6-1997

Molecular recognition of blunt-ended DNA

Michael D. Farwell
Union College - Schenectady, NY

Molecular Recognition of Blunt-Ended DNA

By

Michael D. Farwell

* * * * * * * * * * *

Submitted in partial fulfillment
of the requirements for
Honors in the Department of Chemistry

UNION COLLEGE

June, 1997

ABSTRACT

FARWELL, MICHAEL D. Molecular Recognition of Blunt-Ended DNA. Department of Chemistry, June 1997.

Many important enzymatic and physical processes occur specifically at the ends of DNA molecules. The goal of this project was to develop a molecule that would be an effective probe of the chemistry at the ends of DNA. To do this, ethidium bromide, a well-known intercalator (a molecule that binds between DNA base pairs), was blocked to intercalation with bulky *t*-butyl groups. We proposed that this new molecule would preferentially bind to blunt-ended DNA, where there is much less steric hindrance. The synthesized molecule was characterized by NMR, Fluorescence, UV/Vis, and Mass Spectroscopy. Supporting data were acquired through fluorescence lifetime measurements, which reflect the chemical environment of the fluorophore. Preliminary data indicated that the lifetime of our target molecule lengthened from 3.8 ns to 13.8 ns when the molecule was exposed to short DNA with blunt ends. This effect could not have been due to intercalation, because there was no observed change in the lifetime of the molecule when it was exposed to long DNA with few blunt ends. Unfortunately, we have not yet been able to duplicate the data supporting blunt-end binding, partially because the association between our target molecule and blunt-ended DNA is extremely small. An interesting aside is that the target molecule converts to another form at low pH and low ionic strength. We believe that it might be stacking on itself to form multimers, but the current evidence is inconclusive. In addition to the fluorescence studies, computer modeling was done using Macromodel and Spartan on an SGI workstation. After minimizing the energies of ethidium bromide and the target molecule, several docking studies were done to predict the ability of each molecule to intercalate DNA. Modeling will be used to help select future blunt-end binding target molecules.

For My Parents

Acknowledgment:

I would especially like to thank the entire Chemistry Department at Union College. For a rather small, private college, I truly believe that Union College has one of the top Chemistry Departments in the nation. The high quality of instruction coupled with great facilities and a friendly atmosphere have all helped to mold me into what I am today. It is only by playing that with the best that you truly learn, and I know that I played with some of the best.

In particular, I would like to thank all three of my advisors. Thank you to Professor Adrian: for making me believe in myself and guiding me along the path to graduate school. It is because of you that I came to enjoy Organic Chemistry. Thank you also to Professor Carroll: for showing me that it's not just what you know, but how you explain it that matters. Thank you also to Professor Anderson: for introducing me to the world of computational chemistry.

This research was facilitated by a summer research fellowship from Pfizer, Inc. and a grant from the Internal Education Fund of Union College. Pfizer not only provided a summer stipend, but also contributed to the Chemistry Department.

Table of Contents

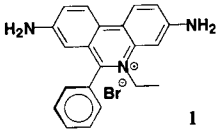
Table of Figures

Chapter I:

Introduction

Many important enzymatic and physical processes occur specifically at the ends of DNA. Examples include the enzymatic degradation of DNA by exonucleases, strand coupling by ligases, terminal phosphate addition through the action of phosphokinases and strand extension by terminal transferases. Also, melting of small DNA molecules and fraying of large molecules are physical processes that begin at the ends. A molecular-level understanding of the chemical rules which control these processes is crucial for a complete understanding of the properties and interactions of DNA.

The objectives of my Senior Research Project were to design, synthesize, and test a capping compound (CC) that will bind to blunt-ended DNA. This CC would be an effective tool for studying events at DNA ends, including both enzyme-mediated and physical processes. The CC was designed using ethidium bromide (EtBr), a well known intercalator[1], as a template.



1

Figure 1. Structure of EtBr

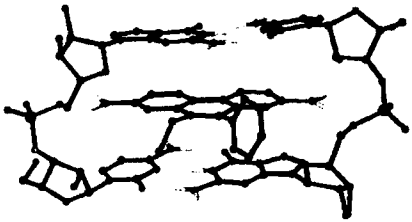


Figure 2. EtBr intercalated between base pairs in an RNA dimer[2]

EtBr's three most important features that allow it to insert (intercalate) between DNA base pairs are: its large flat aromatic surface, its two amino groups, and its positive charge. These features allow EtBr to π-stack with the aromatic base pairs of DNA, form hydrogen bonds with the 5' oxygens on the phosphate backbone, and have electrostatic interactions with the negative phosphate backbone. Veal and Wilson estimated that 80% of the energy of interaction between EtBr and DNA is due to π-π stacking and other van der

Waals interactions, and 7-8 kcal/mol of binding energy is from hydrogen bonding between the amino groups and the 5' oxygens.[3] However, their calculations were at high salt concentrations (standard state), where contributions to the free energy from release of counterions are negligible. Under lower salt, physiological conditions (a state similar to our research conditions), EtBr binds more strongly to DNA, due mostly to the favorable entropy of counterion release for DNA.[3] In other words, having the positively charged EtBr intercalated into DNA allows for the release of positive ions that were associating with the negative phosphate backbone. At low salt concentrations this is favorable, because it allows the few ions in solution to disperse rather than aggregate around the DNA. Therefore, under low salt conditions, electrostatic contributions to the interaction between EtBr and DNA are actually larger than Veal and Wilson calculated.

There are several other points concerning EtBr intercalation into DNA which deserve some attention. First, EtBr always intercalates with both the ethyl and phenyl groups in the minor groove.[4] It is thought that this conformation allows the phenyl group to exhibit some favorable van der Waals interactions with the helix backbone.[4] It seems likely that the narrow minor groove is better able to protect the hydrophobic phenyl group from an aqueous environment compared to the much wider major groove. Also, EtBr obeys the neighbor exclusion principle, according to which it can only occupy every other site along the length of a DNA double helix and is forbidden to occupy neighboring sites simultaneously.[5] In addition, it was found that the amino group at the 3 position on the phenanthridine ring (but not at position 8) is involved in stabilization of the drug-DNA complex.[6] Lastly, EtBr binds more strongly to G-C base pairs than T-A base pairs.[6]

The effects of intercalation on DNA are rather dramatic, causing the distance between adjacent base pairs to increase from 3.4 Å to about 7 Å, and unwinding the helix such that the angle of rotation between two adjacent base pairs changes from 36° to 10°.[7] The DNA-intercalator complex, while not covalently joined, is still difficult to separate. Intercalation happens spontaneously because the intercalator-bp (base pair) complex is

more stable than the bp-bp complex. Thus EtBr can be rather tenacious, and will interfere with DNA replication and transcription (which can lead to cancer). However, the strong affinity of EtBr for DNA, coupled with its fluorescent properties, makes it very useful as a stain for DNA. In this study, we were particularly interested in EtBr's fluorescence lifetime, and the effects of different chemical environments on that lifetime. However, in order to understand the implications of fluorescence lifetime data, it is first helpful to understand the nature of fluorescence.

Fluorescence is one of several processes that can occur after a molecule has been excited by absorbing a photon of light. A Jablonski diagram is shown in Figure 3, which

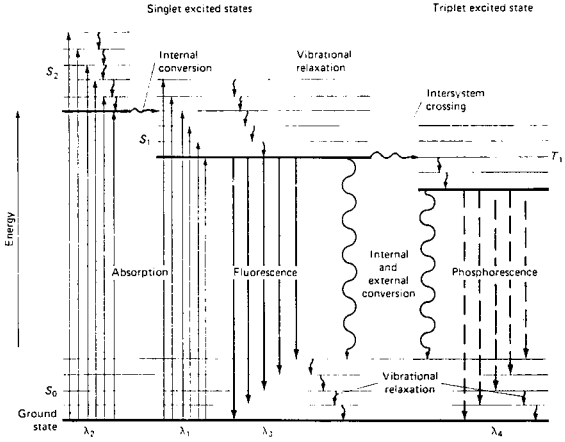


Figure 3. Jablonski diagram (taken from reference 7)

depicts these processes. The heavy upper lines are energy levels for three excited electronic states, and different vibrational energy levels are suggested by the lighter horizontal lines.

Promotion of a molecule to a higher electronic state (S_2) can only occur if radiation with energy corresponding to the difference in energy between the ground state (S_0) and the excited state (S_2) is absorbed. Photons with slightly more energy than $S_2 - S_0$ can promote molecules to higher vibrational states. The time it takes for a photon to be absorbed is on the order of 10^{-14} to 10^{-15} s, whereas the time it takes for fluorescent emission is approximately 10^{-7} to 10^{-9} s.[8] In solution, as soon as the molecule is excited, excess vibrational energy is lost as a consequence of collisions between it and the solvent (taking 10^{-12} s or less).[8] Thus fluorescence always involves a transition from the lowest vibrational level of an excited state. The lifetime of the excited state population is measurable, and the concentration of excited molecules decays exponentially according to the following formula:

$$[F^*] = [F^*_o]\exp(-t/\tau) \qquad \text{Eq. 1}$$

where $[F^*_o]$ is the initial concentration of the excited state, $[F^*]$ is the concentration at time 't', and τ is the fluorescence lifetime. Basically, τ is a statistical measure of the time it takes for the excited population to decay to 1/e (37%) of its initial concentration. Different environments can either stabilize or destabilize the excited state, thereby affecting the decay rate. Thus, a molecule's fluorescence lifetime can be used as a probe of its environment.

One of the reasons why EtBr was chosen as a template for the target CC was because its fluorescence properties have been well characterized. For example, the binding of EtBr to DNA is known to cause an increase in EtBr's fluorescence intensity, and has been shown to produce a clear isosbestic point (indicating a single bound form of the drug).[9] In addition, fluorescence lifetime, quantum yield, and quenching measurements have all been used extensively in studies to monitor EtBr's intercalation into DNA.[1] Our study depends exclusively on fluorescence lifetime measurements to determine if and how a molecule is associating with DNA. For example, τ of free EtBr in solution is 1.7 ns, whereas EtBr bound to DNA has a τ of 23 ns.[10] It seems that the hydrophobic environment inside of DNA helps stabilize EtBr's excited state, perhaps by reducing the

number of collisions with water molecules and by making EtBr more rigid (restricting its ability to vibrate).

One of the advantages of fluorescence lifetime analysis is the ability to resolve signals from multiple species in solution. For example, if an excess amount of EtBr were added to a solution of DNA, then not all of the EtBr would be able to intercalate and some would be left free in solution. Because free and intercalated EtBr have different lifetimes, the fluorescence decay curve of the mixture could be fit to a double exponential curve, so that both lifetimes could be calculated as well as the relative amounts of each species. To make matters more complicated, there are actually four different modes of interaction with DNA, possibly resulting in four different chemical environments (Figure 4).

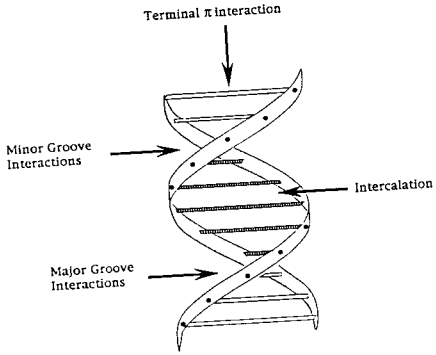


Figure 4. Modes of DNA interaction

It has been demonstrated that EtBr is capable of binding to DNA through all of the possible modes. In crystalline complexes, EtBr both stacks on blunt ends of DNA and intercalates,[10] and in solutions with low ionic strength and excess EtBr, binding occurs on

the outside of DNA.[11] However, under physiological conditions, nearly all of the EtBr is in the intercalated form.

One of the assumptions on which this experiment rests is that an intercalated molecule, an end-capped molecule, and a free molecule in solution will all have different τ's, because their chemical environments will all be different. It seems logical that the τ of the end-capped molecule would fall somewhere between that of the free and that of the bound states, because it is experiencing approximately half of the hydrophobic environment of the intercalated complex.

The motivation for the construction of our target CC came from the fact that bulky *t*-butyl groups can block intercalation by preventing the molecule from fitting between the base pairs of DNA.[12] If we could block EtBr to intercalation, then its number of possible binding modes to DNA would be reduced by one. The *t*-butyl groups used in preventing intercalation shouldn't interfere with blunt-end binding, because there is much less steric hindrance at the ends. Under physiological conditions, we hoped that major and minor groove associations between our target CC and DNA would be very small, thereby allowing our CC to interact exclusively with the terminal ends of DNA. At the blunt end, the CC would still be able to π-stack and hydrogen bond with one set of base pairs. However, the stability of the end-capped state would necessarily be less than that of the intercalated state. This is because intercalation allows almost all of EtBr's surface area to come in contact with the DNA. With blunt-end binding, only 50% (at best) of the substrate would be in contact with DNA. Assuming that Van der Waals, π-π stacking and other surface-area-dependent interactions are the primary means by which an intercalator binds to DNA, then, at best, our capping compound will have an affinity that is half that of EtBr. Thus, additional methods may be required to increase the affinity between the CC and DNA.

One way to increase the amount of blunt-end binding would be to add a groove-binding tail to the CC. This tail would both enhance the overall strength of association

between the CC and DNA, and direct the relative orientation of the CC. The tail would have to be designed such that binding of the tail to DNA would leave the CC dangling at or near the blunt end.

A final tool that will be employed to augment the synthetic and analytical research is molecular modeling. The modeling was done using Macromodel,[13] a software program designed to perform molecular mechanics calculations on large molecules. All of the molecules that were used in the calculations were built using Macromodel, although X-ray structures were used several times as a model for how EtBr orients itself within DNA during intercalation. Because EtBr and our CC were different, their minimized energies couldn't be compared directly. Therefore, in order to draw conclusions about the relative stabilities of EtBr and the CC in their intercalated and end-capped states, the energy differences between the free and bound substances had to be compared.

Chapter II:

Synthetic Experimental

Synthetic Experimental

A) Design of the Capping Compound

The goal of this project was to develop a molecule that would be an effective probe of the chemistry at the ends of DNA. To do this, several potential target molecules were designed using ethidium bromide (EtBr) (**1**), a well-known intercalator[1] (a molecule that binds between DNA base pairs), as a model. Specifically, 3,8-diamino-6-phenyl-phenanthridine (**2**) was used as the precursor in all of our synthetic schemes.

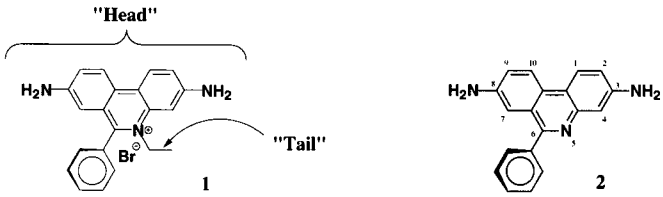
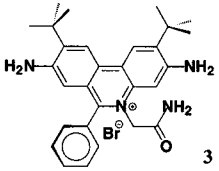


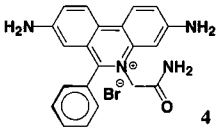
Figure 5. Ethidium Bromide and 3,8-diamino-6-phenyl-phenanthridine

We proposed that EtBr derivatives that were blocked to intercalation would preferentially bind to blunt-ended DNA, where there is much less steric hindrance. The design of these new molecules rested on previous research by Baguley and coworkers which showed that bulky *t*-butyl groups effectively block intercalation.[11] In the case of EtBr, intercalation always occurs "head" first[3] (Figure 5), thereby constraining us to that region for the addition of bulky substituents. The phenyl and "tail" substituents aren't as critical to intercalation, because they are left hanging in the minor groove of DNA.[3]

In the construction of our target molecules from **2**, we decided to use an acetamide group rather than an ethyl group for the "tail" (Figure 5), so that it would better approximate several DNA groove binding tails that may later be attached at that point. Our first proposed target molecule, **3**, had two *t*-butyl groups on positions 2 and 9 of the phenanthridine portion of the molecule. The other target molecule, **4**, was the control: it was expected to intercalate in a manner similar to EtBr. Unfortunately, neither target molecule was realized.



3

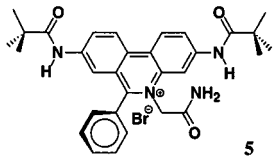


4

Friedel-Crafts alkylation of the phenanthridine portion of **2** failed, as did the selective addition of a "tail" by alkylation. Competition from the amine substituents seemed to be the main problem, so several protecting groups were used to try and quench their reactivity. First we tried using a trifluoroacetamide group (CF_3CONHR) to protect the amines. The resulting complex was resistant to alkylation by 2-bromoacetamide, most likely due to deactivation of the ring from the trifluoro protecting groups. Next we tried a methyl carbamate group ($CH_3OCONHR$). Although this complex was able to be alkylated at the aromatic nitrogen, the reaction to remove the protecting groups (which would have resulted in **4**) seemed to cause decomposition of the complex. Lastly, we tried a benzyl carbamate group ($PhCH_2OCONHR$). The resulting complex couldn't be alkylated at the aromatic nitrogen or the phenanthridine ring system.

Rather than letting the amine substituents impede our progress, we decided to take advantage of their reactivity. A new target molecule (**5**) was drawn up, which had *t*-butyl

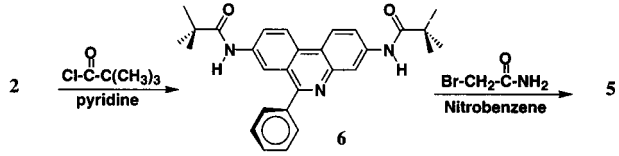
acetyl groups attached to the amine substituents. We reasoned that this molecule would be relatively easy to synthesize, as *t*-butyl acetyl chloride ($(CH_3)_3CCOCl$) should react in a similar manner as the methyl carbamate protecting group.



5

Scheme I illustrates our preparation of **5**. Initially, **2** was acylated with *t*-butyl acetyl chloride to produce **6**. Alkylation of **6** by 2-bromoacetamide then yielded the blocked intercalator **5**.

Scheme 1:



(See Appendix 1 for ^{1}H-NMR spectra of **2**, **5** and **6**, and Mass Spectral data for **5**)

B) General Experimental

Melting points were determined on a Mel-Temp capillary melting point apparatus and are uncorrected. Low and high resolution mass spectra were performed by the Mass Spectrometry Service of the University of Illinois. Proton magnetic resonance spectra were obtained on a Varian Gemini 200 MHz spectrometer. Chemical shifts (^{1}H-NMR) are expressed in parts per million (δ units) downfield from tetramethylsilane (TMS) used as an internal reference. The following abbreviations were used: s=singlet, d=doublet, t=triplet, m=multiplet. Yields are reported based on the amount of isolated material obtained after the indicated procedure. Thin layer chromatography was performed using Whatman® KGF silica gel 60 Å (0.25 mm) analytical glass plates. All starting materials and solvents were purchased from Aldrich and were used as received.

Where indicated, reactions run under nitrogen atmosphere were arranged with a mercury or oil bubbler so that the system could be alternately evacuated and filled with nitrogen and left under positive pressure. Syringes and reaction flasks were dried at least 12 hours in an oven at ≥ 120 °C and cooled in a dessicator over calcium sulfate prior to use.

C) Preparations

3,8-bis(trimethylacetamide)-6-phenyl-phenanthridine (6). To a stirred solution of 0.25 g of **2** (0.88 mmol, 1 eq) in 1.9 mL of pyridine was added 0.30 mL of trimethyl acetyl chloride (2.44 mmol, 2.8 eq). The reaction was stirred at room temperature for 2 hours and then cold water was added. The resulting precipitate was isolated by filtration after 30 minutes and well rinsed with cold water, affording 0.41 g (104%): mp 189-191 °C; ^{1}H-NMR ($CDCl_3$) δ 8.48 (d, 1H, J=9 Hz), 8.35 (d, 1H,

J=9 Hz), 8.21 (dd, 1H, J=9 Hz, J=2 Hz), 8.10 (dd, 1H, J=9 Hz, J=2 Hz), 7.93 (dd, 2H, J=10 Hz, J=2 Hz), 7.73 (m, 3H), 7.59 (m, 4H), 1.36 (d, 18H, J=4 Hz).

5-acetamide-3,8-bis(trimethylacetamide)-6-phenyl-phenanthridinium bromide (5). To a stirred solution of 52 mg (0.115 mmol, 1 eq) of **6** in 1 mL of nitrobenzene was added 32 mg (0.232 mmol, 2 eq) of 2-bromoacetamide. After stirring for 1.5 hours at 100 °C the reaction mixture remained cloudy and yellow. The temperature was then raised to 157 °C over 30 min, and stirred for 2 hours. The color of the suspension gradually went from orange to red during this time. The reaction mixture was allowed to cool, and diluted with 10 mL of Et_2O. The resulting precipitate was isolated by filtration and then dissolved in water. Any undissolved solid was removed by filtration. This solution was continuously extracted with Et_2O for three days to remove excess 2-bromoacetamide. The resulting aqueous solution was lyophilized, affording 20 mg of **5** as a fluffy yellow solid: mp 180-181 °C; ^{1}H-NMR (D_2O) δ 8.73 (d, 1H, J=6 Hz), 8.58 (d, 1H, J=6 Hz), 8.34 (m, 1H), 8.21 (m, 1H), 8.05 (t, 1H, J=7 Hz), 7.92 (m, 1H), 7.72 (m, 3H), 7.35 (m, 2H), 5.44 (s, 2H), 1.16 (d, 18H, J=9 Hz); FAB MS, *m/e* calcd for $C_{31}H_{35}N_4O_3$ $(M-80)^+$ 511.6, measured 511.3.

D) Determination of Purity by ^{1}H-NMR

An NMR sample of **5** was prepared by dissolving 1.8 mg of **5** in 995 μL of D_2O. 5 μL of dioxane was added as an internal concentration reference. By comparing the area of the dioxane singlet at 3.68 ppm to the area of one of the *t*-butyl singlets, the concentration of **5** could then be determined. The relative area of the dioxane singlet to the *t*-butyl singlet was found to be 27.8 : 1. This corresponds to a ratio of 1.1 mg of **5** per 1.8 mg of "crude" material (61% pure, 18% yield). Impurities are presumably due to inorganic salts which are not NMR active.

Chapter III:

Spectroscopic Experimental and Results

Generally, fluorescence is maximized when the sample has an absorbance of about 0.05 at the excitation wavelength. Absorbances greater than 0.05 usually result in decreased fluorescence due to self-absorption. Therefore, before any fluorescence data were collected, UV/Vis absorption data were collected using a Hewlett Packard 8452A Diode Array Spectrophotometer. The concentrations of EtBr and **5** could then be varied so that they gave an absorbance of about 0.05 at a given wavelength. There were three wavelengths of particular interest: 316, 337, and 358 nm, corresponding to the principal emission lines of the nitrogen lamp used to excite the sample in the fluorescence lifetime instrument. UV/Vis absorption spectra (Figures 6 and 7) of both EtBr and **5** suggested that 337 nm might be the best wavelength for excitation because (a) the intensity of the nitrogen lamp at 337 nm was nearly twice that at 316 and 358 nm and (b) the absorbance at 337 nm was relatively close to the absorbances at 316 and 358 nm. In order to maximize fluorescence, we wanted to maximize the number of molecules that were excited. At a given concentration, the greater the absorbance, the more light that is absorbed, and the higher the percentage of molecules that are excited. On the other hand, the greater the intensity of the excitation beam, the greater the fluorescence signal. At 337 nm, we were settling for a slightly lower absorbance value in order to get twice the intensity of the source, so that the fluorescence signal would be maximized.

Once the excitation wavelength was determined, it was possible to determine the optimal concentrations for fluorescence. As stated previously, solutions with absorbances of about 0.05 serve this purpose. For **5**, the optimal concentration was determined to be 1.9×10^{-6} M, and for EtBr, 7.7×10^{-6} M. Although these were the concentrations that were used for all of the fluorescence and fluorescence lifetime measurements, there was an error in the calculation of the concentration of EtBr. In retrospect, it would have been better to use a more dilute solution, perhaps in the 10^{-7} to 10^{-8} M range, which would have given an absorbance closer to 0.05.

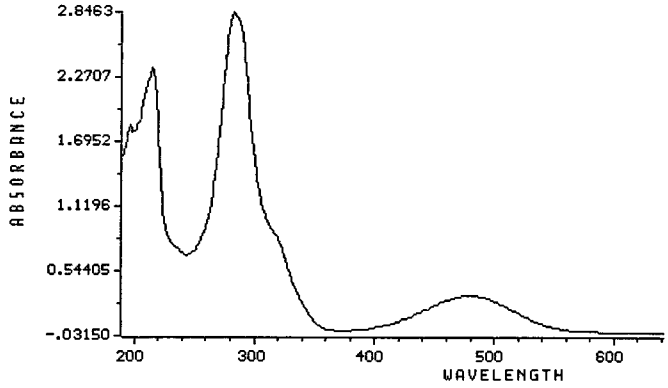


Figure 6. UV/Vis absorption spectrum of EtBr (5×10^{-5} M) in 1×10^{-2} M Tris and 1×10^{-3} M EDTA, measured against a Tris/EDTA blank.

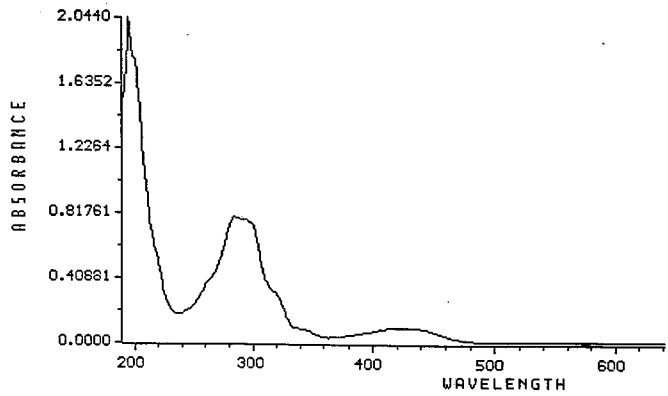


Figure 7. UV/Vis absorption spectrum of **5** (5×10^{-5} M) in DI water, measured against a DI water blank.

After the best conditions for excitation were estimated, a series of fluorescence spectra were recorded using a Perkin-Elmer LS-5B Luminescence Spectrophotometer. These spectra were used to double-check the excitation wavelength and find the optimal emission wavelength for the samples. For EtBr, the maximum emission wavelength was 617 nm (Figure 8), and for **5**, it was 422 nm (Figure 9). All spectra were recorded with 5 nm slits, a response time of 3, a scan speed of 60 nm/min, and a recorder format of 20 nm/cm. All fluorescence data were taken on solutions that contained $1x10^{-2}$ M Tris (pH = 8.5) and $1x10^{-3}$ M EDTA, thereby mimicking the buffers that were used when DNA was added.

A complicating factor is the fact that **5** underwent complete conversion to another species over a period of five weeks. The new species, termed Orange **5** (as opposed to the original material, Yellow **5**) had its fluorescence maximum shifted to 520 nm. This unanticipated conversion prompted the collection of a barrage of new fluorescence spectra, where conditions such as pH, concentration, and ionic strength were varied.

The pH was the first variable manipulated, because we quickly determined that the conversion of **5** was pH dependent. A series of six quartz cuvettes were filled with 1.9 x 10^{-6} M of **5** in DI water. The pH of the solutions in the cuvettes was then changed by adding NaOH or HCl, and then measuring the pH with a glass electrode. Not more than 15 µL of acid or base was ever added to any of the solutions; the total volume of each solution was either 2 or 3 mL. The pH values of the solutions in the six cuvettes were: 3.0, 4.2, 5.7, 7.6, 9.5, and 11.6. The fluorescence curves from the six solutions (Figure 10) showed almost complete conversion from Yellow **5** (high pH) to Orange **5** (low pH). In addition, the curves didn't converge at an isosbestic point, meaning that there were more than two species present in solution. The pH study was repeated and yielded the same results.

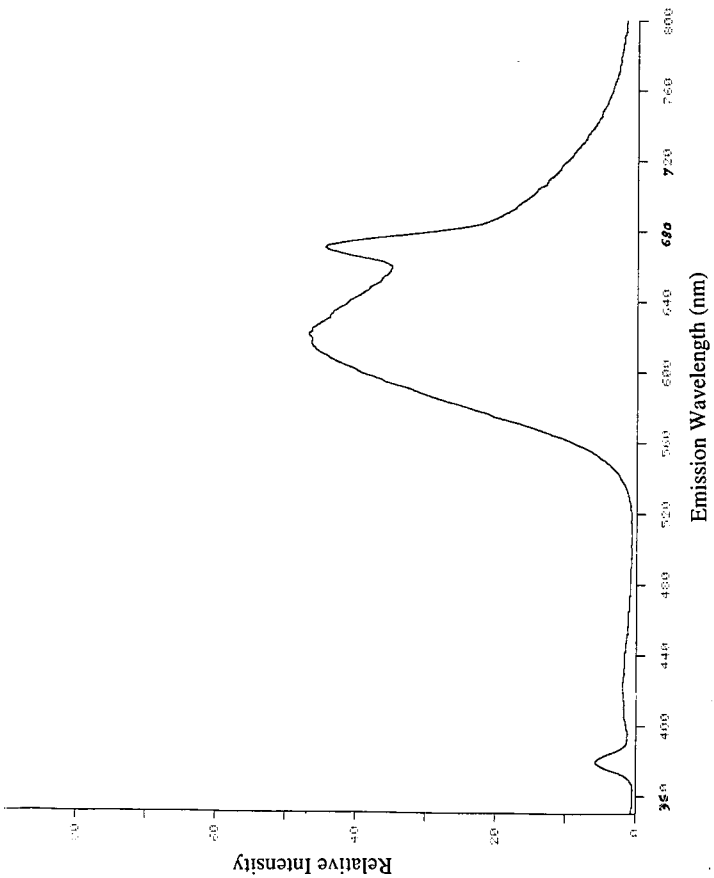


Figure 8. Fluorescence spectrum of EtBr in DI water (Excitation @ 337 nm). Peak at 674 nm is due to scattered light at $2*\lambda_{exc.}$

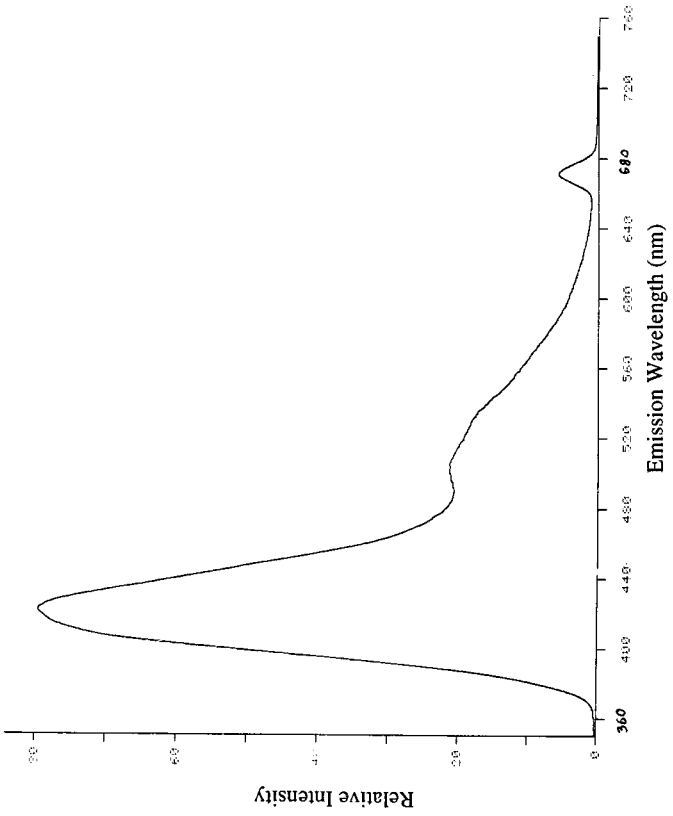


Figure 9. Fluorescence spectrum of **5** in DI water (Excitation @ 337 nm). Peak at 674 nm is due to scattered light at $2*\lambda_{exc.}$

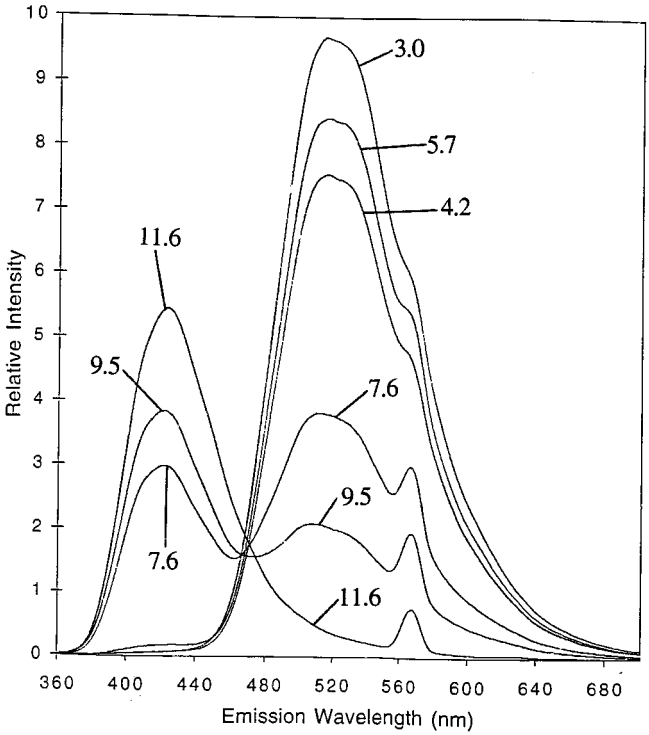


Figure 10. Fluorescence of **5** as a function of pH (Excitation @ 284 nm). Peak at 568 nm is due to scattered light at $2*\lambda_{exc.}$

To examine the effect of ionic strength on the conversion of **5**, two solutions were prepared with **5** at a concentration of 1.9 x 10^{-6} M. The low ionic strength solution was prepared using DI water only, and the high ionic strength solution was prepared so that it was 1 x 10^{-2} M Tris (pH=8.5) and 1 x 10^{-3} M EDTA. The fluorescence spectra of both solutions were recorded (Figure 11), and showed almost complete conversion from the Orange to the Yellow species as the ionic strength increased. Under conditions of low ionic strength, the fluorescence maximum occurred at 520 nm (consistent with the spectral properties of the Orange species), and at high ionic strength, the fluorescence maximum occurred at 420 nm (consistent with the spectral properties of the Yellow species). Although the conversion from the Orange to the Yellow species was partly a result of a change in pH (the high ionic strength solution was buffered at pH 8.5), the complete conversion to the Yellow species had to be due to the high ionic strength. This is because, at pH 8.5, there normally would have been a 1:1 ratio of Orange **5** to Yellow **5** (Figure 10). The fact that there was a complete conversion to the Yellow species means that the high ionic strength had an effect.

To examine the effect of concentration on the conversion of **5**, five solutions of Orange **5** in DI water were prepared in cuvettes: 4.1, 1.5, 0.5, 0.17, and 0.05 x 10^{-6} M. Even after letting the solutions stand for a month, there were no observed spectral shifts from Orange **5** to Yellow **5**. The only change in the spectra was a decrease in intensity, most likely due to degradation of **5**.

A final study was also done, the goal of which was to gather fluorescence and fluorescence lifetime data on a system that was designed to favor aggregation via π-π stacking. To help ensure that stacking would occur, an aromatic compound with a negative "tail" was chosen so that it would aggregate with **5**, an aromatic compound with a positive "tail." 9-anthracene carboxylic acid (9-ACA) seemed an obvious choice, because of its large aromatic surface area and the fact that at pH greater than 4 or 5, the 9-ACA would be deprotonated, giving it a negative charge. In addition, 9-ACA had no measurable

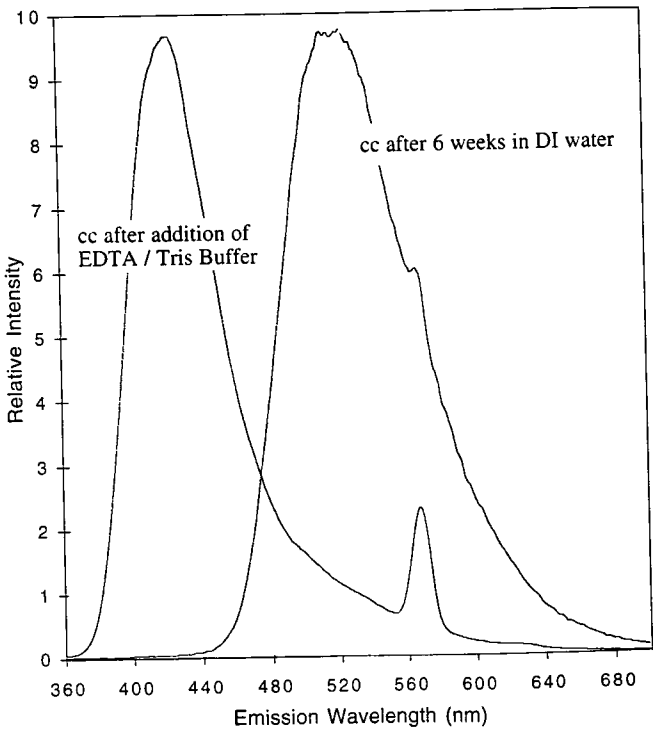


Figure 11. Fluorescence of **5** as a function of ionic strength (Excitation @ 284 nm). Peak at 568 nm is due to scattered light at $2*\lambda_{exc.}$

fluorescence in the region of our study (300 nm to 700 nm) under our experimental conditions.

The 9-ACA was first dissolved in DI water made basic with NaOH, and then added to a cuvette solution such that the concentration of 9-ACA was four times that of **5** (the pH of the solution was 11.1, the concentration of **5**, 1.9×10^{-6} M). The high pH of the solution allowed us to begin with a solution having Yellow **5** as the only species present. A fluorescence spectrum was taken of the solution before any 9-ACA was added (Figure 12), immediately after the addition of 9-ACA (Figure 12) and two days after the addition of 9-ACA (Figure 13). The aggregation of 9-ACA with **5** occurred slowly, and decreased the fluorescence maximum of **5** by a factor of about four. That same solution was then acidified with HCl to a pH of 3.4, thereby protonating the 9-ACA and forcing it to precipitate out of solution. Fluorescence spectra were taken (Figure 14) 2, 9, 17, and 38 days after acidification. The spectra showed a partial conversion to a species emitting at 520 nm (probably the Orange **5** species). After 17 days, the peak at 520 nm stopped growing, and the original peak at 420 nm (Yellow **5** or 9-ACA / **5** aggregate) began decreasing. This implied that there was no further reaction after 17 days, except for degradation of the species emitting at 420 nm.

Finally, the PTI LS-100 fluorescence lifetime instrument was used to analyze for intercalation. Fluorescence lifetime (τ) measurements were first taken for EtBr (7.7×10^{-6} M) so that a comparison could be made with the literature values. λ-DNA was used for this study, because it consists of long strands of DNA with few blunt ends. Our values for the τ of EtBr were 1.4 ns in the free state and 16.9 ns when intercalated in DNA. This was then compared to the literature values taken under similar experimental conditions: 1.7 ns and 23 ns, respectively.[10] Although these values are not numerically equivalent, they exhibit the same trend and lend credibility to our results. After confirming that our methods provided data that was similar to literature values, we studied the fluorescence lifetime of **5** with λ-DNA. Because our molecule was blocked to intercalation, we expected

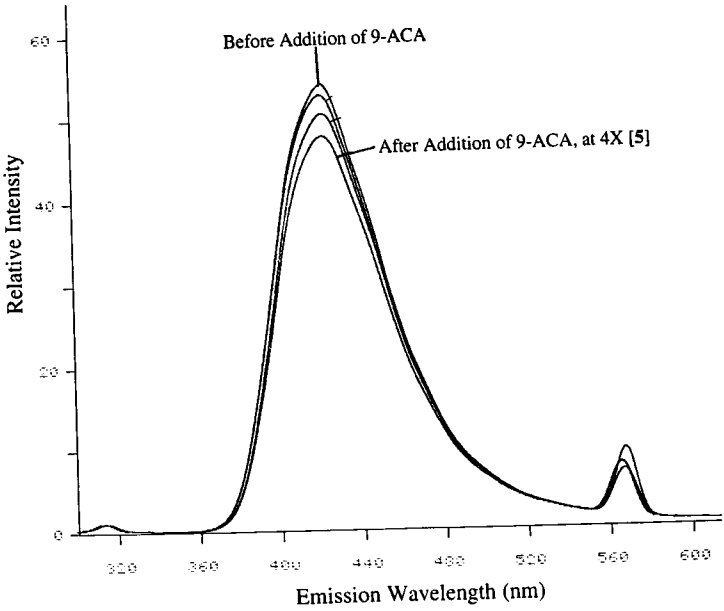


Figure 12. Fluorescence of **5** before addition of 9-ACA, and immediately after addition of 9-ACA (at 4X concentration of **5**); pH = 11.1 (Excitation @ 284 nm). Peak at 568 nm is due to scattered light at $2*\lambda_{exc.}$

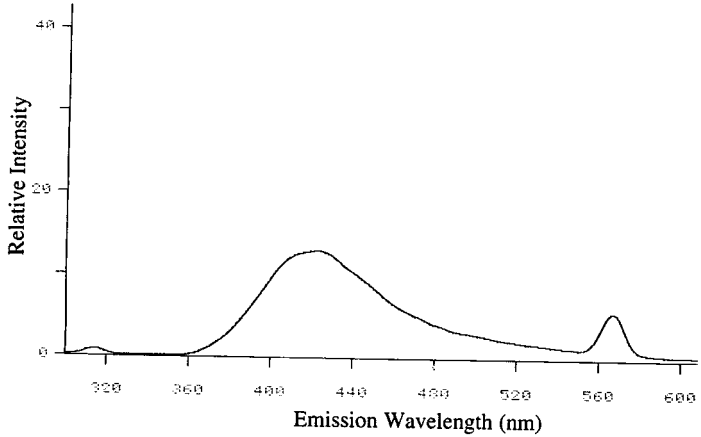


Figure 13. Fluorescence of **5** two days after addition of 9-ACA; pH = 10.0 (Excitation @ 284 nm). Peak at 568 nm is due to scattered light at $2*\lambda_{exc.}$

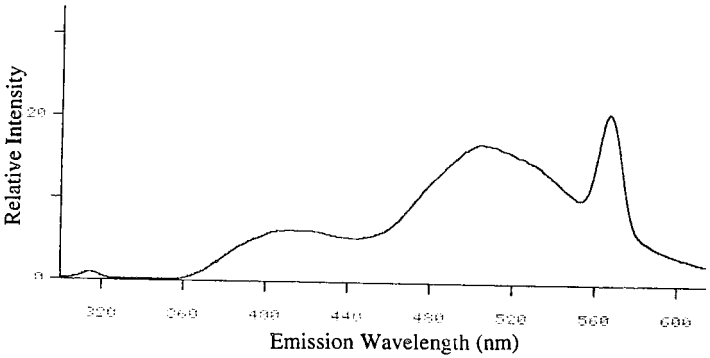


Figure 14. Fluorescence of **5** four days after addition of 9-ACA; two days after acidification to a pH of 3.4 (Excitation @ 284 nm). Peak at 568 nm is due to scattered light at $2*\lambda_{exc.}$

to see no change in its fluorescence lifetime in the presence of λ-DNA. The measured lifetime of the free molecule was 3.8 ns and the measured lifetime of the molecule in the presence of λ-DNA was also 3.8 ns. Therefore, our data strongly suggests that our molecule was, indeed, blocked to intercalation. In all of the studies listed above, there was an excess of DNA base pairs by a factor of 150 : 1.

The second part of the fluorescence lifetime experiments involved measuring the fluorescence lifetime of our molecule in the presence of an excess of blunt ends. To do this, HPLC purified DNA of the sequence: CGCGAATTCGCG was ordered from the University of Pennsylvania Cancer Center. The oligomer was annealed onto itself by dissolving it in 100 μL of aqueous 3.9 M NaCl, heating to 78 °C for 5 min and then cooling to 35 °C over a period of 45 min. The DNA was then frozen at -70 °C for storage. This sequence of dsDNA, also known as the "Dickerson dodecamer," was chosen for two reasons. The main reason was that its crystal structure was already published, and therefore accessible, on the World Wide Web. Thus, the structural data could easily be downloaded for use in various molecular modeling calculations. The second reason this oligomer was chosen was because of the C-G base pairs at either end. EtBr binds more strongly to C-G base pairs than to A-T base pairs, and we hypothesized that our CC might do the same.

Using sterile glassware and solutions, a sample was made up that contained 1.26 μM **5**, 150 μM blunt ends, 0.01 M Tris, and 0.001 M EDTA. Fitting the data to a single exponential curve gave a lifetime of 4.5 ns with a χ^2 value of 2.0; whereas fitting it to a double exponential curve gave two lifetimes with a χ^2 value of 1.5. The double exponential fit was significantly better, suggesting that our CC was in (at least) two different chemical environments: bound and unbound. Since we already knew that the unbound state had a lifetime of 3.8 ns, one of the lifetimes was held fixed at 3.80 ns while the other was calculated. By default, this other lifetime was considered to be that of the

bound state, and it was calculated to be 13.8 ns. The lifetime data is summarized in Table 1. Although this data appeared promising, it wasn't reproducible. Two other

Table 1: Summary Fluorescence Lifetime Data

	EtBr	**5**
Free	1.4 ns	3.8 ns
With λ-DNA	16.9 ns	3.8 ns
With blunt-ended DNA	---------	13.8 ns (?)

experiments were run which produced the following results: the lifetime of free **5** was calculated to be 3.5 and 3.9 ns, and the lifetime of **5** in the presence of blunt ends was calculated, with a single exponential curve fit, to be 4.0 ns in both experiments. In fact, neither experiment produced data that gave a better fit to a double exponential than to a single exponantial curve. The reason why the first experiment produced results that were different than the other two experiments is currently uncertain, although it could simply be due to random error. Further testing is required in order to better establish the interactions between the CC and DNA.

The final part of the spectroscopic studies involved measuring the fluorescence lifetimes of the Orange **5** and the 9-ACA / **5** solution. The fluorescence lifetime of Orange **5**, calculated at 422 nm, resulted in the usual lifetime of 3.8 ns. However, when the emission wavelength was changed to the maximum for Orange **5** (520 nm), the lifetime changed to 11.2 ns. The fluorescence lifetime of Orange **5** wasn't measured in the presence of DNA, because the required buffers would have converted all of the **5** back to the Yellow species. For the second experiment, the previously described 9-ACA / **5** solution was used (1.9×10^{-6} M **5**, 7.6×10^{-6} M 9-ACA, pH 10). Although intercalation or association of molecules with DNA caused their fluorescence lifetimes to lengthen, the interaction of **5** with 9-ACA caused its lifetime to shorten, to 3.2 ns.

Chapter IV:

Molecular Modeling

Before any modeling was started, the Nucleic Acid Database at Rutgers was searched for structures that were similar to those used in our research. The Dickerson dodecamer (CGCGAATTCGCG), our choice of blunt-ended DNA, and several structures of EtBr intercalated into truncated pieces of RNA were found and downloaded. These structures, coming from X-ray crystallography, had the advantage of being truly accurate (at least for the solid-state structure). Initially, we had hoped to use those structures in our calculations. However, because Macromodel[13] was reading imported structures that it didn't create, small differences in code caused two separate errors. Individual atoms were sometimes misread by the software, and all of the double bonds were wiped out. In addition, it was impossible to modify any of the truncated DNA strands either by "growing" more DNA off of them or by fusing two pieces of DNA together. One of Macromodel's functions allows the user to "grow" double-stranded DNA according to all of the usual rules of DNA structure. However, the imported X-ray structures couldn't be modified with the "grow" command because Macromodel couldn't recognize the imported structures as DNA. Fusing DNA sounds much simpler, but Macromodel only allows atoms from two different molecules to be fused into one molecule. Therefore, half of the double helix must be fused together first, and then the other half must be fused. The problem is that fusing averages all of the data from the two atoms being fused, so the resulting DNA is quite contorted. The bottom line is that the modification of the downloaded structures requires time-consuming manipulation of individual bond distances and angles. At least for the collection of preliminary data, building all of the molecules directly on Macromodel (rather than importing them) sufficed. Although the downloaded X-ray structures couldn't be used in the calculations, they did provide excellent references on exactly where EtBr intercalates into DNA.

Macromodel was the only modeling program used because it was better suited for larger molecules, and it had the ability to "grow" DNA. Initially, ethidium bromide, the capping compound, and the Dickerson dodecamer were built. All structures were

minimized using Amber as the force field with water for solvent effects. Because EtBr and our CC were different, their minimized energies couldn't be compared directly. Therefore, in order to draw conclusions about the relative stabilities of EtBr and the CC in their intercalated and end-capped states, the energy differences between the free and bound substances had to be compared.

To model intercalation, EtBr (and the CC) were placed between DNA base pairs in a similar conformation to the X-ray structures. The DNA (which was "grown" to be the Dickerson dodecamer) unwound slightly and seemed to imitate the intercalated structures from the Nucleic Acid Database. When the CC was placed in the intercalated state, its *t*-butyl groups and phenyl ring kept it pinned inside the DNA, so it couldn't be ejected. After the energies of intercalation were calculated, EtBr and the CC were placed at the end of DNA, and a minimization was run. The results from these modeling experiments can be seen below in Table 2.

Table 2: Molecular Modeling Data

	EtBr (kJ/mol)	**5** (kJ/mol)
Change in energy when **Intercalated**	-116.28	-71.73
Change in energy when **End-Capped**	-114.30	-130.55

The fact that intercalated EtBr was calculated to be more stable than end-capped EtBr lends some credibility to these results, because EtBr has been experimentally proven to favor intercalation.[11] In addition, **5** was shown to be a poor intercalator, which corresponds with our experimental results. The only major flaw in our model is that it doesn't accurately portray the chemistry at the terminal ends of DNA. So, although our calculations show that it is thermodynamically favorable for **5** to end-cap DNA, the fact that DNA ends tend to fray[14] make this process less likely. A final, rather interesting result of these calculations, is that **5** was shown to bind more strongly to blunt ends than EtBr. A

further exploration of why **5** binds more strongly could be very useful in the design of future capping compounds.

In other future work, the modeling programs could be used to investigate possible structures for the Orange **5** compound. Energy calculations and emission spectra could be calculated for several protonated complexes, which might hopefully give some insights. Also, dimeric and multimeric interactions could be evaluated. This modeling data could then be compared with experimental results, and an appropriate model and possible mechanism for the conversion could then be proposed. In addition to studying the Orange **5** complex, modeling could be used to assess the viability of using a groove-binding tail on **5** to increase its association with the blunt ends of DNA. Also, docking studies could be done which would involve calculating the energies of **5** as it approaches DNA along an axis suitable for sliding between base pairs. This could then be compared with EtBr under the same circumstances.

Chapter V:

Conclusions

The goal of synthesizing a capping compound that would bind to blunt-ended DNA was not definitively achieved. However, several hurdles were cleared which lay along that path. Most importantly, **5** was shown to be blocked to intercalation. This was very good news, because synthetic difficulties forced us to place the bulky substituents further away from the phenanthridine ring than we had originally planned.

The only negative aspect of the new design was the fact that the amine substituents were converted to amides. This may have reduced the ability of **5** to form hydrogen bonds with the oxygens on the backbone of DNA and, consequently, may have reduced its DNA binding affinity to some extent. On the other hand, our modeling data shows that **5** actually bound better than EtBr at the blunt end. Either way, the hydrogen bonding ability of the CC played only a small part in the association (or lack of association) between **5** and blunt-ended DNA. This is primarily because DNA ends are prone to fraying.[14] Frayed DNA ends do not have a flat topology, so it seems improbable that either EtBr or **5** would bind well to them. The other major reason why blunt-end binding is unfavorable is because DNA ends are very exposed (even assuming that they are not frayed). DNA blunt ends don't possess deep binding pockets which can shelter substrates from the solvent and other kinetically active molecules. As was discussed previously, the strength of association at a blunt-end of DNA is, a priori, at most half that of intercalation. Thus, future capping compounds will require either: (a) modifications to EtBr derivatives to improve their association with DNA (such as a groove binding tail or covalent attachment to DNA), or (b) a new model based on a different intercalator.

Although there was no association between **5** and blunt-ended DNA, interesting data were still produced from fluorescence studies of **5**. The conversion of Yellow **5** to Orange **5** seemed to be a result of π-π stacking to form multimers. On the one hand, aggregation via π-π stacking seems plausible enough, since the nonpolar aromatic regions of **5** would probably want to associate under aqueous conditions. But on the other hand, **5**'s positive charge and its bulky *t*-butyl acetyl substituents would both be a hindrance to

aggregation. Although modeling hasn't yet been done to support or refute this, it would seem that steric effects wouldn't play a large role in preventing aggregation, as **5** could probably stack in a skewed fashion such that the "head" of one molecule was above the "tail" of the other.

Some of our data support the hypothesis that **5** forms multimers in solution. The conversion from Yellow **5** to Orange **5** was shown to be reversible. If covalent modifications were occurring (other than protonation or deprotonation), the reaction would not be reversible. The fluorescence spectra generated at different pH's didn't converge at an isosbestic point, meaning that there were more than two species present in solution. If our hypothesis is correct, then **5** might not only be forming dimers, but might also be forming trimers, tetramers, etc.

Secondly, increasing the ionic strength was shown to convert Orange **5** back to Yellow **5**. Because Yellow **5** is charged, the ionic atmosphere around it increases as the ionic strength increases. Therefore, the higher the ionic strength, the greater the number of ions around Yellow **5**, and the more difficult it becomes to form multimers. In our case, the ionic strength was raised by adding Tris and EDTA, which are two charged organic molecules. These might have interacted slightly with the monomer, especially EDTA, which is negatively charged.

The stacking model was tested with 9-ACA, a molecule which was predisposed to stack with **5**. The stacking of 9-ACA with **5** resulted in fluorescence quenching and an apparent shortening of the fluorescence lifetime compared to Yellow **5**. This is in stark contrast with Orange **5**, which had a greater fluorescence intensity (at a different wavelength) and a longer fluorescence lifetime. The most likely explanation of these results is that the negative 9-ACA and Yellow **5** formed an ion pair and precipitated out of solution, thereby reducing the fluorescence. In this case, the shorter τ must have resulted from random error. However, it is also possible that the 9-ACA / **5** aggregate remained in solution, but was only capable of fluorescence at 420 nm with a reduced intensity. In this

case, the shorter τ could have been a real measurement. As for the Orange **5** aggregate, it is possible that **5** / **5** dimers become more conjugated through π-π stacking, which then allows them to fluoresce at longer wavelengths and with longer lifetimes compared to the monomer.

However, some of our data did not support our stacking hypothesis. For example, when the concentration of Orange **5** was decreased by a factor of about 100, there was no shift to the Yellow **5** form. We expected a shift to the monomer, because aggregation is normally concentration dependent. As the concentration decreases, molecules of **5** encounter other molecules of **5** much less frequently, and so the amount of aggregation decreases. Therefore, it was surprising to find that there was no shift to the monomer. In addition, the pH dependence of **5** can't be explained, because there is no obvious way to protonate or deprotonate **5**. One hypothesis was put forward that the amide at position 3 might be able to lose its hydrogen, thereby making a neutral molecule. But then, one would expect aggregation to occur at high pH (when **5** would be neutral), not low pH's. In addition, if **5** ever did deprotonate, it would become highly insoluble in water. We never noticed a precipitate or cloudiness in our solutions; however, **5** was always in the μM range and only a small amount of precipitate would be expected to form.

In conclusion, there is much more research that needs to be done on DNA end-capping. The large number of variables involved in this research make it likely that a useful capping compound will eventually be discovered, but at the same time an amazing amount of research will be required to get there. For example, some of the variables that could be modified in this research are: the type of intercalator, the type of groove binding tail, the method of increasing the association with DNA, and the method of blocking intercalation. Although this research created more questions than it answered, it hopefully provided some guidance for future work.

References

1) N.E. Geacintov, T. Prusik, J.M. Khosrofian, *J. Am. Chem. Soc.* **98**, 6444-6452 (1976).

2) S.C. Jain, H.M. Sobell, *J. Biomol. Struct. Dyn.* **1**, 1179-1194 (1984).

3) J.M. Veal, W.D. Wilson, *Journal of Biomolecular Structure and Dynamics* **8**, 1119-1145 (1991).

4) T. Lybrand, P. Kollman, *Biopolymers* **24**, 1863-1879 (1985).

5) D.M. Crothers, *Biopolymers* **6**, 575-584 (1968).

6) J. Feigon, W. Leupin, W.A. Denny, D.R. Kearns, *Nucleic Acids Res* **10**, 749-762 (1982).

7) N.H. Hopkins, J.W. Roberts, J.A. Steitz, J.D. Watson, A.M. Weiner, Molecular Biology of the Gene, 4th Ed, 248-255 (1987).

8) D.A. Skoog, J.L. Leary, Principles of Instrumental Analysis, 4th Ed, 174-183 (1992).

9) E.F. Gale, E. Cundiffe, P.E. Reynolds, M.H. Richmond, M.J. Waring, The Molecular Basis of Antibiotic Action, 2nd Ed, Wiley, NY, 1981.

10) C.G. Reinhardt, T.R. Krugh *Biochemistry* **17**, 4845-4854 (1978).

11) D.P. Heller, C.L. Greenstock *Biophys. Chem.* **50**, 305-312 (1994).

12) W.A. Denny, S.J. Twigden, B.C. Baguley, *Anti-Cancer Drug Design* **1**, 125-132 (1986).

13) MacroModel V5.0: F. Mohamadi; N.G.J. Richards; W.C. Guida; R. Liskamp; M. Lipton; C. Caufield; G. Chang; T. Hendrickson; W.C. Still; J. Comput. Chem., **1900**, *11*, 440.

14) S. Nonin, J.L. Leroy, M. Gueron, *Biochemistry* **34**, 10652-10659 (1995).

Appendix 1

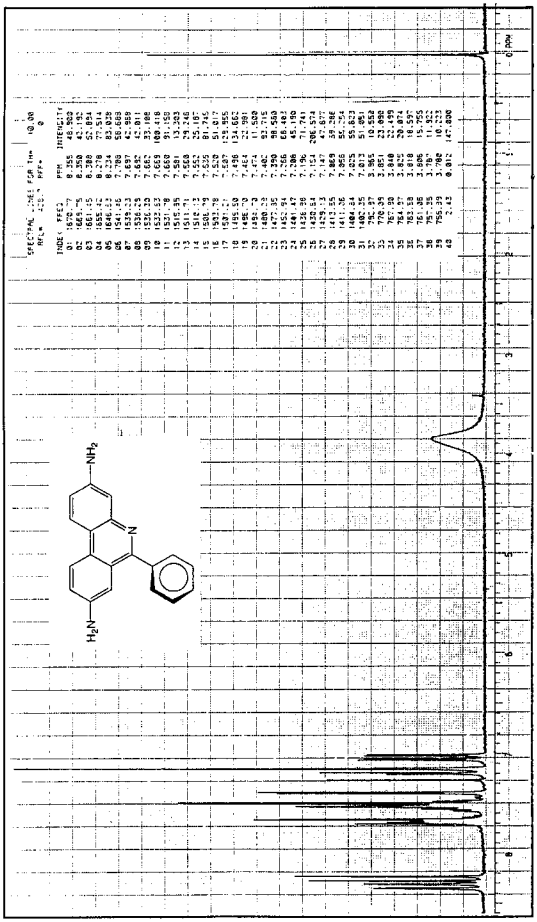


Figure 15. ^{1}H-NMR spctrum of **2** in $CDCl_3$

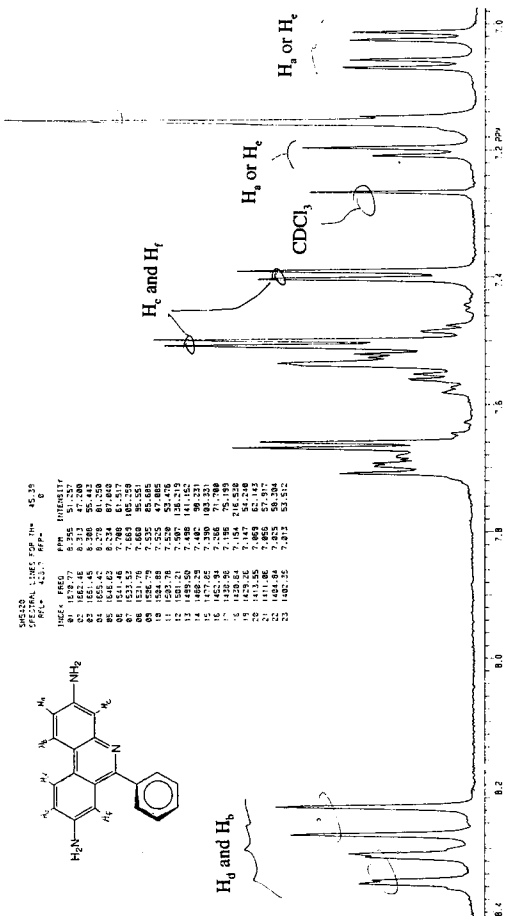


Figure 16. ^{1}H-NMR spectrum of **2** in $CDCl_3$; aromatic region

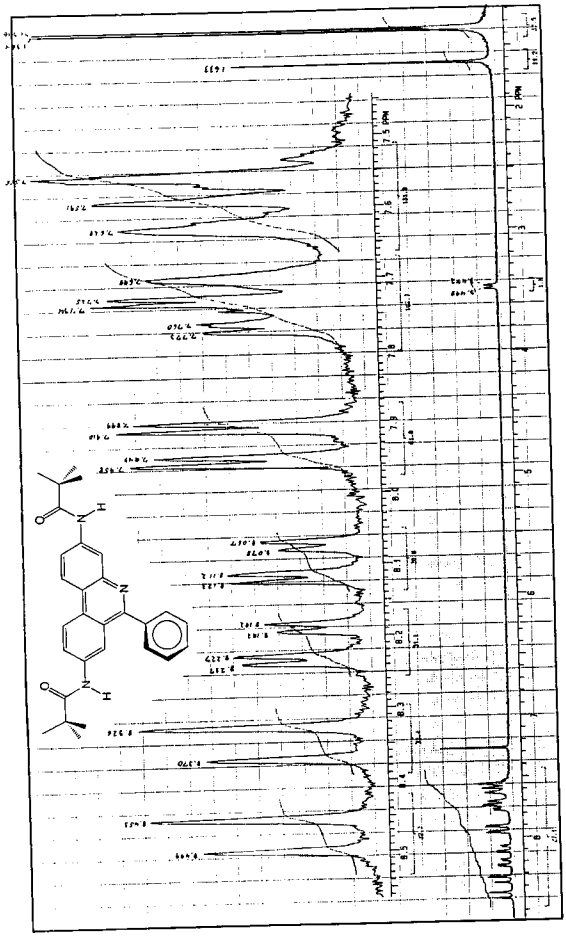


Figure 17. ^{1}H-NMR spectrum of **6** in $CDCl_3$

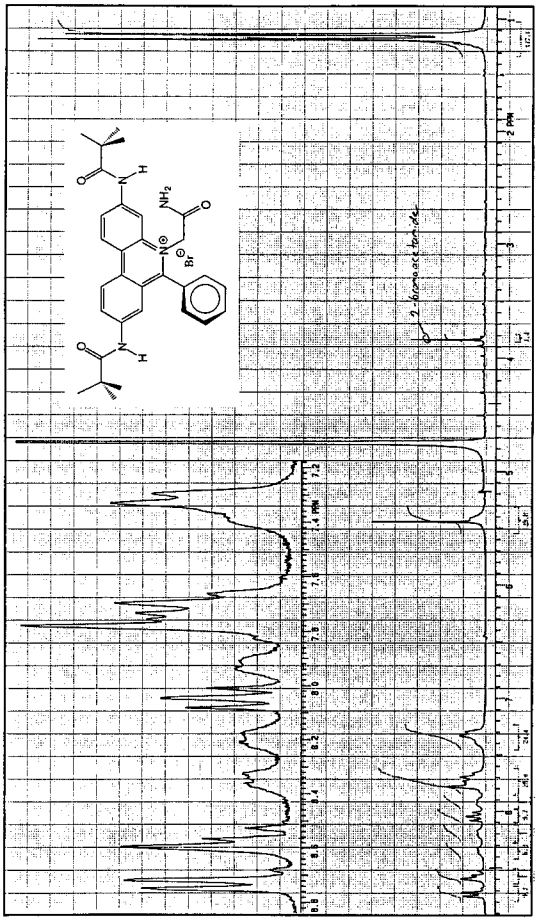


Figure 18. ^{1}H-NMR spectrum of **5** in D_2O

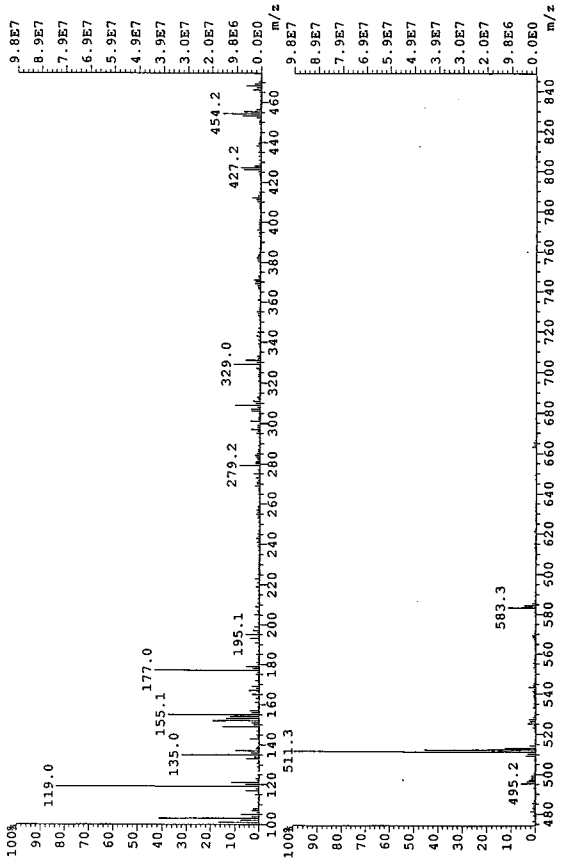


Figure 19. Mass spectrum of **5**