelical Col. E_1 DNA were conducted in Cannon Ubbelohde semimicro dilution viscometers at 28 °C in PIPES buffer as previously described.[20]

Acknowledgment. This work was supported by NSF grant PCM83-09575 and the Georgia State University Research Fund. WDW is a recipient of an American Cancer Society Faculty Research Award.

(35) Duncan, W. G.; Colwell, W. T.; Scott, C. R.; Henry, D. W. *J. Med. Chem.* **1968**, *11*, 1221-1223.
(36) Kusuma, S.; Wilson, W. D.; Boykin, D. W. *J. Heterocycl. Chem.*, in press.
(37) Gorenstein, D. G.; Luxon, B. A.; Goldfield, E. M.; Lai, K.; Vegeais, D. *Biochemistry* **1982**, *21*, 580-589.

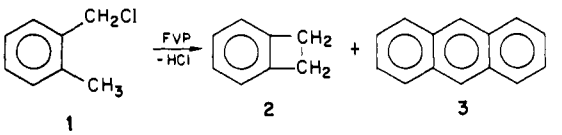
Communications to the Editor

Formation of Anthracenes in the Flash Vacuum Pyrolysis of Benzocyclobutenes and Dimers of *o*-Quinodimethanes

Walter S. Trahanovsky* and Bruce W. Surber

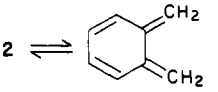
Ames Laboratory and the Department of Chemistry
Iowa State University, Ames, Iowa 50011
Received July 31, 1984

The Flash Vacuum Pyrolysis (FVP) of **1** gives, in addition to the major product benzocyclobutene (**2**),[1] varying amounts of a high molecular weight (MW) material. We have found this material to be primarily anthracene (**3**). Anthracene (**3**) has also been reported as a minor product in the later stages of the static gas-phase pyrolysis of **1** at 430 °C,[2] and **3** and dihydroanthracene



1 **2** **3**

have been obtained in low yields from the FVP of 1,2-bis[(phenylseleno)methyl]benzene, another precursor of **2**.[3] We proposed that these unusual routes to **3** involve the dimerization of *o*-quinodimethane (**4**) followed by loss of two carbon atoms and six



4

hydrogen atoms from the dimer or dimers of **4**. Compound **4** is formed by the thermal opening of the cyclobutene ring of **2** and

(1) Morello, M. J.; Trahanovsky, W. S. *Tetrahedron Lett.* **1979**, 4435 and references cited therein.
(2) Loudon, A. G.; Maccoll, A.; Wong, S. K. *J. Am. Chem. Soc.* **1969**, *91*, 7577-7580.
(3) Higuchi, H.; Sakata, Y.; Misumi, S. *Chem. Lett.* **1981**, 627-630.

0002-7863/85/1507-4995$01.50/0