I. MOLECULAR RECOGNITION OF NUCLEIC ACID BY BMSP.
II. SEQUENCE SPECIFIC B → H → A ALLOSTERIC TRANSITIONS IN DNA

Thesis by

Michael McClellan Becker

In Partial Fulfillment of the Requirements
For The Degree Of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1982

(submitted March 27, 1981)

ACKNOWLEDGEMENTS

I wish to thank Debbie Chester for typing and coordinating the thesis, Heather Marr and Lilian Casler for artwork, Pat Bullard and Peter Dervan for countless administrative favors, Delmer Dill for constructing the competition dialysis cells, Ray Robbins and David Millar for fluorescence lifetime measurements, Kelly Tatchell and Ken Van Holde for nucleosomal DNA, Robert Watson for closed circular PM2 phage DNA, Dan Strauss for helical extension measurements, Professors Norman Davidson, Richard Dickerson, Michael Raftery and Ahmed Zewail for providing laboratory facilities, Horace Drew for many stimulating discussions of DNA structure, Richard Scheller for constant advice, support, and encouragement and Professors John Hopfield and Norman Davidson for critical comments on the constant X binding isotherm technique.

I am also deeply indebted to my research advisor, Peter Dervan, for his enthusiasm and support and to my undergraduate research advisor, Douglas Raber, for enabling me to come to Caltech. Financial support by the National Institutes of Health, Biomedical Research Support Funds, Eli Lilly Company and the American Cancer Society, as well as the receipt of an NIH predoctoral traineeship are also gratefully acknowledged.

ABSTRACTS

Chapter One

Bis(methidium)spermine (BMSp), a dimer of two intercalating monomers of ethidium bromide (EB) connected by a spermine link has been synthesized and characterized. The results of these studies clearly demonstrate that both monomers of BMSp simultaneously intercalate nucleic acid, substantially enhancing both its binding affinity and specificity. Under physiological conditions both the binding affinity and specificity of BMSp are similar to DNA binding regulatory proteins. Thus BMSp represents one of the first rationally designed drugs which may selectively inhibit or alter gene expression.

The binding affinity, binding cooperativity, binding site size, and visible spectrum of BMSp are found to vary with nucleic acid conformation. Both BMSp and EB are shown to bind H and A conformations of nucleic acid much more tightly than B conformations. As a result, both compounds induce sequence specific B → H → A allosteric transitions in DNA.

Chapter Two

A sensitive experimental technique which can accurately estimate equilibrium binding isotherms is described. Ligand-macromolecule interactions are monitored by classical indirect techniques over a broad ratio of L_T/M, where L_T is the total ligand concentration and M is the macromolecule concentration. When analyzed at constant X, where X is some physical property of ligand which is proportional to its concentration, the dependence of L_T on M can be used to estimate the binding densities and free ligand concentrations characterizing the ligand-macromolecule interaction. In contrast to classical indirect and direct techniques, accurate binding isotherms can be estimated over a wide range of binding densities for tightly or weakly bound ligands. When the binding of ethidium bromide to polyd(C-G) and polydCdG is examined by this technique, previously undetected allosteric transitions are revealed.

Chapter Three

Evidence for a new conformational family of Watson-Crick DNA is presented. Termed H or hybrid DNA, such DNA is postulated to be an intermediate in the interfamily B → A transition. Hybrid DNA is characterized by a 2'-endo (3' → 5') 3'-endo alteration in sugar pucker every base pair and may also be an intermediate in DNA melting, DNA kinking, and drug intercalation. The ease with which DNA undergoes a B → H → A transition is found to vary greatly with its sequence. On the basis of these results, the equilibrium stability, rather than the structure of Watson-Crick DNA is postulated to vary greatly with base sequence.

LIST OF ABBREVIATIONS

BMSp - bis(methidium)spermine

EB - ethidium bromide

M^+ - monovalent cation concentration

ν, r - binding density (bound drug/base pair)

L_F, C_F - free ligand concentration

MSp - mono(methidium)spermine

TABLE OF CONTENTS

CHAPTER I

Bis(methidium)spermine, A Polyintercalating Molecule Exhibiting High Affinity and Specificity for Nucleic Acid

It is widely recognized that nucleic acids constitute the molecular basis of life. As a result, drugs which interfere with nucleic acid functions can exhibit profound pharmacological properties. During the last 40 years such drugs have found important applications in both cancer[1] and antibiotic chemotherapy,[2] and in addition have served as highly useful probes of nucleic acid structure[3] and function.[4]

Since the double helix description of DNA in 1953,[5] a large mass of data regarding both the static and dynamic structure of nucleic acids has emerged. As a result, nucleic acids now represent one of the few biologically important receptors whose characterization is of sufficient detail as to encourage a rational approach to drug design. Many drugs which bind to nucleic acids do so by a process called intercalation, the insertion of a flat aromatic molecule between the base pairs of nucleic acids.[6] Since intercalating drugs bind individual subunits (base pairs) of the nucleic acid polymer, substantial increases in affinity and specificity are anticipated for compounds which are capable of <u>simultaneously</u> inserting two or more intercalating drugs into the double helix.[7] In the absence of unfavorable entropic or steric constraints, polyintercalators[7,8] may bind with a free energy approaching the sum of the

individual free energy contributions of the monomeric intercalating drugs. Since monomeric intercalators bind nucleic acid with significant specificity and affinity, ($K \sim 10^5$) polyintercalators are expected to bind with affinities and specificities approaching those of regulatory proteins, thereby greatly enhancing their biological activity as well as their usefulness as biological probes.

We report the nucleic acid binding properties of bis(methidium)spermine (BMSp), a polyintercalator consisting of two monomers of ethidium bromide (EB) (Fig. 1) Ethidium bromide was chosen because of its potent biological

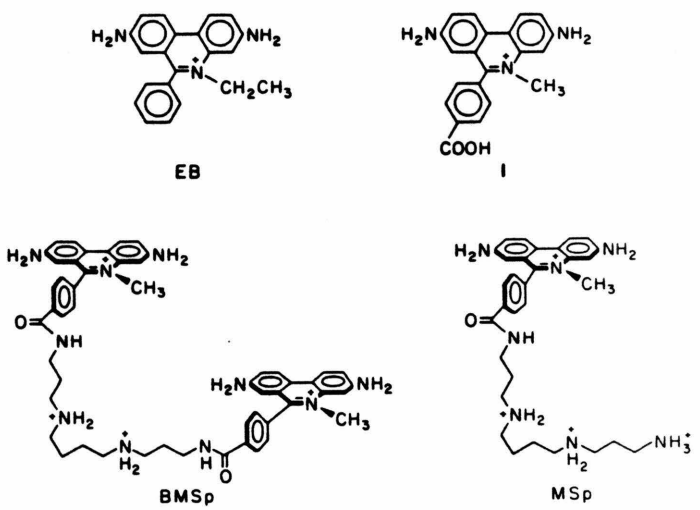


Figure 1
Structures of bis(methidium)spermine (BMSp), monomethidium)spermine (MSp), ethidium bromide (EB), and p-carboxyl methidium chloride (1).

properties[9] and its widespread use as a probe of nucleic acid structure[3] and function.[4] The synthetic strategy employed preserves the major structural attributes of the ethidium monomer as an intercalator. The tetramine, spermine, was chosen to link the intercalators because of its known affinity for nucleic acid[10] and its length, which allows a geometry sufficient to reach non-adjacent intercalation sites in accordance with a nearest neighbor exclusion binding mode.[11] (Fig. 2)

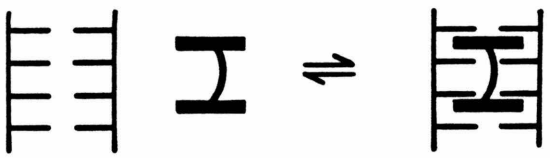


Figure 2
Nearest neighbor intercalation of DNA by a bis-intercalator.

For a bisintercalated species the polyamine linker should lie intimately in the groove of the DNA helix. Structural modifications of any linker with respect to charge, chirality, length, flexibility, and functionality are expected to play an important role in controlling the stability and the nature of these polyintercalator-nucleic acid complexes. Comparison of the binding

properties of BMSp and EB obtained under nearly identical conditions clearly demonstrates that both EB moieties of BMSp simultaneously intercalate nucleic acid, substantially enhancing both its binding affinity and specificity.

Synthesis

BMSp was prepared by condensing 0.5 equivalents of spermine with the acylimidazole activated ester of p-carboxylmethidium chloride, 1. Activation of compound 1 followed by condensation with spermine in excess yielded the corresponding monointercalator, mono(methidium)-spermine (MSp) (Fig. 1). Compound 1 was prepared in six steps from O-amino-biphenyl by modification of previous techniques (Fig. 3). Nitration of O-aminobiphenyl,[12] condensation with p-cyanobenzoyl chloride and cyclization yielded 6-(4-cyanophenyl)-3,8-dinitrophenantridine. Successive methylation, hydrolysis and reduction afforded maroon crystals of p-carboxylmethidium chloride in an overall yield of 16%.[13] The infrared and NMR spectra of compound 1 were identical with those of an authentic sample.[13]

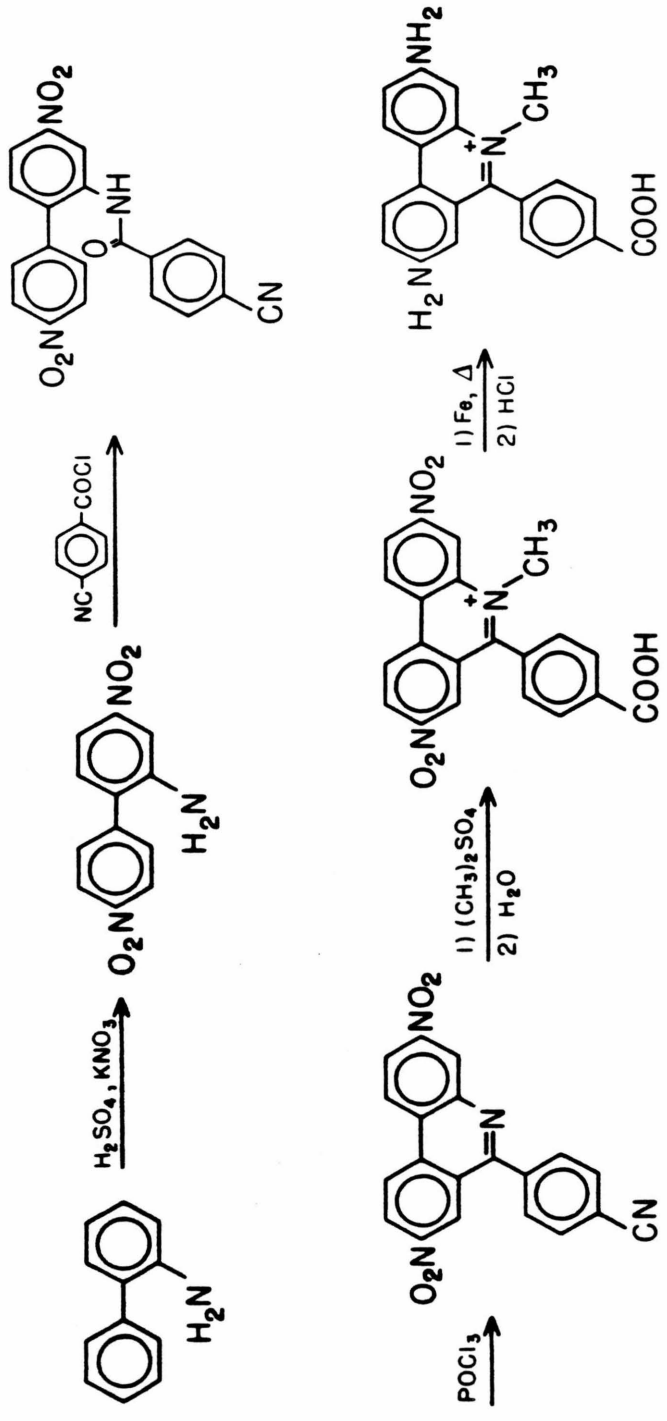


Figure 3
Synthesis of p-carboxylmethidium chloride from 2-aminobiphenyl.

Physical Properties

BMSp is an extremely water soluble (0.1 M) tetracation at pH 7. Under physiological conditions both chromophores of BMSp are electronically coupled into a dimer state since the maximum of its visible absorbance, relative to the spermine analog MSp, is red-shifted 17 nm (Fig. 4). Enhancing protonation of the spermine linker of BMSp by lowering the pH to 3 weakens this dimer as reflected by a hyperchromic shift to lower wavelengths of its visible maximum (ε_{max} = 9850, 520 nm, pH 7; ε_{max} = 10,200, 500 nm, pH 3.0). Decreasing the pH further protonates the phenanthridine chromophores, further disrupting the dimer (ε_{max} = 9510, 496 nm, pH 2.1). Although the maximum of BMSp's visible absorbance does not shift with a change in ionic strength, an increase in ionic strength does markedly decrease the extinction coefficient of this maximum (Fig. 5).

Taken together, the above optical measurements indicate that in solution BMSp exists as an intramolecularly stacked dimer. Whereas decreases in pH apparently disrupt this dimer, changes in salt alter its stacking geometry. Presumably, electrostatic repulsions between phenanthridine chromophores and the spermine linker are responsible for this behavior. Although the equilibrium constant between stacked and un-

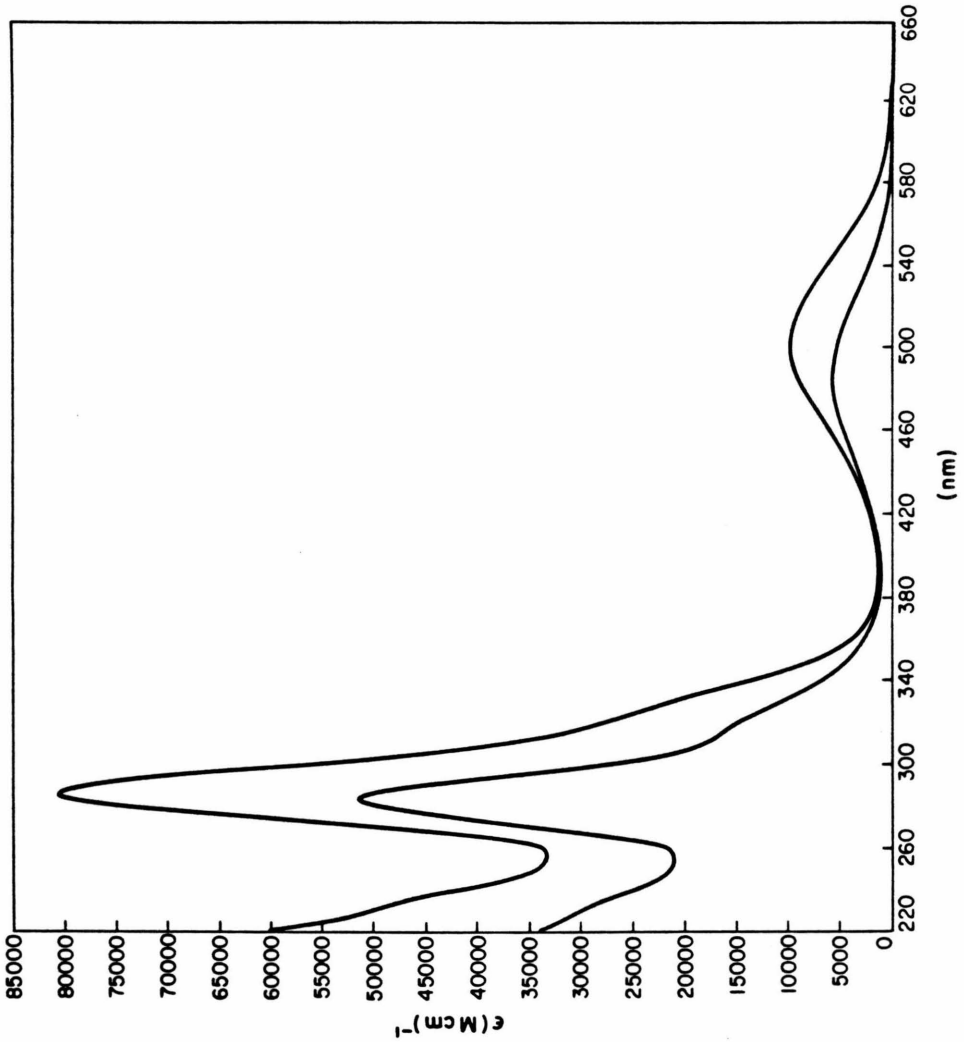


Figure 4
Comparison of the absorption spectrum of BMSp (upper curve) and MSp at M^+ = 0.075 M.

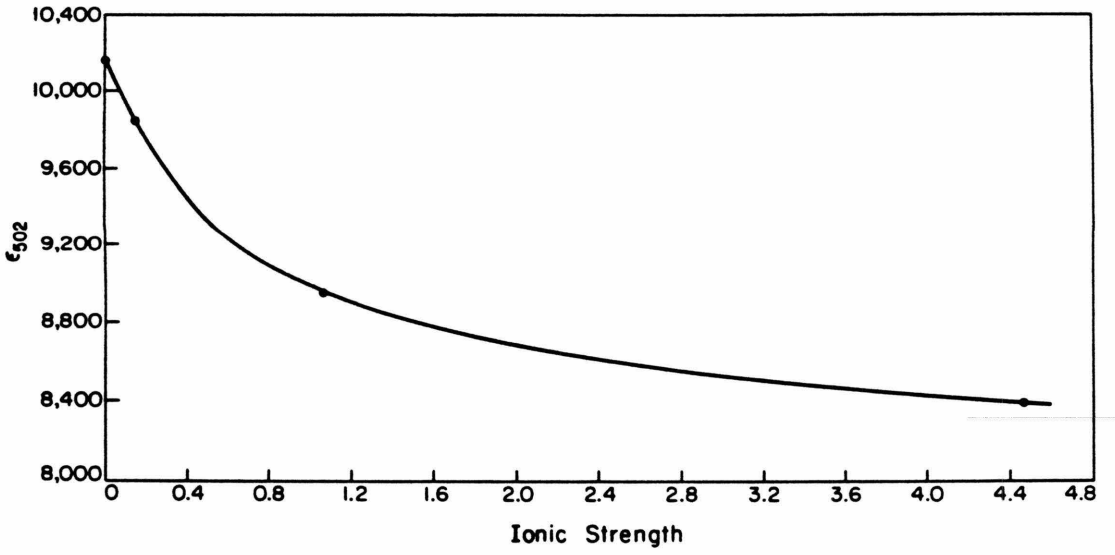


Figure 5
Dependence of the extinction coefficient of unbound BMSp on ionic strength at 502 nm.

stacked forms of BMSp is unknown, under identical conditions EB dimerizes with an equilibrium constant of 950 (Fig. 6). Unlike EB, we do not observe any optical evidence for intermolecular stacking by BMSp, intramolecular association apparently being more favorable.

Although the flourescence emission spectrum of BMSp is also red-shifted (638 nm) relative to EB (603 nm), this behavior results from incorporation of amide linkages in BMSp since MSp also exhibits the same red-shift. The visible excitation fluorescence spectra of BMSp, EB, and MSp appear identical.

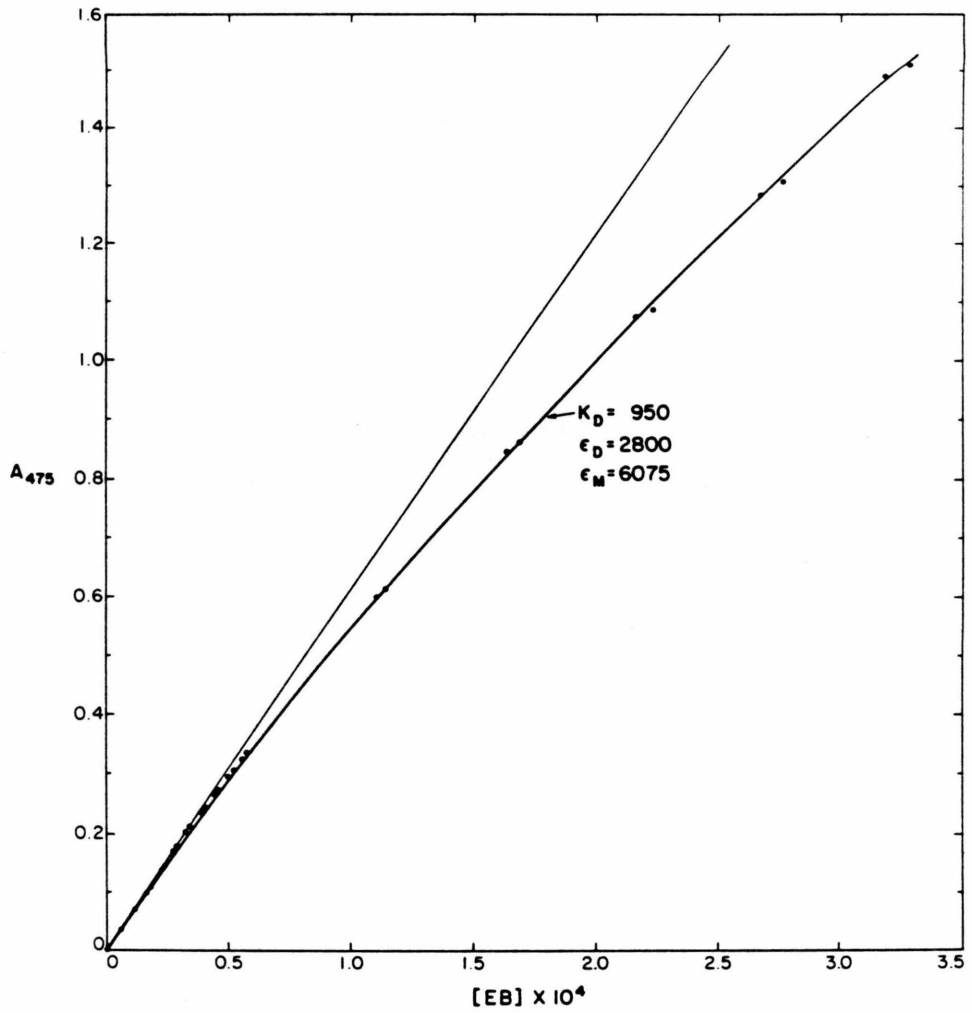


Figure 6
Dimerization of ethidium bromide at $[M^+]$ = 1.0 as determined by absorption spectroscopy. The curved line is the best fit to the experimental data using the indicated extinction coefficients and dimerization constant.

Binding of BMSp to Sonicated Calf Thymus DNA at Low Salt

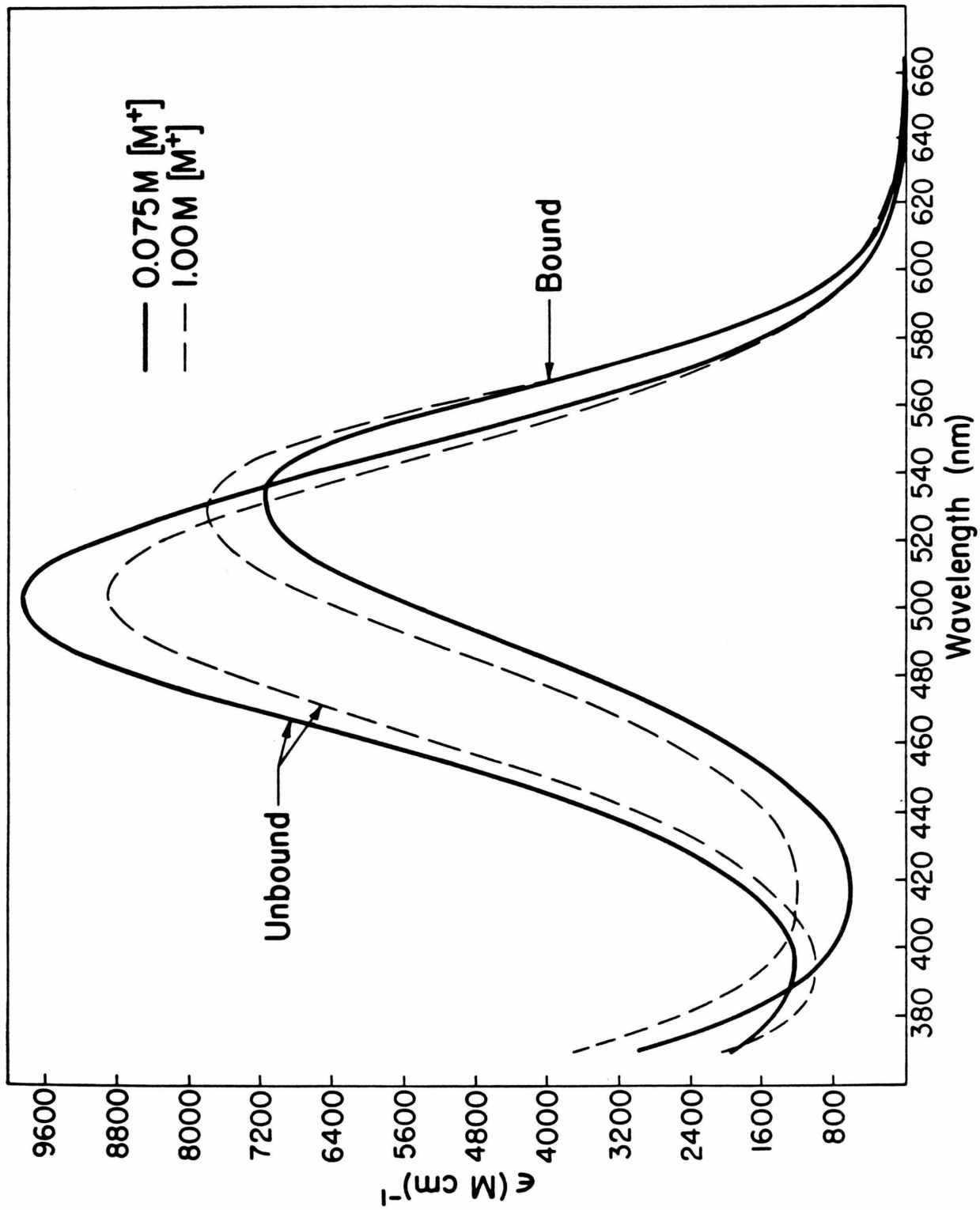
The interaction of BMSp with nucleic acid is expected to involve electrostatic as well as non-electrostatic interactions. We have therefore investigated its interaction with nucleic acid under conditions of both low ($M^+ = 0.075$) and high salt ($M^+ = 1.0$). To reduce the effect of sequence dependence, we first investigated BMSp binding to calf thymus DNA which had been sonicated to a double strand length of 200 base pairs. The following hydrodynamic, spectrophotometric, and fluorescence data indicate that under low or high salt conditions, both chromophores of BMSp prefer to simultaneously intercalate nucleic acid. Although the intercalation geometry is found to be modified by electrostatic interactions between the spermine linker and the DNA helix under conditions of low salt, little or no distortion of the double helix is induced upon binding.

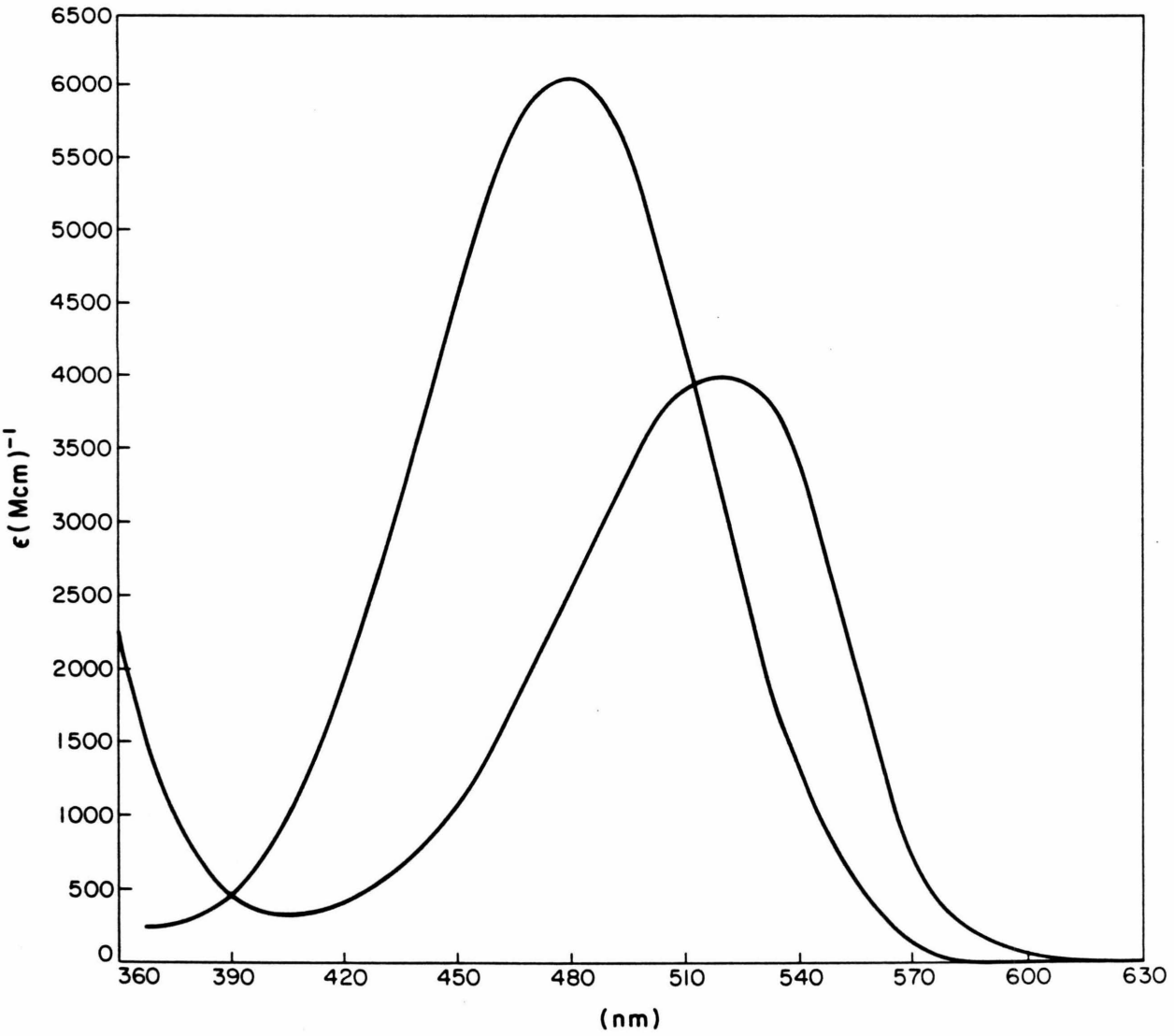
The binding of BMSp to calf thymus DNA can be monitored by absorption and fluorescence spectroscopy because, like ethidium, a metachromic shift[14] (Fig. 7) and quantum yield enhancement[15] result when BMSp intercalates nucleic acid. The binding of BMSp to calf thymus DNA as monitored by absorption spectroscopy at 490 nm, a wavelength where the extinction coefficients of bound

Figure 7

(A) Comparison of the visible spectrum of BMSp unbound and bound to sonicated calf thymus DNA at low (0.075 M^+) and high (1.0 M^+) salt.

(B) Comparison of the visible spectrum of unbound EB (upper curve) and EB bound to sonicated calf thymus DNA (lower curve) at 1.0 M^+.


0.075M [M+]
1.00M [M+]
Bound
Unbound
Wavelength (nm)
ε (M cm)-1
380 400 420 440 460 480 500 520 540 560 580 600 620 640 660
800 1600 2400 3200 4000 4800 5600 6400 7200 8000 8800 9600


6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
ε(Mcm)⁻¹
360
390
420
450
480
510
540
570
600
630
(nm)

and unbound BMSp differ most, is shown in Fig. 8. The change in absorbance (ΔA_{490}) is linear up to a BMSp/base pair ratio (BMSp/BP) of 0.25 (region A). After BMSp/BP = 0.25 there is a sharp break in the observed ΔA_{490}. The observation that the slope in this region is greater than that exhibited by unbound BMSp but less than that exhibited in region A reflects the appearance of (an) additional BMSp species distinct from those formed in region A.

The binding of BMSp to calf thymus DNA as monitored by fluorescence spectroscopy is shown in Fig. 9. The increase in the fluorescence of BMSp in the presence of DNA (I_1) minus the fluorescence of an equivalent solution of BMSp in the absence of DNA (I_0) is plotted against the ratio BMSp/BP.[15] We find in agreement with the spectrophotometric titration, that there are at least two bound forms of BMSp. A highly fluorescent complex[15] is formed for BMSp/BP ratios up to 0.25 and the fluorescence of this species is diminshed by additional bound forms for BMSp/BP ratios > 0.18.

The bound species observed for BMSp/BP ratios up to 0.25 is attributed to a <u>bisintercalated</u> species on the basis of its binding stoichiometry and spectroscopic properties. The metachromic shift induced upon

Figure 8

Spectrophotometric titration of sonicated calf thymus DNA with BMSp. The concentration of DNA was 1.726×10^{-6} M in base pairs (BP) and the results of two separate titrations are shown for Region A (BMSp/BP $\leq$ 0.25).

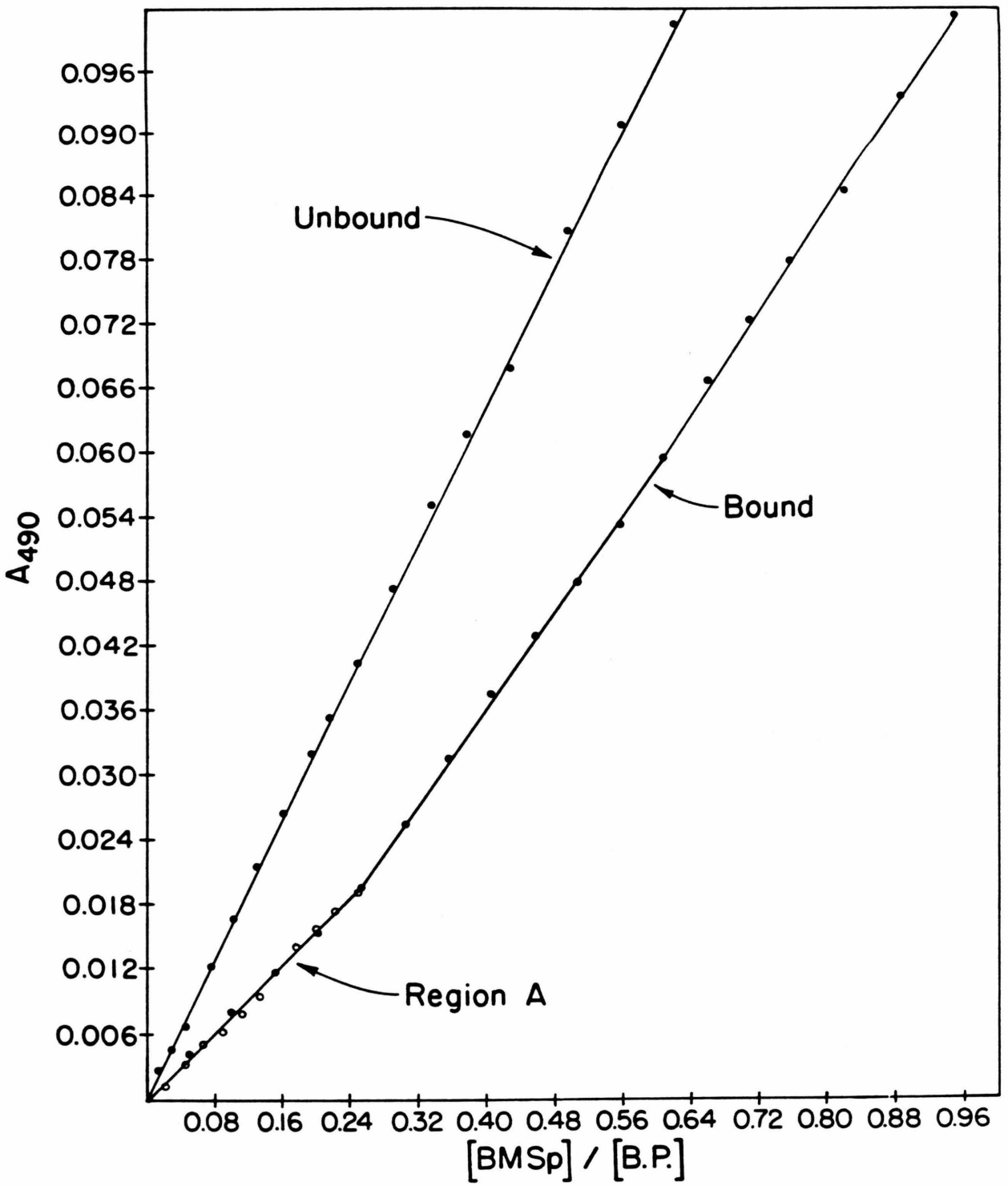
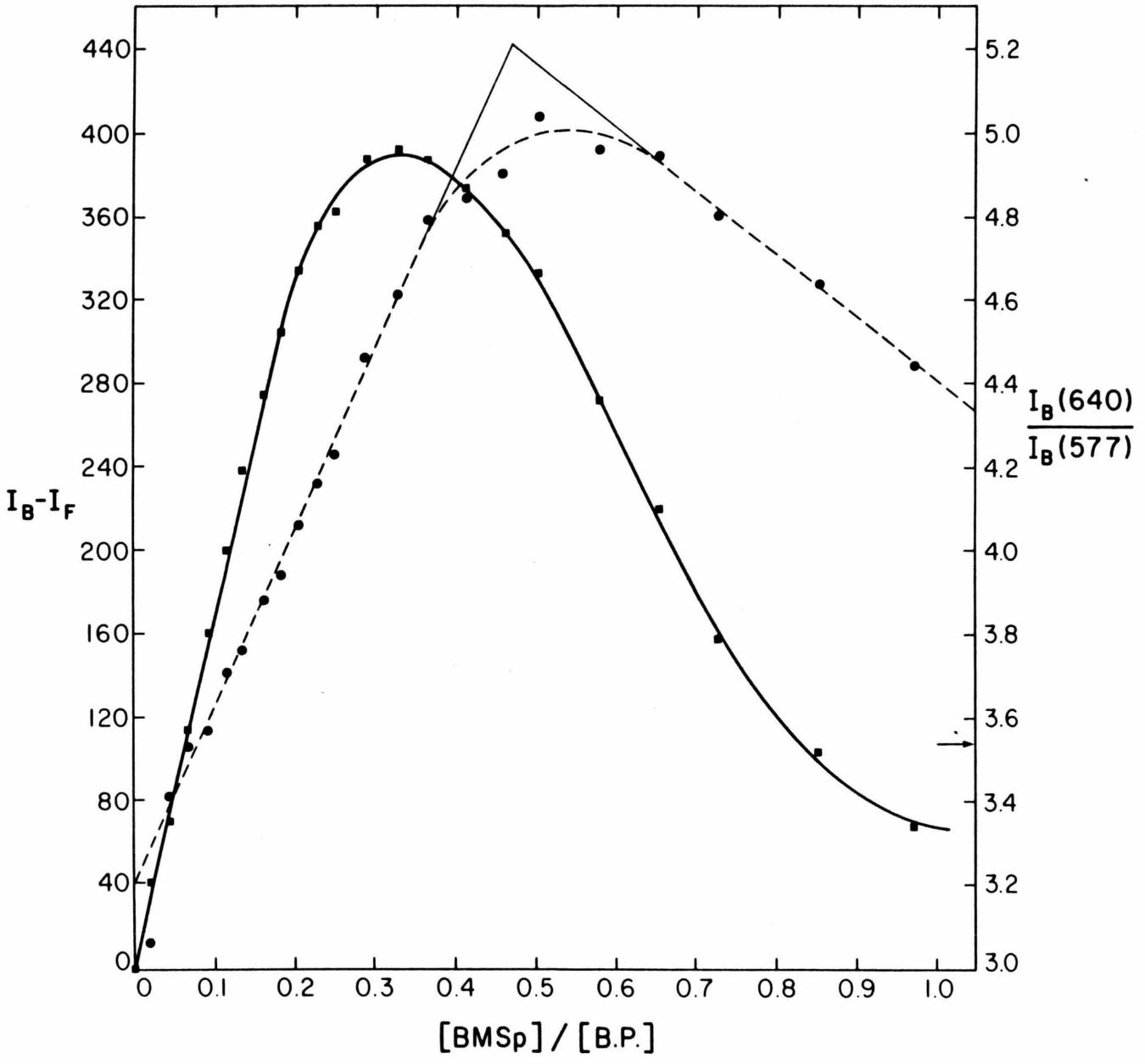

0.096
0.090
0.084
0.078
0.072
0.066
0.060
0.054
0.048
0.042
0.036
0.030
0.024
0.018
0.012
0.006
A490
Unbound
Bound
Region A
0.08 0.16 0.24 0.32 0.40 0.48 0.56 0.64 0.72 0.80 0.88 0.96
[BMSp] / [B.P.]

Figure 9

Fluorescence titration of sonicated calf thymus DNA (1.726×10^{-6} M BP) with BMSp at $[M^+] = 0.075$. The increase in the fluorescence of BMSp in the presence of DNA (I_B) minus the fluorescence of an equivalent solution of BMSp in the absence of DNA (I_F) is plotted (solid line) against the ratio BMSp/BP. The ratio of the fluorescence emission intensity at 640 nm relative to the same emission intensity at 577 nm (dashed line) is also plotted against the ratio BMSp/BP. The value of this emission ratio for unbound BMSp is indicated by the arrow.


440
400
360
320
280
240
200
160
120
80
40
0
$I_B - I_F$
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
$\frac{I_B(640)}{I_B(577)}$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
[BMSp] / [B.P.]

binding is found to be identical in magnitude to the metachromic shift experienced by EB when intercalated, a finding which is consistent with intercalation of both chromophores of BMSp into the DNA helix (Fig. 7). The large enhancement of the quantum yield and fluorescence lifetime of this bound BMSp species, (Table I) is also consistent with bisintercalation based on the

BMSp	Q_B/Q_F	τ(nsec)
Unbound; 1 M^+	—	2.81 ± 0.02
Calf thymus; 0.075 M^+	42 ±1	—
Calf thymus; 1.0 M^+	—	15.8 ± 0.2*
d(C-G); 1.0 M^+	80	—
dAdT; 1.0 M	—	14.7 ± 0.2*

*Determined at a binding density of 2.5×10^{-3} bound BMSp per base pair.

known mechanism of fluorescent enhancement fo ethidium by DNA.[16] Moreover, the stoichiometry observed from absorption measurements, one BMSp per four base pairs, is identical with the stoichiometry that would be predicted for a bisintercalated species based on the known

stoichiometry of ethidium (one ethidium/two BP)[17] and in accordance with the nearest neighbor exclusion model.[18]

In addition to a quantum yield and fluorescence lifetime enhancement, the fluorescence emission of BMSp red-shifts when it binds calf thymus DNA (Fig. 9). This behavior is reflected by an increase in the emission measured at 640 nm relative to that measured at 577 nm. The emission red-shifts linearly with increasing saturation of the DNA lattice up to a BMSp/BP ratio of 0.18. With further addition of BMSp the emission ceases to red-shift and eventually begins to reverse its red-shift. This behavior can be attributed to two effects: a red-shift up to a BMSp/BP ratio of 0.5 and a reversal of this red-shift for BMSp/BP ratios greater than 0.5. A red-shift reflects interaction between identical chromophores which couple to form a state of lower energy (coupling is intermolecular as opposed to intramolecular since it depends on the degree of lattice saturation). This red-shift is attributed to intercalative binding since, by rigidly holding each chromophore parallel to one another, the DNA helix allows efficient electronic coupling between BMSp chromophores. The reversal of this red-shift detected at higher binding density is then attributed to nonintercalatively bound

species, which by virtue of residing in a different molecular environment, transfer rather than exchange energy with intercalated chromophores.

Although BMSp is expected to bisintercalate up to a BMSp/BP ratio of 0.25, the observation that the fluorescent enhancement of BMSp decreases for BMSp/BP ratios greater than 0.18 indicates that above this value a second form of BMSp also binds DNA. This species is apparently intercalated since the emission of BMSp can continue to red-shift up to a BSMp/BP ratio of 0.5. Since the theoretical saturation limit of a monointercalated species is also 0.5, it may be suggested that BMSp bisintercalates up to a BMSp/BP ratio of 0.18 after which point it also begins to monointercalate. The appearance of monointercalated BMSp can be attributed to entropic constraints imposed by the DNA lattice. For example, as shown in Fig. 10, when the BMSp/BP ratio approaches 0.166, gaps between bound BMSp molecules will be too small for bisintercalation but large enough for monointercalation. Since further bisintercalation would require a redistribution of bound ligands, the apparent affinity of a bisintercalation mode decreases, enabling monointercalation modes to compete more effectively for available binding sites.

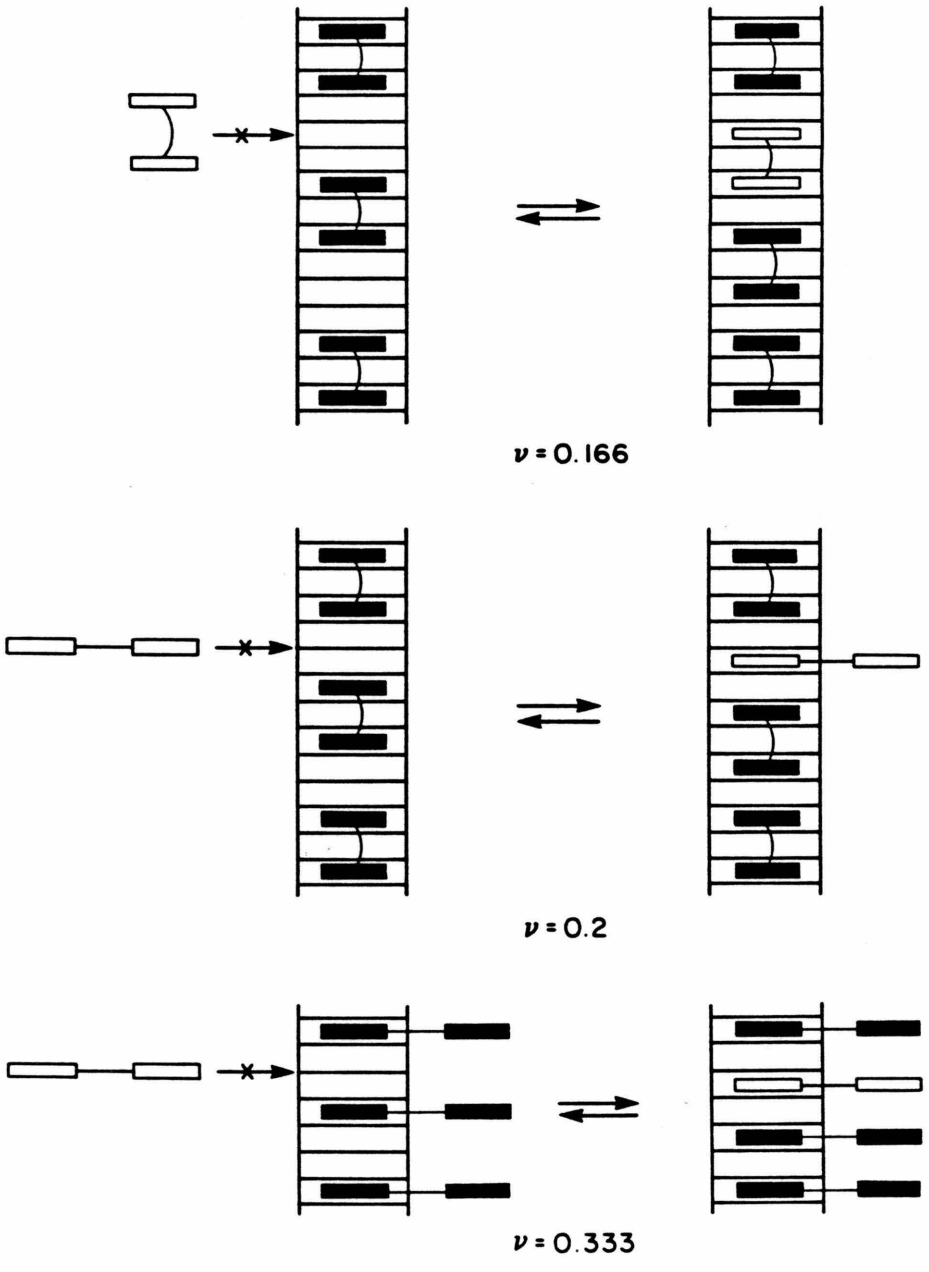


Figure 10
Change in the binding mode of BMSp with increased binding density, ν([bound BMSp]/[base pair]).

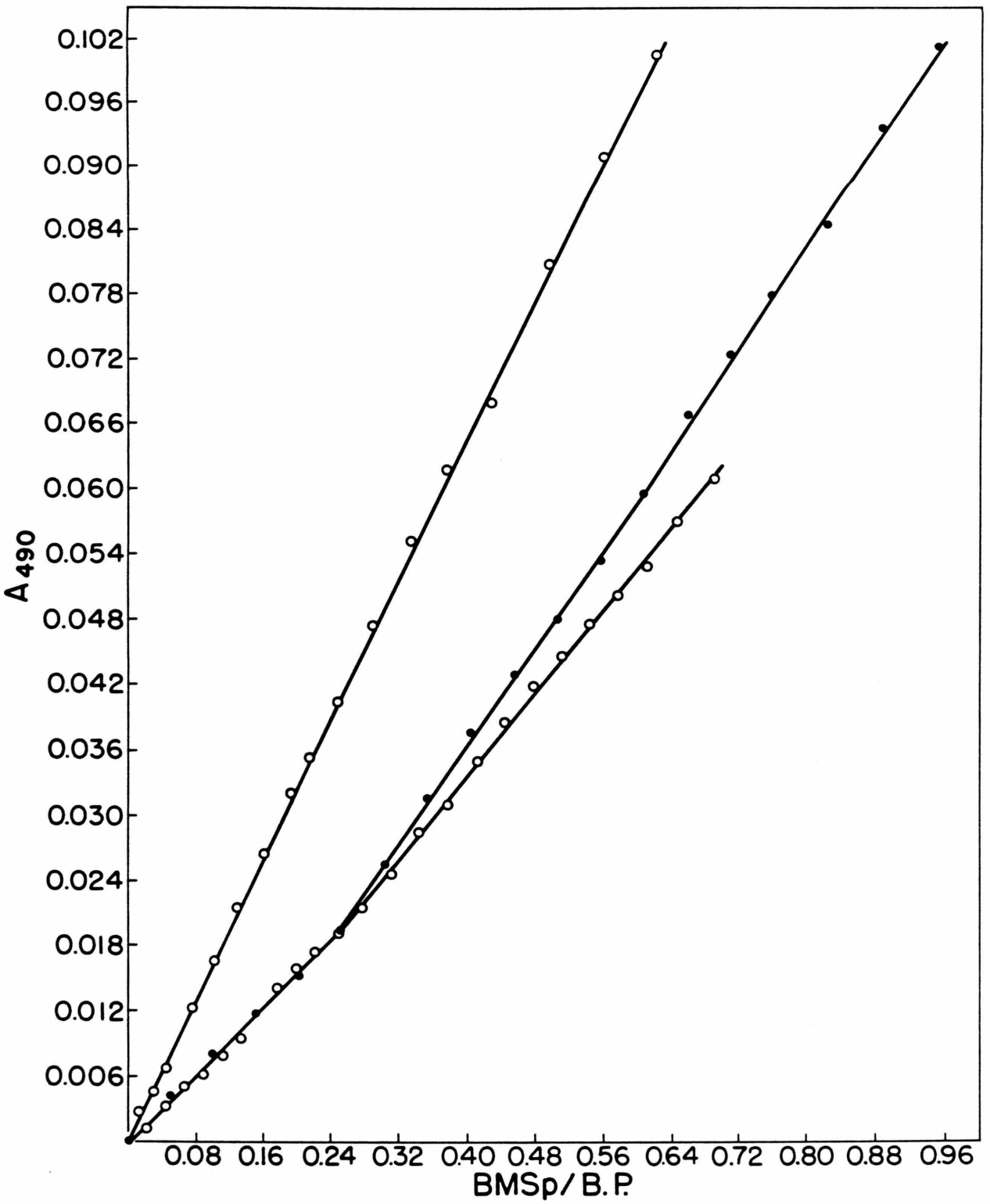
The emission of BMSp continues to red-shift up to a BMSp/BP ratio of 0.37 after which point a reversal in the red-shift occurs. Commensurate with this reversal is a decrease in the fluorescence of bound BMSp. Both the reversal of the red-shift and the decreased fluorescence of bound BMSp are consistent with a nonintercalated binding mode for BMSp for a BMSp/BP ratio of 0.37 to approximately 1. Owing to a larger binding site size (one versus four base pairs) nonintercalated species displace previously intercalated molecules, thereby decreasing the observed fluorescence.[19]

The appearance of nonintercalatively bound species can also be attributed to entropic constraints imposed by the DNA lattice since as the BMSp/BP ratio approaches 0.333, gaps between monointercalated BMSp molecules will be too small to prevent further intercalation without ligand redistribution (Fig. 10). As a result, outside binding begins to compete effectively with intercalative binding.

When monointercalated, one chromophore of BMSp will be free to either bind DNA in a nonintercalated fashion or extend freely into solution. In the latter case, intermolecular crosslinking of DNA helices would be possible. We have observed that after a BMSp/BP ratio

Figure 11

Spectrophotometric titration of sonicated calf thymus DNA with BMSp. The concentration of DNA was 1.726×10^{-6} M and the results of two titrations are shown. Since the time elapsed between measurements is the same in both titrations ($\sim$ 30 mins), the time required to reach any BMSp/BP ratio is twice as long for the titration represented by open circles. The time dependence observed for a BMSp/BP ratio greater than 0.25 apparently results from a slow BMSp induced precipitation of DNA.


0.102
0.096
0.090
0.084
0.078
0.072
0.066
0.060
0.054
0.048
0.042
0.036
0.030
0.024
0.018
0.012
0.006
A490
0.08 0.16 0.24 0.32 0.40 0.48 0.56 0.64 0.72 0.80 0.88 0.96
BMSp/ B. P.

of 0.25, a slow time dependent decrease in the absorbance of bound BMSp takes place (Fig. 11). Since this decrease in absorbance is accompanied by a BMSp induced precipitation of DNA, crosslinking of DNA molecules is suggested. Although monointercalated species begin to form at a BMSp/BP ratio > 0.18, they apparently do not crosslink until a BMSp/BP ratio of 0.25 since we observe no evidence for DNA precipitation below this value. Presumably, the unintercalated chromophore can bind to the outside of the DNA helix when BMSp/BP < 0.25 but is forced to extend into solution when this ratio is greater than 0.25.

Unwinding and Helical Extension of DNA by BMSp and EB

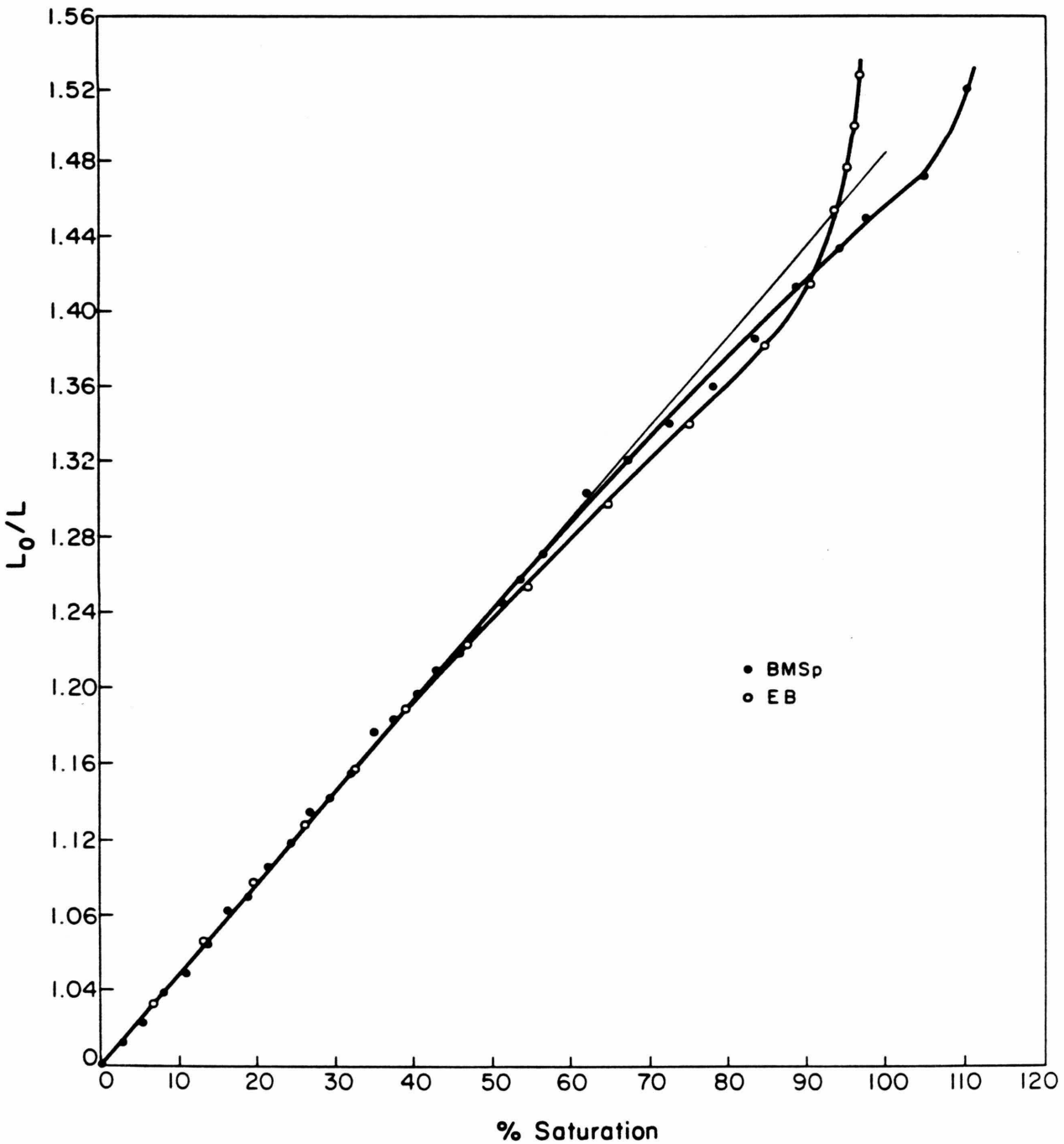
The absorption and fluorescence measurements described above provide indirect evidence that the preferred binding mode of BMSp to a heterogeneous DNA sequence involves simultaneous intercalation of both ethidium chromophores. More direct evidence for this conclusion is provided by helical extension[6,20] and unwinding measurements.[21] If both EB chromophores simultaneously intercalate, BMSp would be expected to extend and unwind DNA twice as much as EB.

The ability of both drugs to increase the length of DNA has been investigated by viscometry. Since viscosity increases are associated with both lengthening and stiffing of DNA upon intercalation,[20,22] we chose as our DNA for these studies DNA isolated by nuclease digestion of H-1 depleted chicken erythrocyte chromatin.[23] Because this DNA is short (150 BP) and exhibits a low incidence of single strand nicks (< 2%), stiffening effects are minimized thus allowing a less biased estimation of length extension.

For these small DNA molecules, the approach of Cohen and Eisenberg[24] is expected to accurately relate viscosity increases to helical extension. If it is assumed that intercalation of the double helix corresponds to an increase in the length of the DNA by one base pair, than the quantity L/L_0 (where L is the length of DNA in the presence of drug and L_0 its length in the absence of drug) should increase linearly with intercalation.[24] As shown in Fig. 12, BMSp intercalation extends DNA exactly twice as much as EB (6.6 versus 3.3 Å).

Figure 12

Increase in the length of 150 base pair DNA upon the binding of BMSp and EB. DNA was obtained by micrococcal nuclease digestion of H-1 depleted chicken erythrocyte chromatin. The length of DNA in the presence of bound ligand L_0, divided by its length in the absence of ligand, L, is plotted against the percent saturation of available intercalation sites. EB is assumed to cover two base pairs when bound and BMSp four base pairs. Measurements were obtained by viscometry at 20.0°C and $M^+ = 0.075$.

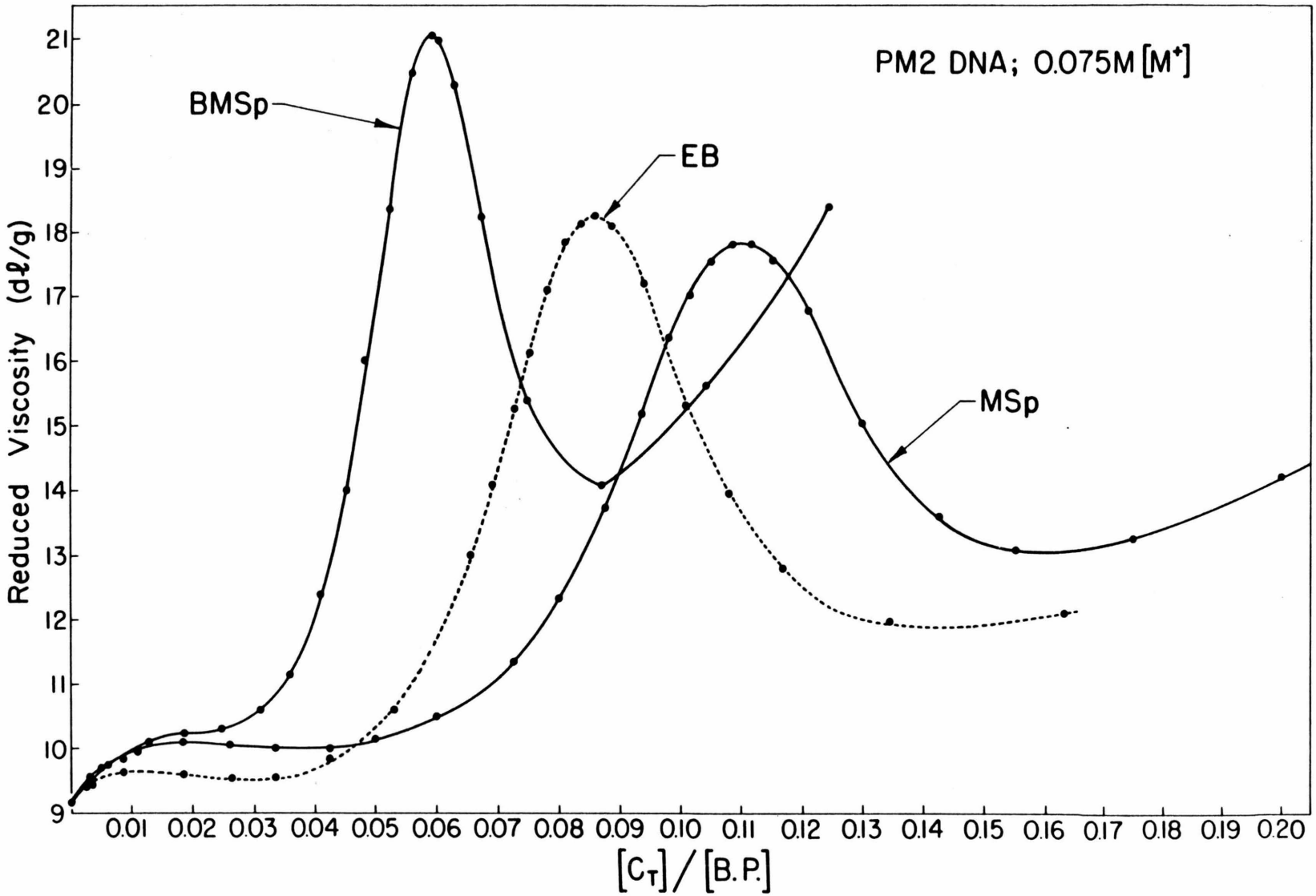

1.56
1.52
1.48
1.44
1.40
1.36
1.32
1.28
1.24
1.20
1.16
1.12
1.06
1.04
0
L_o/L
0
10
20
30
40
50
60
70
80
90
100
110
120
% Saturation
• BMSp
○ EB

In addition to length increases these viscometric measurements may also monitor helix bending, since bending is expected to introduce downward curvature into a L/L_0 plot.[22] Although both EB and BMSp plots curve as saturation of available intercalation sites is approached, this behavior could be attributed to the binding of non- or monointercalated forms which extend DNA less than intercalated forms. Since our fluorescence binding measurements indicate that for BMSp these later forms begin to bind at approximately 70% saturation, examination of Fig. 12 suggests that any bending induced by EB and BMSp, if present, is small.

Additional direct evidence for the simultaneous intercalation of both chromophores of BMSp into DNA comes from helical unwinding measurements. Previous measurements have shown that EB intercalation unwinds DNA, thereby removing and reserving the supercoiling of closed circular DNA.[25] To test for simultaneous intercalation of both BMSp chromophores, viscometric titrations of closed circular PM2 DNA with BMSp and EB at low salt were carried out[26] (Fig. 13). From the observed drug/base pair ratios at the maximum of the titration and the known unwinding angle of EB (26°),[27] the unwinding angle of BMSp is calculated to be $38^{\circ} \pm 1^{\circ}$.[55] Since the value of the unwinding angle of BMSp is only 1.5 times the unwinding of EB, the

Figure 13

Viscometric titration of closed circular PM2 DNA with BMSp, EB, and MSp at $M^+ = 0.075$. The reduced viscosity in deciliter/gram is plotted against the ratio of total ligand added per base pair (C_T/BP). Under the conditions of the experiment, all ligand is bound at the maxima.


PM2 DNA; 0.075M [M+]
BMSp
EB
MSp
Reduced Viscosity (dl/g)
21
20
19
18
17
16
15
14
13
12
11
10
9
0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.10 0.11 0.12 0.13 0.14 0.15 0.16 0.17 0.18 0.19 0.20
[C_T] / [B.P.]

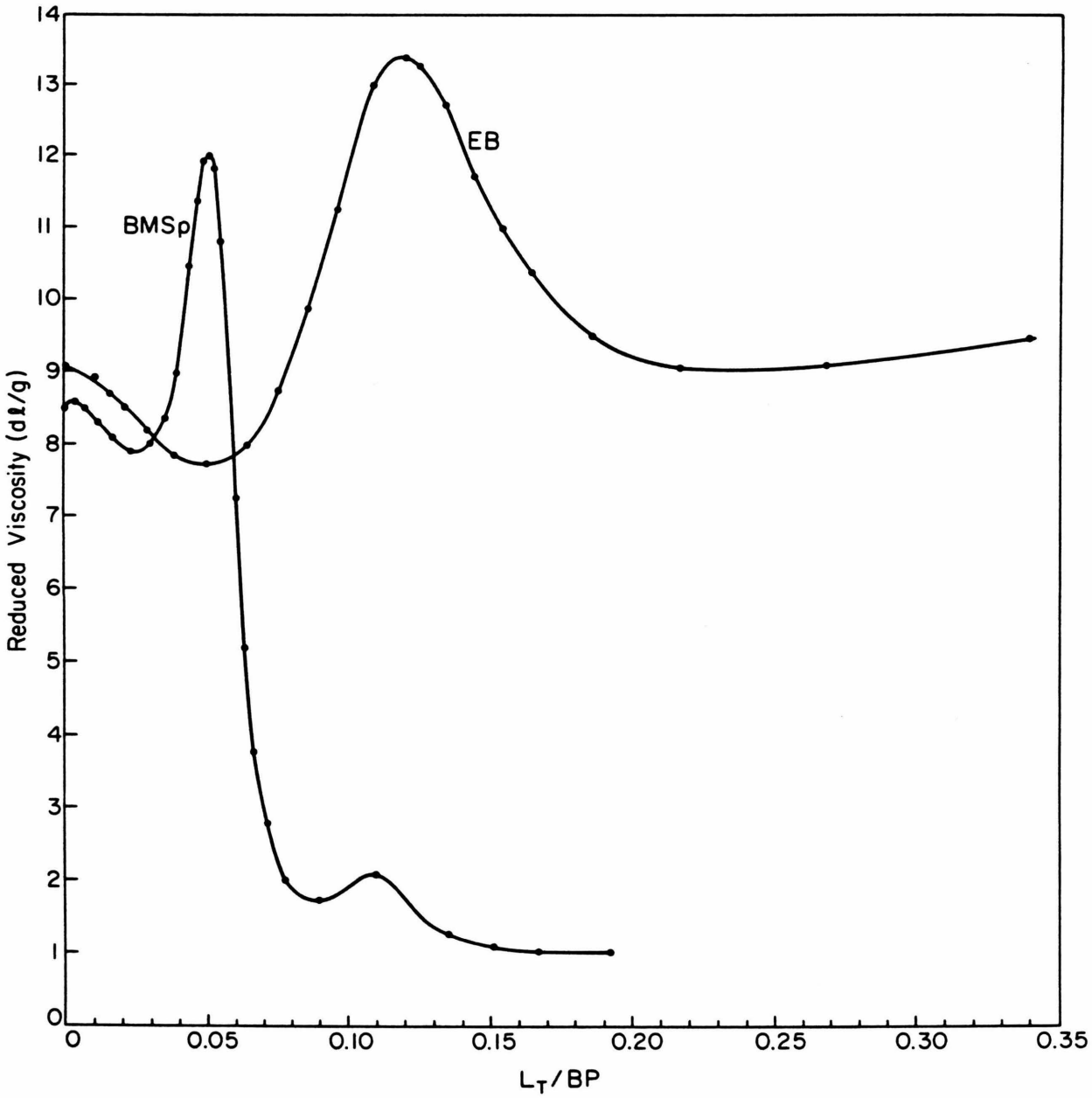
observed unwinding could reflect contributions from mono- and bisintercalated species. However, at these drug/base pair ratios (BMSp/BP < 0.06), our absorption, fluorescence, and helical extension measurements reveal no significant monointercalation component. The low unwinding angle must therefore reflect an altered intercalation geometry induced by low salt.

Two additional experiments demonstrate that electrostatic interactions between the spermine linker of BMSp and the DNA backbone are responsible for the reduction of BMSp's unwinding angle from an anticipated 52^{o} to 38^{o}. First, when the salt concentration is raised from low salt to high salt conditions (1 M Na^{+}), the unwinding angle of BMSp increases from 38^{o} to $49^{o} \pm 3^{o}$ or twice that observed for EB under similar conditions (Fig. 14). Secondly, under conditions of low salt, the monointercalating derivative of BMSp, MSp (Fig. 1), unwinds DNA only (20 ± 2^{o}) or one-half as much as BMSp (Fig. 13).

How is the intercalation geometry of BMSp modified by electrostatic interactions between its spermine linker and DNA under conditions of low salt? Two limiting models may be proposed. In the first model, both chromophores of BMSp intercalate in a fashion identical to EB

Figure 14

Viscometric titration of closed circular PM2 DNA with BMSp and EB at 1.0 M^+. The reduced viscosity in deciliter/gram is plotted against the ratio of total ligand added per base pair concentration.


Reduced Viscosity (dl/g)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
L_T/BP
BMSp
EB

unwinding DNA 52°; however, electrostatic interactions between DNA and the spermine linker then wind the helix 14°. Alternatively, electrostatic interactions between the spermine linker and DNA could modify the intercalation geometry of each chromophore so that each only unwinds 19°. It has been noted experimentally, and demonstrated theoretically, that for DNA at least two major intercalation geometries exist: one which unwinds 26° and another which only unwinds 18°.[28] Apparently, each intercalation geometry extends DNA equally since the helical extensions induced by the binding of proflavin (18° unwinding)[29] and EB (26°) are identical.[30] Since we do not observe a decrease in helical extension in conjunction with the decrease in helical unwinding, a modified (18°) unwinding geometry for both BMSp chromophores at low salt is suggested. As shown in Fig. 7, the visible spectrum of BMSp also appears to sense this change in intercalation geometry with salt. Presumably as the salt concentration is lowered, the orientation of one chromophore relative to the other changes, inducing a change in the transition dipole moment overlap and hence strength of the visible absorbance of BMSp. Similar behavior is exhibited by the intramolecularly stacked dimer of unbound BMSp (Fig. 5).

Since the intercalation geometry of BMSp is expected to vary with nucleic acid conformation (see later discussion), the visible absorption of BMSp may also sense changes in DNA conformation. A comparison of the absorption spectrum of BMSp bound to dAdT and rAdT, which under the conditions of these experiments exist in B and A conformations[31] respectively, lends support to this hypothesis (Fig. 15). The insensitivity of the visible absorption of EB to nucleic acid conformation stresses the importance of the relative orientation of BMSp chromophores in reporting differences in nucleic acid conformation.[40] Although a change in DNA conformation is expected to alter the visible absorbance of BMSp, a change in absorbance does not necessarily imply a change in DNA conformation. For example, although the intercalation geometry and visible absorbance of BMSp bound to sonicated calf thymus DNA changes with salt, DNA is expected to remain in the same conformational family over this range of salt concentration.

For a heterogeneous DNA sequence, the extinction coefficient of the visible maximum of bisintercalated BMSp is found to increase as the unwinding angle is increased. Although the difference in the visible spectrum of BMSp bound to dAdT and rAdT suggests that the later nucleic

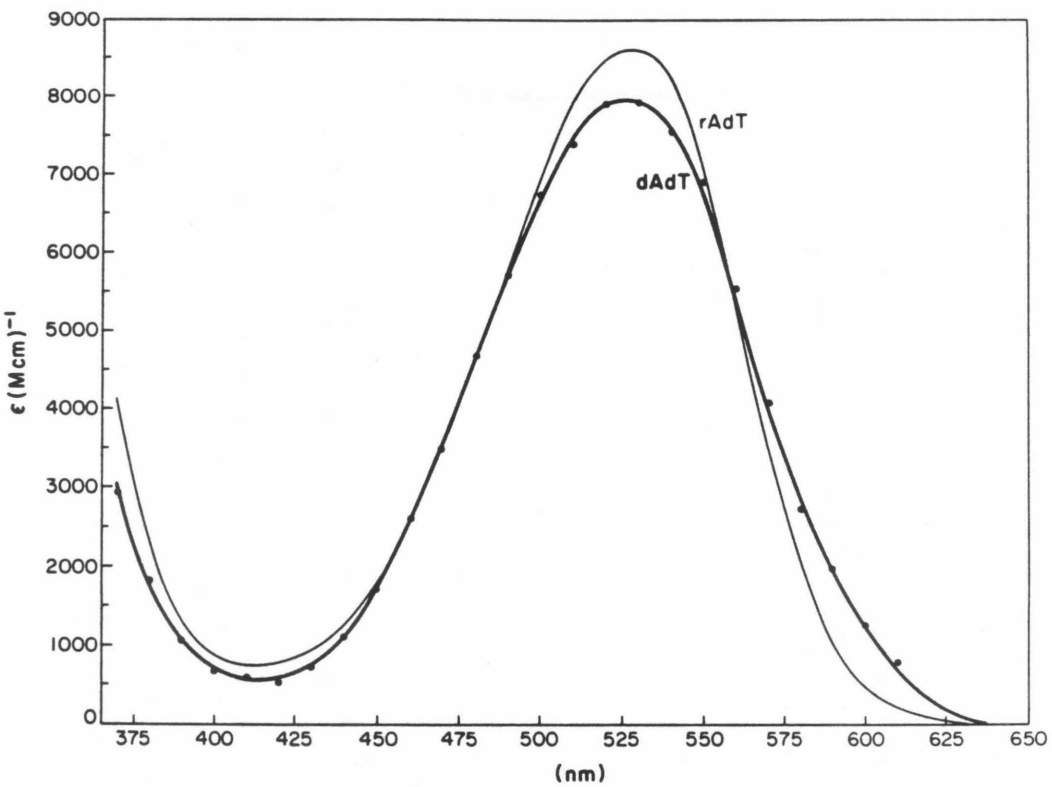


Figure 15

Comparison of the absorption spectrum of BMSp bound to polyrAdT and polydAdT at 1.0 M^+. For polyrAdT, the bound spectrum was obtained at a binding density $\nu < 0.07$ whereas for polydAdT it was obtained at $\nu < 0.007$.

acid may also be more unwound when bound by BMSp, the distance between intercalated chromophores is also different, (Table III) thus precluding a simple correlation.

Binding of BMSp and EB to Nucleic Acid at 1 M^+: Determination of Binding Affinities.

The binding affinity of BMSp and EB to nucleic acid has been measured at a monovalent cation concentration of 1 M because at this concentration electrostatic contributions to the observed binding affinity (and intercalation geometry) are minimized. Thus the nucleic acid affinities measured for BMSp, a tetracation at pH 7, can be directly compared to EB a monocation. This conclusion follows from the polyelectrolyte theory of DNA[32] which postulates that the double helical conformation of DNA is electrostatically unstable and therefore "condenses" counter ions on its surface. The concentration of these "condensed" ions is apparently 1 M and independent of the salt concentration in solution.[32,33] Record has noted that when a charged ligand binds double helical DNA, some of these condensed ions must be lost.[34] Consequently, by mass-action, if the salt concentration in solution is less than 1 M, condensed ion release will enhance ligand binding whereas if it is more than 1 M, it will disfavor ligand binding.

If K_1 is the binding affinity of ligand for DNA at 1 M^+, its binding affinity at another salt concentration, K, is given by:

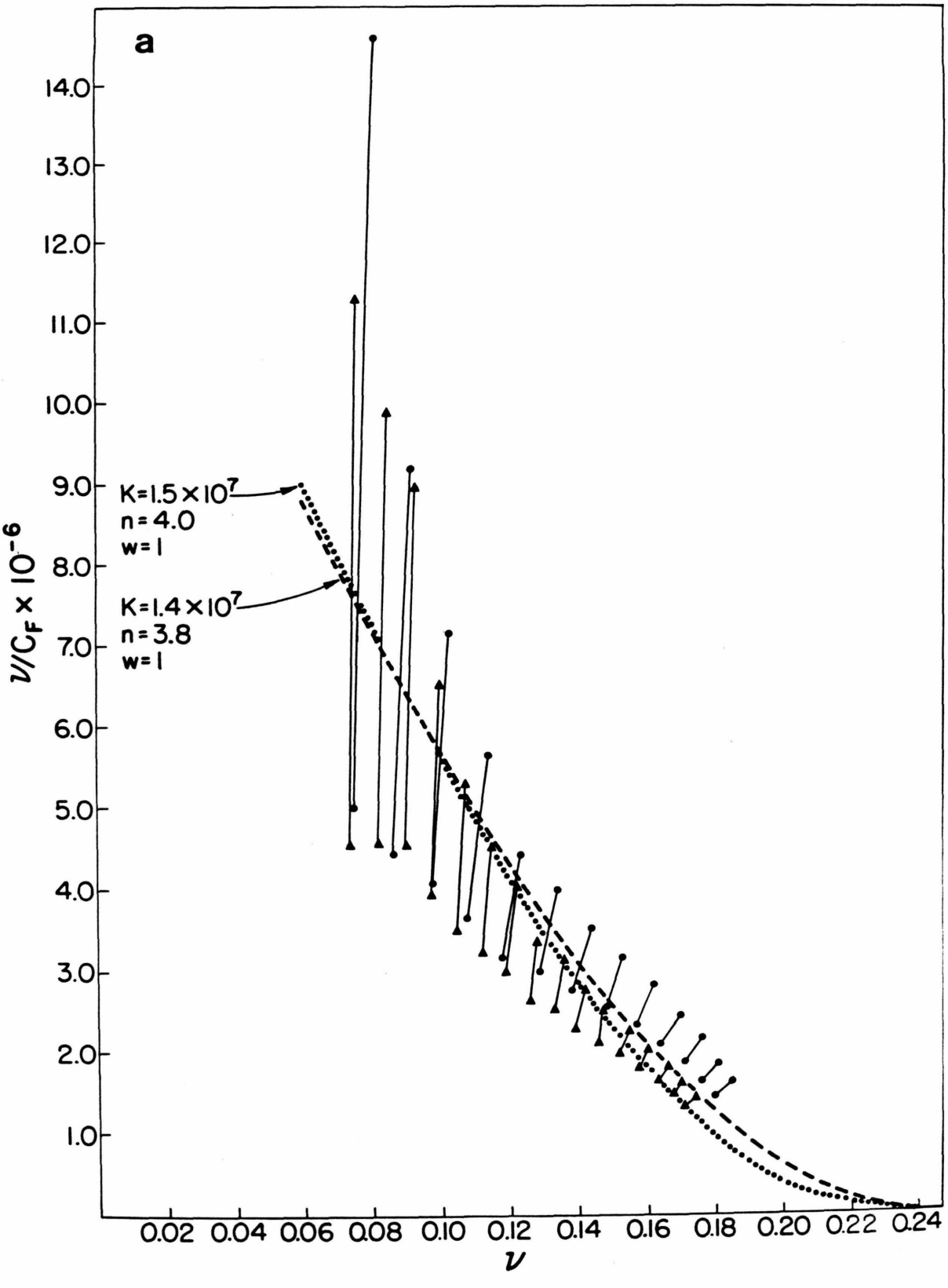
$$K = K_1[M^+]^{-n\psi}$$

where M^+ is the monovalent cation concentration, (n) is the number of ion-pair interactions which ligand makes with nucleic acid, and ψ is the charge density parameter which is known for a wide variety of nucleic acids.[34] This relationship has been shown by LePecq to accurately describe the binding of several mono- and bisintercalators to nucleic acid where n is the number of positive charges of the respective intercalator.[35]

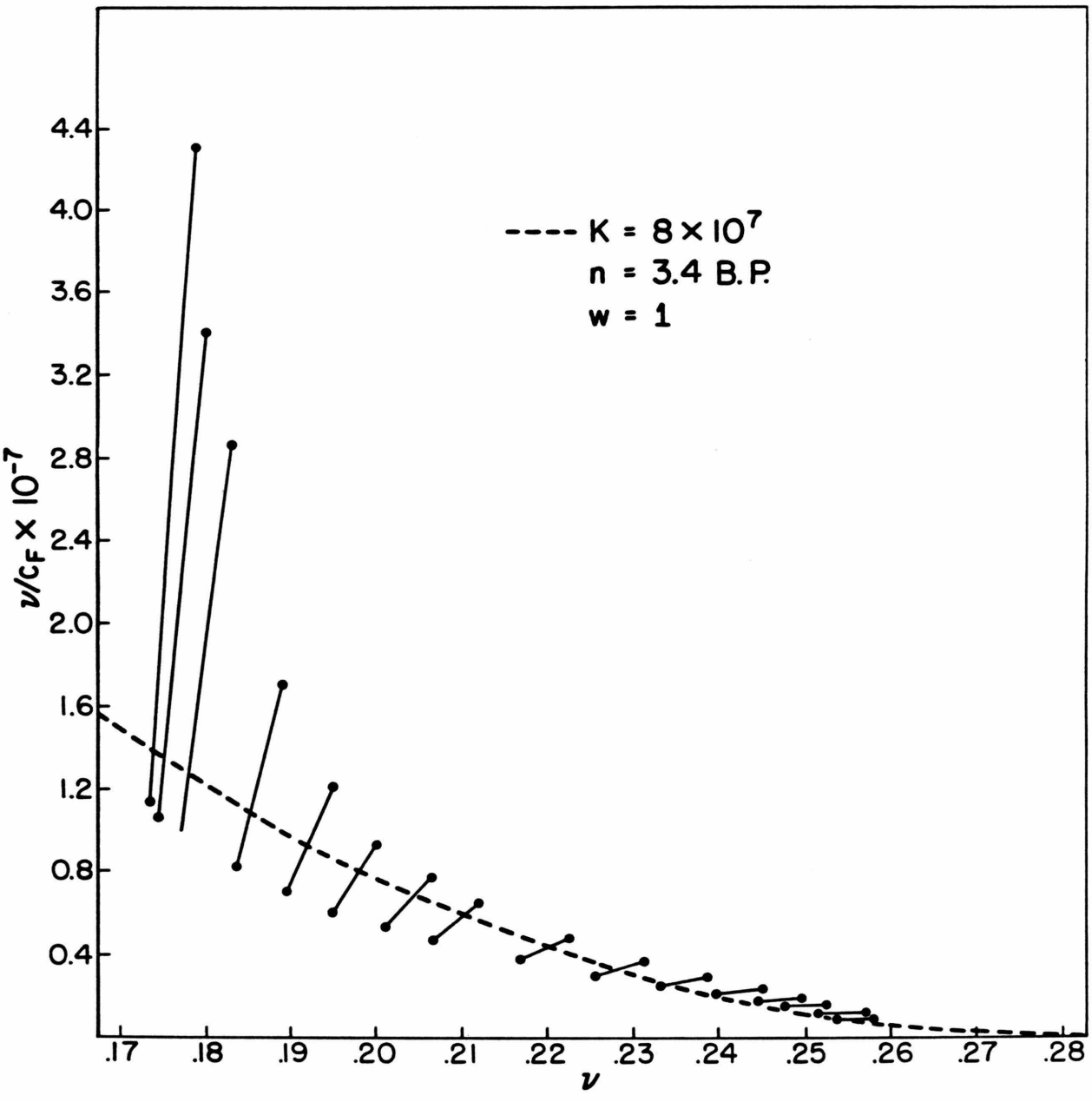
Four different techniques have been utilized to determine the binding affinity of BMSp and/or EB for different nucleic acids. The first two methods take advantage of changes in the absorption[14] and fluorescence properties[15] of BMSp when it binds nucleic acid. Concentrations of bound and free BMSp can be determined by these indirect methods assuming the formation of one bound complex. These data is then plotted in terms of a Scatchard plot.[36] Comparison of the experimentally observed Scatchard plot to theoretical plots generated by the binding equations of von Hippel-McGhee allows estimation of the binding affinity.[37] (Fig. 16)

Figure 16

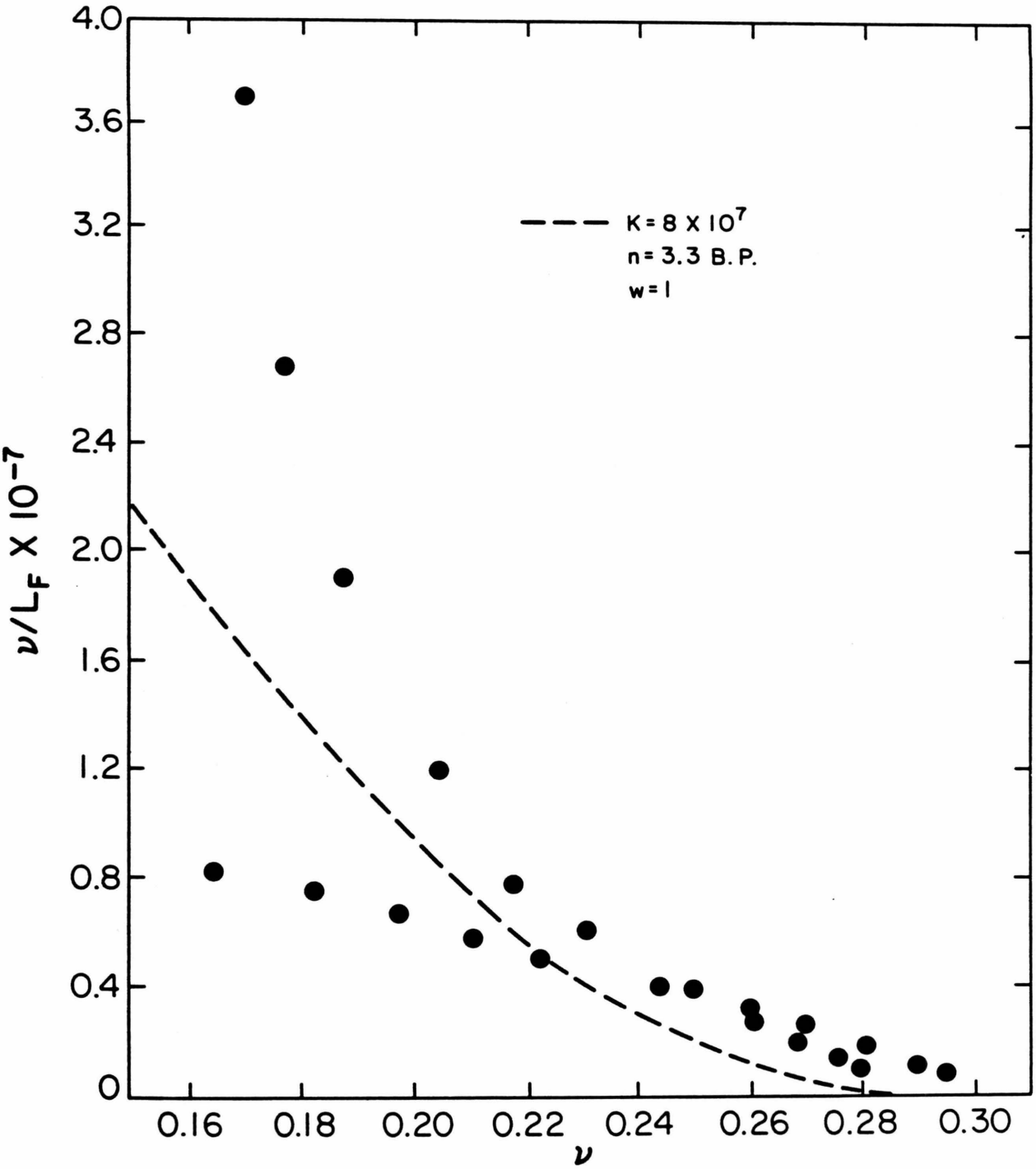
Scatchard plots of the binding of BMSp to (a) sonicated calf thymus DNA[56] and (b) polyd(C-G)[57] at 1 M^+ determined by absorption spectroscopy. The results of two titrations (●—● , ▲—▲) are shown for calf thymus DNA. The binding density, concentration of bound drug per base pair (ν), is plotted against the ratio ν/C_F where C_F equals the concentration of free drug. Dotted lines are theoretical plots generated by the von Hippel-McGhee binding equation for the indicated binding affinity, K, binding site size n, and binding cooperativity ω. For noncooperative binding, $\omega = 1$. (c) Scatchard plot of the binding of BMSp to polyd(C-G) at 1 M^+ as determined by fluorescence spectroscopy. Experimental conditions are the same as in (b) and the results of two titrations are shown. (d) Scatchard plot of the binding of EB to sonicated calf thymus DNA at $[M^+] = 1.0$ determined by absorption spectroscopy. The results of two titrations at DNA concentrations of 1.545 and 5.213 x 10^{-5} BP are shown. Dimerization of free drug (Fig. 6) has been taken into account.

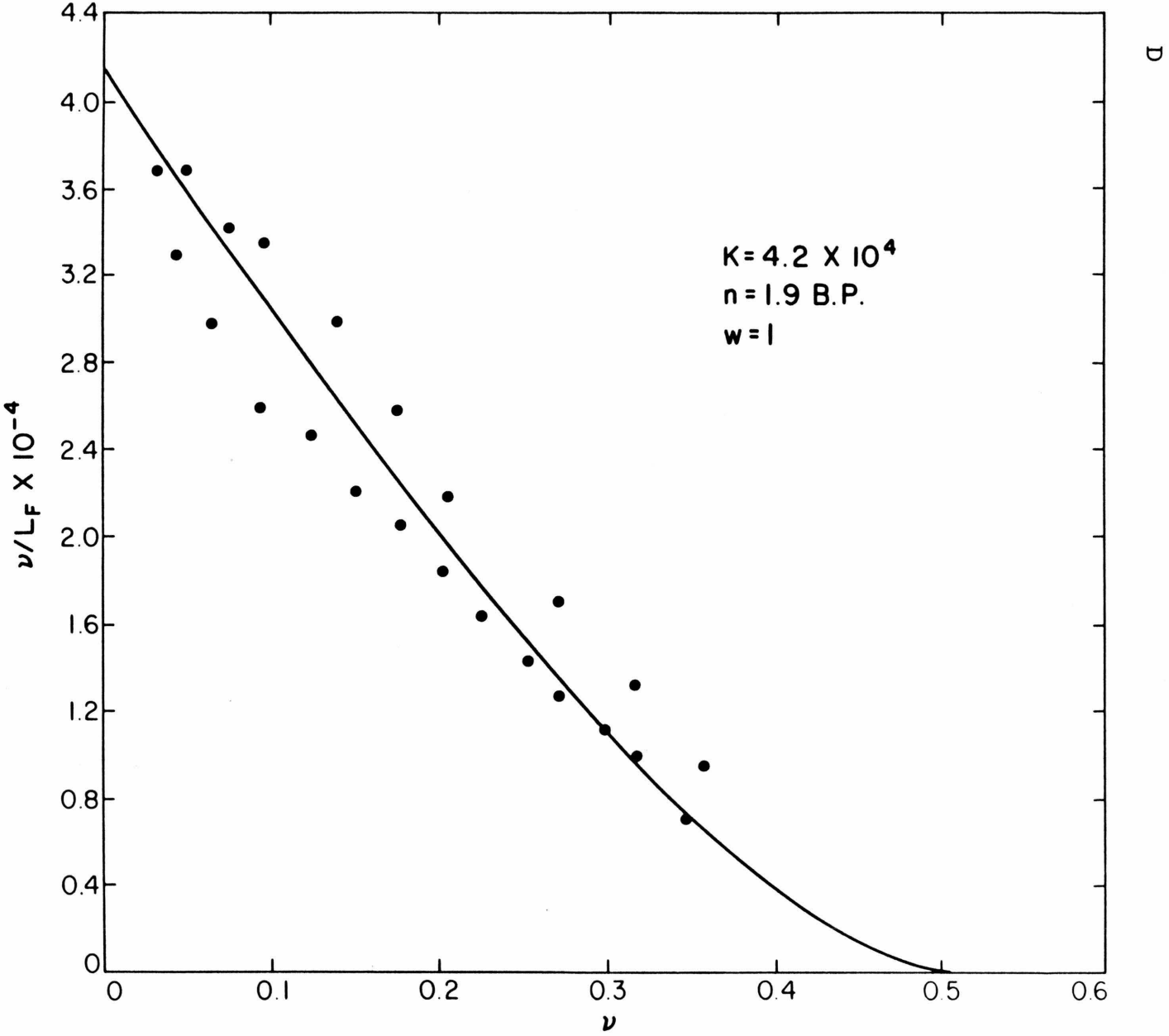

a
14.0
13.0
12.0
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
$\nu/C_F \times 10^{-6}$
K=1.5×10^7
n=4.0
w=1
K=1.4×10^7
n=3.8
w=1
0.02 0.04 0.06 0.08 0.10 0.12 0.14 0.16 0.18 0.20 0.22 0.24
ν

B


K = 8 × 10^7
n = 3.4 B.P.
w = 1
ν/c_F × 10^-7
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
.17
.18
.19
.20
.21
.22
.23
.24
.25
.26
.27
.28
ν

c


4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
$\nu/L_F \times 10^{-7}$
$K = 8 \times 10^7$
n = 3.3 B. P.
w = 1
0.16
0.18
0.20
0.22
0.24
0.26
0.28
0.30
ν

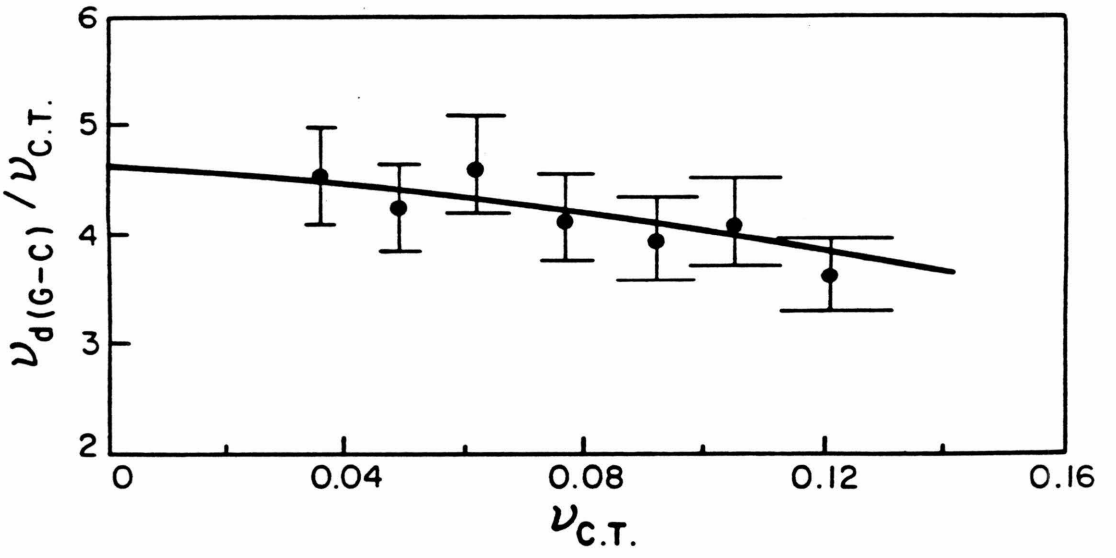

4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
0
0.1
0.2
0.3
0.4
0.5
0.6
$\nu/L_F \times 10^{-4}$
ν
K = 4.2 X 10^4
n = 1.9 B.P.
w = 1

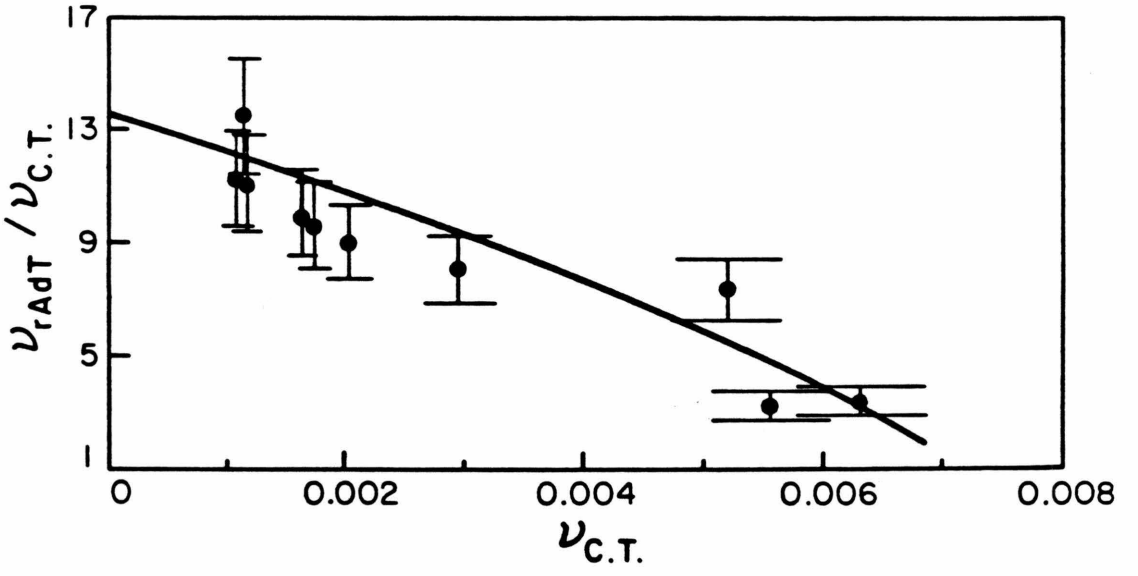
A third technique used to determine the binding affinity of BMSp to nucleic acid involves competition equilibrium dialysis.[38] In this method, a three-part dialysis cell containing two different nucleic acids in the outer compartments is stirred to equilibrium. At equilibrium the two nucleic acid-BMSp complexes are in simultaneous equilibrium with the same concentration of free BMSp. The ratio of BMSp association constants for two nucleic acids i and j (K_i/K_j) can be determined directly from the observed ratio of binding densities (ν_i/ν_j) in the limit of low binding density ($\nu_i, \nu_j \rightarrow 0$).[39] Unlike Scatchard plots, the competition approach allows relative binding affinities to be determined without specifying the binding site size and binding cooperativity of each complex.

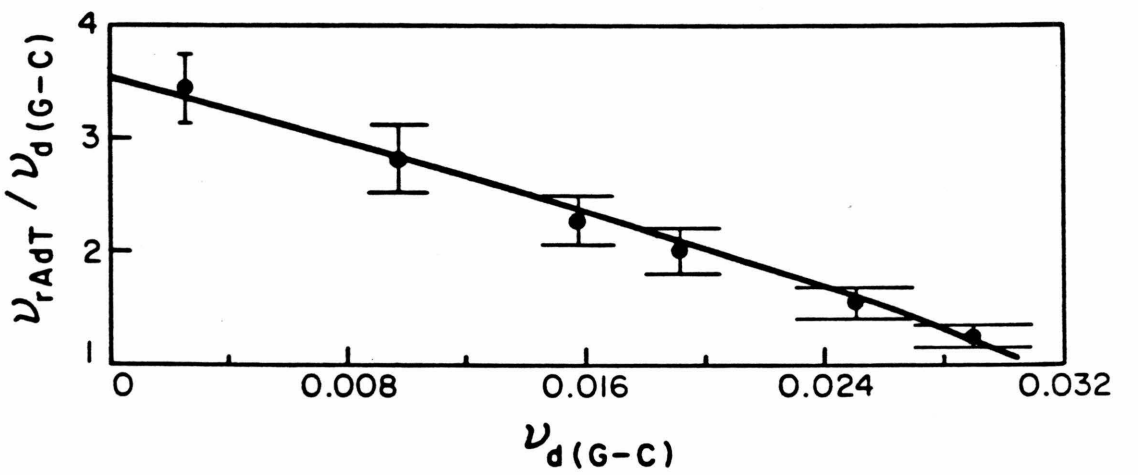
We have used this competition dialysis approach to determine the relative binding affinity of BMSp to polyd(G-C), polydC·polydG, calf thymus DNA, and polyrA·polydT (Fig. 17). By conducting competition dialysis between several different pairs of nucleic acids, we have been able to internally verify binding affinities obtained by the competition approach as well as those obtained by the indirect spectroscopic techniques. Combination of the two methods allows an estimation of

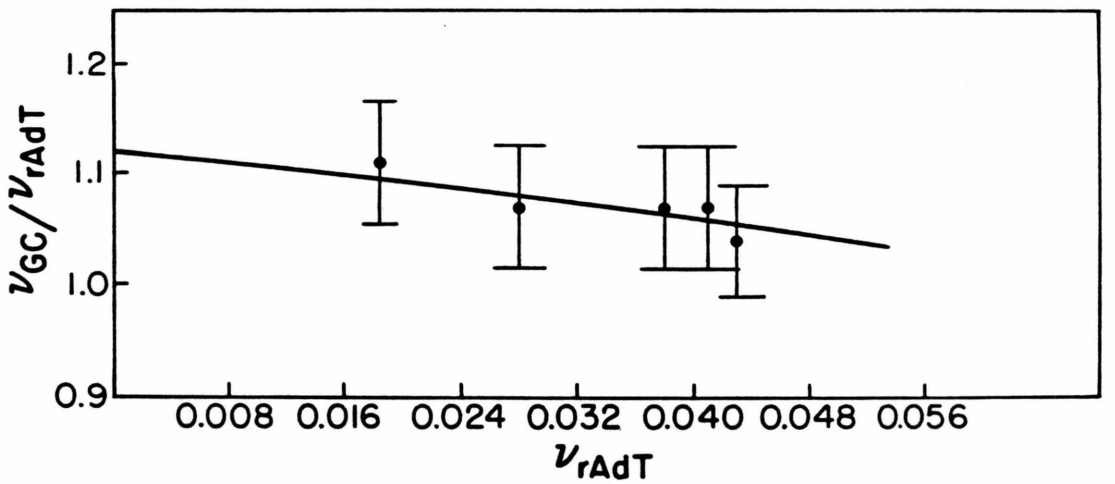
Figure 17

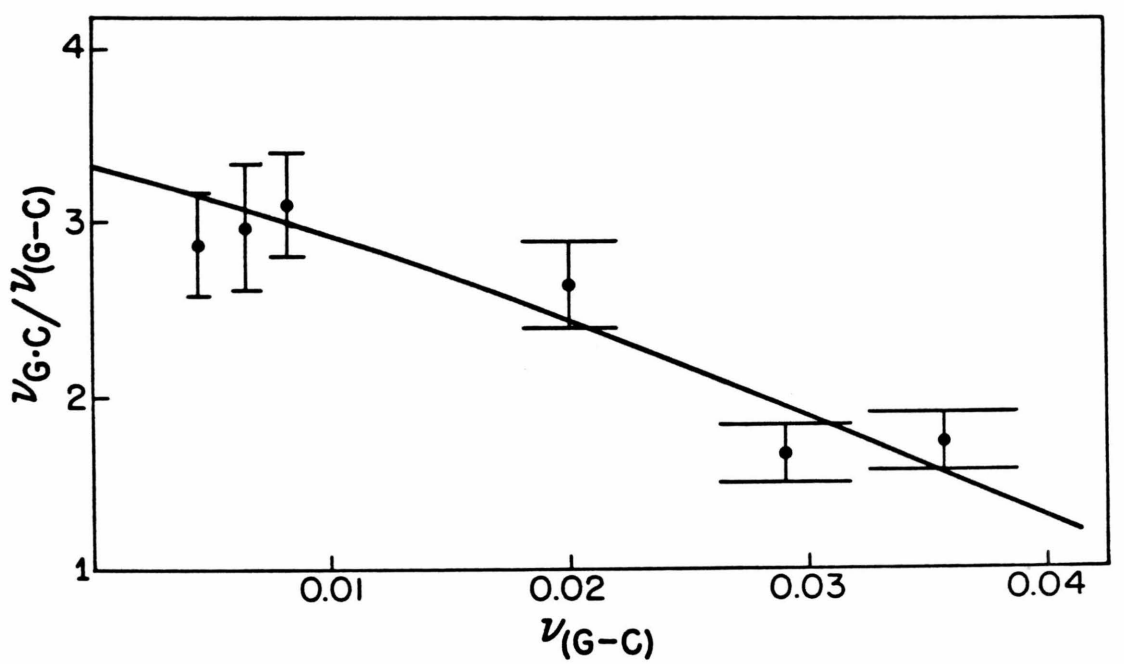
Equilibrium dialysis competiton plots at $[M^+]$ = 1.0 M. The BMSp binding density of calf thymus DNA: ν_{CT}, poly(dC-dG): $\nu_{d(G-C)}$, poly(rA:dT): ν_{rAdT} or poly(dC:dG): ν_{dCdG} is plotted against the binding density ratio for the indicated nucleic acids.


$\nu_{d(G-C)} / \nu_{C.T.}$
6
5
4
3
2
0
0.04
0.08
0.12
0.16
$\nu_{C.T.}$


$\nu_{rAdT} / \nu_{C.T.}$
17
13
9
5
1
0
0.002
0.004
0.006
0.008
$\nu_{C.T.}$


$\nu_{rAdT} / \nu_{d(G-C)}$
4
3
2
1
0
0.008
0.016
0.024
0.032
$\nu_{d(G-C)}$


1.2
1.1
1.0
0.9
ν_{GC}/ν_{rAdT}
0.008 0.016 0.024 0.032 0.040 0.048 0.056
ν_{rAdT}


4
3
2
1
$\nu_{G\cdot C}/\nu_{(G-C)}$
0.01 0.02 0.03 0.04
$\nu_{(G-C)}$

the binding affinity to within an accuracy of ± 10% for calf thymus DNA, d(C-G), dCdG, and rAdT.

The fourth technique used to estimate binding affinities monitors ligand binding by indirect techniques over a broad ratio of (L_T/M), where L_T is the total ligand concentration and M is the macromolecule concentration.[31] If two or more solutions have the same value of X, where X is some physical property of ligand which is proportional to its concentration, than a plot of L_T/M versus 1/M will be a straight-line with intercept $(q_B-q_F/q_F)\nu$ and slope L_F provided that M is significantly less than the inverse of the equilibrium constant (q_B and q_F are the extinction coefficients of bound and free ligand, respectively). Construction of a Scatchard plot from these ν and L_F values allows one to estimate the affinity of ligand for macromolecule. If M is not significantly less than 1/K, a plot of L_T/M versus 1/M will not be linear unless a narrow macromolecule concentration range is investigated. In this case, the intercept and slope can be used to construct approximate Scatchard plots which, in most cases, can be further corrected to yield accurate estimates of the ligand-macromolecule association constant. We have utilized this technique to monitor the binding of BMSp to polydAdT and of EB to polydCdG and poly(dC-dG) (see Chapter II and Chapter III for corresponding measurements).

The binding affinities of BMSp and EB measured by these four techniques are summarized in Table II. We find that for the five nucleic acids investigated:sonicated calf thymus DNA, poly(C-G), polydCdG, polydAdT and polyrAdT the free energy of BMSp binding, ΔG_{BMSp}, is equal to 1.5-1.6 ΔG_{EB}, where ΔG_{EB} is the free energy of EB binding for the same nucleic acid. If the salt concentration is lowered from 1.0 M^+ to 0.075 M^+, thus enhancing electrostatic interactions, the enhancement of BMSp's binding affinity becomes much larger. For a heterogeneous DNA sequence (sonicated calf thymus DNA) we observe $K_{EB} = 4 \times 10^4$ and $K_{BMSp} = 1.5 \times 10^7$ at 1.0 M^+ whereas at 0.075 M^+ $K_{EB} = 2 \times 10^5$ and $K_{BMSp} \geq 2 \times 10^{11}$. The binding affinity of BMSp for sonicated calf thymus DNA at low salt is estimated from the spectrophotometric titration data shown in Fig. 7 and the value so obtained agrees well with the value, 1×10^{11}, extrapolated from the measured affinity at 1 M^+ (Table II) by the polyelectrolyte analysis of Record (Eq. 1).

Binding Specificity of BMSp and EB at 1 M^+

In addition to binding affinity, the binding specificity of BMSp relative to EB is substantially enhanced. Bresloff and Crothers have shown that EB binds the RNA-DNA duplex rAdT 100 times more tightly than the DNA-DNA duplex dAdT.[58] This 100-fold specificity exhibited by

EB increases to 1500 for BMSp (Table II). Since the only difference between rAdT and dAdT is the presence of a 2' hydroxyl group on the sugar ring and not base sequence, these results indicate that the specificity which BMSp and EB exhibit for certain nucleic acids can arise from preferential recognition of different nucleic acid conformations.[39]

As discussed in Chapter III, the conformational specificity exhibited by EB and BMSp is not restricted to the copolymers dAdT and rAdT. Instead, evidence is presented which suggests that regardless of base sequence, BMSp and EB prefer to bind H and A conformations of nucleic acid much more tightly than B conformations.[31] Since RNA[41] and RNA-DNA[42] hybrids adopt A conformations as their native structure, BMSp and EB bind these nucleic acids more tightly than corresponding DNA duplexes. Furthermore, since DNA adopts a native B conformation,[43] BMSp and EB can cooperatively induce sequence specific B → H → A allosteric[44] transitions in DNA. Although both compounds show litte preference for GC at AT base pairs,[15] certain base sequences bind them more readily than others[46] because the ability of DNA to undergo a B → H → A transition varies greatly with base sequence.[31]

Table II - Comparison of the binding affinity of ethidium bromide (EB) and bis(methidium)-spermine (BMSp).

		EB			BMSp		
Nucleic Acid	$[M^+]$	Conformation	K	Technique	Conformation	K	Technique
Sonicated Calf Thymus	0.075	H	1.4×10^5	a	H	$\geq 2 \times 10^{11}$ 1×10^{11}	a d
Sonicated Calf Thymus	1.0	H	4.5×10^4	a	H	$1.5 \pm .2 \times 10^7$	a
Poly d(C-G)	1.0	B	3×10^3	b	H	8×10^7 8×10^7	a e
		H	1×10^5			7×10^7	f
Poly dCdG	1.0	B	5×10^3	b	H	2.8×10^8	f
		H	1×10^5			2.5×10^8	f
Poly dAdT	1.0	B	2×10^3	c	B	$1.6 \pm .7 \times 10^5$	b
Poly rAdT	1.0	A	2×10^5	c	A	2.1×10^8 2.6×10^8	f f

a) spectrophotometric; b) constant X; c) dialysis[40]; d) polyelectrolytic extrapolation; e) fluorescence; f) competition.

Binding Site Size Measurements

The number of base pairs covered by BMSp when it binds a nucleic acid, n, the binding site size, has been determined in two ways. For calf thymus DNA at low salt, and for dCdG as well as rAdT at 1 M^+, n can be determined directly from absorbance measurements (Figs. 8, 18, & 19). In these cases the binding affinity of BMSp is sufficiently high as to render binding vritually stoichiometric. For all other nucleic acids, n has been determined from corresponding Scatchard plots (Fig. 16). As expected for a bisintercalation binding mode, we find that the binding site size of BMSp for four of five nucleic acids investigated is twice that observed for EB(Table III).

The one exception to this behavior, polyd(C-G), exhibits a binding stoichiometry of 3.4 base pairs at 1 M^+ (Fig. 16). This nonintegral binding site size implicates a mixture of binding modes, as for example bisintercalation with n = 4 and monointercalation with n = 2; alternatively, both n = 4 and n = 2 binding species could be bisintercalated. In any event, the observation that the fluorescence of BMSp begins to decrease at high binding density when the concentration of polyd(C-G) is increased demonstrates that two bound species with different binding stoichiometries are formed (Fig. 20). At

Figure 18

Determination of the binding stoichiometry of BMSp to polyrAdT at 1 M^+. The upper curve was obtained at a rAdT concentration of 2.04×10^{-6} M BP and the lower curve at 1.361×10^{-6} M BP. The absorbance at 485 nm (path length 10 cm) for increasing additions of BMSp is plotted against the corresponding BMSp/BP ratio.

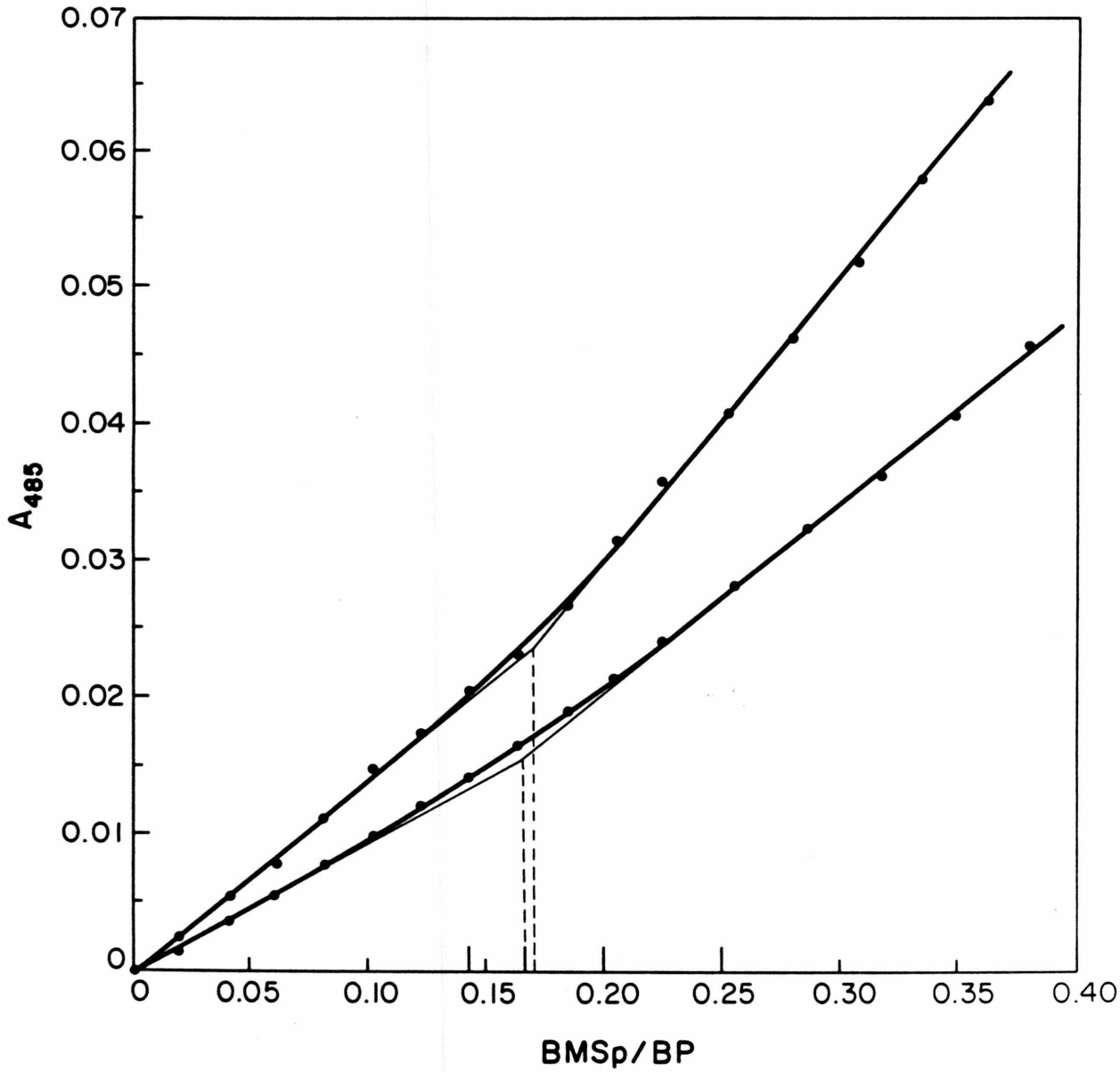

0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
A485
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
BMSp/BP

Figure 19

Determination of the binding stoichiometry of BMSp to poly(dC;dG) (1.712×10^{-6} M BP) at 1 M^{+}. The absorbance at 485 nm (path length 10 cm) for increasing additions of BMSp is plotted against the corresponding BMSp/BP ratio.

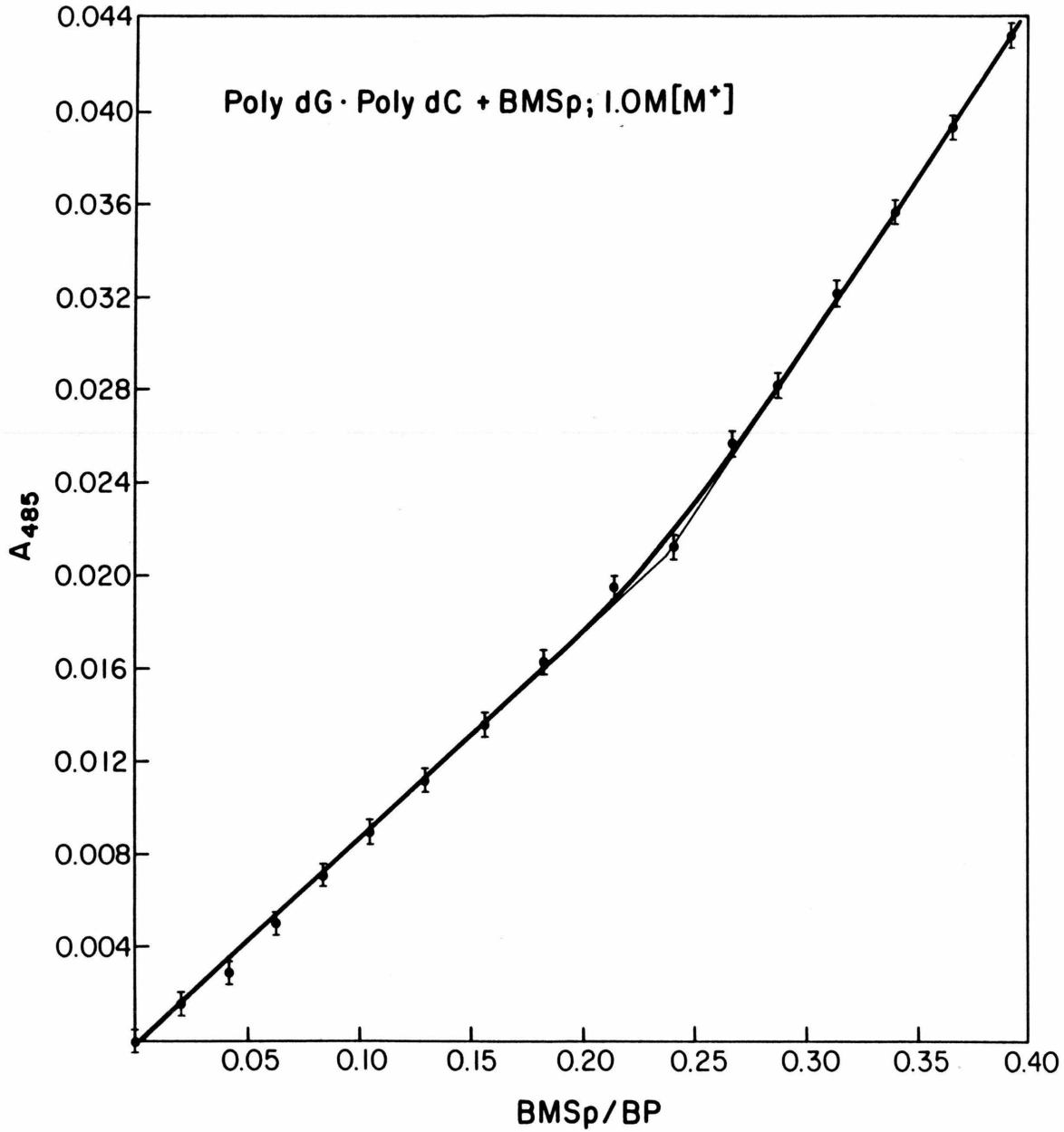

Poly dG · Poly dC + BMSp; 1.0M[M+]
A485
0.044
0.040
0.036
0.032
0.028
0.024
0.020
0.016
0.012
0.008
0.004
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
BMSp/BP

Table III - Comparison of the binding site size (in base pairs) of BMSp and EB bound to various nucleic acids.

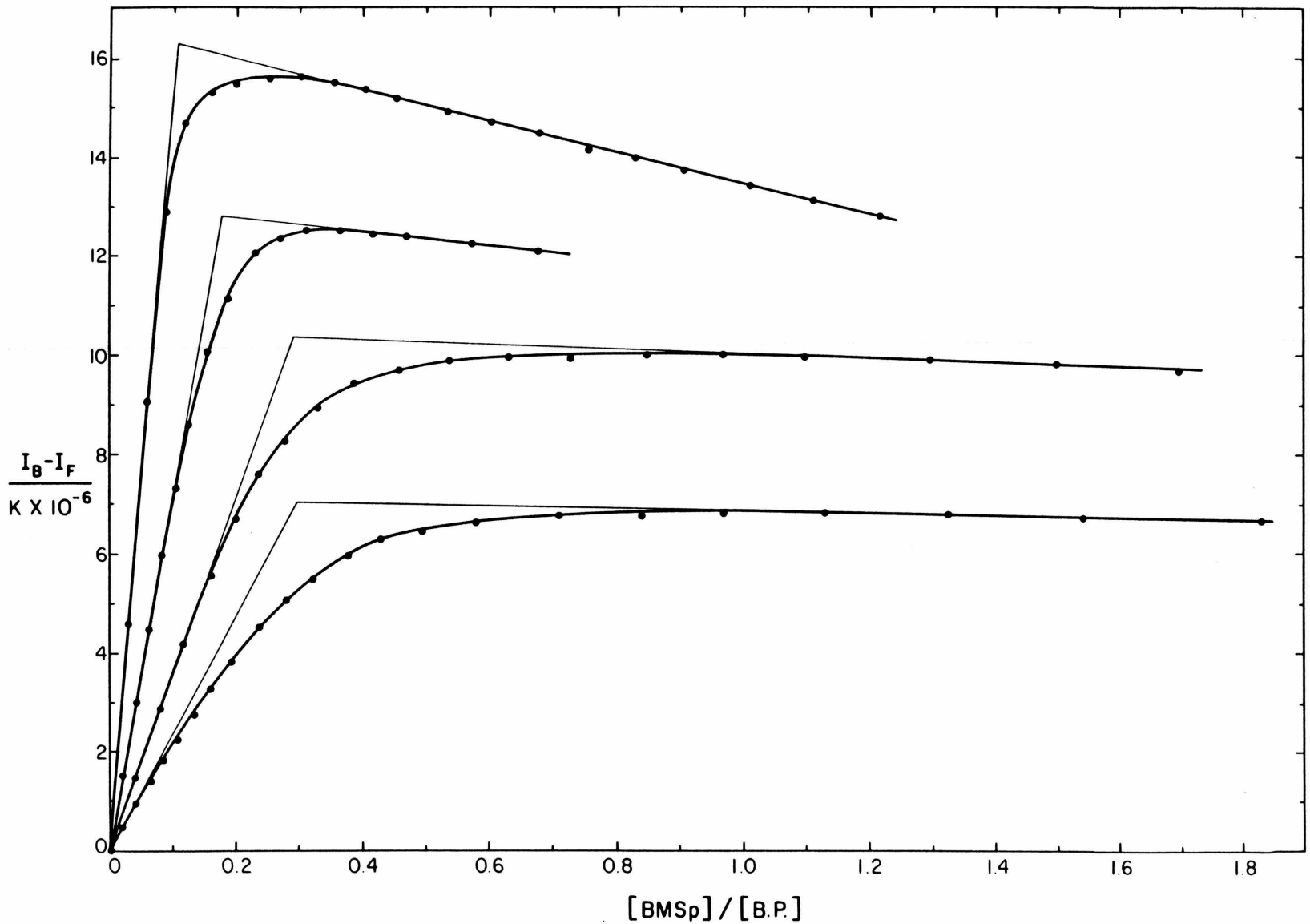
Nucleic Acid	[M⁺]	Binding Site Size	
		BMSp	EB
Calf thymus	1.0	4	2
	0.075	4	2
d(C-G)	1.0	3.3(4,2)	2
dC·dG	1.0	4	2
rA·dT	1.0	6	3

least one species must be intercalated since the quantum yield of BMSp is substantially enhanced upon binding (Table I). Although we do not know if the second species is bisintercalated, the equations of Shafer[45] indicate that the binding affinities of both species are similar (6.5 versus 1.5 x 10^7).

Although mono- or nonintercalated BMSp species are expected to bind as saturation of available bisintercalation sites is approached, polyd(C-G) is the only nucleic acid investigated which exhibits a significant departure from an integral binding site size. Although the reason for this be-

Figure 20

Fluorescence titration of poly(dC-dG) with BMSp at $[M^+] = 1.0$. The increase in the fluorescence of BMSp in the presence of nucleic acid (I_B) minus the fluorescence of an equivalent solution of BMSp in the absence of nucleic acid (I_f) is plotted against the ratio BMSp/BP. The difference $I_B - I_f$ is normalized by dividing by k where $I_F = k[BMSp]$. The concentration of poly(dC-dG) was 3.464, 1.748, 0.867, and 0.578 x 10^{-6} M BP (proceeding from top to bottom).


$\frac{I_B - I_F}{K \times 10^{-6}}$
0
2
4
6
8
10
12
14
16
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
[BMSp] / [B.P.]

havior is presently unknown, it is of interest to note that a naturally occurring bisintercalator, echinomycin, exhibits similar behavior.[45]

The binding site size of BMSp and EB appear to sense changes in nucleic acid structure. For example, the measurements of Bresloff have demonstrated that when EB binds DNA n = 2 whereas when it binds RNA or RNA-DNA duplexes n = 3.[40] Similarly BMSp covers four base pairs when bound to DNA and six base pairs when bound to the RNA-DNA duplex rAdT (Table III).

We have demonstrated that the binding site size of ethidium also varies with nucleic acid conformation. For example, we find that like RNA and RNA-DNA duplexes (known A conformations), the binding site size of ethidium to an A conformation of DNA is three base pairs.[31] Thus the presence of a 2' OH group on one or both strands is not necessary for ethidium to exhibit a three base pair binding site size. This observation suggests that the binding site size of BMSp may also vary with nucleic acid conformation. A particularly interesting example in this respect is the binding of BMSp to dAdT. At low BMSp concentrations dAdT adopts a B conformation whereas at higher BMSp concentrations it adopts an H conformation.[31] The binding

site size observed for the H conformation, four base pairs, is the same as observed for dCdG, calf thymus DNA and d(C-G), and presumably arises from steric restrictions imposed by the sugar pucker alternation of the H conformation.[18] When bound to the B conformation of dAdT, the half-width of the visible spectrum of BMSp is 98 nm whereas it narrows to 90 nm when bound to calf thymus, polyd(C-G), polydCdG or polyrAdT (Figs. 7 and 21; Table IV). Since unbound BMSp also exhibits a visible half-width of 98 nm whereas the corresponding monomers ethidium and MSp exhibit a 90 nm value, it may be suggested that both chromophores of BMSp are in close proximity and electronically coupled when bound to the B conformation of dAdT. Although binding of an intramolecularly stacked dimer of BMSp to the outside of the helix could exhibit similar behavior, the fluorescence lifetime of bound BMSp argues that it is intercalated into the B conformation of dAdT (Table I). Because the visible spectrum of BMSp does not broaden when two or three base pairs separate its intercalated chromophores, it may be suggested that only one base pair separates both chromophores of BMSp when it is bound to the B conformation of dAdT.

Table IV

Comparison of selected optical properties of BMSp, EB and MSp. The extinction coefficient of the visible maximum (ε), wavelength of the visible maximum (ε_{max}), and half-width of the visible absorption spectrum are shown for cases where drug is either unbound or bound to nucleic acid. Measurements of BMSp bound to calf thymus DNA, polyd(C-G), polydCdG and polyrAdT were obtained at 1 M^+.

BMSp	ϵ	$\epsilon_{(max)}$	½ W
H_2O	10,160	502	99
0.075 M^+	9,850	502	98
1.0 M^+	8,900	502	97
4.4 M^+	8,400	502	98
pH 3.0	10,200	500	98
DMSO	–	539	96
MeOH	–	531	97
dAdT	7,880	526	98
C.T. 0.075 M^+	7,140	533	90
C.T.	7,800	530	92
(dC–dG)	7,240	529	90
dC : dG	7,880	530	92
rAdT	8,600	530	89
MSp			
0.075 M^+	–	488	92
C.T. 0.075 M^+	–	529	91
EB			
1.0 M^+	6,060	480	87
C.T. 1.0 M^+	4,000	520	85

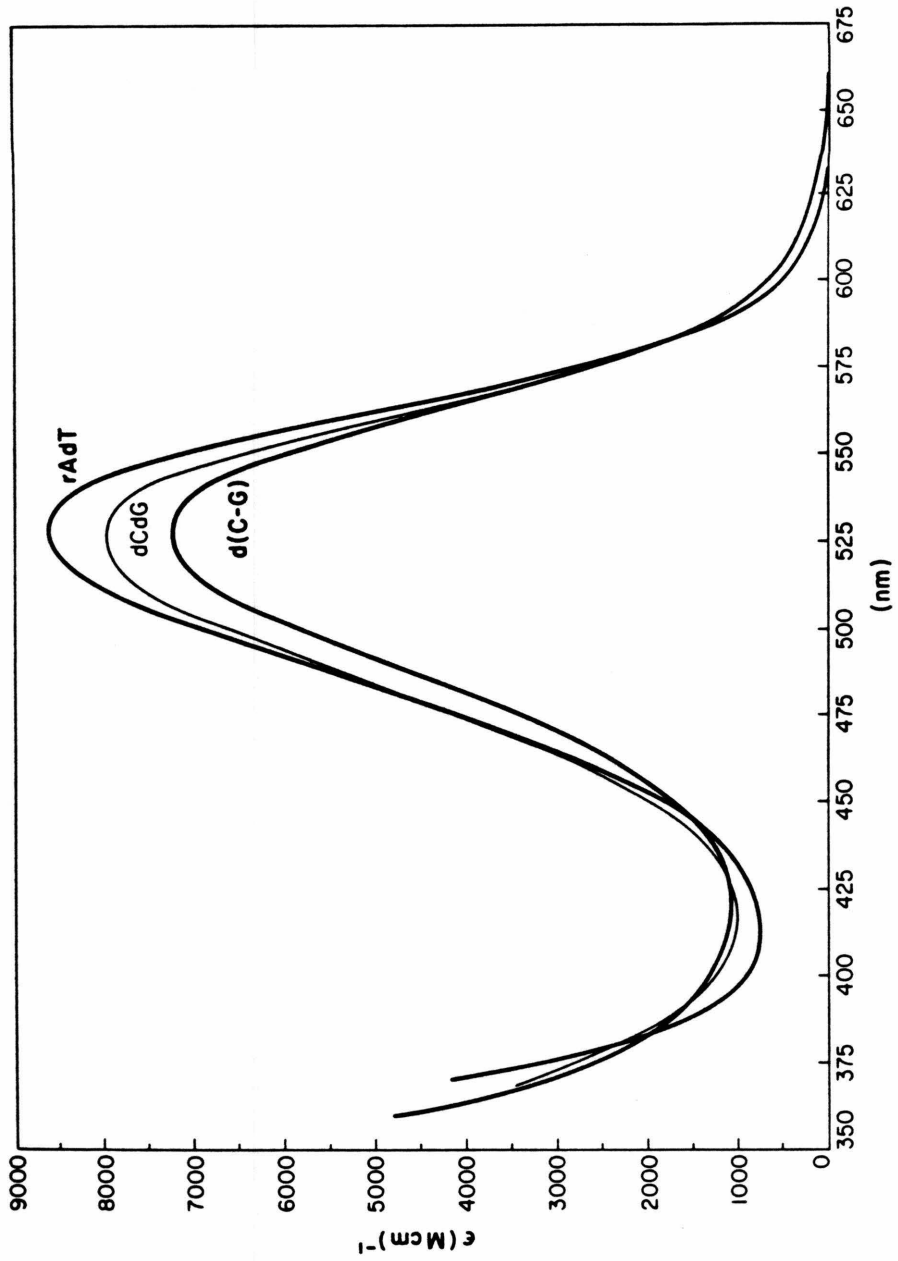


Figure 21
Comparison of the absorption spectrum of BMSp bound to polyrAdT, polydCdG, and polyd(C-G) at 1.0 M^+.

In conclusion, these stoichiometry measurements demonstrate that the binding site size of BMSp varies with both the structure and conformation of nucleic acid. The preferred binding site size of BMSp is two base pairs when bound to a B conformation of DNA, four base pairs when bound to an H conformation of DNA and six base pairs when bound to an A conformation of an RNA or RNA-DNA duplex (and perhaps a DNA duplex) (Fig. 22). Thus changes in linker length and flexibility may alter the conformational specificity of BMSp and related compounds.

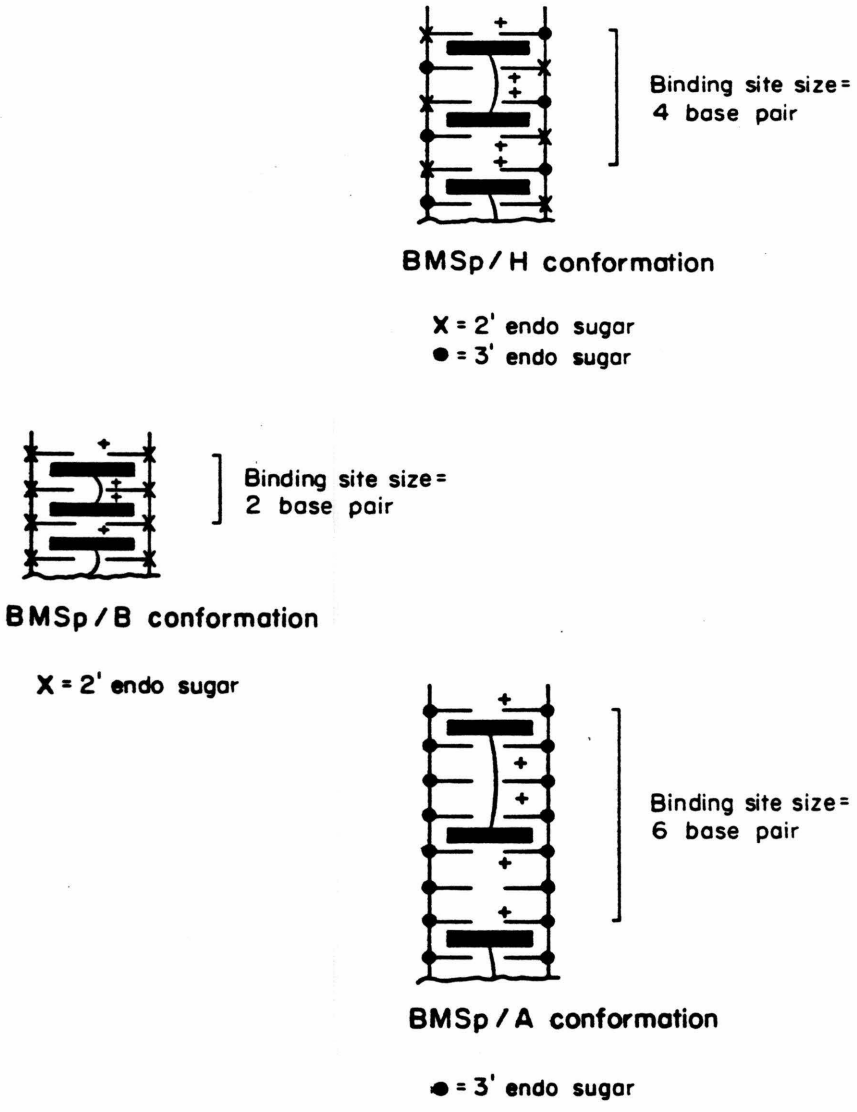


Figure 22
Proposed dependence of BMSp's binding site size on nucleic acid conformation. Although absorption and fluorescence measurements indicate that intercalated chromophores of BMSp are separated by one base pair in the B conformation, the actual binding site may be larger than two base pairs since binding may distort DNA (e.g. compare to BMSp bound to the A conformation.)

SUMMARY

The results of our BMSp-nucleic acid binding studies can be summarized as follows:

1) The preferred binding mode of BMSp to nucleic acid involves simultaneous intercalation of both EB moieties into the double helix (bisintercalation). Although less important, the relative contributions of mono- and nonintercalated species increases as saturation of bisintercalation sites is approached.

2) Under conditions of low salt (0.075 M^+), BMSp extends DNA twice as much as EB, causes little or no bending of the double helix, and unwinds DNA 38^o or 1.5 times as much as EB. The failure of BMSp to unwind DNA twice as much as EB is shown to arise from electrostatic interactions between the spermine linker of BMSp and DNA. Although these electrostatic interactions do not affect the ability of BMSp to extend DNA, they do alter slightly the unwinding geometry of each intercalated chromophore from an ethidium (26^o) to an acridine like 19^o geometry. Increasing the salt concentration to 1.0 M^+ decreases electrostatic interactions between the spermine linker of BMSp and DNA resulting in an increase in its unwinding angle from 38 to 49 $\pm$ 3^o, or twice that of EB under similar conditions.

3) At 1 M^+, where electrostatic contributions to the observed binding affinity and intercalation geometry are small, the free energy of BMSp binding to five different nucleic acids is found to equal 1.5-1.6 ΔG_{EB}, where ΔG_{EB} is the free energy of EB binding to the same nucleic acid. When the salt concentration is lowered from 1 M^+ to 0.075 M^+, thus enhancing electrostatic interactions, the enhancement of BMSp's binding affinity becomes much larger. For a heterogeneous DNA sequence at 1 M^+ $K_{EB} = 4 \times 10^4$ and $K_{BMSp} = 1.5 \times 10^7$ whereas at 0.075 M^+ $K_{EB} = 2 \times 10^5$ and $K_{BMSp} \geq 2 \times 10^{11}$. The estimate of K_{BMSp} at low salt agrees well with the value 1×10^{11} extrapolated from the measured affinity at 1 M^+ by the polyelectrolyte analysis of Record.[34]

4) In addition to binding affinity, the binding specificity of BMSp relative to EB is substantially enhanced. The 100-fold preference of EB for the RNA-DNA hybrid rAdT over its DNA counterpart, dAdT, is found to increase to 1500 for BMSp. Since the only difference between rAdT and dAdT is the presence of a 2'-hydroxyl group on the sugar ring and not base sequence, these results indicate that the specificity which BMSp and EB exhibit for certain nucleic acids can arise from preferential recognition of different nucleic acid conformations.

5) As discussed in Chapter III, the conformational specificity exhibited by EB and BMSp is not restricted to the copolymers dAdT and rAdT. Evidence is presented which suggests that regardless of base sequence, both compounds bind H and A conformations of nucleic acid much more tightly than B conformations. Since RNA and RNA-DNA hybrids adopt A conformations as their native structure, BMSp and EB bind these nucleic acids more tightly than corresponding DNA duplexes (eg., rAdT versus dAdT). Furthermore, since DNA adopts a native B conformation, BMSp and EB can cooperatively induce sequence specific B → H → A allosteric transitions in DNA.

6) In addition to binding affinity and binding cooperativity, the binding site size, and visible spectrum of BMSp are found to vary with nucleic acid conformation. Whereas B DNA binds one BMSp every two base pairs and H DNA one BMSp every four base pairs, A conformations can bind one BMSp every six base pairs. Thus control of linker length and flexibility may be used to predictably alter the conformational specificity of BMSP and related compounds.

BIOLOGICAL IMPLICATIONS

1) The observation that ethidium is both a potent and selective inhibitor of nucleic acid function[9] suggests that polyintercalators like BMSp may exhibit remarkable inhibitory properties. Of particular importance in this regard is the finding that under physiological conditions, both the binding affinity and specificity of BMSp are similar to DNA binding regulatory proteins. Thus BMSp represents one of the first rationally designed drugs which may selectively inhibit or alter gene expression. The finding that BMSp preferentially inhibits certain sequences on closed circular pBR-322 from restriction digestion, whereas the monomer ethidium shows no selectivity, demonstrates the possibility of such behavior.[46]

2) The very slow dissociation rate of BMSp coupled with its high affinity and specificity for nucleic acid conformation could be used to selectively alter nucleic acid structure or function by delivering additional functionality to the nucleic acid backbone.

3) The greatly enhanced affinity of BMSp for A and H conformations of nucleic acid should significantly enhance both the detection and characterization of B → H → A transitions in nucleic acid (eg. see Propositions 4 and 5).

4) The high binding affinity and specificity of BMSp should significantly enhance its ability to fractionate nucleic acids. For example, the proposal that gene activation necessitates enhancement of a B → H transition in DNA (Proposition 1) suggests that BMSp will selectively inhibit activated or active genes. Similarly, BMSp could also be used to specifically localize and probe these genes (eg. see Proposition 5).

5) Through competition methods, the high affinity and specificity of BMSp could be used to probe the binding properties of other tightly binding ligands such as nucleic acid binding regulatory proteins, an otherwise difficult experimental task.

MATERIALS AND METHODS

Nucleic Acids: Highly polymerized calf thymus DNA (Sigma) was phenol extracted, sonicated at 4°C under nitrogen (medium probe of a Branson ultrasonicator at maximum power) and fractionated on Sepharose 4B to yield double stranded DNA 150-300 base pairs in length. PolydAdT, polyrAdT, polyd(C-G), polyd(A-T), polydG and polydT were purchased from PL Biochemicals. Each polymer was exhaustively dialyzed against low salt phosphate buffer (0.025 M KH_2PO_4, 0.025 M Na_2HPO_4) containing 2 mM EDTA followed by exhaustive dialysis against either low or high salt phosphate buffers containing no EDTA. PolydCdG was obtained from Boehringer Mannheim and was carefully reannealed as previously described.[47] PM2 phage DNA was a gift from R. Watson and contained 80% supercoils, 20% nicked circles and < 1% linear molecules as determined by agarose gel electrophoresis. Chicken erythrocyte nucleosome core DNA was a gift of K. Tachell. The sample used was 150 $\pm$ 5 bp with < 2% of the single strands exhibiting a nick as determined by gel electrophoresis. All nucleic acid solutions were sterilized by passage through 0.45 μ millipore filters and stored frozen. DNA concentrations were determined spectrophotometrically using the following extinction coefficients: calf thymus, chicken erythrocyte, PM2 (6600 at 260 nm);

dCdG (7400 at 255 nm);[48] dAdT (600 at 258 nm);[48] rAdT (6900 at 258 nm);[49] and d(C-G) (7100 at 256 nm).[50]

Buffers: The following buffers, designated by their total monovalent cation concentration, were prepared by adding the appropriate amount of sodium chloride to a solution containing 0.025 M Na_2PO_4 and 0.025 M $KHPO_4$. (1) $[M^+]$ = 0.075, No NaCl; (2) $[M^+]$ = 0.21, NaCl = 0.25 M; (3) $[M^+]$ = 1.0, 0.9 M NaCl; (4) $[M^+]$ = 4.4, 4.3 M NaCl. The pH was then adjusted to 7.0 with solid NaOH. The corresponding ionic strengths are 0.14, 0.27, 1.05, and 4.45, respectively.

Competition Equilibrium Dialysis: Competition equilibrium dialysis was conducted in the dark using a specially prepared three compartment dialysis cell (Fig. 23). Each compartment, containing 25 ml of buffer, was vigorously stirred to equilibrium for 36 hr and the concentration of drug determined by comparing observed absorption spectra with appropriate "bound" spectrum of BMSp. Typical absorbances ranged from 0-0.02 units for the 10 cm cell path length. Nucleic acid concentrations used in each cell were such that at equilibrium all drug was bound and equally distributed between outer compartments.

Figure 23

Competition dialysis cells. Additional notes:

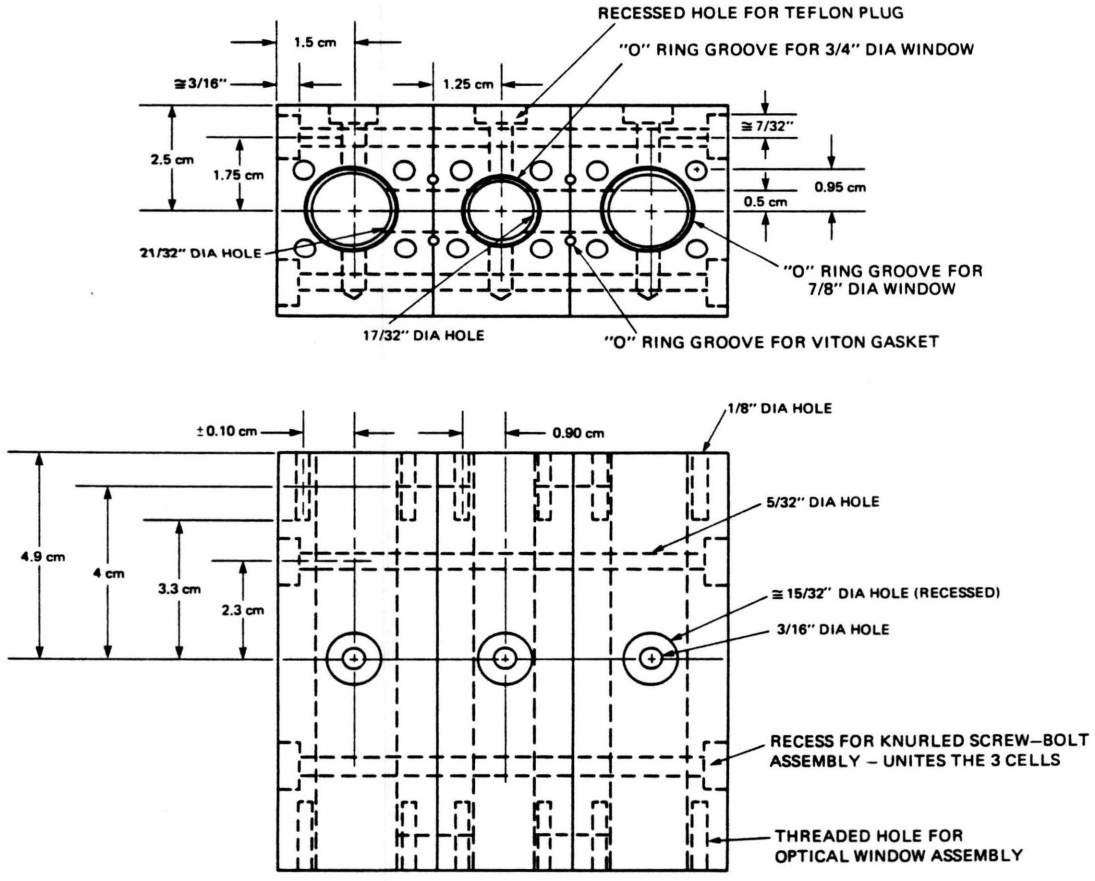
1. The three compartment dialysis cell was constructed from polished lucite. Cells were sealed with Viton O-rings.

2. Metal optical window assemblies were painted flat black to reduce scattered light. Suprasil optical windows were also painted flat black except for a 3/8" wide slit down the center of each window.

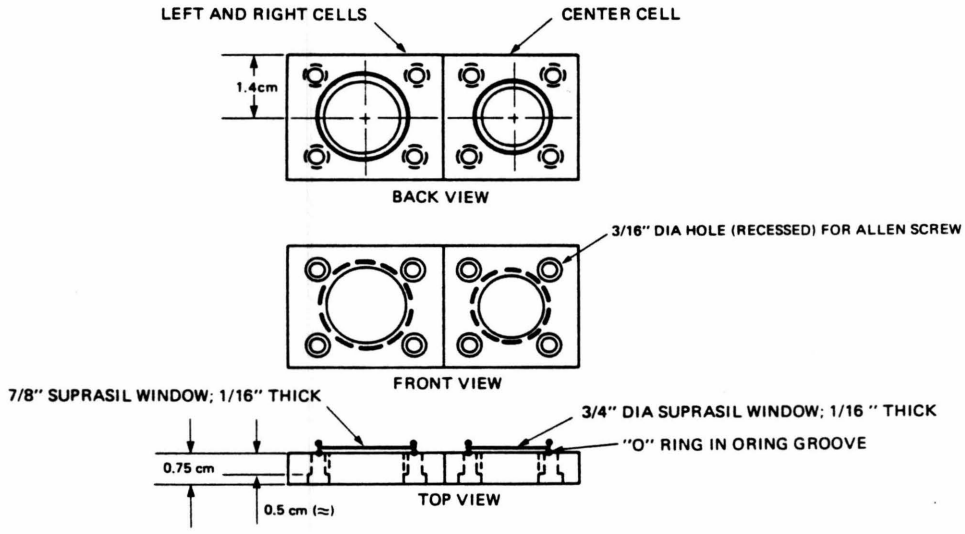
3. Solutions were stirred with teflon magnetic stirbars (13 x 3 x 3 mm) vertically positioned in the well at the bottom of each dialysis cell. All three dialysis cells can be stirred simultaneously on a Corning PL351 magnetic stirrer.

4. Nucleic acid and drug solutions were introduced by syringe through a small hole introduced in the middle of each teflon plug used to seal each cell.

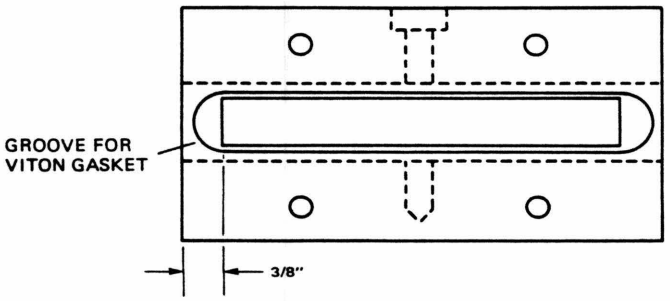
5. Cells were extensively rinsed with 0.45 μ millipore filtered buffer before introducing millipore filtered solutions.



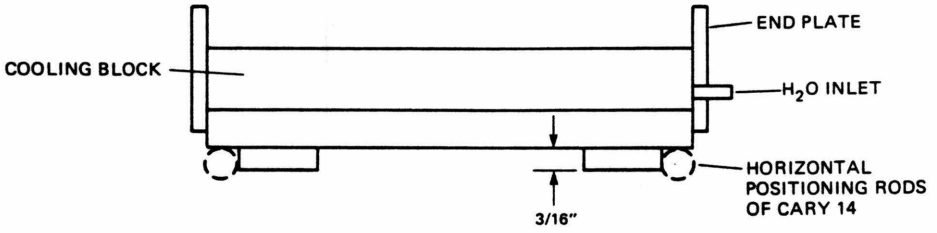
DIALYSIS CELLS: FRONT AND TOP VIEW; OPTICAL WINDOW ASSEMBLY REMOVED



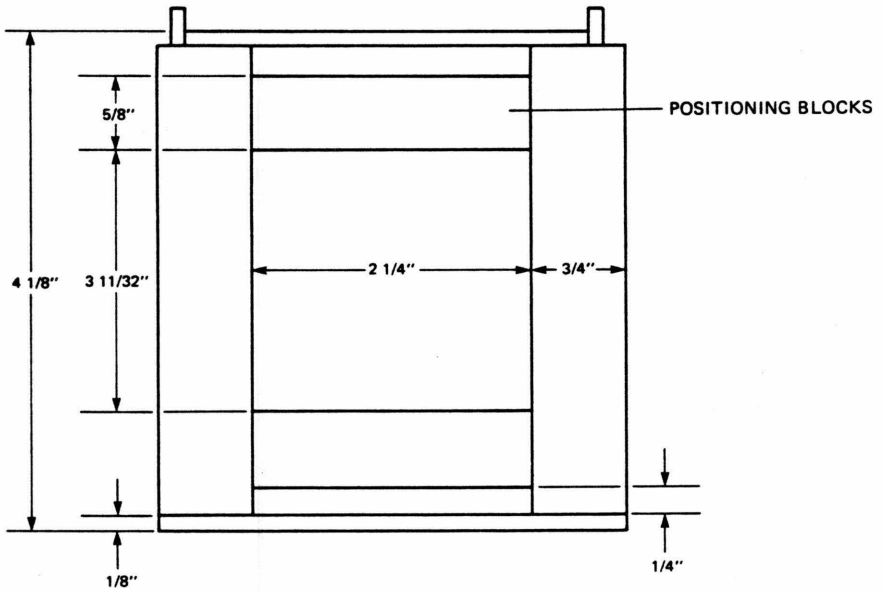
DIALYSIS CELLS: OPTICAL WINDOW ASSEMBLY



ONE DIALYSIS CELL: SIDE VIEW



DIALYSIS CELL HOLDER: SIDE VIEW



DIALYSIS CELL HOLDER: BOTTOM VIEW

Spectrophotometric titrations were performed either in 10 cm long cells (25 ml) on a Cary 14 spectrophotometer at 23 ± 1°C or in 1 and 0.5 cm long cells on a Beckman Model 25 spectrometer at 21 ± 0.5°C. Increasing amounts of BMSp were added to a known quantity of DNA and the absorbance recorded until equilibrium was attained (15 min for EB and 30-60 min for BMSp). Each cell was rendered dust free before adding DNA and drug solutions by extensive rinsing with millipore filtered (0.45 μm) buffer. Absorbance measurements at 10 cm were reproducible to ± 0.0005 absorbance units; for 0.5 and 1.0 cm cells reproducibility was 0.001 absorbance units.

Fluorescence Measurements: Fluorescence measurements were conducted on a Perkin-Elmer MPF 4 fluorescence spectrophotometer operating in the energy mode with the chopper on. The spectrophotometer was equipped with a circulating bath which was used to regulate the sample temperature (25°C). Fluctuations in the Xenon lamp source and recorder response were monitored and adjusted if necessary by alternatively measuring a standard solution of ethidium bromide. Fluorescence emission was monitored at 640 nm and/or 577 nm through a 515 nm (calf thymus) or 610 nm d(C-G) cutoff filter with a 10 nm slit width. Excitation was at 482 nm through a 455 nm cutoff filter with a 4 nm excitation slit. 1 cm Silica cuvetts were silanized with $SiCl_2(CH_3)_2$ before use. Buffers were pre-

pared from triply distilled > 99% D_2O. Samples of BMSp were introduced via a 10 μℓ Hamiltonian syringe and the contents of the cell were gently mixed by bubbling air into the solution through a millipore dust filter.

Viscometry Measurements: Viscometry measurements were performed in the same apparatus as previously described.[51] Temperature control was maintained to within ± 0.01°C and flow times recorded by photoelectric detection. Concentrated drug solutions were added via a Mannostate microliter pipette to 1.00 ml of nucleic acid solution. After each addition of titrant solution, the solution was thoroughly mixed by bubbling air gently through the solution via a millipore dust filter. Flow times were recorded until three successive determinations agreed to ≤ (± 20 msec).

Miscellaneous: All solutions were millipore filtered (0.45 μm) before use. Dialysis tubing was boiled in EDTA before use.

CALCULATIONS

A) Estimation of the Binding Constant of BMSp to Calf Thymus DNA Under Conditions of Low Salt.

When ligand binds macromolecule noncooperatively and forms only one bound complex, the binding density (ν) is given by:[37]

$$\nu/L_F = K(1-n\nu)\left(\frac{1-n\nu}{1-(n-1)\nu}\right)^{n-1} \qquad (1)$$

or

$$K = \nu/L_F(1-n\nu)\left(\frac{1-n\nu}{1-(n-1)\nu}\right)^{1-n} \qquad (2)$$

The minimum amount of free drug which could have escaped detection at saturation was determined to be 3% of the total drug added (based on a detection error of $\pm$ 6 x 10^{-5} A units cm^{-1}). Since the DNA concentration employed was 1.727 x 10^{-6} BP/L, the minimum ν and maximum L_F at saturation are 0.2425 and 1.3 x 10^{-8} M/L respectively. (Fig.7) Substituting these values into equation 2 and taking n as 4 yields a minimum binding constant, K, of $\geq$ 2 x 10^{11}. This value must be considered approximate since fluorescence measurements indicate a mixture of mono- and bis-intercalated forms for ν > 0.18. (Fig. 9)

B) Relative Quantum Yields

The quantum yield of bound BMSp relative to free BMSp was obtained by simultaneously titrating a solution of DNA and buffer, respectively, with BMSp under identical conditions. The dependence of the fluorescent intensity on BMSp concentration was found to be linear for unbound

BMSp as well as for bound BMSp at low degrees of saturation. The ratio of these bound and unbound slopes was multiplied by the inverse ratio of the corresponding extinction coefficients to yield relative quantum yields. No correction for polarization effects were made.

C) Binding of EB to Calf Thymus DNA at 1 M Na^+.

Under the 1 M^+ conditions employed EB was found to dimerize (Fig. 6). The dimerization constant and extinction coefficient were calculated as described by Bresloff.[40] The concentrations of bound and free EB in the presence of calf thymus DNA were then determined from the equation:

$$A = L_T(\frac{\varepsilon_D}{2}) + L_B(\varepsilon_B-\varepsilon_D/2) + (\varepsilon_M-\varepsilon_D/2)\left(\frac{-1 + \sqrt{1-8K_D(L_B-L_T)}}{K_D}\right)$$

where ε_B = 2250, ε_M = 6075, and ε_D = 2800 (at 475 nm) and K_D = 950. For different L_T values, a range of A values was generated from a range of L_B values. Comparison of the observed to calculated A values for a given L_T yields L_B and L_F.

D) Calculation of Bound Spectra

Spectra of BMSp or EB bound to nucleic acid were determined in two ways. For polyrAdT and polydCdG at 1 M^+ and calf thymus DNA at low salt, BMSp spectra obtained at low levels

of saturation (see Figs. 8, 18, 19) were taken to be 100% bound, an assumption compatible with their high affinities. For calf thymus DNA, d(C-G) and dAdT, bound spectra were obtained by analyzing the dependence of their absorption spectrum on nucleic acid concentration as previously described;[52] this technique was also utilized to obtain the spectrum of EB bound to calf thymus DNA. (Fig. 24)

E) Helical Extension of 150 Base Pair DNA by Drug Binding

The relationship between intrinsic viscosity ($[\eta]_{INT}$) and DNA length (L) is given by:

$$[\eta]_{INT} = CL^3 \left\{ \frac{3}{\ell Np-A} + \frac{1}{\ell N2p-B} \right\} \qquad (3)$$

for axial ratios (p) greater than one.[53] For a 150 base pair DNA molecule, p equals 25 (20 Å diameter), A = 1.5 B = 2.5 and C equals a constant.[53] If L_0 equals the length of DNA in the absence of ligand and L equals its length in the presence of ligand we can write:[53]

$$\frac{L}{L_0} = \left\{ [\eta]_{INT} f(p)_0 / f(p)[\eta_0]_{INT} \right\}^{1/3} \qquad (4)$$

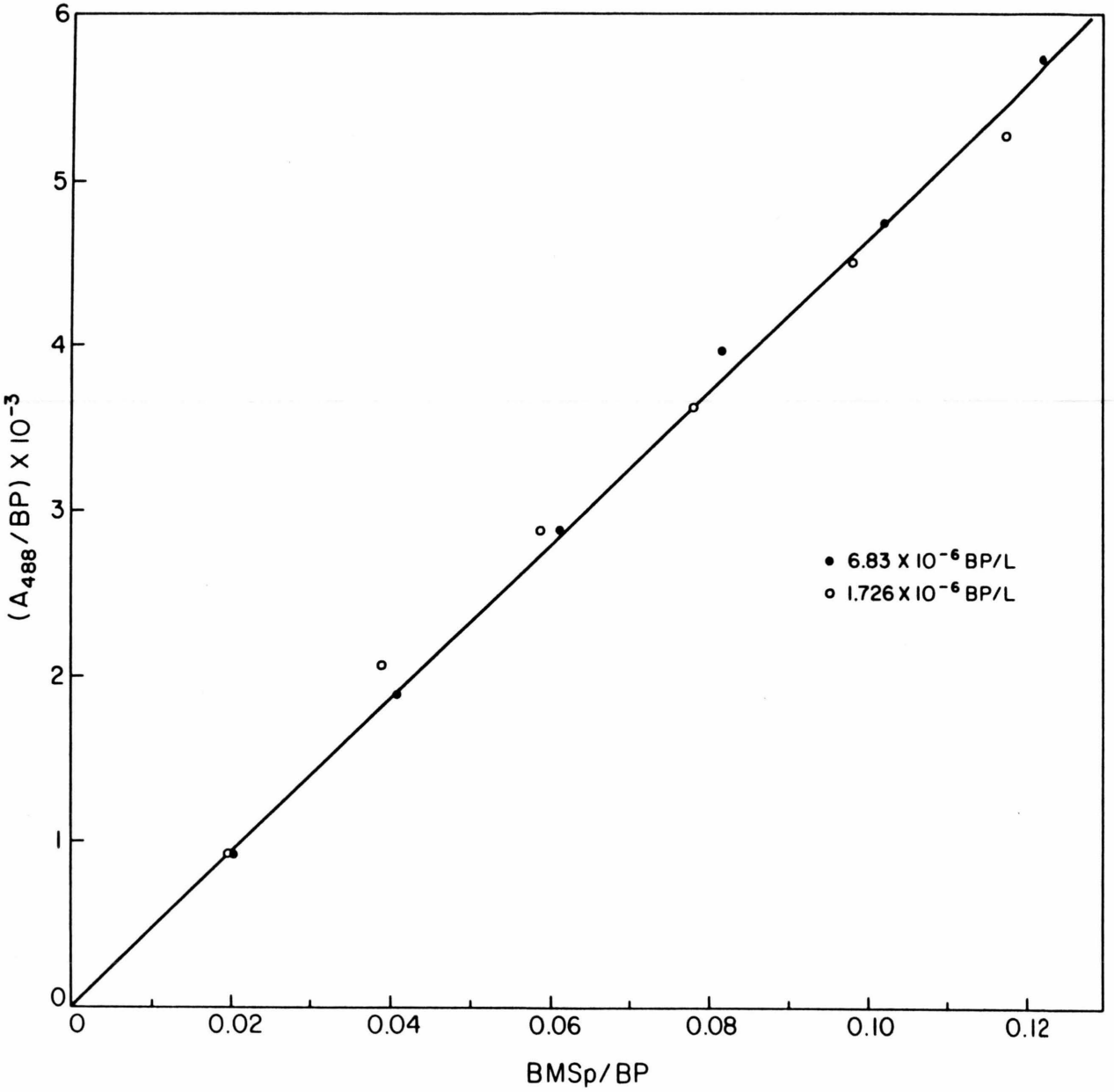
where f(p) is given by:

$$f(p) = \left\{ \frac{3}{\ell Np-A} + \frac{1}{\ell Np-B} \right\} \qquad (5)$$

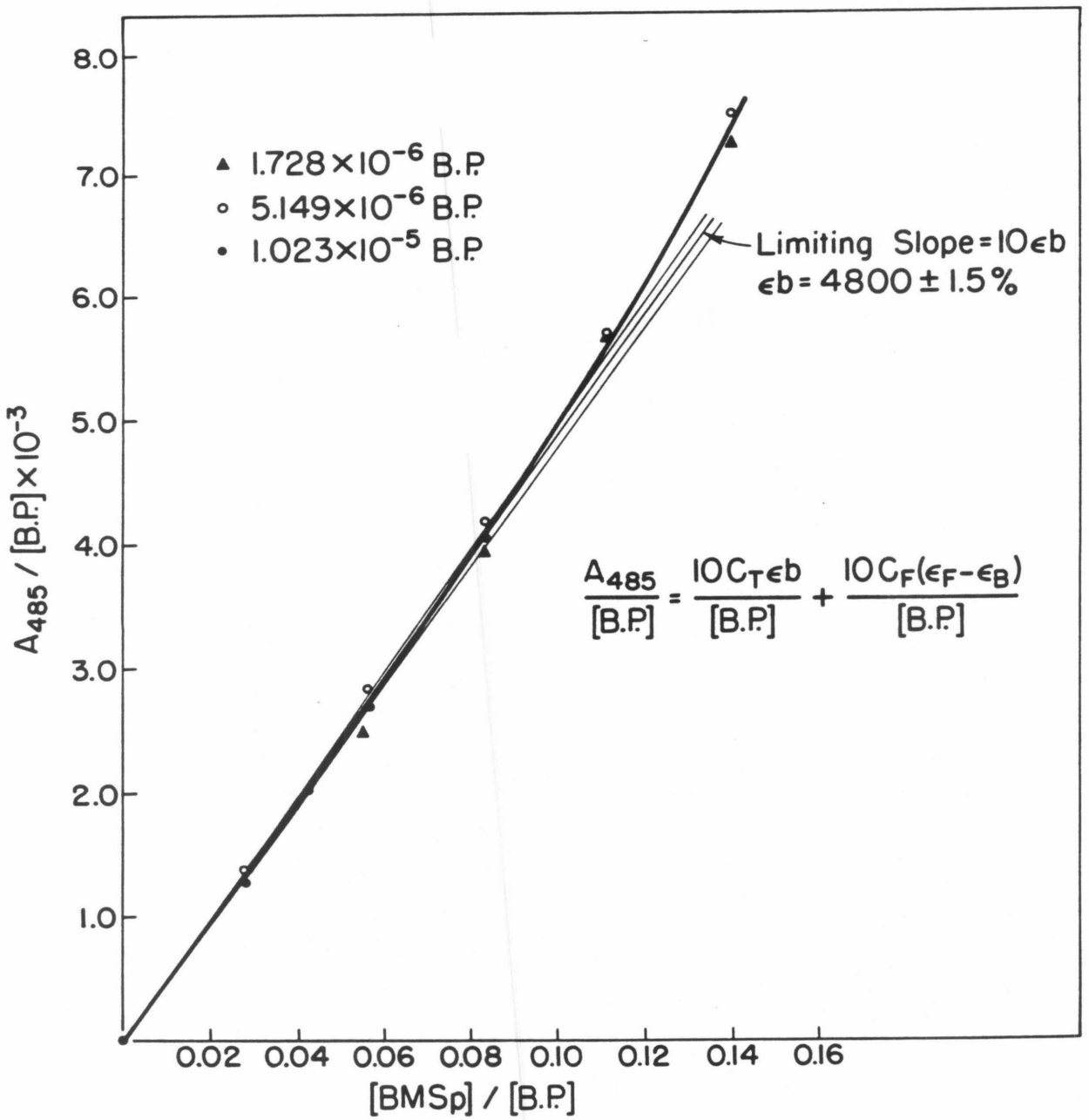
Figure 24

Determination of the extinction coefficient of BMSp bound to (A) polyd(C-G); (B) sonicated calf thymus DNA and (C,D) polydAdT at 1 M^+.

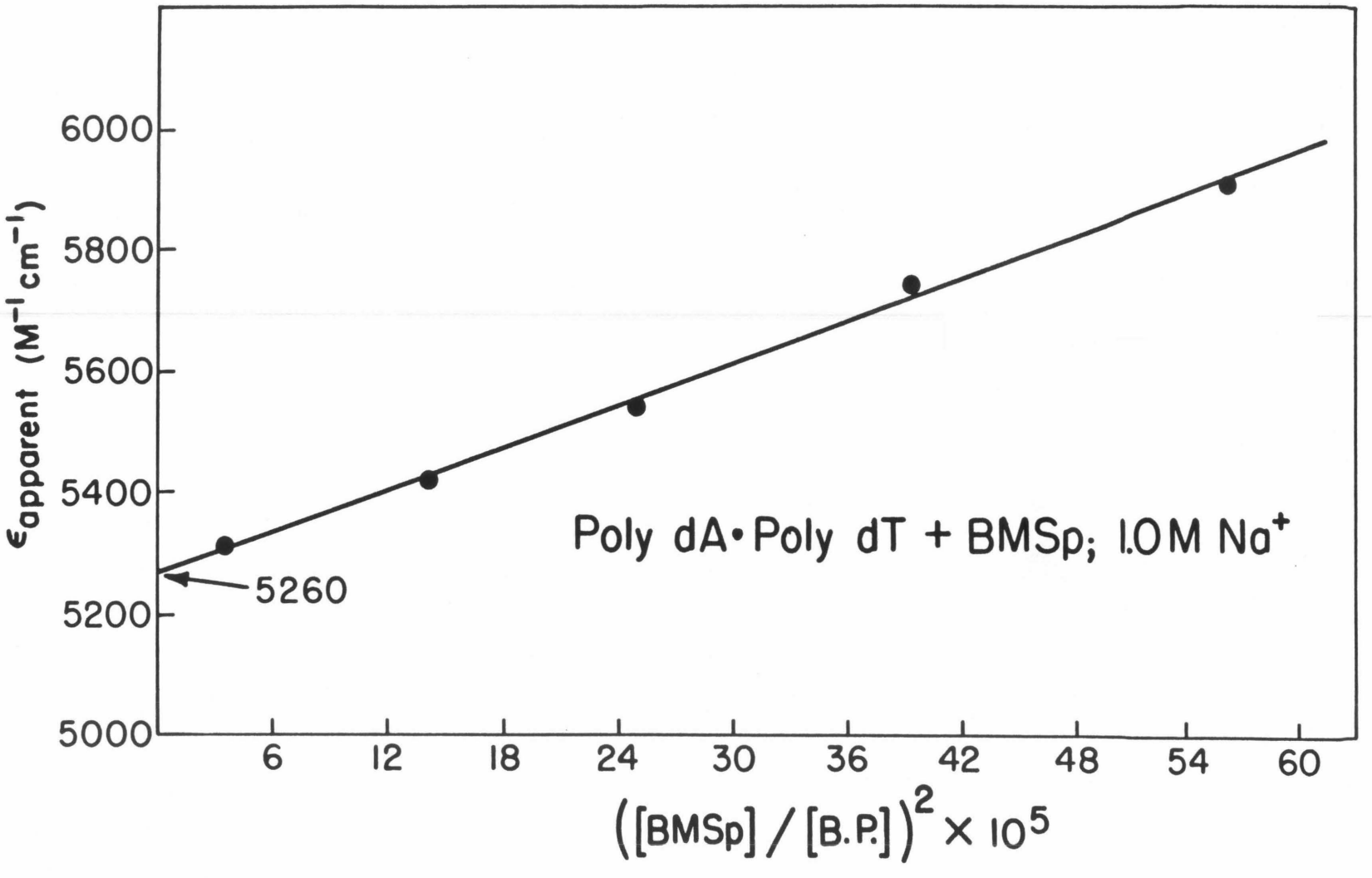
A


6
5
4
3
2
1
0
$(A_{488}/BP) \times 10^{-3}$
0
0.02
0.04
0.06
0.08
0.10
0.12
BMSp/BP
• 6.83 X 10^{-6} BP/L
o 1.726 X 10^{-6} BP/L

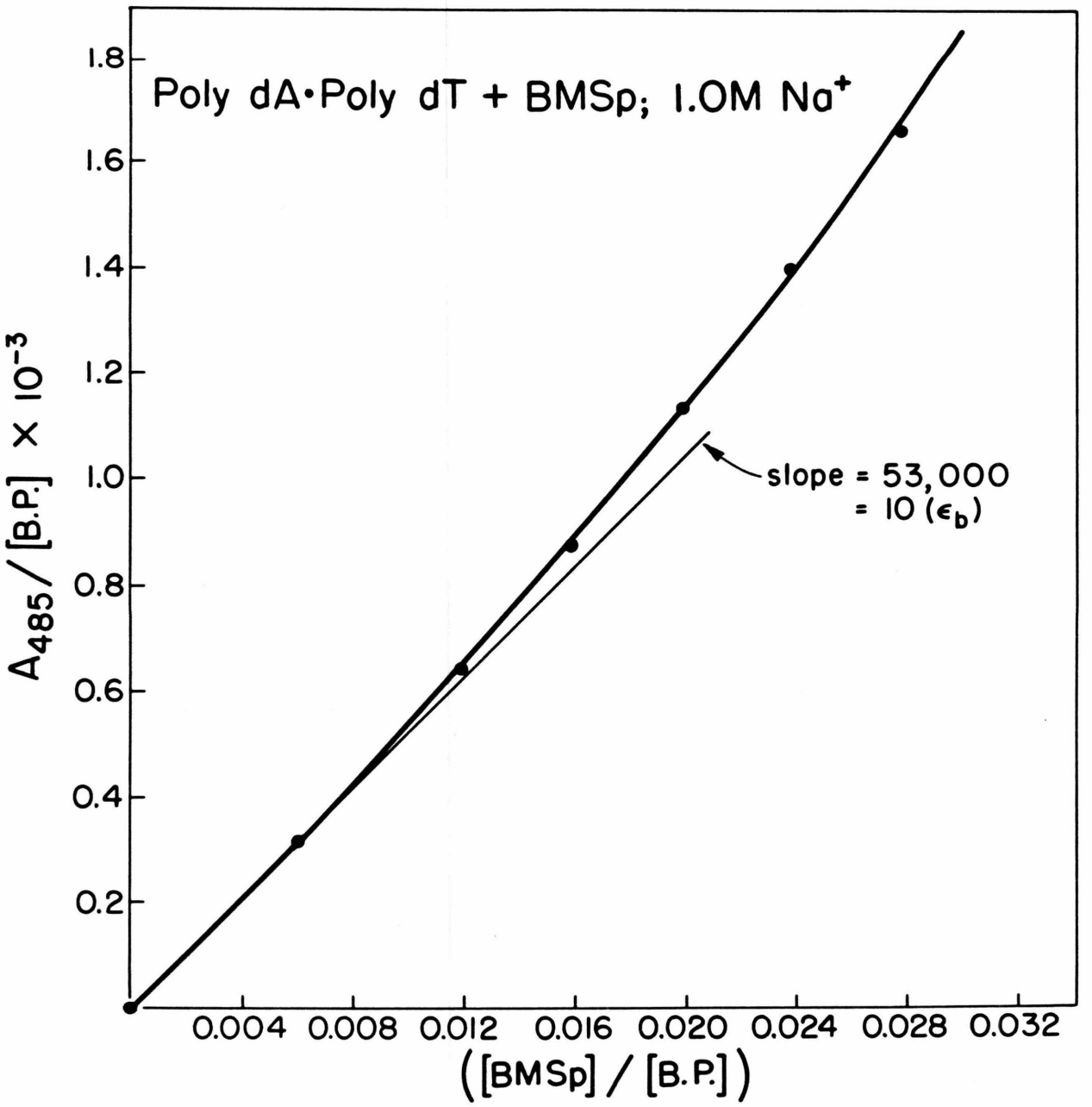
B


8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
A_{485} / [B.P.] × 10^{-3}
▲ 1.728 × 10^{-6} B.P.
○ 5.149 × 10^{-6} B.P.
● 1.023 × 10^{-5} B.P.
Limiting Slope = 10εb
εb = 4800 ± 1.5%
$\frac{A_{485}}{[B.P.]} = \frac{10C_T\epsilon b}{[B.P.]} + \frac{10C_F(\epsilon_F - \epsilon_B)}{[B.P.]}$
0.02 0.04 0.06 0.08 0.10 0.12 0.14 0.16
[BMSp] / [B.P.]

c


6000
5800
5600
5400
5200
5000
$\epsilon_{apparent}$ ($M^{-1}cm^{-1}$)
5260
Poly dA•Poly dT + BMSp; 1.0M Na^+
6
12
18
24
30
36
42
48
54
60
$([BMSp]/[B.P.])^2 \times 10^5$

D


Poly dA•Poly dT + BMSp; 1.0M Na+
slope = 53,000
= 10 (εb)
A485 / [B.P.] × 10-3
([BMSp] / [B.P.])
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.004
0.008
0.012
0.016
0.020
0.024
0.028
0.032

Since (p) changes with drug binding, f(p) must be corrected to yield the desired ratio L/L_0. Assuming that one bound BMSp extends 6.8 Å and one bound EB 3.4 Å, equation 5 can be rewritten as:

$$\frac{L}{L_0} = \left\{\frac{[\eta]_{INT}}{[\eta_0]_{INT}}\right\}^{1/3} \left(1 + 0.2\ \frac{\%\ sat}{100}\right)^{1/3} \qquad (6)$$

where (% sat) equals the percent saturation of available intercalation sites. At saturation, the correction of f(p) for drug extension amounts to a 6% increase in L/L_0.

The intrinsic viscosity is related to the relative viscosity by:

$$\eta_{rel} = [\eta]_{INT} + K[\eta]^2_{INT}\ C \qquad (7)$$

where C equals the DNA concentration in g/dl and K is the Huggins constant.[54] For the 150 base pair DNA used in this study, K has been experimentally determined to equal 5 and the intrinsic viscosity of DNA in the absence of drug equal to 0.392 dl/g.[55] Since K is a function of molecular weight, ligand binding will continually alter its value. However, at saturation (one drug/two base pairs) the effect of this correction on the ratio L/L_0 is less than 1% and therefore is ignored.

F) Competition Equilibrium Dialysis: Estimation of Relative Binding Affinities

When ligand forms one bound complex with macromolecule, the relationship between binding density (ν) and free ligand concentration (L_F) is given by:

$$\nu/L_F = K(1-n\nu)\left(\frac{1-n\nu}{1-(n-1)\nu}\right)^{n-1} \qquad (1)$$

for noncooperative binding and

$$\nu/L_F = K(1-n\nu)\left[\frac{(2\omega-1)(1-n\nu)+\nu-R}{2(\omega-1)(1-n\nu)}\right]^{n-1}\left[\frac{1-(n+1)\nu+R}{2(1-n\nu)}\right]^{2} \qquad (2)$$

where $R = ([1-(n+1)\nu]^2 + 4\omega\nu(1-n\nu))^{1/2}$

for cooperative binding.[37] In the limit of low binding density, equations (1) and (2) reduce to:

$$\lim_{\nu\to 0} (\nu/L_F) = K \qquad (3)$$

If $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$

then $\lim_{x\to a}\frac{f(x)}{g(x)} = L/M$. Thus for two nucleic acids (i) and (j) we can write

$$\frac{\lim_{\nu_i\to 0} (\nu_i/L_F)}{\lim_{\nu_j\to 0} (\nu_j/L_F)} = K_i/K_j \qquad (4)$$

When ligand forms more than one bound complex and binding is noncompetitive we have:[37]

$$\frac{\Sigma\nu_i}{L_F} = \sum_i \left\{ K_i(1-n_i\nu_i) \left(\frac{1-n_i\nu_i}{1-(n_i-1)\nu_i} \right)^{n_i-1} \right\} \qquad (5)$$

Thus for two nucleic acids (i) and (j) we can write:

$$\frac{\underset{\Sigma\nu_i \to o}{\text{Limit}} \frac{\Sigma\nu_i}{L_F}}{\underset{\Sigma\nu_j \to o}{\text{Limit}} \frac{\Sigma\nu_j}{L_F}} = \Sigma K_i / \Sigma K_j \qquad (6)$$

When ligand forms more than one bound complex and binding is competitive we have:[37]

$$\frac{\Sigma\nu_i}{L_F} = \sum_i \left\{ K_i(1-\Sigma n_i\nu_i) \left(\frac{1-\Sigma n_i\nu_i}{1-\Sigma(n_i-1)\nu_i} \right)^{n_i-1} \right\} \qquad (7)$$

Thus for two nucleic acids (i) and (j) we have:

$$\frac{\underset{\Sigma\nu_i \to o}{\text{Limit}} \frac{\Sigma\nu_i}{L_F}}{\underset{\Sigma\nu_j \to o}{\text{Limit}} \frac{\Sigma\nu_j}{L_F}} = \Sigma K_i / \Sigma K_j \qquad (8)$$

When ligand binding induces an allosteric transition in macromolecule, we have:

$$\nu = \frac{n'Lc\alpha(1+c\alpha)^{n'-1} + n'\ (1+\alpha)^{n'-1}}{n'L(1+c\alpha)^{n'} + n'(1+\alpha)^{n'}} \qquad (9)$$

where n' equals the number of macromolecule subunits, c equals the ratio of ligand association constants for the two forms of macromolecule, L equals the allosteric constant and α equals $L_F K_R$, where K_R is the ligand association constant for macromolecule in the R configuration.[59] In the limit of low binding density we can write:[59]

$$\lim_{\nu \to o} \nu/L_F = \left(\frac{1}{1+L}\right)K_R + \left(\frac{L}{1+L}\right)K_T \qquad (10)$$

which, for $N \gtrsim 4$, reduces to

$$\lim_{\nu \to o} \nu/L_F \simeq K_R \qquad (11)$$

Thus for two nucleic acids (i) and (j) we can write:

$$\frac{\lim_{\nu \to o} \frac{\nu_i}{L_F}}{\lim_{\nu \to o} \frac{\nu_j}{L_F}} \simeq {K_R}^i/{K_R}^j \qquad (12)$$

when $N \geq 4$.

PREPARATION OF COMPOUNDS

Preparation of 4,4'-dinitro-2-aminobiphenyl[12]

In a 2 liter 3-necked round bottom flask fitted with Hershberg stirrer, immersion thermometer, and drying tube was place 810 ml of reagent grade concentrated sulfuric acid. The flask was immersed in a cooling bath and cooled to 0°C. To the rapidly stirring solution was added 70g of 2-aminobiphenyl (Aldrich reagent grade distilled *in vacuo* before use: 129°C/1.2 mm) and stirring continued until the biphenylamine dissolved. To the rapidly stirring white solution was slowly added, portion-wise, 84.7g of dry powdered potassium nitrate over a period of two hours. During the addition, the solution was maintained at -2 to -5°C; addition resulted in an initial green solution which later turned olive and then brown in color. Stirring at 2°C was continued for an additional 15 h. The mixture was poured onto 4000 ml of ice and the resulting orange solid isolated by suction filtration. This orange solid was transferred to a 2 liter beaker and stirred for 30 min with 1.2 liters of 5 N sodium hydroxide. After isolation by suction filtration the orange solid was washed with water until the washings were no longer basic to litmus. Recrystallization from 2-ehtoxyethanol followed by drying *in vacuo* at 25°C yielded 74.7g of 4,4'-dinitro-

2-aminobiphenyl (70%): mp 206-208°C (lit. 205-206°); NMR (δ, DMSO-d_6), 8.65 to 8.35 (m, ar), 8.08 to 7.28 (m, ar), 5.8 (s, broad-NH_2, exchanges in D_2O).

Preparation of 4,4'-dinitro-2-(p-cyanobenzamide)biphenyl

In a three-necked, 2 liter round-bottom flask fitted with Hershberg stirrer, reflux condensor, and nitrogen gas inlet was placed 132.5g of dry 4,4'-dinitro-2-amino-biphenyl, 84.7g of p-cyanobenzoyl chloride (Aldrich, reagent grade) and 650 ml of dry chlorobenzene (dried over 4A Linde sieves). The nitrogen inlet was positioned just above the solution and a steady stream of nitrogen was passed through the reaction flask and into an acid trap. The mixture was vigorously stirred to reflux until no further hydrogen chloride was evolved. After cooling overnight, the resulting yellow crystals were isolated by suction filtration and washed with dry chlorobenzene (∿ 500 ml). Recrystallization from dry reagent grade pyridine-absolute ethanol followed by drying at 25°C in vacuo yielded 171.3g (86%) of 4,4'-dinitro-2-(p-cyanobenzamide)biphenyl as yellow crystals: mp (247-249°C); IR (KBr) 3430, 3100, 3075, 2230, 1680, 1620 to 1585, 1530, 1510, 1455, 1420, 1390, 1340, 1315, 1305, 1105, 1080, 1015, 1005, 910, 890, 870, 865, 830, 820, 755, 740, 695, 675, 560.

Preparation of 6-(4-cyanophenyl)-3,8-dinitrophenanthridine[13]

In a three-necked 2 liter round-bottom flask fitted with reflux condenser, Hershberg stirrer, and 250 ml dropping funnel was placed 171g of dry 4,4'-dinitro-2-(p-cyanobenzamide)biphenyl and 650 ml of dry nitrobenzene (distilled *in vacuuo* from P_2O_5). A stream of nitrogen was slowly passed through the dropping funnel and exited into an acid trap. The mixture was vigorously stirred to reflux and a solution consisting of 56 ml of dry phosphous oxychloride (distilled under nitrogen and dried over 4A Linde sieves) and 60 ml of dry nitrobenzene was added dropwise (very exothermic) over a period of one minute. Reflux with vigorous stirring was maintained for an additional three hours and the solution concentrated to one-half volume under reduced pressure. After cooling overnight, the resulting solid was isolated by suction filtration and washed with 300 ml of cold acetone. After air drying the tan solid was dissolved in a minimum amount of hot dry nitrobenzene and filtered. The filtrate upon cooling deposited cream colored crystals which were dried at 25°C *in vacuuo* to yield 105.3g (65%) of the desired product; IR (KBr) 3120-3000, 2230, 1630-1400, 1330 (NO_2), 1090, 1070, 820, 730.

Preparation of 6-(4-carboxylphenyl)-3,8-dinitro-5-methyl-phenanthridinium methosulfate[13]

In a three-necked, 2 liter round-bottom flask fitted with Hershberg stirrer, reflux condenser, immersion thermometer, and nitrogen inlet was heated to 160°C 192 ml of dry dimethyl sulfate (4A Linde sieves). To the rapidly stirring solution was added 75g of 6-(4-cyanophenyl)-3,8-dinitro-phenanthridine. After stirring at 106°C for one hour the mixture was allowed to cool to 50°C and 750 ml of water added all at once. The resulting orange solution was refluxed for four hours, rapidly suction filtered, and the filtrate concentrated to one-half volume under reduced pressure. Storage at 0°C overnight resulted in light brown crystals which were isolated by suction filtration and saved. The filtrate was transferred to a 500 ml round-bottom flask and stirred overnight with 5g of lithium hydroxide. The resulting solution was adjusted to approximately pH 8 with ammonium hydroxide yielding a gummy brown solid which was isolated by suction filtration and air dried. The brown solid was then suspended in water and made weakly acidic. Filtration of the mixture yielded a brown solid which was combined with the brown solid previously isolated. After washing with cold acetone and air drying, 79g (76%) of the desired product was obtained as a light brown solid: IR (KBr) 3420, 3120-3000, 2960,

1715, 1610, 1605, 1590, 1545, 1520, 1345, 1225, 1110, 1070, 1020, 870, 850, 820, 740, 600.

Preparation of 6-(4-caboxyphenyl)-3,8-diaminophenanthridium-5-methylchloride monohydrochloride monohydrate[13]

In a three-necked 2 liter round-bottom flask fitted with Hershberg stirrer, stopper, and reflux condenser were placed 45g of reduced iron powder and 180 ml of distilled water. To the rapidly stirring mixture at reflux was added portionwise over 10 mins, a mixture of 45g of 6-(4-carboxyphenyl)-3,8-dinitro-5-methylphenanthridium methosulfate in 900 ml of boiling water. Reflux was maintained for an additional 30 min and the purple solution rapidly filtered through a coarse glass filter. The filtrate was concentrated to one-half volume and 45 ml of hot concentrated hydrochloric acid added to the boiling solution. The red solution was immediately stored at $0^{o}C$ and cooling overnight yielded dark purple crystals. The purple crystals were isolated by suction filtration, dissolved in approximately 250 ml of hot 1 N hydrochloric acid and filtered through Whatman #1 filter paper. The filtrate was concentrated to a viscous red solution which deposited 20.3g (54%) of the desired product as maroon crystals upon storage at $0^{o}C$ overnight: IR (KBr) 3340,

3200, 3120-3000, 2590, 1790, 1630, 1490, 1475, 1440, 1405, 1315, 1280, 1240, 1180, 1110, 1020, 825, 690. NMR (DMSO-d_6) 4.72-4.62 (m, 9H, ar + phenyl), 3.5 (s, 1H, -H7), 2.87 (s, exchangeable), 2.08 (s, 3H, Me^+).

Preparation of bis-methidium spermine hydrochloride

In a dry 50 ml two-necked round-bottom flask fitted with nitrogen inlet, rubber inlet septum, and magnetic stirrer was placed 500 mg of dry paracarboxy methidium chloride (recrystallized from 1 N HCl), 20 ml of dry dimethyl sulfoxide (distilled from calcium hydride *in vacuuo*) and 155 µl of dry N-ethylmorpholine (stirred over calcium hydride overnight; distilled *in vacuuo* from lithium aluminum hydride). To the purple solution at 25°C was added, all at once, 200 mg of acyldimidazole (freshly sublimed *in vacuuo* before use) in 5 ml of dry dimethyl sulfoxide. After stirring at room temperature for one hour, 126 mg of spermine (Aldrich, dried *in vacuuo* at 25°C) dissolved in 1 ml of dry dimethyl sulfoxide was added. The reaction mixture was allowed to stir overnight and the purple solution upon concentration *in vacuuo* at 25°C yielded a purple solid. This solid was dissolved in a minimum amount of dry methanol (distilled from sodium metal) and chromatographed on silica gel

60 (70-230 ASTM) which had been extensively washed with the elution solvent (100 mg/500g of silica). The elution solvent consisted of dry acidic methanol generated by adding 7 ml of acetylchloride to 1000 ml of dry methanol. The elution pattern consisted of two faint orange bands followed by a dark orange band. The dark orange band was collected, filtered through a fine glass frit (4-4.5 ASTM) and concentrated in vacuo at 25°C. The resulting dark orange solid was dissolved in a minimum amount of 1 N HCl at 25°C in a test tube and precipitated in the cold upon the addition of tetrahydrofuran. This orange solid was washed several times with tetrahydrofuran and upon drying at 110°C in vacuo over P_2O_5 yielded the desired product as maroon crystals. Rechromatography and isolation of the maroon crystals as described above yielded 450 mg (70%) of analytically pure solid: IR (KBr) 3400-3300, 3200, 3120-3000, 2950, 2800, 1630, 1540, 1490, 1475, 1320, 1260, 1235, 1160, 1015, 825. Elemental analysis: Found C, 60.07; H, 5.76; N, 13.41; Cl, 16.82. Calculated ($C_{52}H_{61}N_{10}O_2Cl_5$) C, 60.32; H, 5.94; N, 13.53; Cl, 17.12.

Preparation of mono(methidium)spermine

In a dry two-necked 25 ml round-bottom flask fitted with magnetic stirrer, rubber inlet septum, and nitrogen

inlet were placed 250 mg of dry p-carboxy methidium hydrochloride (recrystalized from 1 N HCl), 15 ml of dry dimethyl sulfoxide, and 78 μℓ of dry N-ethyl morpholine. To the stirring solution at 25°C was added 100 mg of freshly sublimed acyldiimidazole dissolved in 2.5 ml of dry dimethyl sulfoxide. After stirring for one hour, the contents of the flask were transferred under nitrogen with a syringe to the dropping funnel of reaction flask two. Reaction flask two consisted of a 50 ml three-necked round-bottom flask fitted with a rubber inlet septum, nitrogen inlet, magnetic stirrer, and 25 ml dropping funnel. The contents of the dropping funnel were added dropwise, over a period of two hours, to a stirring solution of 610 mg of spermine dissolved in 3 ml of dry dimethyl sulfoxide. Stirring under nitrogen was maintained for an additional 24 h. The purple solution was concentrated to a purple solid at 25°C *in vacuo* and chromatographed in the same way as bis(methidium)-spermine. The elution pattern consisted of a faint orange band (compound 1) followed by a major dark orange band. The dark orange band was collected, filtered through a glass frit (4-5.5 ASTM) and concentrated at 25°C *in vacuo*. After drying *in vacuo* at 110°C the orange solid was rechromatographed and isolated as above. The resulting orange solid was dissolved in 1 *N* ammonium hy-

droxide and extracted with isoamyl alcohol. The isoamyl alcohol was concentrated at 25°C *in vacuo* to yield a dark purple solution which precipitated a purple solid upon the addition to cold ethyl acetate. Thin layer chromatography (cellulose 0.1 mm plates without fluorescent indicator) in 0.025 M phosphate buffer (50%) - 50% 2 propanol indicated the absence of bis(methidium)-spermine, approximately 1% p-carboxyl(methidium)chloride, 8% of a derivative of p-carboxyl(methidium)chloride (presumably the secondary amine condensation product), and 90% of the desired product. The sample contained less than 2% of unreacted spermine as determined by TLC analysis of fluram derivatives. The UV-visible spectrum exhibited maxima at 486, 269, and 243 nm. IR (KBr) 3400, 3200, 3100, 2950, 1630, 1540, 1490, 1475, 1440, 1420, 1320, 1260, 1160, 1015, 965, 885, 825.

REFERENCES AND NOTES

1) a) Sartorelli, A.C., ed. "Cancer Chemotherapy," ACS Symposium series No. 30; American Chemical Society: Washington, DC, 1976; b) Conners, T.A.; Roberts, J.J., eds. "Platinum Coordination Compounds in Cancer Chemotherapy," Springer-Verlag: Heidelberg, 1974; c) LePecq, J.-B.; Xuong, N.D.; Gosse, C.; Paoletti, C., Proc. Natl. Acad. Sci. USA, 1974, 71, 5078; d) Stagnet, M, "Adriamycin Review," EORTC International Symposium European Press: Ghent, Belgium, 1975.

2) For example see: a) Gate, E.F.; Candliffe, E.; Reynolds, P.E., "The Molecular Basis of Antibiotic Action," Wiley: New York, 1972; b) Corcoran, J.W.; Hahn, F.E., eds., "Mechanism of Action of Anti-microbial and Antitumor Agents," Springer-Verlag: New York, 1975; c) Hahn, F.E., ed. "Progress in Molecular and Submolecular Biology," Vol. 2; Springer-Verlag: New York, 1971

3) For representative examples employing ethidium bromide see: a) Hudson, B.; Upholt, U.B.; Devinny, J.; Vinograd, J., Proc. Natl. Acad. Sci. USA, 1969, 62, 813; b) Bittman, R., J. Mol. Biol., 1969, 46, 251; c) Urbanke, C.; Romer, R.; Maass, G., Eur. J.

Biochem., 1973, 33, 511; d) Gray, P.D.; Saunders, G.F., Biochim. Biophys. Acta, 1971, 254, 60; e) Feunteun, J.; Monier, R.; Garrett, R.; LeBret, M.; LePecq, J.-B., J. Mol. Biol., 1975, 93, 535; f) Ide, T.; Baserga, R., Biochemistry, 1976, 15, 600; g) Lawrence, J.J.; Daune, M., Biochemistry, 1976, 15, 3301; h) Wahl, P.H.; Paoletti, J.; LePecq, J.-B., Proc. Natl. Acad. Sci. USA, 1970, 65, 417; i) Millar, D.P.; Robbins, R.J.; Zewail, A.H., Proc. Natl. Acad. Sci. USA, 1980, 77, 5593; j) Bauer, W.; Vinograd, J., J. Mol. Biol., 1968, 33, 141; k) Jones, C.R.; Bolton, P.H.; Kearns, D.R., Biochemistry, 1978, 17, 601.

4) For representative examples employing ethidium bromide see: a) Rigler, R.; Cornvall, E.; Ehrenberg, M.; Pachmann, U.; Hirsch, R.; Zachau, H.G., FEBS Lett., 1971, 18, 193; b) Paoletti, C.; LePecq, J-B.; Lehman, I.R., J. Mol. Biol., 1971, 55, 75; c) LePecq, J.-B., "Methods of Biochemical Analysis," Vol. 20, 41, 1971; d) Ballestra, J.P.G.; Waring, M.J.; Vazquez, D., Nucl. Acid Res., 1976, 3, 1307; e) Andre, J.; Pfeiffer, A.; Rochefort, H., Biochemistry, 1976, 15, 2964; f) Feunteun, J.; Monier, R.,; Garrett, R.; LeBret, M.; LePecq, J-B., J. Mol. Biol. 1975, 93, 535; g) Paoletti, J.; Magee, B.B.; Magee,

P.T., Biochemistry, 1977, 16, 351.

5) Watson, J.D.; Crick, F.H.C., Nature, 1953, 171, 737.

6) Lerman, L.S., J. Mol. Biol., 1961, 3, 18.

7) LePecq, J.-B.; LeBret, M.; Barbet, J.; Rogues, B., Proc. Natl. Acad. Sci. USA, 1975, 72, 2915.

8) a) Waring, M.J., Biochem. J., 1976, 157, 721; b) Canellakis, E.S.; Shaw, Y.H.; Hanners, W.E.; Schwartz, R.A., Biochim. Biophys. Acta, 1976, 418, 277; c) Fico, R.M.; Chen, T.K.; Canellakis, G.S., Science, 1977, 198, 53; d) Waring, M.J.; Wakelin, F.P.G., Nature (London), 1974, 252, 653; e) Dervan, P.B.; Becker, M.M., J. Am. Chem. Soc., 1978, 100, 1968; f) Becker, M.M.; Dervan, P.B., J. Am. Chem. Soc., 1979, 101, 3664; g) Kuhlman, K.F.; Charbencau, N.J.; Mosher, C.W., Nucl. Acid Res., 1978, 5, 2629; h) Wakelin, L.P.G.; Romanos, M.; Chen, T.K.; Glaubiger, D.; Canellakis, E.S.; Waring, M.J., Biochemistry, 1978, 17, 5057; i) Cain, B.F.; Baguley, B.C.; Denny, W.A., J. Med. Chem., 1978, 21, 658; j) Gaugain, B.; Barbet, J.; Oberlin, R.; Roques, B.P; LePecq, J.-B., Biochemistry, 1978, 17, 5071; k) Gaugain, B.; Barbet, J.; Capelle, N.; Roques, B.P.;LePecq, J.-B., Biochemistry, 1978, 17, 5078; l) Lown, J.W.; Gunn, B.C.; Majumdar, K.C.;

McGoran, E., Can. J. Chem., 1979, 57, 2305; m) Wakelin, L.P.G.; Romanos,M.; Cannellakis, E.S.; Waring, M.J., Studia Biophysica, 1976, 60, 111; n) Chen, T.K.; Fico, R.; Cannellakis, E.S., J. Med. Chem., 1978, 21, 868.

9) For examples see: a) LePecq, J-B., "Methods of Biochemical Analysis," Vol. 20, 1971, 41; b) Balis, E., "Antagonists and Nucleic Acids," North Holland Research Monograph, Frontiers of Biology, Vol. 10, Amsterdam, 1968; c) Goldin, A.; Hawkins, F., eds., "Advances in Chemotherapy," Vol. 1, Academic Press, New York, 1964; d) Gottlief, D.; Shaw, P.D.; Corcoran, J.W., eds. "Antibiotics," Vol. 3, Springer-Verlag, New York, 1975; e) Aktipis, S.; Kindelis, A., FEBS Lett., 1974, 49, 149. f) Nishiwahi, H.; Miura, M.; Imai, K.; Ohno, R.; Kawashima, K.; Ezaki, K.; Ueda, R.; Yoshikawa, H.; Nagata, K.; Takeyama, H.; Yamada, K., Cancer Res., 1974, 34, 2699 g) Birkmayer, G.D.; Balda, B.R.; Miller, F., Eur. J. Cancer, 1973, 9, 859; h) Heiner, E., Biochem. Pharmacol., 1974, 23, 1569; i) Balda, B.R.; Birkmayer, G.D., Yale J. Biol. Med., 1973, 46, 464; j) Slonimski, P.P.; Perrodin, G.; Croft, J.H., Biochem. Biophys. Res. Commun., 1968, 30, 232; k) Bouanchaud, D.H.; Scavizzi, M.R.; Chabbert, Y.A., J. Gen. Microbiol.,

1968, 54, 417; 1) Monales, N.M.; Schaefer, F.W.; Keller, S.J.; Meyer, R.R., J. Protozool., 1972, 19 667.

10) a) Hirschman, S.Z.; Leng, M.; Felsenfeld, G., Biopolymers, 1967, 5, 227; b) Stevens, L., Biol. Rev., 1970, 45, 1.

11) a) Bauer, W.; Vinograd, J., J. Mol. Biol., 1970, 47, 419; b) Crothers, D.M., Biopolymers, 1968, 6, 575; c) Bond, P.J.; Langridge, R.; Jennette, K.W.; Lippar, S.J., Proc. Natl. Acad. Sci. USA, 1975, 72, 4825.

12) Cymerman, J.; Short, W.F., J. Chem. Soc., 1949, 1, 703.

13) May and Baker Ltd., Nottingham, England.

14) Waring, M.J., J. Mol. Biol., 1965, 13, 269.

15) LePecq, J-B.; Paoletti, C., J. Mol. Biol., 1967, 27, 87.

16) Olmstead, J.; Kearns, D., Biochemistry, 1977, 16, 3647.

17) Bauer, W.; Vinograd, J., J. Mol. Biol., 1970, 47, 419.

18) Sobell, H.M.; Tsai, C.C.; Jain, S.C., Gilbert, S.G., J. Mol. Biol., 1977, 114, 333.

19) See reference 8k.

20) a) Cohen, G.; Eisenberg, H., Biopolymers, 1969, 8, 45; b) Hogan, M.; Dattagupta, N.; Crothers, D.M., Biochemistry, 1979, 18, 280; c) Reinert, K.E., Biochim. Biophys. Acta, 1973, 319, 135; d) Saucier, J.M.; Festy, B.; LePecq, J.-B., Biochemie, 1971, 53, 973; e) Muller, W.; Crothers, D.M., J. Mol. Biol., 1968, 35, 251.

21) a) see reference 11a; b) Waring, M.J., J. Mol. Biol., 1970, 54, 247.

22) Reinert, K.E., J. Mol. Biol., 1972, 72, 593.

23) Tatchell, K.; Van Holde, K.E., Biochemistry, 1977, 16, 5295.

24) See reference 20a.

25) Vinograd, J.; Bauer, W., J. Mol. Biol., 1968, 33, 141.

26) Revet, B.; Schmir, M.; Vinograd, J., Nature New Biol., 1971, 229, 10.

27) Wang, J.C., J. Mol. Biol., 1974, 89, 783.

28) Miller, K.J.; Pycior, J.F., Biopolymers, 1979, 18, 2863.

29) Saucier, J.M.; Festy, B.; LePecq, J.-B., Biochemie, 1971, 53, 973.

30) See reference 20b.

31) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

32) Manning, G.S., Quart. Rev. Biophys. II, 1978, 2, 179.

33) Anderson, C.F.; Record, M.T.; Hart, P.A., Biophys. Chem., 1978, 7, 301.

34) a) Record, M.T.; Anderson, C.F.; Lohman, T.M., Quart. Rev. Biophys. II, 1978, 2, 103; b) Record, M.T.; Lohman, T.M.; DeHaseth, P., J. Mol. Biol., 1976, 107, 145.

35) See reference 8k.

36) Scatchard, G., Ann. N.Y. Acad. Sci., 1949, 51, 660.

37) McGhee, J.D.; von Hippel, P.H., J. Mol. Biol., 1974, 86, 469.

38) a) Müller, W.; Crothers, D.M.; Waring, M.J., Eur. J. Biochem., 1973, 39, 223; b) Müller, W.; Crothers, D.M., ibid, 1975, 54, 267.

39) Becker, M.M.; Dervan, P.B., J. Am. Chem. Soc., 1979, 101, 3664.

40) Bresloff, J., Ph.D. Dissertation, Yale University, 1974.

41) Milman, G.; Langride, R.; Chamberlin, M.J., Proc. Natl. Acad. Sci. USA, 1967 57, 1804.

42) Arnott, S.; Fuller, W.; Hodgson, A.; Prutton, I., Nature, 1968, 220, 561.

43) a) Langride, R.; Seeds, W.E.; Wilson, H.R.; Hooper, C.W.; Wilkins, M.H.F.; Hamilton, L.D., J. Biophysc. and Biochem. Cytol., 1957, 3, 767; b) Hamilton, L.D.;

Barclay, R.K.; Wilkins, M.H.F.; Brown, G.L.; Wilson, H.R.; Marvin, D.A.; Ephrussi-Taylor, H.; Simmons, N.J., J. Biophys. Biochem. Cytol., 1959, 5, 397; c) Bram, S., J. Mol. Biol., 1971, 58, 277; d) Wang, J.C., Proc. Natl. Acad. Sci. USA, 1979, 76, 200; e) Arnott, S.; Hukins, D.W.L., Biochem. Biophys. Res. Commun., 1972, 47, 1504.

44) Monod, J.; Wyman, J.; Changeux, J.P., J. Mol. Biol., 1965, 12, 88.

45) a) Shäfer, R.H., Biopolymers, 1980, 19, 419; b) Shafer, R.H.; Waring, M.J., Biopolymers, 1980, 18, 431.

46) Ikeda, R., Caltech, unpublished results.

47) Wartell, R.M.; Larson, J.E.; Wells, R.D., J. Biol. Chem., 1974, 249, 6719.

48) Wells, R.D.; Larson, J.E.; Grant, R.C.; Shortle, B.E.; Cantor, C.R., J. Mol. Biol., 1970, 54, 465.

49) Riley, M.; Maling, B.; Chamberlin, M.J., J. Mol. Biol., 1966, 20, 359.

50) Pohl, F.M.; Jovin, T.M., J. Mol. Biol., 1972, 67, 375.

51) Watson, R.; Bauer, W.R., Biopolymers, 1977, 16, 1343.

52) Bontemps, J.; Fredericq, E., Biophys. Chem., 1974, 2, 1.

53) Cohen, G.; Eisenberg, H., Biopolymers, 1966, 4, 429.

54) Godfrey, J.E., Biophys. Chem., 1976, 5, 285.

55) D. Straus, Caltech, unpublished work.

55) Based on two separate determinations. The second measurement is shown in the Appendix Fig. 1.

56) Raw absorbance data are shown in Fig. 2 of the Appendix.

57) Raw absorbance data are shown in Fig. 3 of the Appendix.

58) Bresloff, J.; Crothers, D.M., unpublished results.

59) Gibson, R.E.; Levin, S.A., Proc. Natl. Acad. Sci. USA, 1977, 74, 139.

CHAPTER II

Experimental Determination of Equilibrium Binding Isotherms at Constant X: A New Sensitive Technique

INTRODUCTION

Ligand-macromolecule interactions often constitute the molecular basis of biological function. Their complete description necessitates specification of a binding: site size, affinity, and cooperativity for each type of bound ligand. In addition, the presence of multiple forms of free ligand species as well as multiple binding forms of macromolecule must be considered. The most widely used technique for extracting these binding parameters from experimental data entails plotting the calculated binding density ([bound ligand]/[macromolecule subunits]) versus the same quantity divided by the corresponding free concentration of ligand. This plot, termed a Scatchard plot,[1] owes its wide application to the sensitivity which it shows toward changes in binding parameters. Although theoretical interpretation of Scatchard plots is possible for many different binding situations,[2-6] acquisition of accurate data for their construction is often difficult.

Experimentally, ligand-macromolecule interactions are monitored by indirect or direct methods. Equilibrium dialysis of radioactively labeled ligand is an example of a direct method because the concentration of bound and free ligand(s) can be calculated without specifying species

specific proportionality factors. In contrast, indirect methods such as spectroscopic techniques necessitate specification of the extinction coefficient, quantum yield, Beer's Law behavior, etc. of each species before their concentrations can be calculated.

Although less dependent on assumptions, direct methods are generally less applicable, less sensitive, and more subject to error than indirect methods. For example, consider how concentrations of free and bound ligand are determined by equilibrium dialysis, a direct technique. Radioactive ligand is added to both compartments of a two-part dialysis cell which contains the macromolecule of interest in one cell. A membrane impermeable to macromolecule separates both compartments. At equilibrium, the compartment with macromolecule contains both bound and free ligand whereas the other compartment contains only free ligand. At low binding density the concentrations of free and bound ligands are very small. Since the concentration of bound ligand is determined by subtraction of two quantities of <u>similar</u> magnitude (either $[L_B + L_F] - L_F$ or $L_T - L_F$, where L_T equals the total input ligand concentration), the error in L_B will be large. Thus binding parameters at low binding density cannot be accurately measured by direct methods. Similar behavior will be exhibited at high

binding densities since as the concentration of L_F relative to L_B becomes large, L_B will be determined by the subtraction of two quantities of similar magnitude.

Additional complications arise when ligand-macromolecule binding is tight or weak. When binding is tight, very dilute concentrations of ligand and macromolecule must be employed to assure detection of free ligand. This restriction introduces two complications. First, the concentration of bound and free ligand will often be too small for accurate detection by either direct or indirect methods. Secondly, since dilute concentrations of macromolecule must be employed, nonspecific binding of ligand and macromolecule to cell walls and/or membrane becomes important. Therefore, calculation of binding densities by either method will require determination of bound ligand and macromolecule concentrations actually present in solution. This requirement will enhance the error in the corresponding binding density since it is a ratio of these quantities.

When ligand-macromolecule binding is weak, large concentrations of macromolecule must be employed to assure formation of bound ligand. Often, however, this requirement cannot be tolerated experimentally since the complex between ligand and macromolecule will be too insoluble at these high concentrations.

In light of these deficiencies we propose a new method for estimating accurate ligand binding parameters by indirect measurements. In contrast to other indirect techniques or direct measurements, accurate binding parameters can be estimated over a wide range of binding densities for tight or weakly binding ligands in most cases by this method.

Theory

When added to a solution of macromolecule, some ligand will bind and some will remain free in solution. The total ligand concentration in the presence of macromolecule can therefore be written as:

$$(1) \qquad L_T = \sum_i L_B^i + \sum_i L_F^i$$

where L_B^i and L_F^i equal the concentration of bound and free ligand species i, respectively. If some physical property X of ligand is monitored during binding, its value will be given by:

$$(2) \qquad X = \sum_i q_B^i L_B^i + \sum_i q_F^i L_F^i$$

where q_B^i and q_F^i are species specific proportionality factors of bound and free ligands, respectively.

For any given concentration of macromolecule M, associated with each value X is an unknown binding density $\Sigma L_B^i/M$ and free ligand concentration ($\Sigma\ L_F^i$) whose values we wish to determine. These quantities can be related to the experimental observables X, M, and L_T/M in the following manner. Dividing equation (1) by M and letting the total free ligand concentration equal L_F^T yields:

$$(3) \qquad L_T/M = \sum_i L_B^i/M + L_F^T/M$$

Since L_B^i is proportional to M at constant L_F (see later discussion) equation (3) can be rewritten as:

$$(4) \qquad L_T/M = \Sigma\ \nu_i + L_F^T/M$$

where $\sum_i L_B^i/M$ is now a constant ($\sum_i \nu_i$). Since ν and L_F^T are independent of M, (4) is a straight-line equation.[7] Therefore, if two or more solutions have the same ν and L_F^T, the dependence of L_T/M on M for these solutions can be used to determine their total binding density (intercept) and total free ligand concentration (slope). As discussed below, two or more solutions will have the same binding density and free ligand concentration when their values of X are identical and a) all ligand is essentially free or b) their macromolecule concentrations are similar.

Consider a ligand-macromolecule equilibrium in which each bound ligand covers one macromolecule site and one type of bound and free ligand are formed. The relationship between ν and L_F is then given by:

$$\nu/L_F = K(1 - \nu) \tag{5}$$

where K is the ligand-macromolecule association constant.[1] Substituting equation (5) into equation (2) and rearranging yields:

$$\frac{X}{M} = q_B\nu + \frac{q_F}{MK}\left(\frac{\nu}{1-\nu}\right) \tag{6}$$

If $M << 1/K$ (or equivalently $MK << 1$), than for any ν, equation (6) can be rewritten as

$$\frac{X}{M} = \frac{q_F}{KM}\left(\frac{\nu}{1-\nu}\right) \tag{7}$$

or equivalently, from equation (5):

$$X \simeq q_F L_F \tag{8}$$

Thus when $M << 1/K$, two or more solutions with the same value of X will have similar free ligand concentrations. Furthermore, as required by equation (5), these solutions will also have similar binding densities (ν).

Recalling equation (2) and solving for L_F^T yields

$$L_F^T = \frac{(X - q_B L_B)}{q_F} \tag{9}$$

Substituting this expression into equation (4) yields:

$$\frac{L_T}{M} = \frac{X - q_B L_B}{q_F} \frac{1}{M} + \nu \qquad (10)$$

or

$$\frac{L_T}{M} = \left(\frac{X}{q_F}\right) \frac{1}{M} + \left(\frac{q_F - q_B}{q_F}\right)\nu \qquad (11)$$

Since $X \simeq q_F L_F$ when M is significantly less than 1/K (equation 8), equation (11) can be rewritten:

$$\frac{L_T}{M} \simeq \frac{L_F}{M} + \frac{q_F - q_B}{q_F} \nu \qquad (12)$$

Thus if two or more solutions have the same value of X and $M << 1/K$, a plot of L_T/M versus $1/M$ for these solutions will be a straight-line with slope L_F and intercept $(q_F - q_B/q_F)/\nu$.

When M is not significantly less than 1/K the value of X will be a function of both bound and free ligand. From equation (2) we have:

$$X = q_F L_F + q_B \nu M \qquad (13)$$

Since ν is now a function of M at constant X, equation (4) will not be a straight-line with slope L_F and intercept ν. However, if a narrow range of macromolecule concentration is investigated, solutions with the same value of X will have nearly identical values of ν and L_F. Since equation (4) can be linear over this range of M, values of

ν and L_F may be estimated from the corresponding intercept and slope respectively. For example, reconsider the non-cooperative binding of a ligand to macromolecule in which each bound ligand covers one lattice site. For this simple case, the behavior of equation (4) at constant X for any range of M can be calculated directly from equation (11) by assigning known values to X, M, ν and L_F. Letting $q_B = 2300$, $q_F = 6000$ and $K = 1 \times 10^4$, the dependence of ν, L_F and L_T/BP on M at constant X is given in Table I. As expected from equations (5) and (11), when M is not less than 1/K, the value of ν and L_F decrease with an increase in M at constant X. As a result, a Scatchard plot constructed from these estimated ν,L_F values underestimates the true Scatchard plot (see Fig. 1). However, the following points should be noted:

a) Both estimated and true Scatchard plots are linear - that is, little or no distortion (curvature) is introduced if macromolecule concentrations not significantly less than 1/K are investigated;

b) Although estimated values of K and ν will be obtained, the underestimation is not bad; eg. K(estimated) = 3.7×10^3 whereas K(true) = 1×10^4;

c) Better estimated Scatchard plots are obtained when the range of M considered is more dilute than 1/K; when $M \ll 1/K$ only the value of ν is significantly underestimated, its value being $(q_F-q_B/q_F)\nu$.

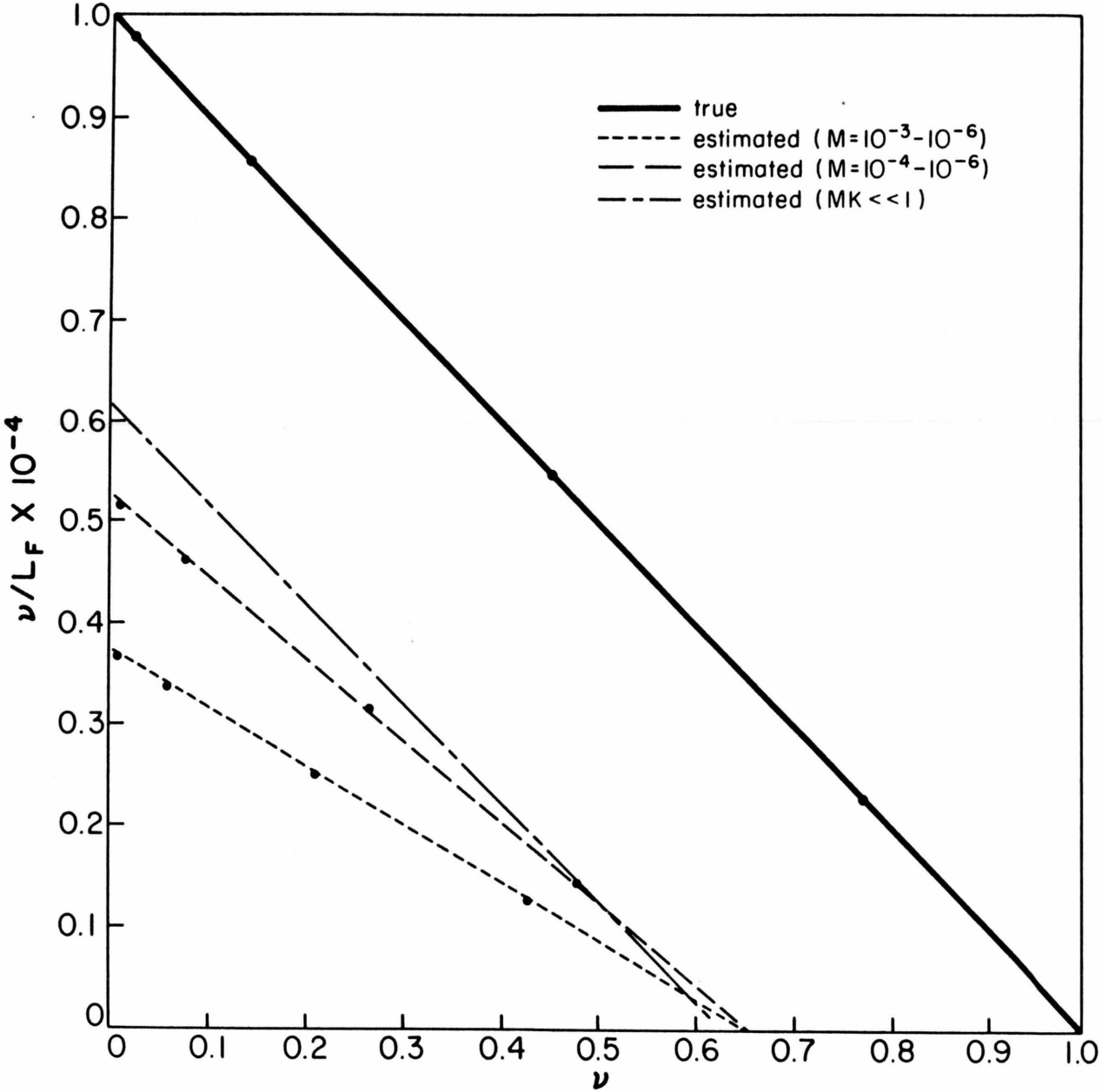
Table I

Dependence of binding density (ν) and free ligand concentration (L_F) on macromolecule concentration (M) at constant absorbance (A). Bound ligand is assumed to interact noncooperatively and cover one macromolecule subunit. The association constant is 1×10^4 and the extinction coefficients of bound and free ligand are 2300 and 6000 M^{-1} cm^{-1} respectively.

		A = 0.01		A = 0.1		A = 0.5		A = 2.0	
[M]	$[M]^{-1}$	υ	L_F	υ	L_F	υ	L_F	υ	L_F
5×10^{-2}	2×10^{1}	0.0000865	8.651×10^{-9}	0.000865	8.657×10^{-7}	0.004325	4.3437×10^{-7}	0.01729	1.759×10^{-6}
1×10^{-2}	1×10^{2}	0.000424	4.2388×10^{-8}	0.004237	4.255×10^{-7}	0.02117	2.163×10^{-6}	0.08455	9.236×10^{-6}
5×10^{-3}	1×10^{2}	0.000826	8.271×10^{-8}	0.00826	8.33×10^{-7}	0.04123	4.3×10^{-6}	0.1637	1.957×10^{-5}
1×10^{-3}	1×10^{3}	0.00345	3.45×10^{-7}	0.0342	3.54×10^{-6}	0.1656	1.984×10^{-5}	0.515	1.06×10^{-4}
5×10^{-4}	2×10^{3}	0.0057	5.73×10^{-7}	0.056	5.93×10^{-6}	0.2556	3.434×10^{-5}	0.672	2.05×10^{-4}
1×10^{-4}	1×10^{4}	0.01194	1.21×10^{-6}	0.11	1.236×10^{-5}	0.4042	6.78×10^{-5}	0.753	3.04×10^{-4}
5×10^{-5}	2×10^{4}	0.0138	1.4×10^{-6}	0.1249	1.423×10^{-5}	0.429	7.513×10^{-5}	0.761	3.18×10^{-4}
1×10^{-5}	1×10^{5}	0.0158	1.61×10^{-6}	0.139	1.613×10^{-5}	0.4494	8.162×10^{-5}	0.768	3.303×10^{-4}
2×10^{-6}	5×10^{5}	0.0163	1.65×10^{-6}	0.142	1.65×10^{-5}	0.4535	8.298×10^{-5}	0.769	3.327×10^{-4}
1×10^{-6}	1×10^{6}	0.01633	1.66×10^{-6}	0.1424	1.66×10^{-5}	0.454	8.315×10^{-5}	0.769	3.33×10^{-4}
1×10^{-7}	1×10^{7}	0.0164	1.67×10^{-6}	0.1428	1.666×10^{-5}	0.4545	8.33×10^{-5}	0.769	3.33×10^{-4}

Figure 1

Scatchard plot describing the noncooperative binding of a ligand to a macromolecule with association constant 1 x 10^4 and a binding site size of one macromolecule subunit (denoted as true). When ligand binding is monitored at constant absorbance, the binding density (ν) and free ligand concentration (L_F) determined by equation (4) for three ranges of macromolecule concentration: MK << 1, (10^{-3}-10^{-5}) and (10^{-4}-10^{-6} macromolecule subunits/liter) yield the indicated "estimated" Scatchard plots.


true
estimated (M=10^-3-10^-6)
estimated (M=10^-4-10^-6)
estimated (MK<<1)
ν/L_F X 10^-4
ν
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Although only the simplest ligand-macromolecule interaction possible has been considered, eq. (4) can be applied at constant X to <u>any</u> ligand-macromolecule equilibrium. For example, as noted previously, when $M << 1/K$ all ligand is essentially free. Since this statement will be true regardless of the number or type of bound complexes formed, at constant X and for $M << 1/K$ eq. (4) will be a straight-line with intercept $(q_F-q_B/q_F)\nu$ and slope L_F for any ligand-macromolecule interaction. When M is not significantly less than 1/K, eq. (4) can still be applied to any ligand-macromolecule interaction since, regardless of the number or nature of bound forms, solutions with similar M will exhibit similar ν and L_F values at constant X. For non-cooperative (Fig. 1) and cooperative (Fig. 2) ligand binding, the values of ν and L_F so obtained underestimate the true ν, L_F values characterizing the ligand-macromolecule equilibrium.

In summary, for any ligand-macromolecule interaction, if two or more solutions have the same value of X, and $M << 1/K$, than a plot of L_T/M versus $1/M$ will be a straight-line with intercept $(q_F-q_B/q_F)\nu$ and slope L_F. If M is not less than 1/K, a plot of L_T/M versus $1/M$ will not be linear. However, if a narrow macromolecule concentration range is investigated, equation (4) will be linear and its intercept and slope will yield good estimates of the values of ν and L_F, respectively.

Application of equation (4) requires that L_B be proportional to M at constant L_F. Two general binding situations must be considered: ligand binding is (A) noncooperative or (B) cooperative. Noncooperative ligand binding is considered first. When the binding of ligand to macromolecule results in the formation of only one type of bound species, the relationship of the binding density to the free ligand concentration is given by:

$$(14) \qquad \frac{L_B}{M} = \nu = K(1-n\nu)L_F \left[\frac{1-n\nu}{1-(n-1)\nu}\right]^{n-1}$$

where K is the association constant and (n) is the number of macromolecule subunits covered by one bound ligand.[2] Rearranging equation (12) yields:

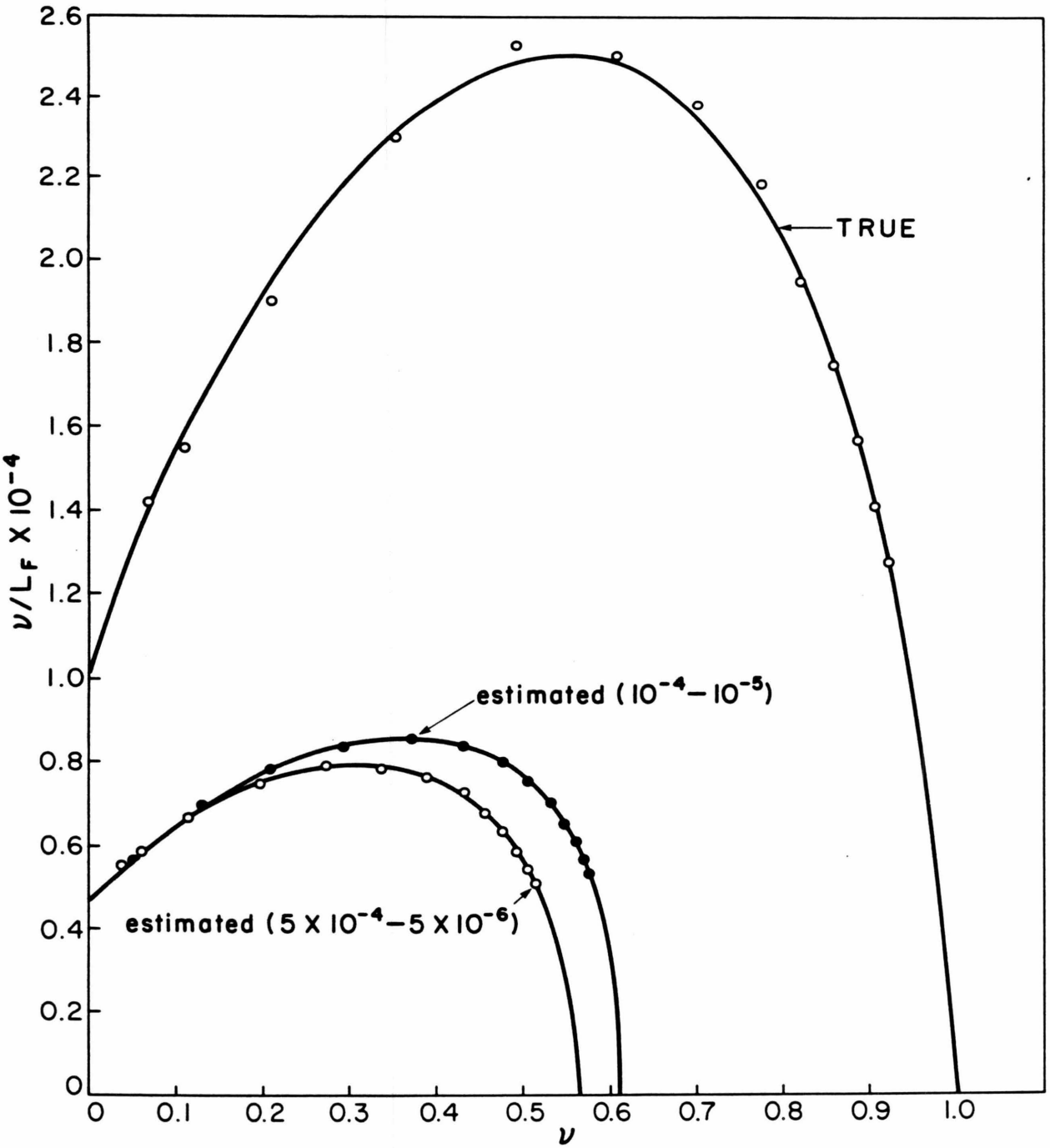
$$(15) \qquad -KL_F(1-n\nu)^n + \nu[1-(n-1)\nu]^{n-1} = 0$$

Since (n) is integral, equation (13) is polynomial and its solution, ν, will equal a constant as desired.

When two or more bound complexes are formed and binding is noncompetitive, $\sum_i \nu_i$ is given by:[2]

Figure 2

Scatchard plot describing the cooperative binding of a ligand to a macromolecule with association constant 1×10^4, a binding site size of one macromolecule subunit, and a cooperativity factor[2] equal to 5 (denoted as "true"). When ligand binding is monitored at constant absorbance, the binding density (ν) and free ligand concentration (L_F) determined by equation (4) for two ranges of lattice concentration (5×10^{-4}-5×10^{-6}) and (1×10^{-4}-1×10^{-6}) macromolecule subunits/liter) yield the indicated "estimated plots". Ligand is assumed to have an extinction coefficient of 2300 M^{-1} cm^{-1} (bound), 6000 cm^{-1} M^{-1} (unbound) and points are calculated for an absorbance range of 0.06-0.65. Points distributed about the "true" Scatchard plot are obtained by correcting the estimated Scatchard plot (5×10^{-4}-5×10^{-6}) according to equation (21) with $M^* = 7 \times 10^{-5}$ and $\nu = 1.8\nu'$. See text for further details.


2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
ν/L_F X 10^-4
ν
TRUE
estimated (10^-4 – 10^-5)
estimated (5 X 10^-4 – 5 X 10^-6)

$$(16)\quad \sum_i \nu_i = L_F \sum_i \left\{ K_i (1-n_i\nu_i) \left(\frac{1-n_i\nu_i}{1-(n_i-1)\nu_i} \right)^{n_i-1} \right\}$$

whereas when binding is competitive $\sum_i \nu_i$ is given by:

$$(17)\quad \sum_i \nu_i = L_F \sum_i \left\{ K_i (1-\sum_i n_i\nu_i) \left(\frac{1-\sum_i n_i\nu_i}{1-\sum_i (n_i-1)\nu_i} \right)^{n_i-1} \right\}$$

Since $\sum_i \nu_i / L_F$ is measured experimentally, when two bound species are formed equations (14 and 15) will be the sum of two polynomial equations containing one unknown variable, ν_1. Such an equation is also polynomial and its solution, $\sum_i \nu_i$, will be equal to a constant as desired.

When ligand binds cooperatively as through ligand-ligand contact, ν is given by:

$$(18)\quad \nu = L_F K (1-n\nu) \left[\frac{(2\omega-1)(1-n\nu)+\nu-R}{2(\omega-1)(1-n\nu)} \right]^{n-1} \cdot \left[\frac{1-(n+1)\nu+R}{2(1-n\nu)} \right]^2$$

where $R = [(1-(n+1)\nu)^2 + 4\omega\nu(1-n\nu)]^{1/2}$ and ω is the cooperativity constant.[2] As with equation (12), (16) is a polynomial equation whose solution, ν is a constant.

Finally, when cooperative ligand binding arises from a ligand induced allosteric transition in macromolecule, the binding density is given by:

$$(19) \quad \nu = \frac{n'Lc\alpha(1+c\alpha)^{n'-1} + n'\alpha(1+\alpha)^{n'-1}}{n'L(1+c\alpha)^{n'} + n'(1+\alpha)^{n'}}$$

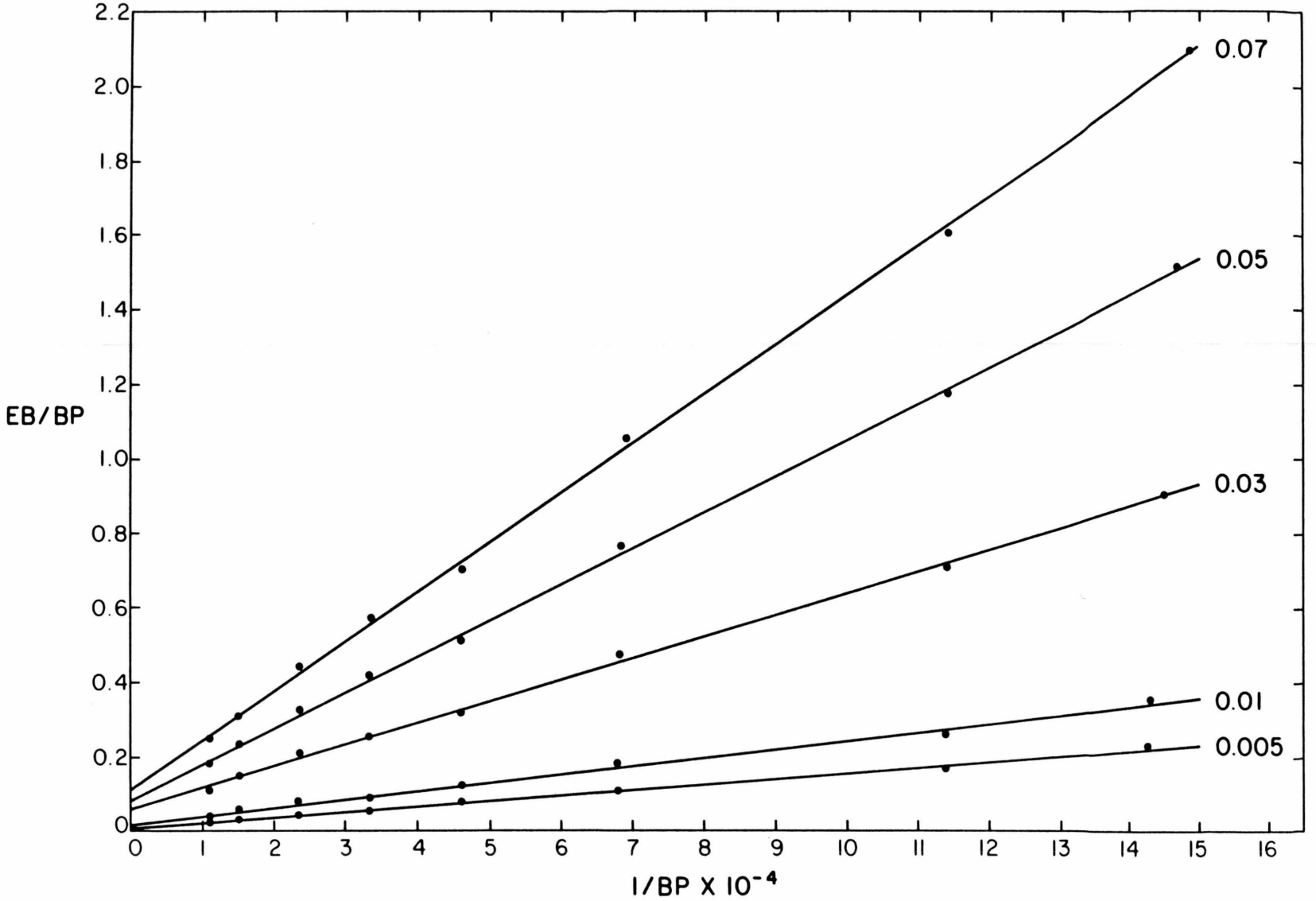
where n' equals the number of macromolecule subunits, c equals the ratio of ligand association constants for the two forms of macromolecule, L equals the allosteric constant and α equals $(L_F)K_R$ where K_R is the ligand association constant for macromolecule in the R configuration.[5] Since the right side of equation (17) is a constant at constant L_F, ν will also be a constant.

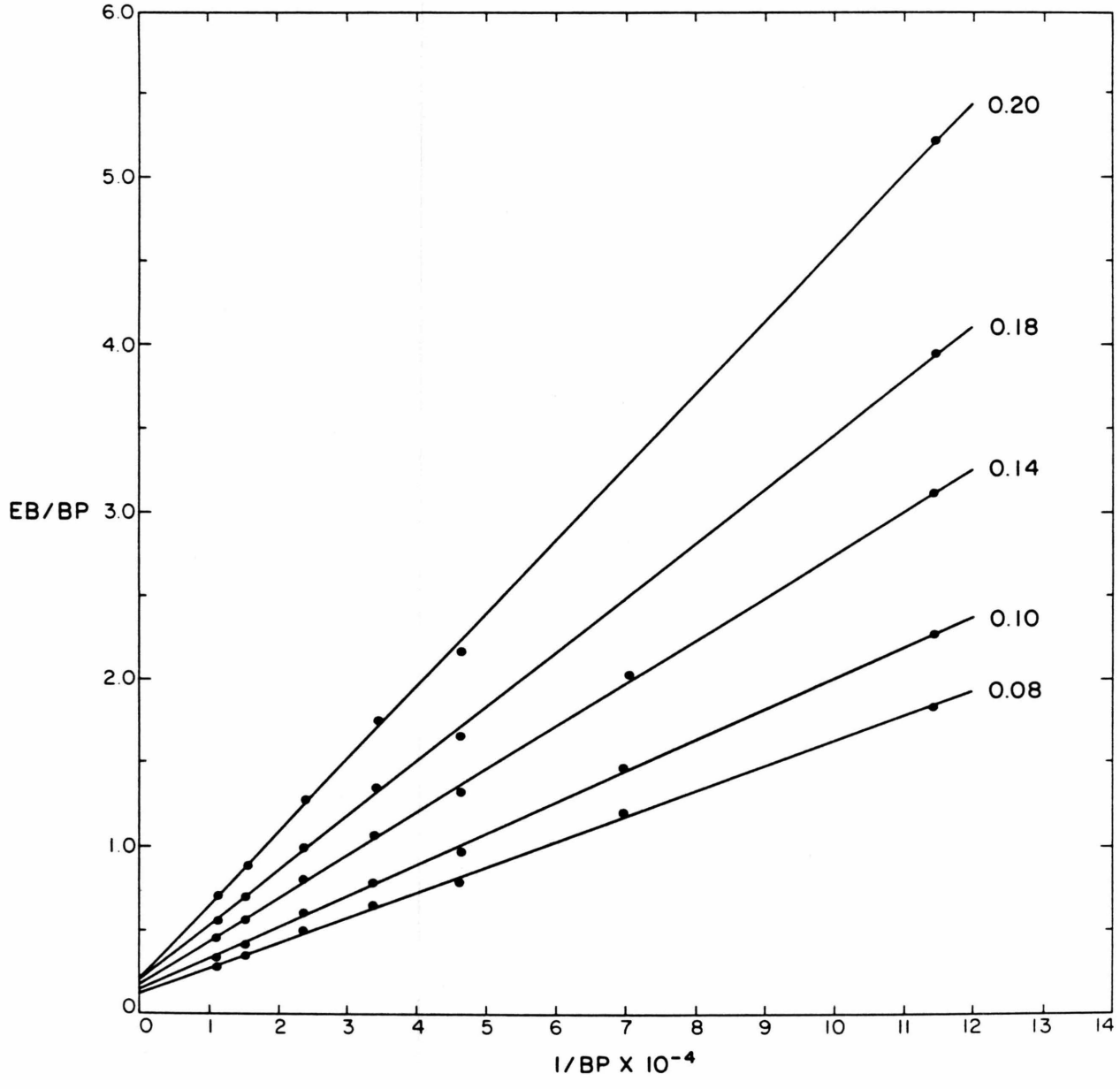
APPLICATIONS

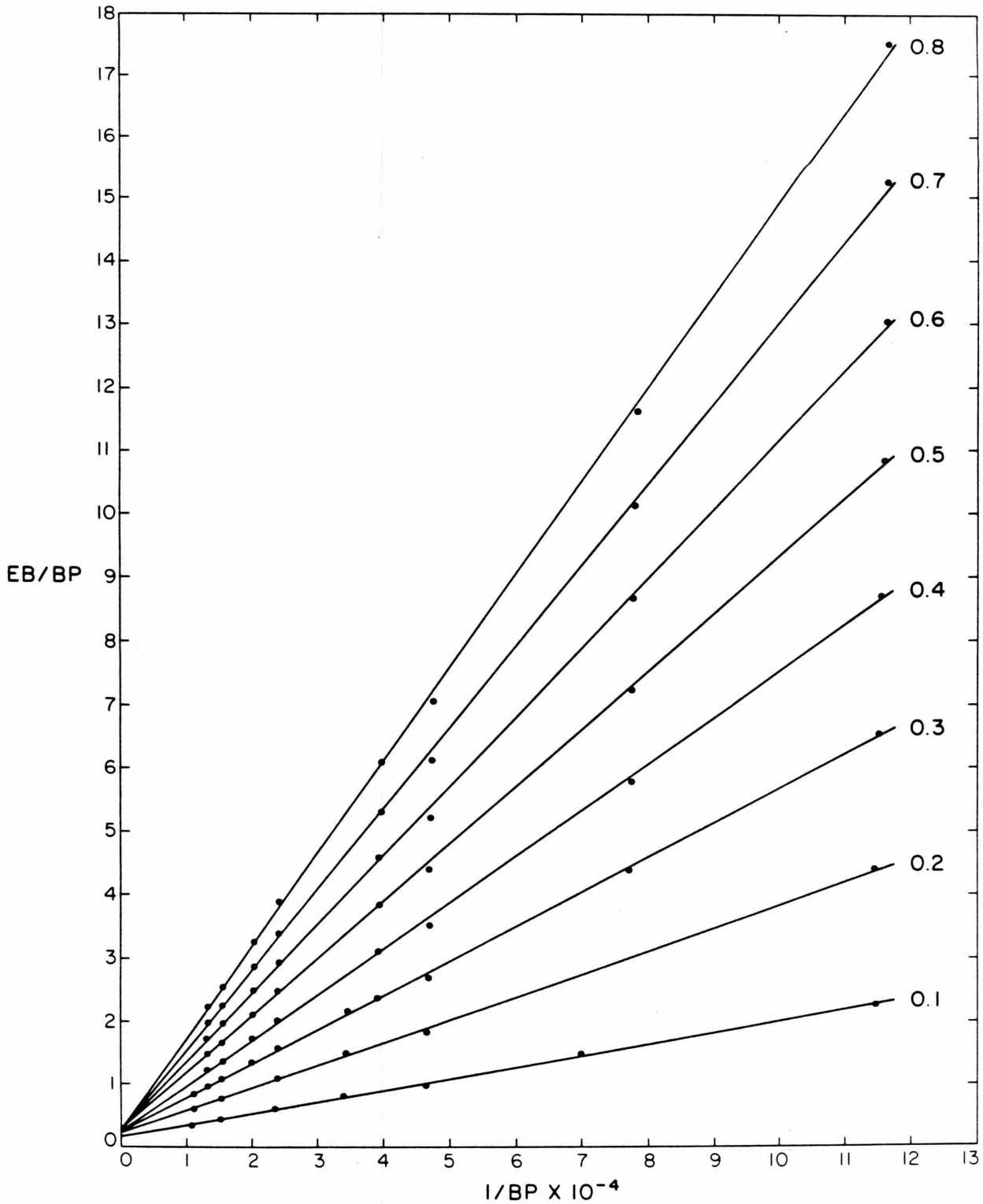
The binding of the intercalating dye ethidium bromide to polyd(G-C) has been investigated by absorption spectroscopy at 12 different polymer concentrations. When the dependence of L_T/M on M at constant X is analyzed by eq. (4) (see Fig. 3), no systematic curvature of the data is detected over the range of macromolecule concentration investigated. Approximate values of ν and L_F can thus be determined from the corresponding intercept and slopes yielding the Scatchard plot shown in Fig. 4 (see later discussion for the error limits of these measurements). As described elsewhere,[8] the three limbs of this Scatchard plot represents the progressive induction of a cooperative B → H → A conformational change in polyd(G-C) by ethidium binding.

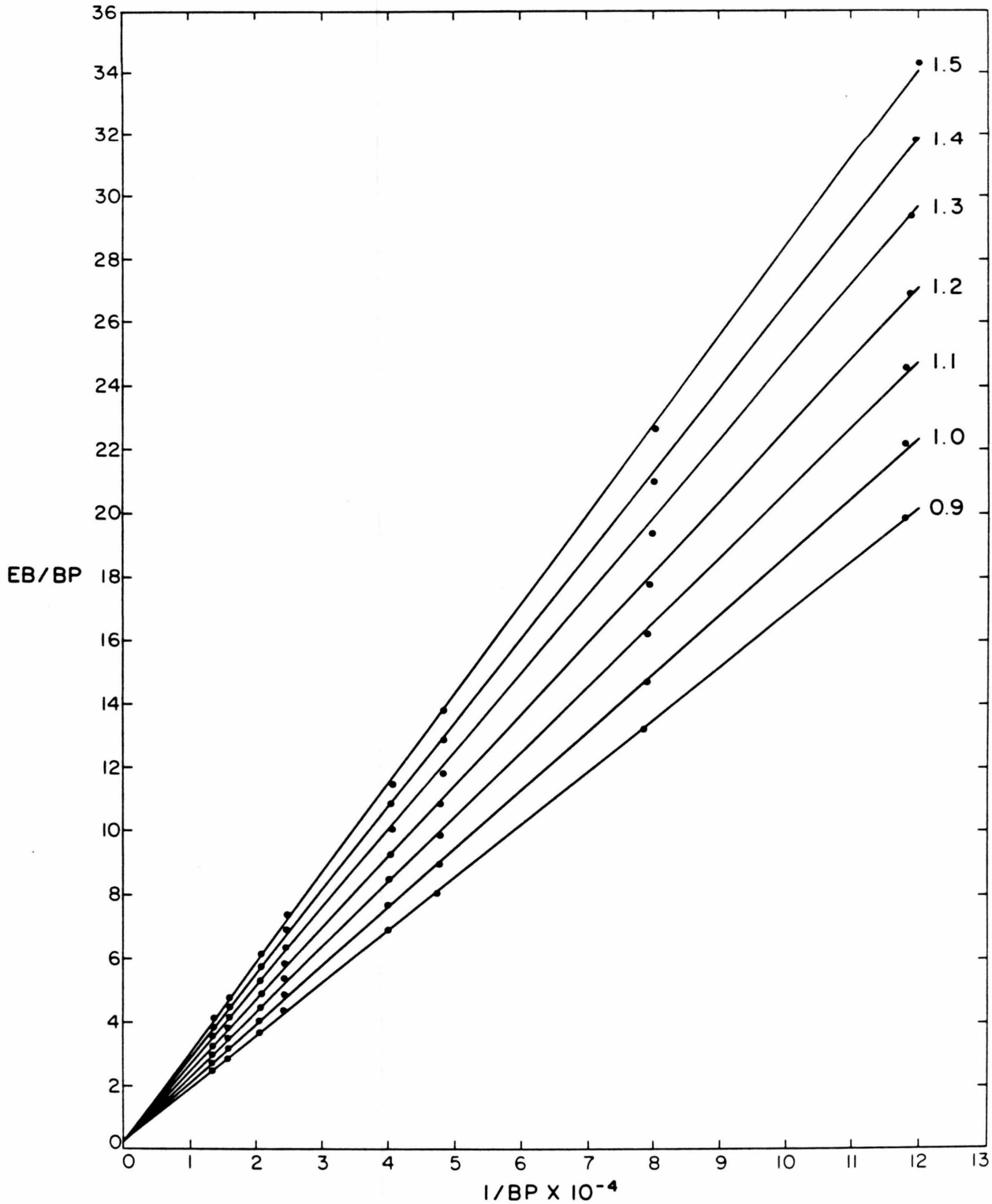
Figure 3

Binding of EB to polyd(C-G) at 1 M^+ as monitored by absorption spectroscopy at 475 nm.[11] The total EB concentration, EB, divided by the concentration of d(C-G) in base pairs is plotted against the inverse of the d(C-G) concentration at constant absorbance. The absorbance of the solution (1 cm path length) is indicated adjacent to each series of measurements.


EB/BP
1/BP X 10⁻⁴
0.07
0.05
0.03
0.01
0.005
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16


6.0
5.0
4.0
3.0
2.0
1.0
0
EB/BP
0.20
0.18
0.14
0.10
0.08
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
1/BP X 10⁻⁴


EB/BP
1/BP X 10⁻⁴
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1


EB/BP
1/BP X 10⁻⁴
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
0
1
2
3
4
5
6
7
8
9
10
11
12
13

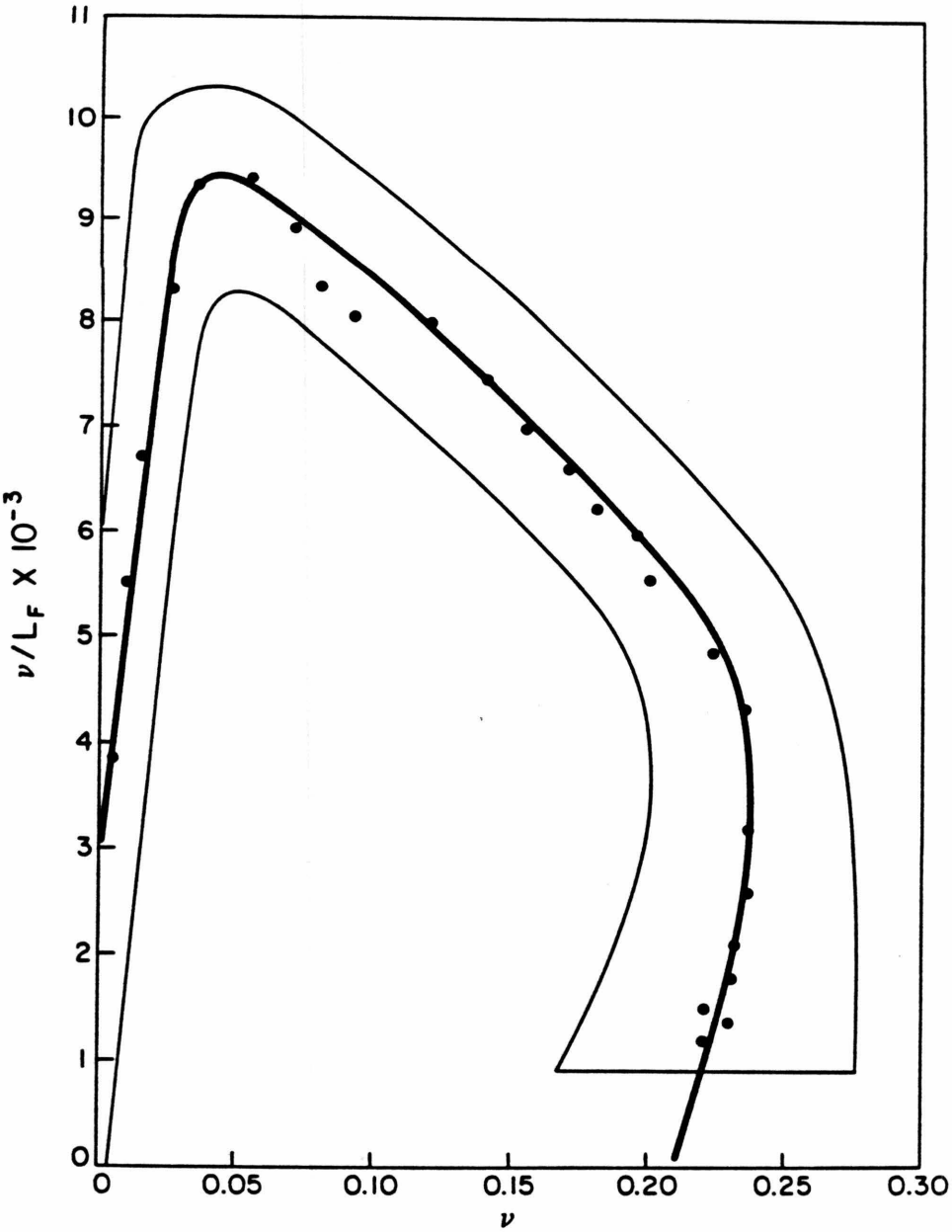
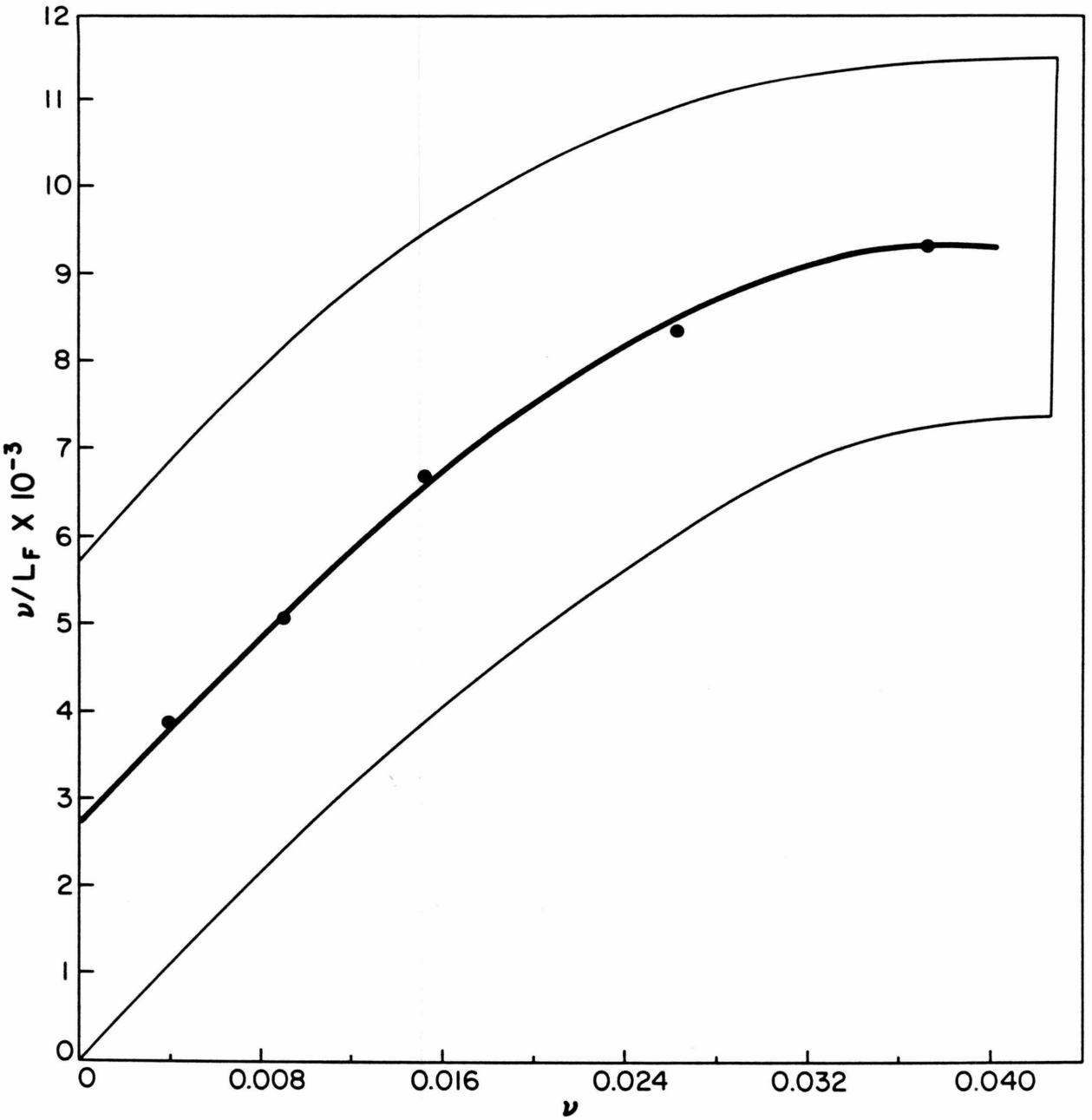


Figure 4
(a) Scatchard plot of the binding of EB to poyld(C-G) at 1 M^+. The values of ν and L_F are obtained from linear regression analysis of the data in Fig. 3 by eq. (4). That region of the Scatchard plot defined by $\pm$1 standard deviation error limit in the value of ν and ν/L_F is also outlined. (b) Same measurements at low binding density. Under these conditions, ε_B = 2250[9] and ε_F = 6000 M^{-1} cm^{-1}.[20]


12
11
10
9
8
7
6
5
4
3
2
1
0
$\nu/L_F \times 10^{-3}$
0
0.008
0.016
0.024
0.032
0.040
ν

Previous measurements[9] by two different techniques have shown that the binding stoichiometry, n, (inverse of the ν axis intercept of a Scatchard plot) of the second species is two base pairs. However, the Scatchard plot obtained at constant X exhibits a stoichiometry of 2.5 base pairs. This discrepancy indicates that the values of ν deduced from eq. (4) at constant X underestimate the true values. As previously discussed, this behavior is expected when the concentration of macromolecule investigated is not significantly less than 1/K. Although the values of ν and L_F extrapolated from these measurements are approximate, the qualitative shape (whether a slope is positive or negative) or the Scatchard plot is unaffected.

To correct Scatchard plots obtained at constant X, one or more physical properties of ligand must be known. For example, if binding is evaluated at a concentration of macromolecule such that MK < 1, the value of L_F obtained at constant X will be close to the true value of L_F and the value of ν will be close to the quantity: $(q_F-q_B/q_F)\nu$ (e.g. equation 12 and Table I). Thus if q_B and q_F are known, one can correct ν and hence the Scatchard plot obtained at constant X. The quantites q_B and q_F may be estimated directly from the experimentally measured dependence of X on L_T and M by conventional techniques[13,23]

or they may be obtained from other measurements.

If the quantity (MK) is not significantly less than one, the value of L_F obtained at constant X must also be corrected. To correct L_F, recall that the values of ν and L_F obtained at constant X obey an equation analogous to equation (4):

$$(20) \qquad (L_T/M)' = \nu' + L_F'/M$$

Similarly, for the true dependency:

$$(21) \qquad L_T/M = \nu + L_F/M$$

At some macromolecule concentration, M*, L_T/M will equal L_T/M and one can write:

$$(22) \qquad \nu' + L_F'/M^* = \nu + L_F/M^*$$

which upon rearrangement yields:

$$(23) \qquad M^*(\nu' - \nu + L_F'M^*) = L_F$$

Calculation of L_F from equation (21) requires knowledge of two unknowns: ν and M*. If one or more ligand binding parameters are known, such as K(ν/L_F axis intercept), n(ν axis intercept) or some other portion of the Scatchard plot, a relationship between ν and ν' can be readily established thus allowing estimation of ν.

Other measurements may be used to establish one or more ligand binding parameters or they may be estimated directly from the experimentally observed dependence of X on L_T and M.

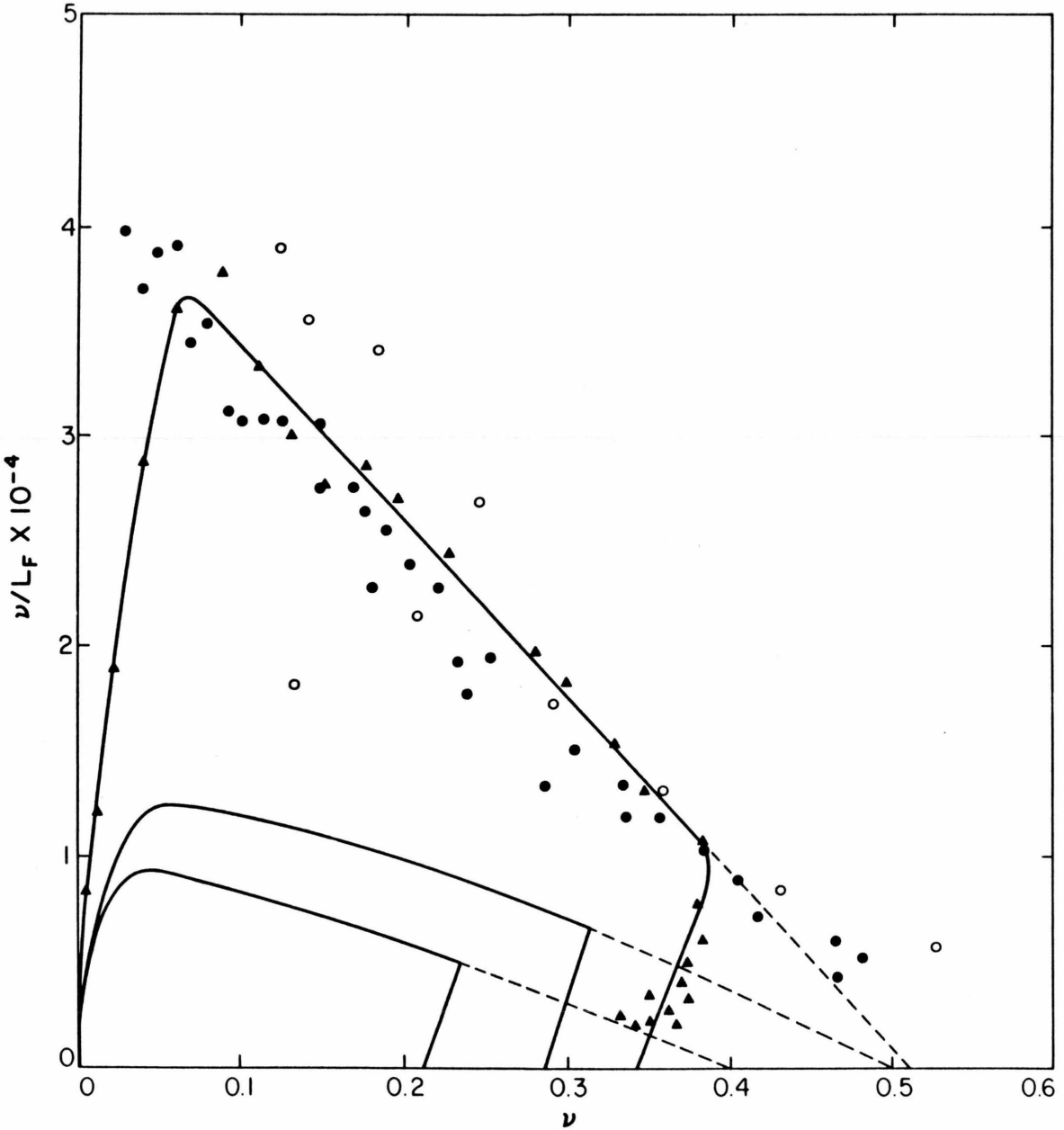
In the later case, ligand is assumed to form one type of bound complex. Using known values of q_B and q_F, a Scatchard plot is constructed from the experimentally measured dependence of X on L_T and M by conventional procedures. Comparison of this Scatchard plot to the Scatchard plot obtained at constant X allows one to estimate a value of ν for each ν'. Having estimated a value of ν, the value of L_F can be then be estimated from equation (21) by adjusting M* until the corrected values of ν and L_F agree with one or more ligand binding parameters. Since M* may vary with X, values of L_F calculated from equation (21) in this manner are approximate. Use of these various correction procedures to correct the EB-polyd(C-G) Scatchard plot is demonstrated in Figure 6 (see also Fig. 2).

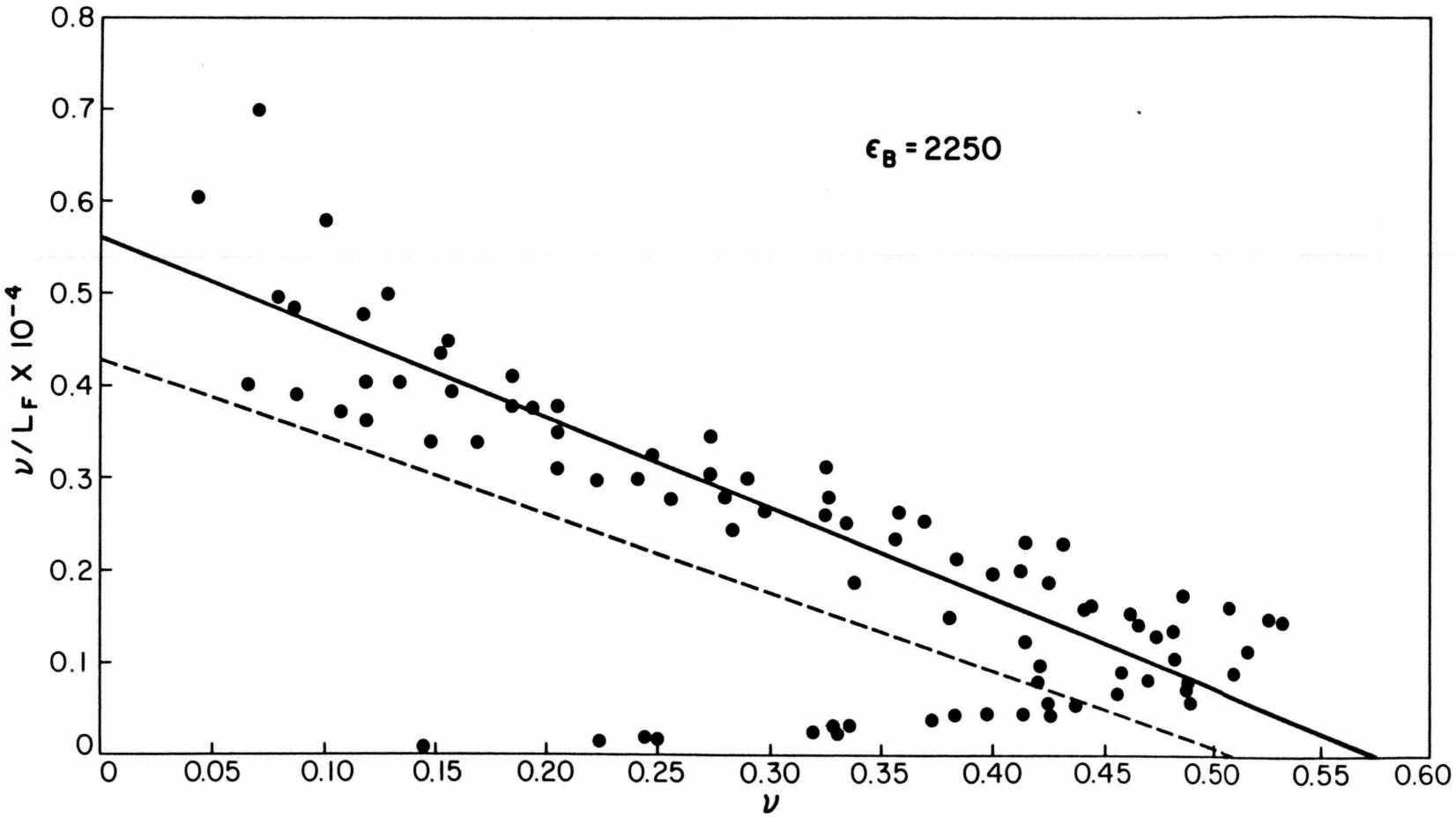
At low and high binding densities the EB-polyd(C-G) Scatchard plot obtained at constant X differs from previous measurements.[9] The failure of classical indirect techniques (compare Fig. 6b to 4) or direct (Fig. 6a) techniques to accurately describe the binding at these ligand

Figure 6

a) Comparison of ethidium bromide - polyd(G-C) Scatchard plot deduced from indirect optical[16] (●) and more direct sedimentation (o) measurements of Pohl, et al.[9] with the values obtained by correcting the Scatchard plot in Fig. 4a (▲). The first (smallest) curve is the uncorrected data of Fig. 4a ($\nu = \nu'$, $M^* = 0$). The second curve is obtained from the first by letting $1/M^* = 1 \times 10^4$ and $\nu = 1.33\ \nu'$. The third curve is obtained in an analogous manner with $1/M^* = 12 \times 10^4$ and $\nu = 1.63\ \nu'$.

b) Scatchard plot of the binding of EB to polyd(C-G) at 1 M^+. Measurements are from Fig. 3. EB is assumed to form one bound complex with extinction coefficient equal to 2250 M^{-1} cm^{-1} at 475 nm.[16] Corrections for unbound EB dimerization have been applied. The same data, when analyzed by the constant X technique, yield the Scatchard plot shown in Fig. 4. The optical measurements of Jovin from (a) are represented by the dashed line.


5
4
3
2
1
0
$\nu/L_F \times 10^{-4}$
0
0.1
0.2
0.3
0.4
0.5
0.6
ν


$\epsilon_B = 2250$
$\nu / L_F \times 10^{-4}$
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
0.60
ν

densities stems from two sources. First, these measurements assume that all bound species have identical physical properties. Our measurements indicate, however, that at least three different bound species are formed. Secondly, as discussed in the introduction, the indirect and direct techniques utilized in these measurements are intrinsically inaccurate at low and high binding densities. The constant X technique avoids both of these errors because no assumptions regarding ligand physical properties are made and the values of ν and L_F are extrapolated rather than measured.

Error limits associated with each value of ν and L_F can be determined from eq. (4) by linear regression analysis (Figs. 4,8,12,14). With increasing X, the error in ν and L_F increase because larger values of L_T must be evaluated. However, as one proceeds down the last limb of the polyd(C-G) (Fig. 4) or polydCdG (Fig. 5) Scatchard plot, the error in ν and L_F abruptly increase with increasing X. This behavior may result from the presence of multiple forms of bound and/or free ligand since their concentrations will not be uniquely defined by eq. (2) at constant X. The formation of dimers of unbound EB at these values of X (Fig. 6, Chapter I) as well as the ability of polyd(C-G) and polydCdG to form multiple bound complexes with EB (Chapter III), support this hypothesis.

In addition to the copolymer polyd(C-G) we have also measured by absorption spectroscopy the binding of EB to polydCdG at 1 M^+ (Fig. 7). When these data are analyzed by eq. (4) at constant X, the Scatchard plot shown in Fig. 8 is obtained. Alternatively, if one corrects for free drug dimerization, and assumes the formation of only one bound complex, the same data yield the Scatchard plot in Fig. 9. As previously discussed, this second Scatchard plot can be used together with eq. (20) to correct the Scatchard plot obtained at constant X. The Scatchard plot corrected in this manner is shown in Fig. 10 and discussed at more length in Chapter III.

The binding of polyintercalator BMSp to polydAdT has also been measured by absorption spectroscopy at 1 M^+ (Fig. 11). When these data is analyzed by eq. (4) at constant X, the Scatchard plot shown in Fig. 12 is obtained. Repeating these measurements over a more dilute range of macromolecule concentration (Fig. 13) yields the Scatchard plots shown in Figs. 14 and 15. Each of these Scatchard plots is discussed in more detail in Chapter III.

Dependence of the Constant X Technique on Concentration Proportionality Factors

As noted previously, at constant X, the maximum value of the intercept of equation (4) equals $(q_F-q_B/q_F)\nu$. Thus if binding is evaluated at a wavelength where q_F equals q_B (isobestic point), no information regarding the ligand-macromolecule interaction will be obtained. When two bound ligand species are formed, the maximum value of this intercept equals: $(q_F-q_B^1/q_F)\nu_1 + (q_F-q_B^2/q_F)\nu_2$. In this case, if binding is monitored at a wavelength where q_B^1 and q_B^2 do not equal q_F, the Scatchard plot obtained at constant X will describe the binding of both ligands to macromolecule. Alternatively, if binding is also monitored at a wavelength where $q_B^2 = q_F$, only the binding of bound species one will be detected. Therefore, when ligand forms more than one type of bound complex, it may be possible by the constant X technique to obtain separate Scatchard plots for each individual bound species, an otherwise difficult task.

In the EB-polyd(C-G) and EB-polydCdG Scatchard plots obtained at constant X, the value of ν increases and then decreases. Since EB binding is postulated to change the conformation of both DNA's, ν may decrease because the

corresponding extinction coefficient, q_B, changes with macromolecule conformation (see equation 12). However, at the wavelengths used to monitor binding, q_B is insensitive to changes in macromolecule concentration. For example, at the wavelength used to monitor BMSp binding to dAdT (485 nm), the value of q_B for the B and A conformations are identical (Fig. 15, Chapter 1). Similarly, at the wavelength utilized to monitor EB binding (475 nm), q_B does not change significantly with nucleic acid conformation:

Nucleic Acid	Conformation	$q_B(M^{-1}\ cm^{-1})$
rArU	A	2400[22]
dAdT	B	2100[22]
d(A-T)	H	2250[22]
d(C-G)	H	2250[9]
calf thymus	H	2250

Thus it is unlikely that ligand induced changes in macromolecule conformations cause q_B and hence ν to decrease.

If ligand is assumed to form one bound and one free complex, the experimentally observed dependence of X on L_T and M yields the Scatchard plots in Figures 6 and 9. Like the corresponding Scatchard plots obtained at constant X, the value of ν increases and then decreases. Since ligand is assumed to form only one bound complex, the de-

crease in ν observed in these and corresponding constant X. Scatchard plots cannot result from free drug dimerization (Fig. 6, Chapter 1) or formation of multiple bound forms of ligand. For example, when ligand is assumed to form one bound and one free species, the concentration of bound ligand, L_B, is given by:

$$L_B = \frac{L_T\varepsilon_F - X}{\varepsilon_F - \varepsilon_B} \qquad (24)$$

where

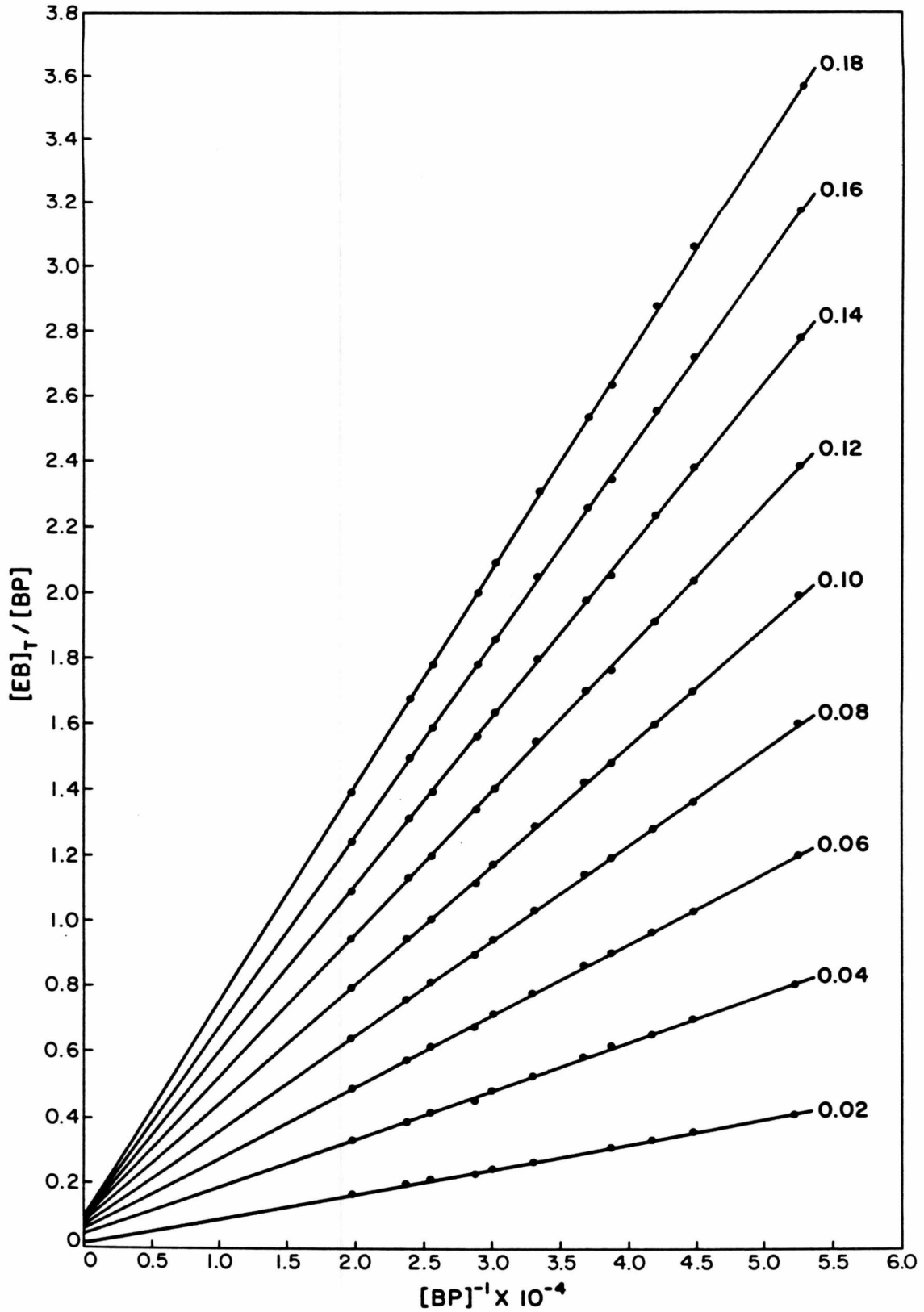
$$X = L_F\varepsilon_F + L_B\varepsilon_B$$

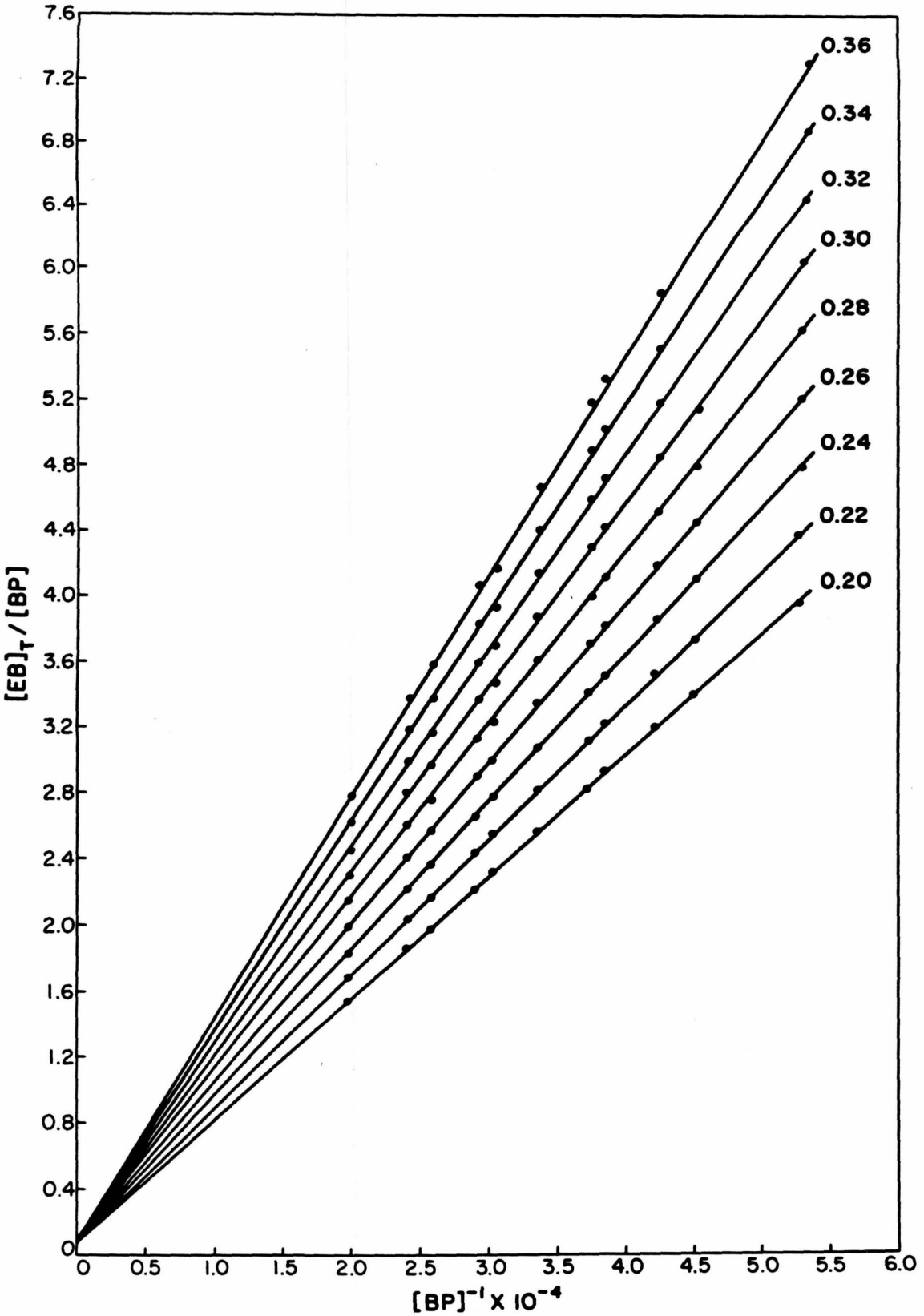
For a given concentration of ligand, L_T, free ligand dimerization or multiple bound species will decrease the observed value of X. As a result, the corresponding values of L_B and L_F will be underestimated and overestimated respectively, causing the Scatchard plot to exhibit anamolously high values of ν and ν/L_F. Since low rather than high values of ν and ν/L_F are observed at high degrees of binding, the unusual bending back of the EB-polyd(C-G) and EB-polydCdG Scatchard plots cannot be due to multiple forms of bound or free ligand.

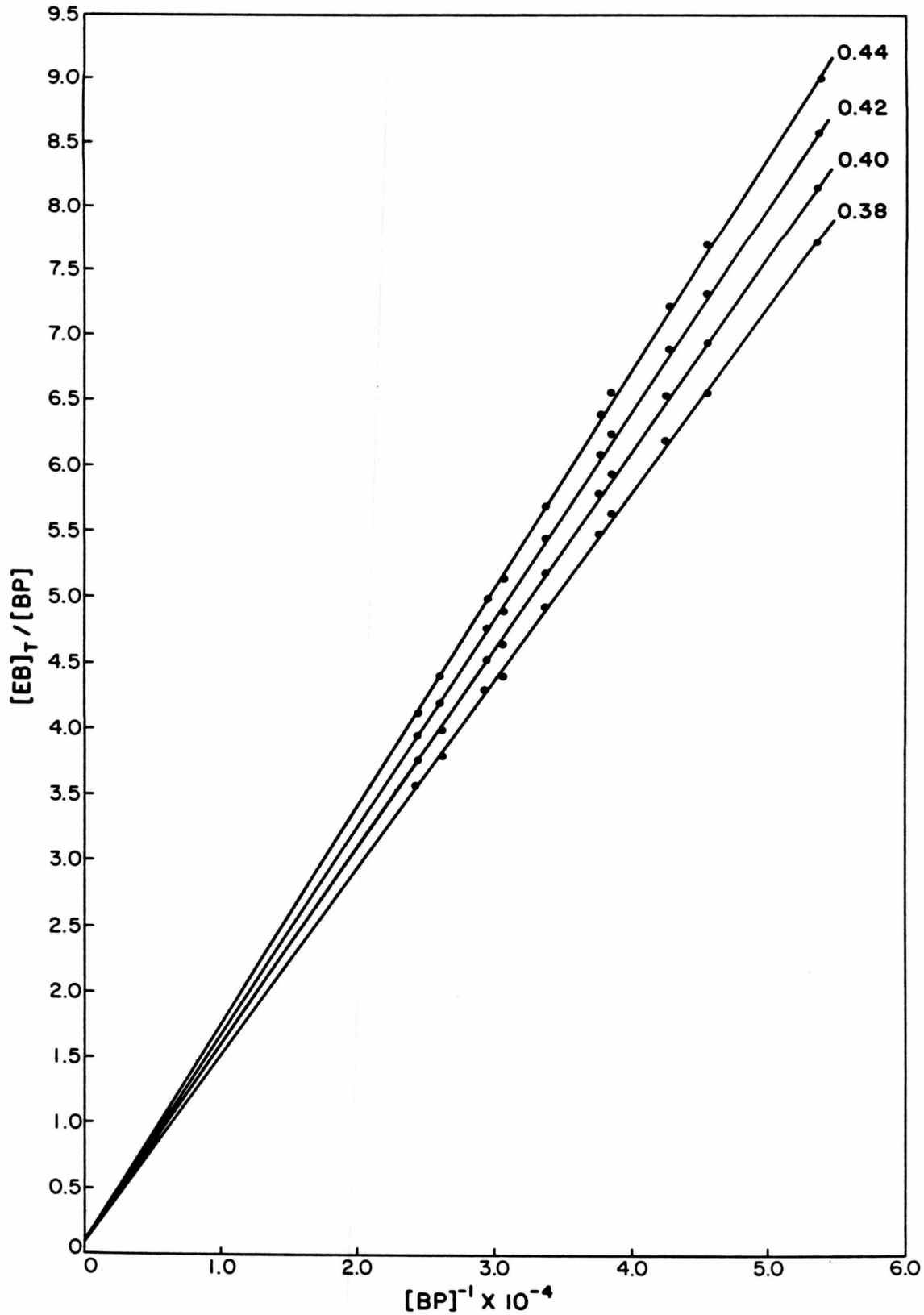
According to Hopfield, equilibrium chemistry of dilute systems does not allow Scatchard plots to have two values of L_F for a single ν, regardless of the complicated interactions which may take place on binding. Thus some artifact (other than free ligand dimerization or changing extinction coefficient of bound ligand which we have ruled out) may be responsible for the apparent retrograde third limb of the ethidium bromide, polydG·dC and polyd(G-C)·d(G-C) Scatchard plots.

Figure 7

Binding of EB to polydCdG at 1 M^+ as monitored by absorption spectroscopy at 475 nm.[12] The total EB concentration, EB, divided by the concentration of dCdG in base pairs is plotted against the inverse of the dCdG concentration at constant absorbance. The absorbance (0.5 cm path length) is indicated adjacent to each series of measurements.


[EB]T / [BP]
[BP]-1 X 10-4
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
3.6
3.8
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0


7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
[EB]T / [BP]
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
[BP]-1 X 10-4


9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
0.44
0.42
0.40
0.38
[EB]T / [BP]
0
1.0
2.0
3.0
4.0
5.0
6.0
[BP]-1 X 10-4

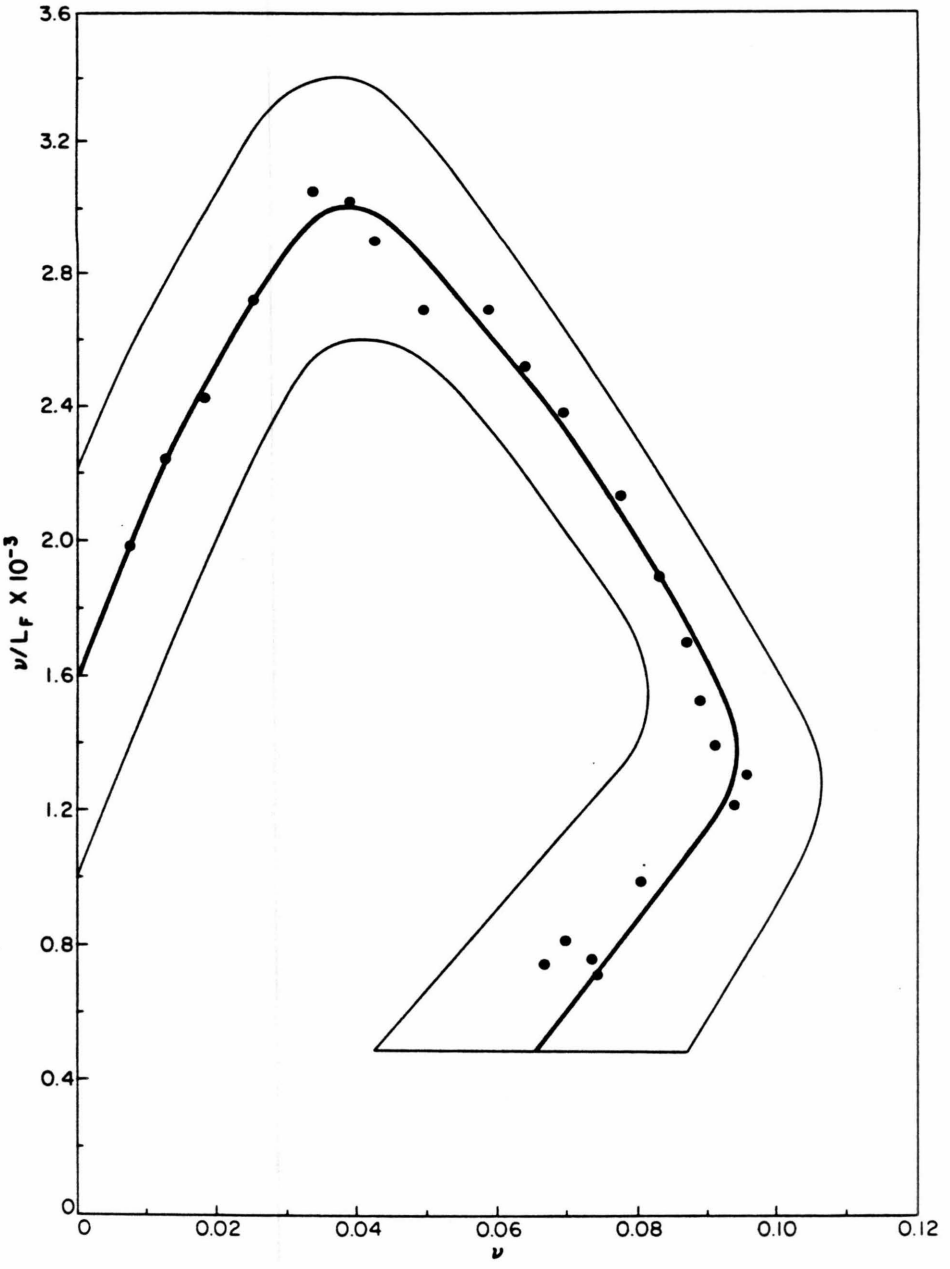


Figure 8
Scatchard plot of the binding of EB to polydCdG at 1 M^+. The values of ν and L_F are obtained from linear regression analysis of the data in Fig. 7 by eq. (4). That region of the Scatchard plot defined by a $\pm$ 1 standard deviation error limit in the value of $\bar{\nu}$ and ν/L_F is also outlined. Under these conditions $\varepsilon_B = 3100$[13] and $\varepsilon_F = 6000\ M^{-1}\ cm^{-1}$.[20]

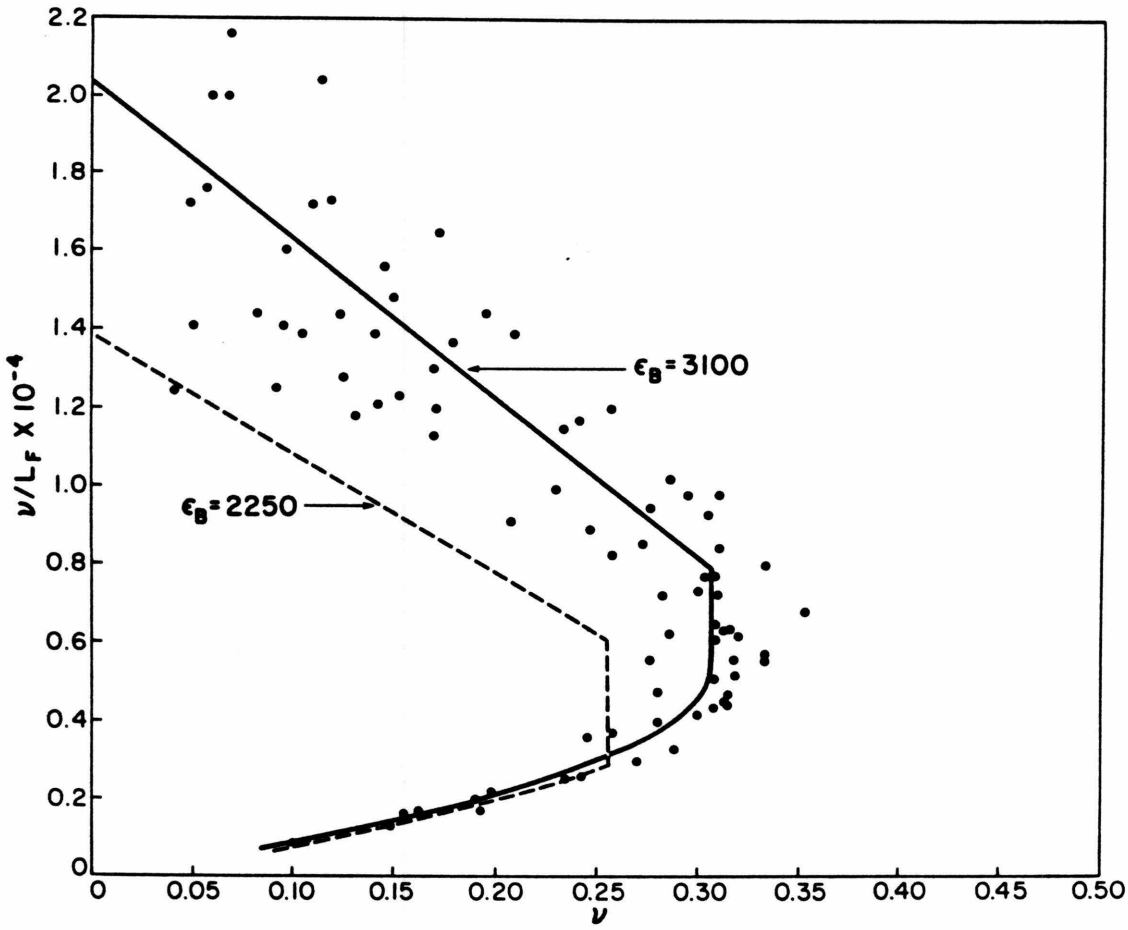


Figure 9
Scatchard plot of the binding of EB to polydCdG at 1 M^+ determined by absorption spectroscopy. Experimental data are taken from Fig. 7. EB is assumed to form one bound complex with extinction coefficient equal to 2250 M^{-1} cm^{-1} or 3100 M^{-1} cm^{-1} [13] at 475 nm. Corrections for free EB dimerization have been applied. The same data, when analyzed by the constant X technique, yield the Scatchard plot shown in Fig. 8.

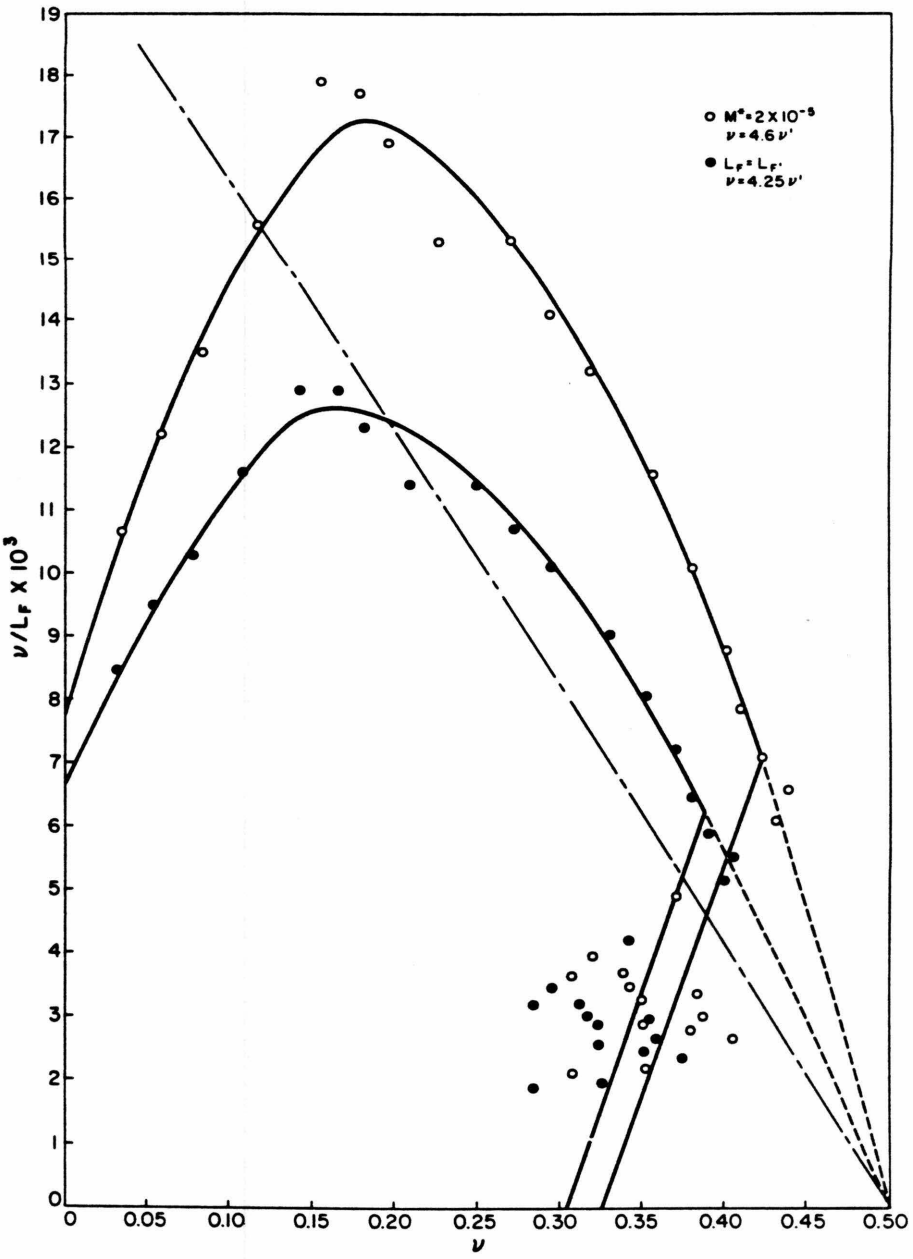
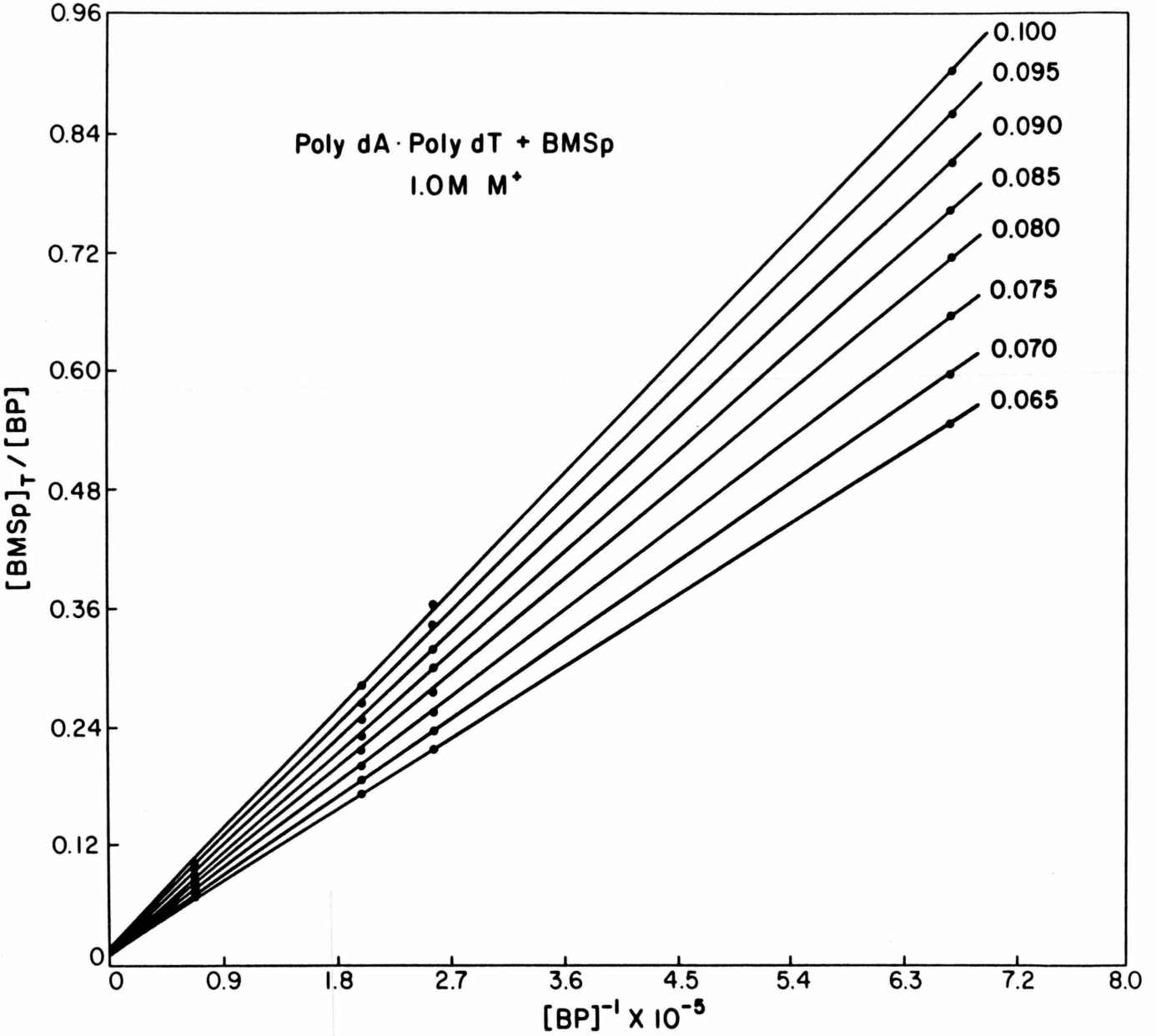
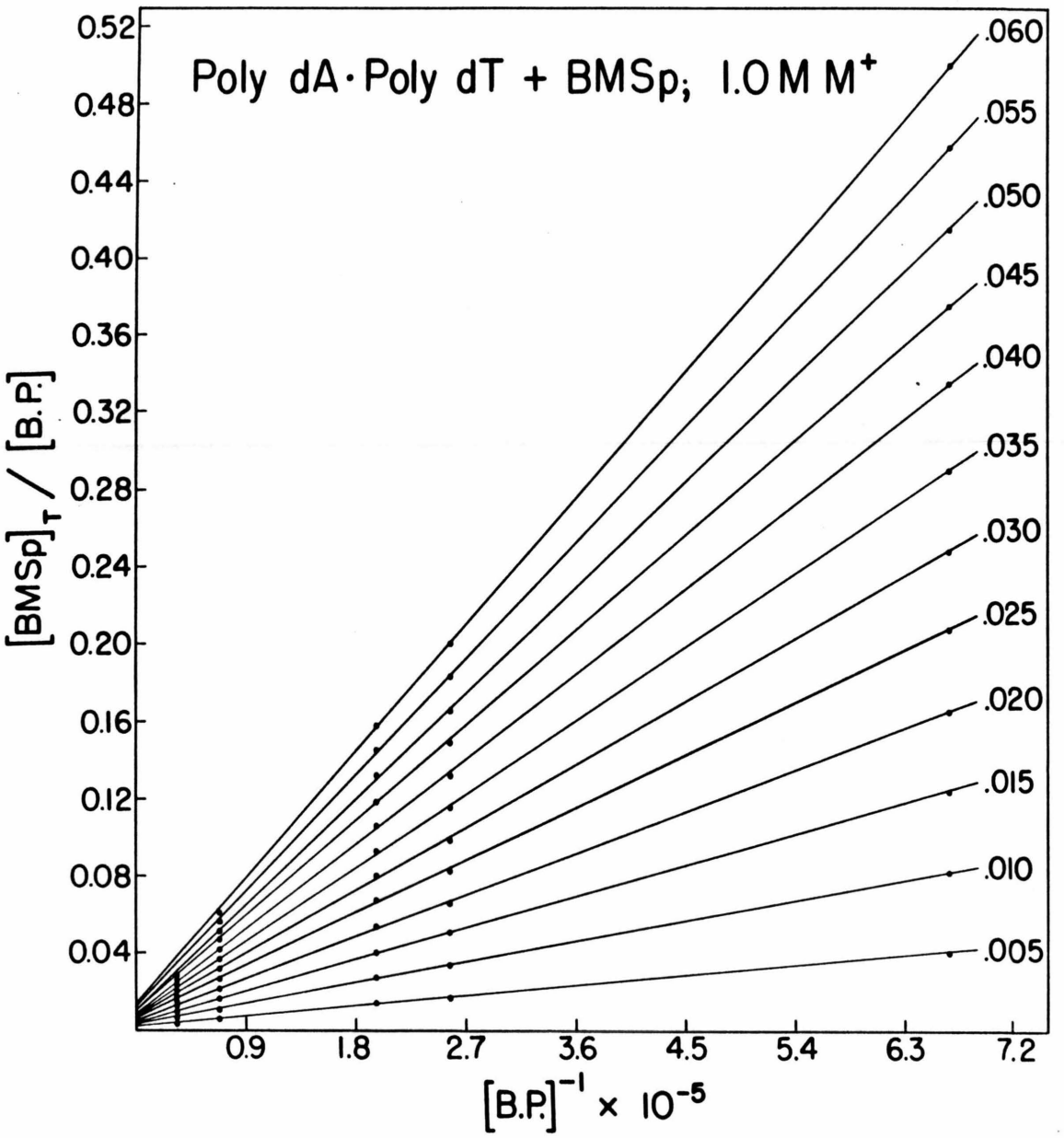


Figure 10
Scatchard plots of the binding of EB to polydCdG at 1 M^+ as corrected by eq. (20) with (○) $M^+ = 2 \times 10^{-5}$, $\nu = 4.6\ \nu'$ or (●) $L_F = L'_F$, $= 4.25\ \nu'$. The line denoted (—— - ——) represents the Scatchard plot obtained by indirect measurements (Fig. 9) with $\varepsilon_B = 3100$.

Figure 11

Binding of BMSp to polydAdT at 1 M^+ as monitored by absorption spectroscopy at 485 nm.[14] The total BMSp concentration, BMSp, divided by the concentration of dAdT in base pairs is plotted against the inverse of the dAdT concentration at constant absorbance. The absorbance (10 cm path length) is indicated adjacent to each series of measurements.


Poly dA · Poly dT + BMSp
1.0M M+
0.100
0.095
0.090
0.085
0.080
0.075
0.070
0.065
[BMSp]T / [BP]
0.96
0.84
0.72
0.60
0.48
0.36
0.24
0.12
0
0
0.9
1.8
2.7
3.6
4.5
5.4
6.3
7.2
8.0
[BP]-1 X 10-5


Poly dA·Poly dT + BMSp; 1.0 M M+
[BMSp]T / [B.P.]
[B.P.]-1 × 10-5
0.52
0.48
0.44
0.40
0.36
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0.9
1.8
2.7
3.6
4.5
5.4
6.3
7.2
.060
.055
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005

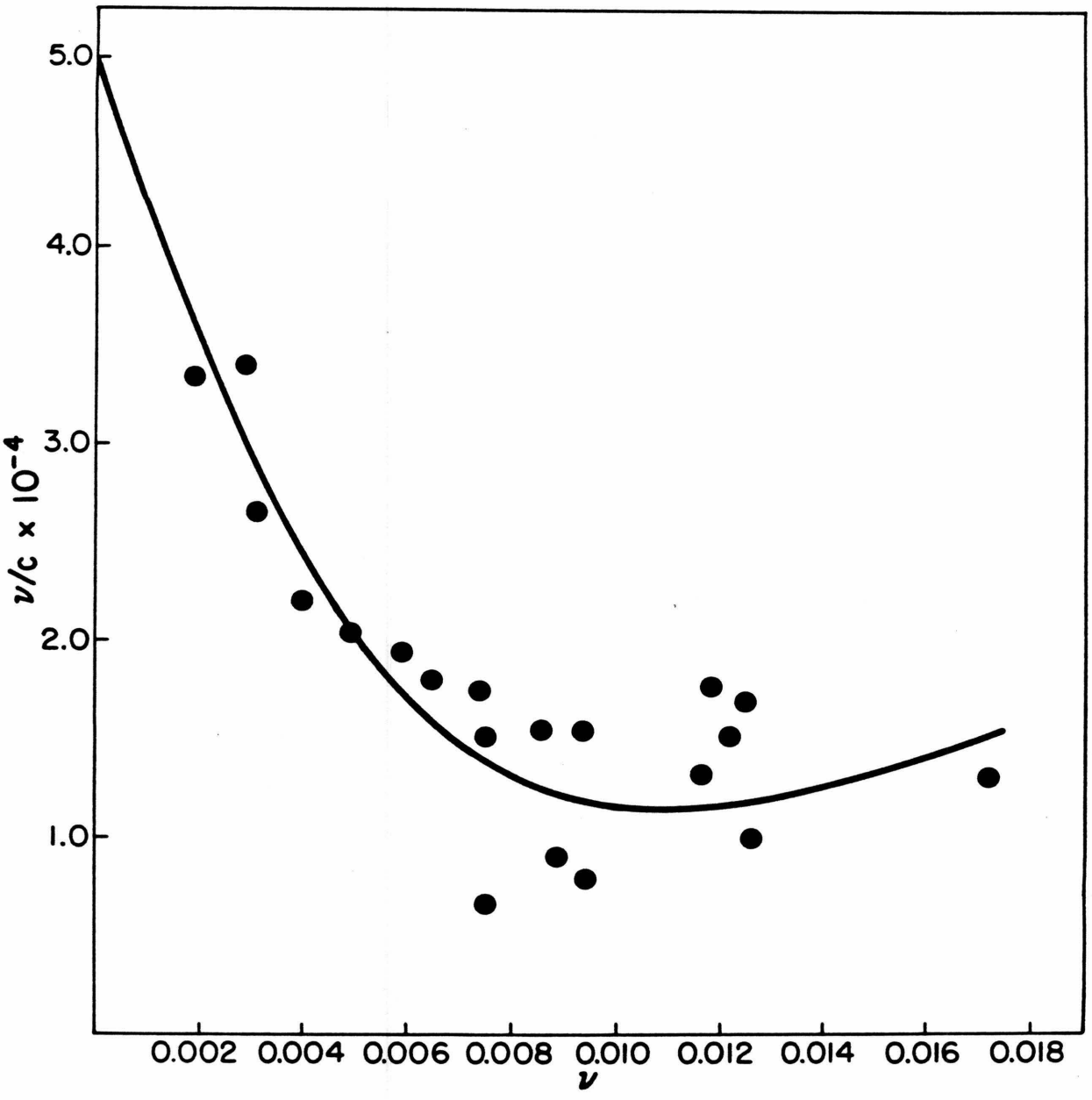


Figure 12
Scatchard plot of the binding of BMSp to polydAdT at 1 M^+ (A). The values of ν and L_F are obtained from linear regression analysis of the data in Fig. 11 by eq. (4). That region of the Scatchard plot defined by a $\pm$ 1 standard deviation error limit in the value of ν and ν/L_F is also outlined. Under these conditions ε_B = 5240 and ε_F = 8000 M^{-1} cm^{-1}.

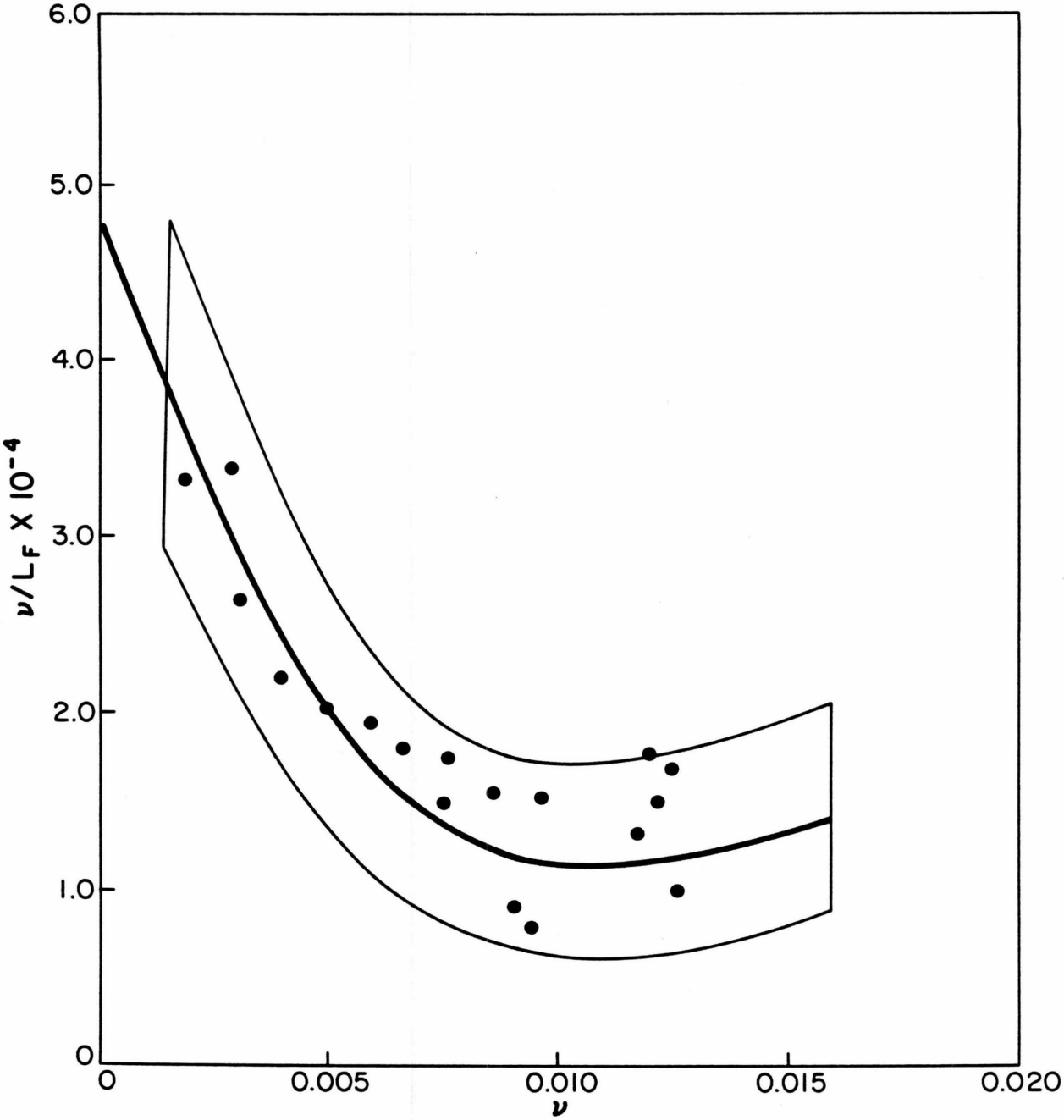
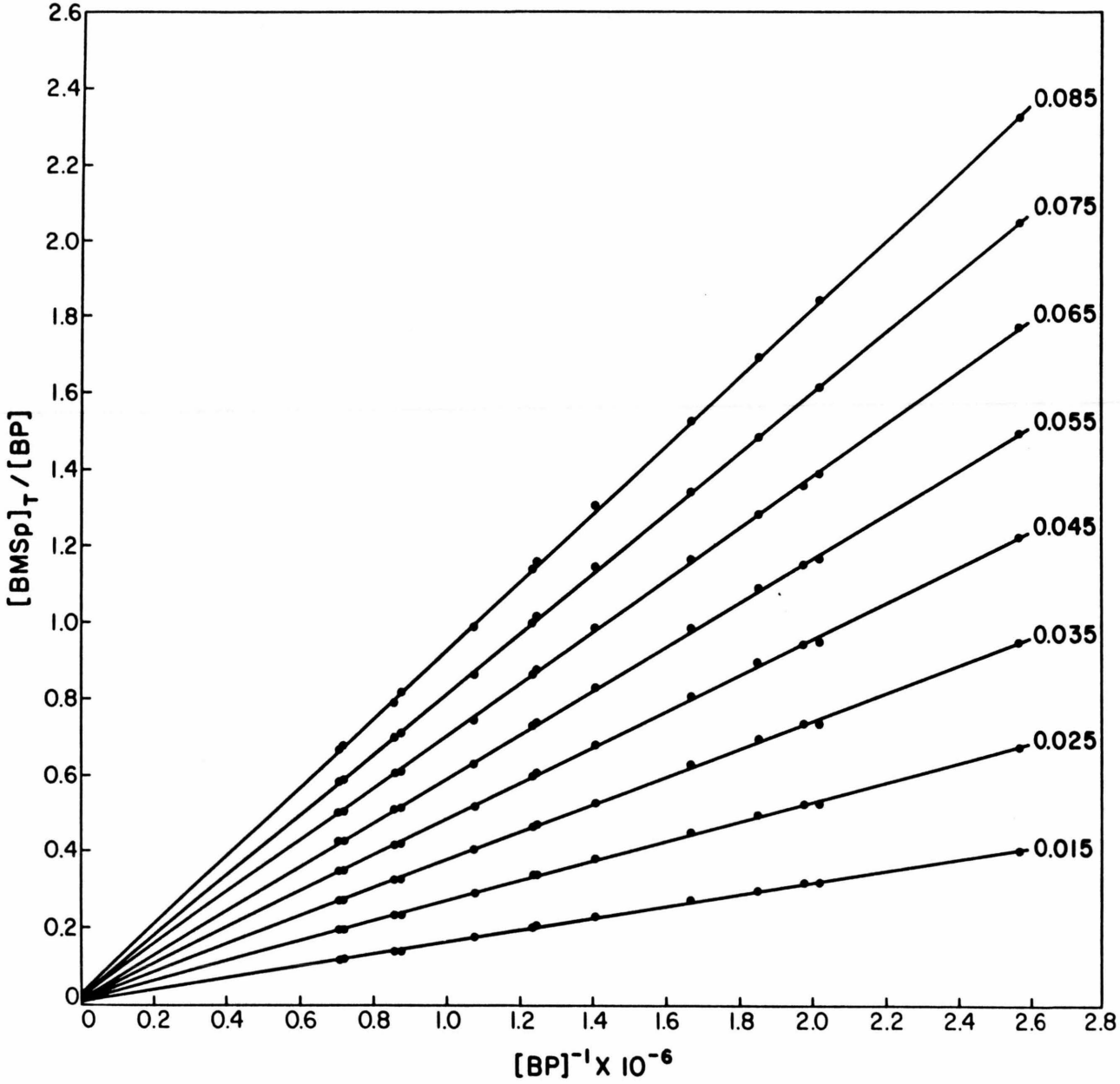
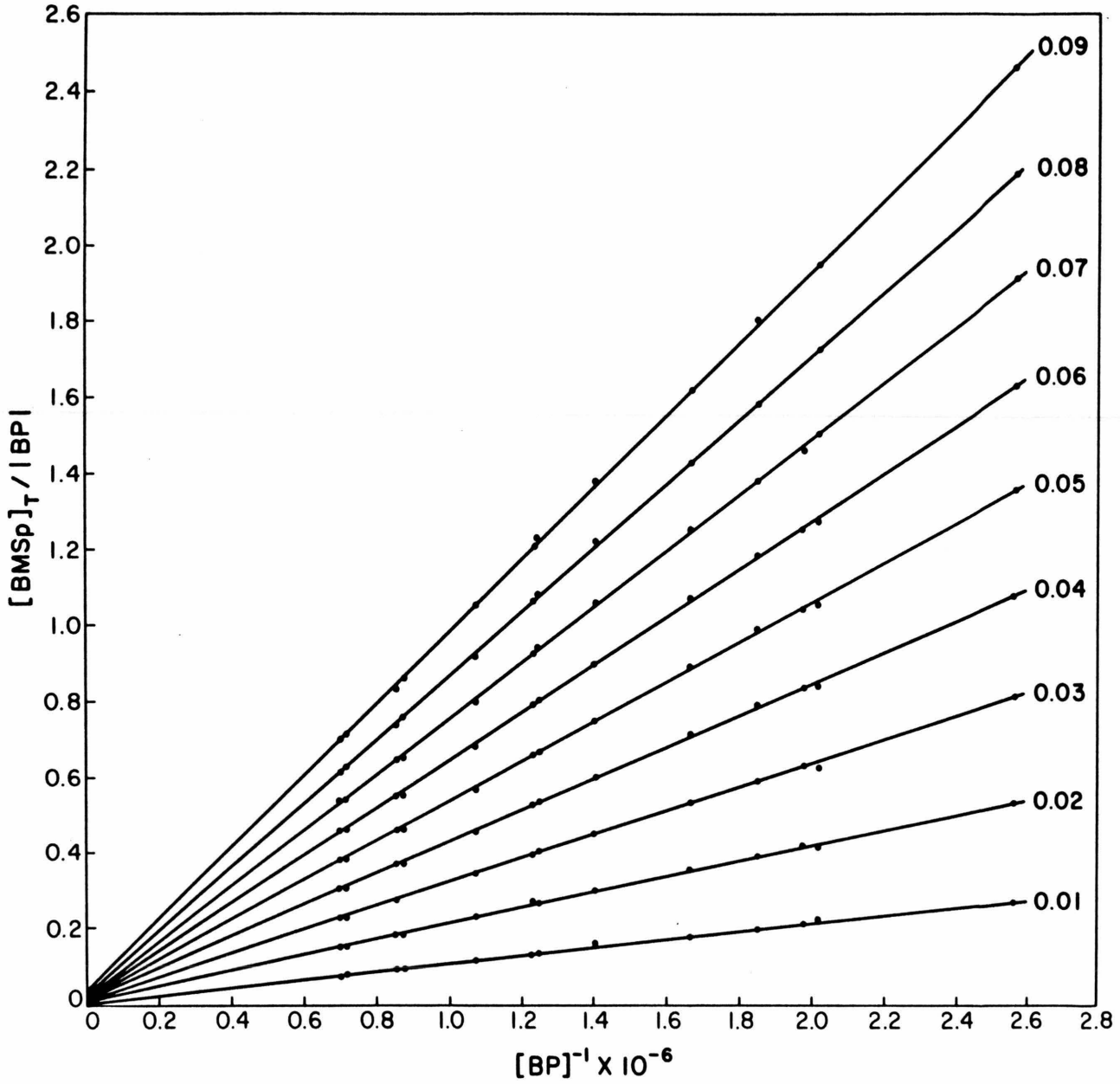
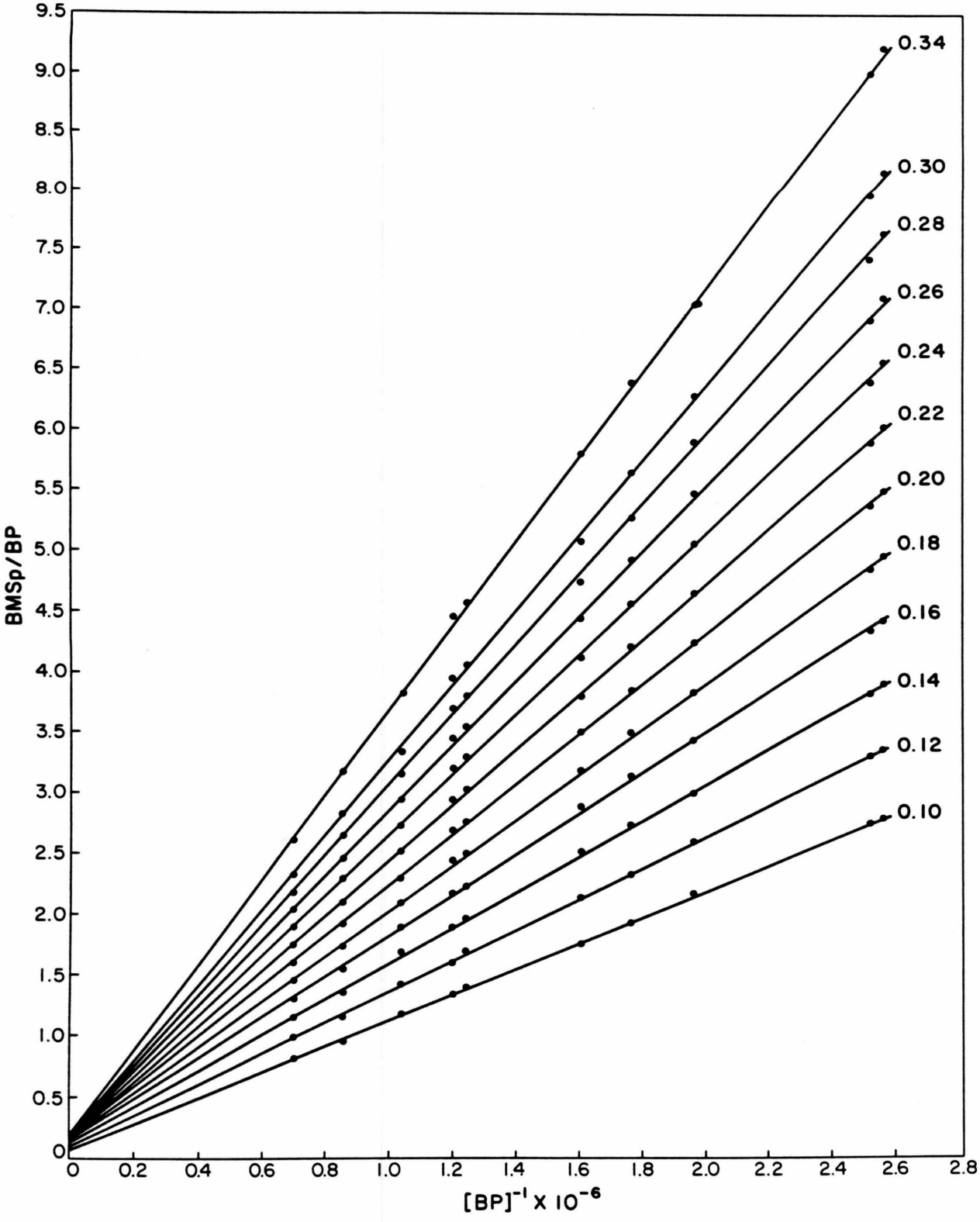

6.0
5.0
4.0
3.0
2.0
1.0
0
$\nu/L_F \times 10^{-4}$
0
0.005
0.010
0.015
0.020
ν

Figure 13

Binding of BMSp to dAdT at 1 M^+ as monitored by absorption spectroscopy at 485 nm.[15] The total BMSp concentration, BMSp, divided by the concentration of dAdT in base pairs is plotted against the inverse of the dAdT concentration at constant absorbance. The absorbance (10 cm path length) is indicated adjacent to each series of measurements.


2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
[BMSp]T / [BP]
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
[BP]-1 X 10-6
0.085
0.075
0.065
0.055
0.045
0.035
0.025
0.015


2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
[BMSp]T / [BP]
[BP]-1 X 10-6


BMSp/BP
[BP]⁻¹ X 10⁻⁶
0.34
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10

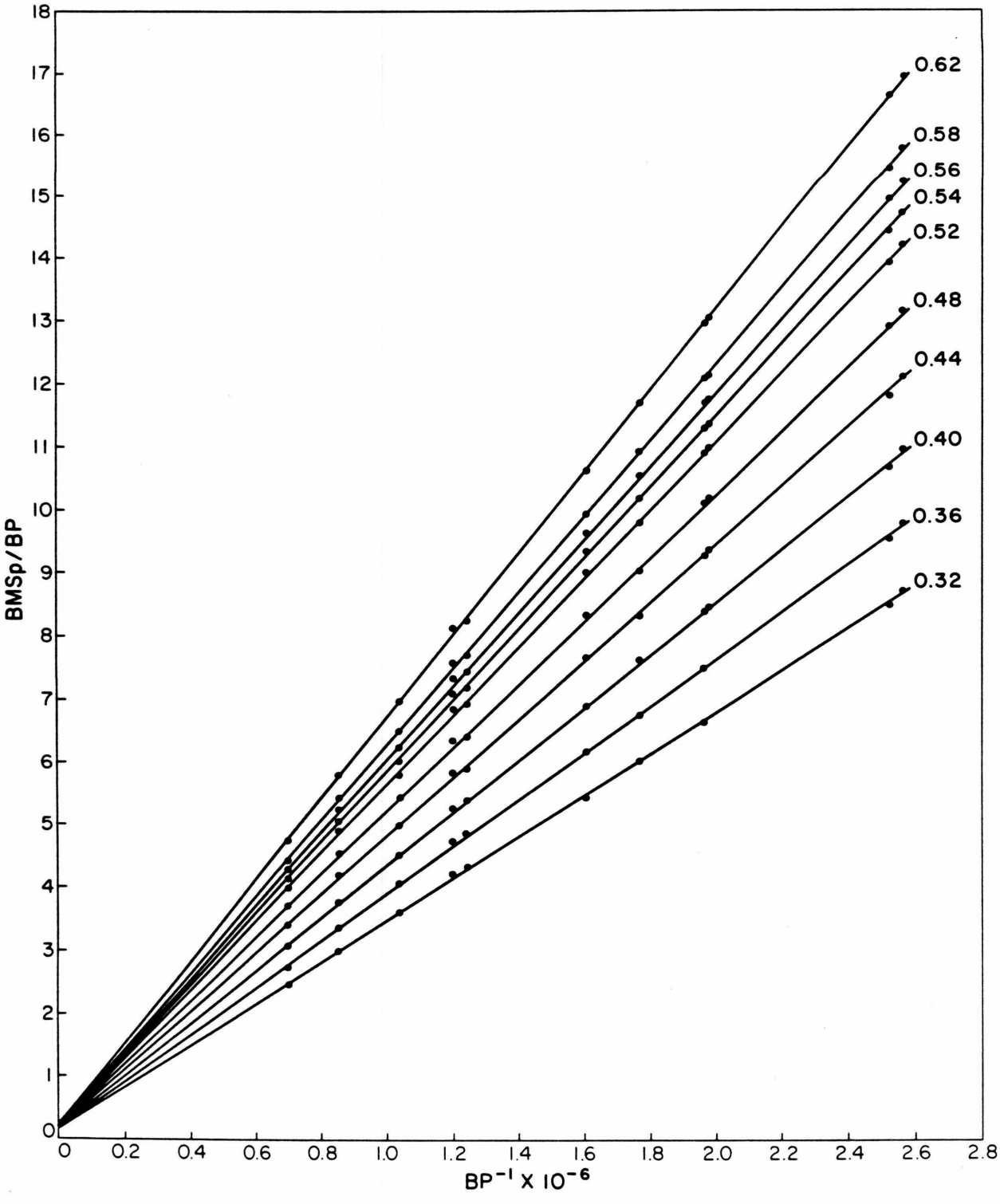
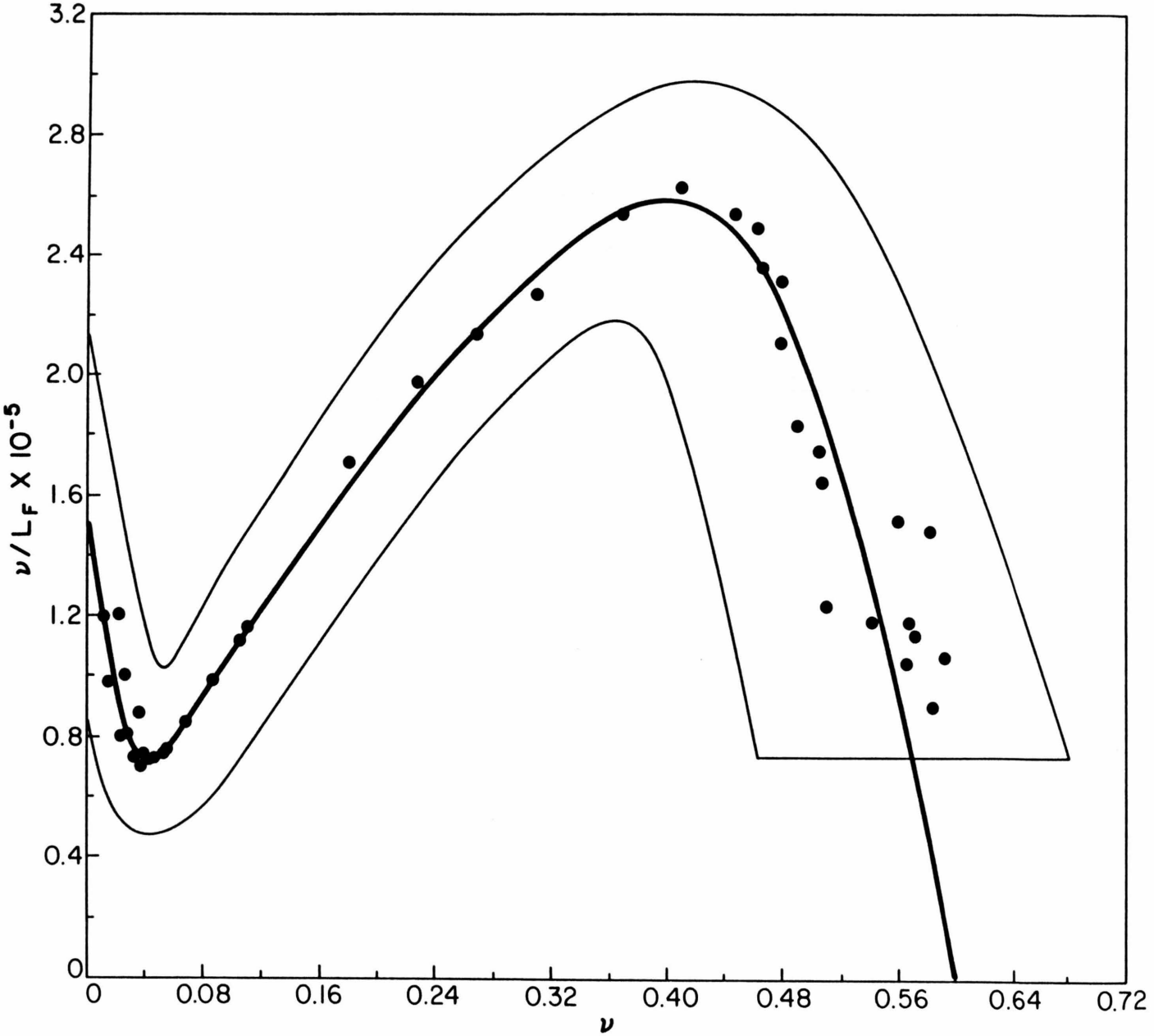
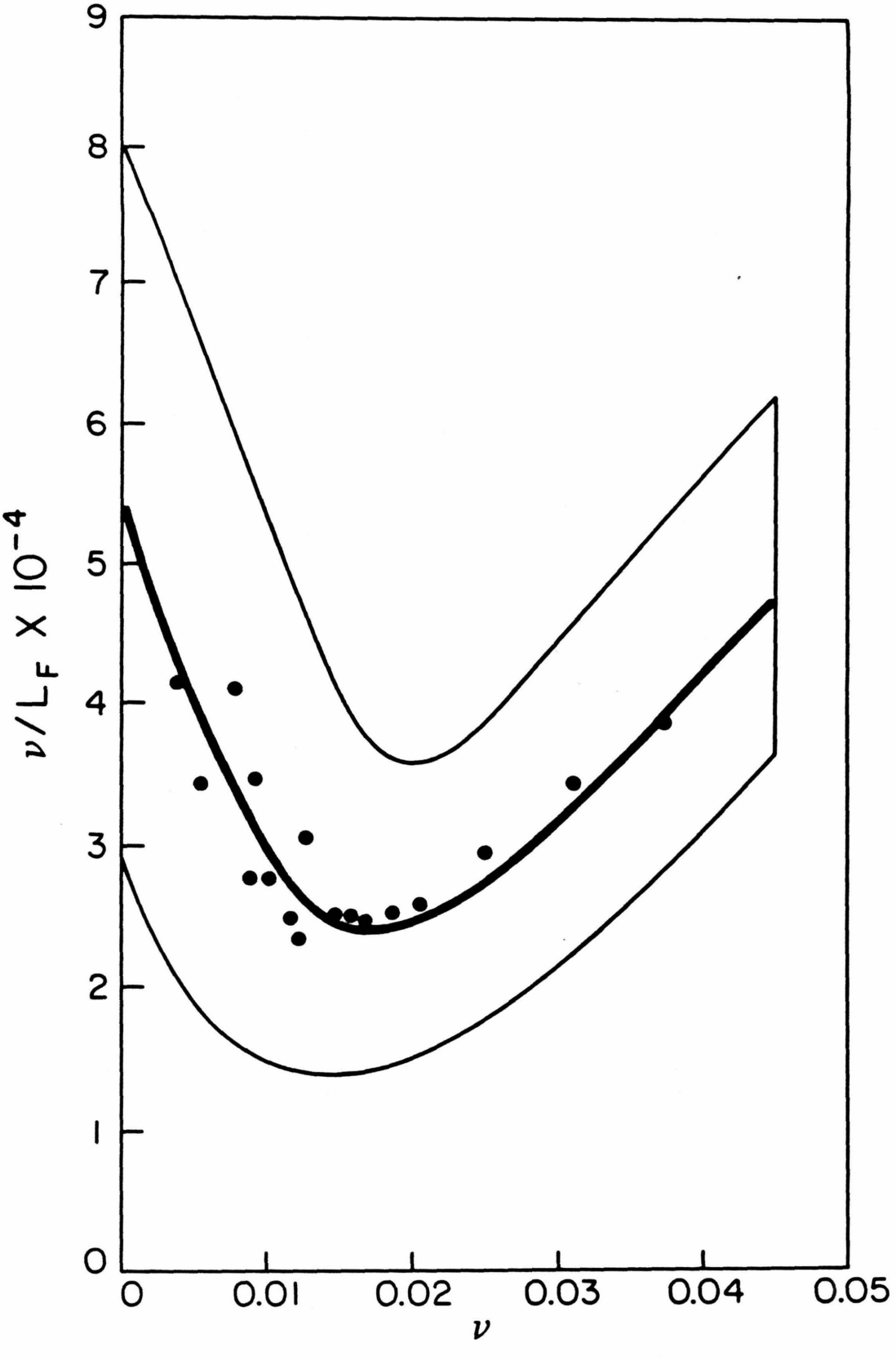

BMSp/BP
BP⁻¹ X 10⁻⁶
0.62
0.58
0.56
0.54
0.52
0.48
0.44
0.40
0.36
0.32

Figure 14

Scatchard plot of the binding of BMSp to polydAdT at 1 M^+. The values of ν and L_F are those obtained from linear regression analysis of the data in Fig. 13 and ν is corrected using $\varepsilon_B = 5240$[21] and $\varepsilon_F = 8000\ M^{-1}\ cm^{-1}$.[20] That region of the Scatchard plot defined by a $\pm$ 1 standard deviation error limit in the value of ν and ν/L_F is also outlined. Measurements at low binding density are shown separately in Fig. 13b.


3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
$\nu/L_F \times 10^{-5}$
0
0.08
0.16
0.24
0.32
0.40
0.48
0.56
0.64
0.72
ν


9
8
7
6
5
4
3
2
1
0
$\nu/L_F \times 10^{-4}$
0
0.01
0.02
0.03
0.04
0.05
ν

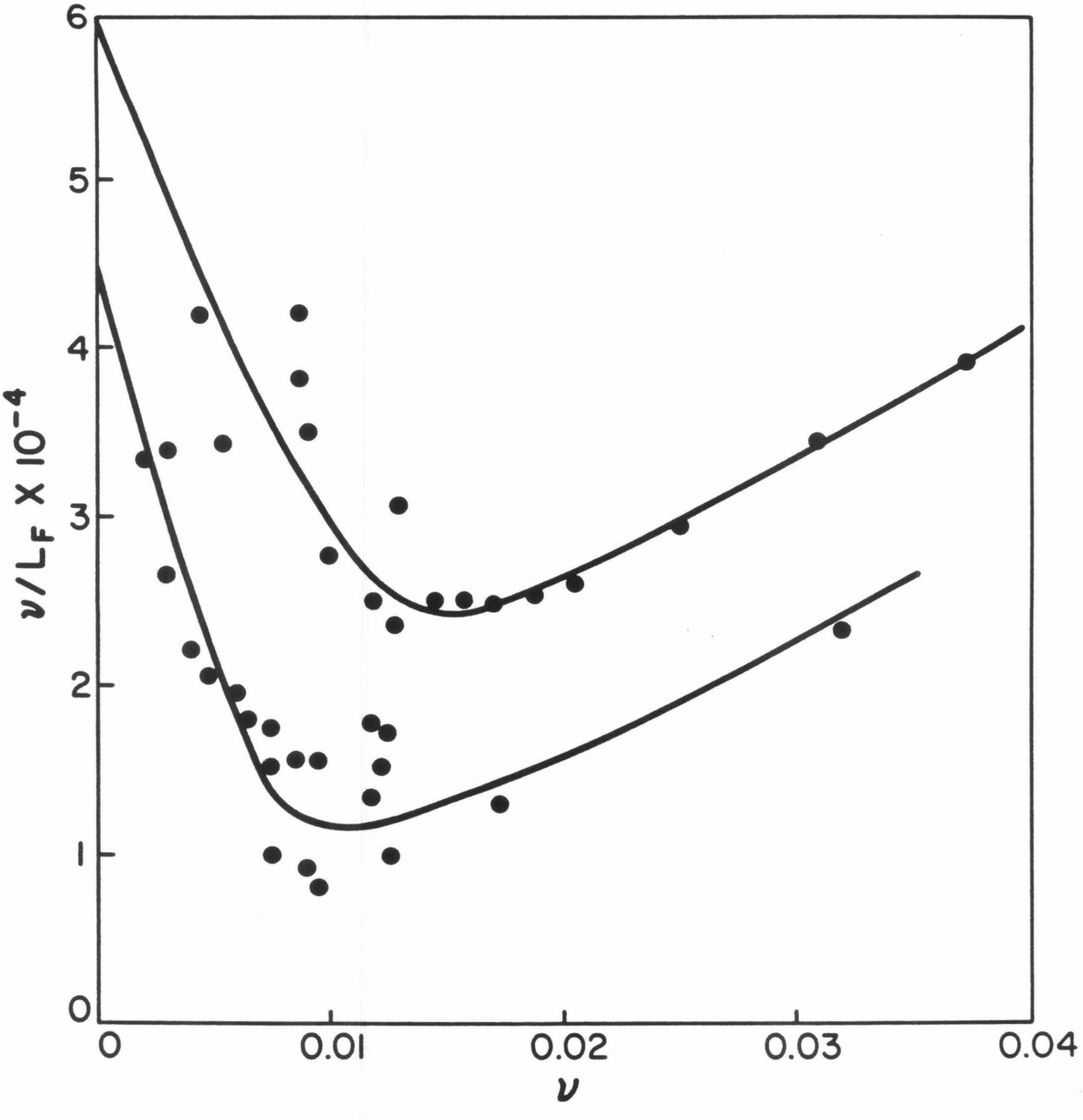


Figure 15
Scatchard plots of the binding of BMSp to polydAdT at 1 M^+ obtained at constant absorbance. The upper curve is from Fig. 14 and the lower curve from Fig. 12.

Comparison With Previous Work

Several groups have previously suggested the use of equations similar to eq. (4) to determine ligand binding parameters.[10,17-19] In Hélène's approach, the dependence of L_T/M on $1/M$ is monitored at constant X/X_0 where X is the circular dichroism of bound macromolecule and X_0 is the circular dichroism of unbound macromolecule.[18] Halfman and Nishida monitor ligand binding at constant X/X_0 where X is some physical property of ligand[17] whereas Baguley and Falkenhaug monitor binding at constant X/M where X is the fluorescence of ligand at some macromolecule concentration M.[19] Unfortunately, since these techniques evaluate ligand binding at different ratios of macromolecule or ligand physical properties, their sensitivity is no better than conventional indirect or direct techniques. In contrast, because the constant X technique evaluates ligand binding at absolute values of ligand or macromolecule physical properties, ligand-macromolecule interactions which are normally unmeasurable by other techniques can be investigated (see Introduction). (Although unknown to these authors,[17-19] like the present technique, estimates rather than true values of ν and L_F will be obtained).

REFERENCES

1) Scatchard, G., Ann. N.Y. Acad. Sci., 1949, 51, 660.

2) McGhee, J.D.; von Hippel, P. H., J. Mol. Biol., 1974, 86, 469.

3) Koshland, D.E.; Nemethy, G.; Filmer, D., Biochemistry, 1966, 5, 365.

4) Cornish-Bowden, A.; Koshland, D.E., Biochemistry, 1970, 9, 3325.

5) Monod, J.; Wyman, J.; Changeux, J-P., J. Mol. Biol., 1965, 12, 88.

6) Gibson, R.G.; Lewin, S.A., Proc. Natl. Acad. Sci. USA, 1977, 74, 139.

7) Tanford, C., "Physical Chemistry of Macromolecules," Wiley, New York, 1961, p. 528.

8) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

9) Pohl, F.M.; Jovin, T.M.; Baehr, W.; Holbrook, J.J., Proc. Natl. Acad. Sci. USA, 1972, 69, 3805.

10) Revet, M.J.; Schmir, M.; Vinograd, J., Nature, New Biology, 1971, 229, 10.

11) Raw absorbance data are shown in Figs. 6-11 of the Appendix.

12) Raw absorbance data are shown in Figs. 12-14 of the Appendix.

13) See Figs. 28 and 29 of the Appendix.

14) Raw absorbance data are shown in Figs. 15-19 of the Appendix.

15) Raw absorbance data are shown in Figs. 20-27 of the Appendix.

16) See Fig. 30 and reference 9.

17) Halfman, C.J.; Nishida, T. Biochemistry, 1972, 11 , 3492.

18) Durand, M.; Maurizot, J-C.; Borazan, H.N.; Hélène, C., Biochemistry, 1975, 13, 563.

19) Baguley, B.C.; Falkenhaug, E-M., Nucl. Acid Res., 1978, 5, 161.

20) See Table I, Appendix.

21) See Figure 15 of Chapter 1.

22) Bresloff, J., Ph.D. Dissertation, Yale University, 1974.

23) Bontemps, J.; Fredericq, E., Biophys. Chem., 1974, 2, 1.

CHAPTER III

Sequence Specific B → H → A Allosteric Transitions in DNA: A New Conformational Theory

We have previously reported the synthesis and characterization of a dimer of ethidium bromide termed BMSp or bis(methidium)spermine. Comparison of the binding properties of BMSp to ethidium under identical conditions clearly demonstrate that both ethidium moieties of BMSp simultaneously intercalate nucleic acid, substantially enhancing both its binding affinity and specificity.[1-3] Since BMSp is a tetracation at pH 7 whereas ethidium is a monocation, the binding of both compounds was studied under conditions which minimize electrostatic contributions to the observed binding affinity and intercalation geometry (pH 7.0 and a monovalent cation concentration of 1.0 M).[2] We found that for the five nucleic acids investigated:sonicated calf thymus DNA, polyd(C-G), polydCdG, polydAdT and polyrAdT, the free energy of BMSp binding, ΔG_{BMSp}, was equal to 1.5-1.6 ΔG_{EB}, where ΔG_{EB} is the free energy of EB binding to the same nucleic acid. If the salt concentration is lowered from 1.0 to 0.075 M thereby enhancing electrostatic interactions, the enhancement of BMSp's binding affinity relative to ethidium becomes much larger. For a heterogeneous DNA sequence (sonicated calf thymus DNA), we observed $K_{EB} = 4 \times 10^4$ and $K_{BMSp} = 1.5 \times 10^7$ at 1.0 M^+ whereas at 0.075 M^+ $K_{EB} = 2 \times 10^5$ and $K_{BMSp} \simeq 2 \times 10^{11}$.

In addition to binding affinity, the binding specificity of BMSp is also substantially enhanced relative to ethidium. Bresloff and Crothers have demonstrated that ethidium binds the RNA-DNA duplex polyrAdT approximately 100 times more tightly than the corresponding DNA duplex polydAdT.[4] This 100-fold specificity exhibited by ethidium was found to increase to 1500 for BMSp.[2,3] Since the only difference between rAdT and dAdT is the presence of a 2'hydroxyl group on the sugar ring and not base sequence, these studies indicated that the specificity which BMSp and EB exhibit for certain nucleic acids could arise from preferential recognition of different nucleic acid conformations.[2,3] Furthermore, since the binding affinity of EB is not altered by changes in the AT or GC content[5] of DNA, this conformational specificity is apparently much larger than any direct base specificity.

In this paper, both the molecular basis and generality of the conformational specificity of EB and BMSp are examined in more detail. On the basis of these measurements, evidence for a new conformational family of right-handed Watson-Crick DNA is presented. The ability of small sequence changes to greatly alter the conformational stability of this DNA suggests a new conformational theory of DNA which has important implica-

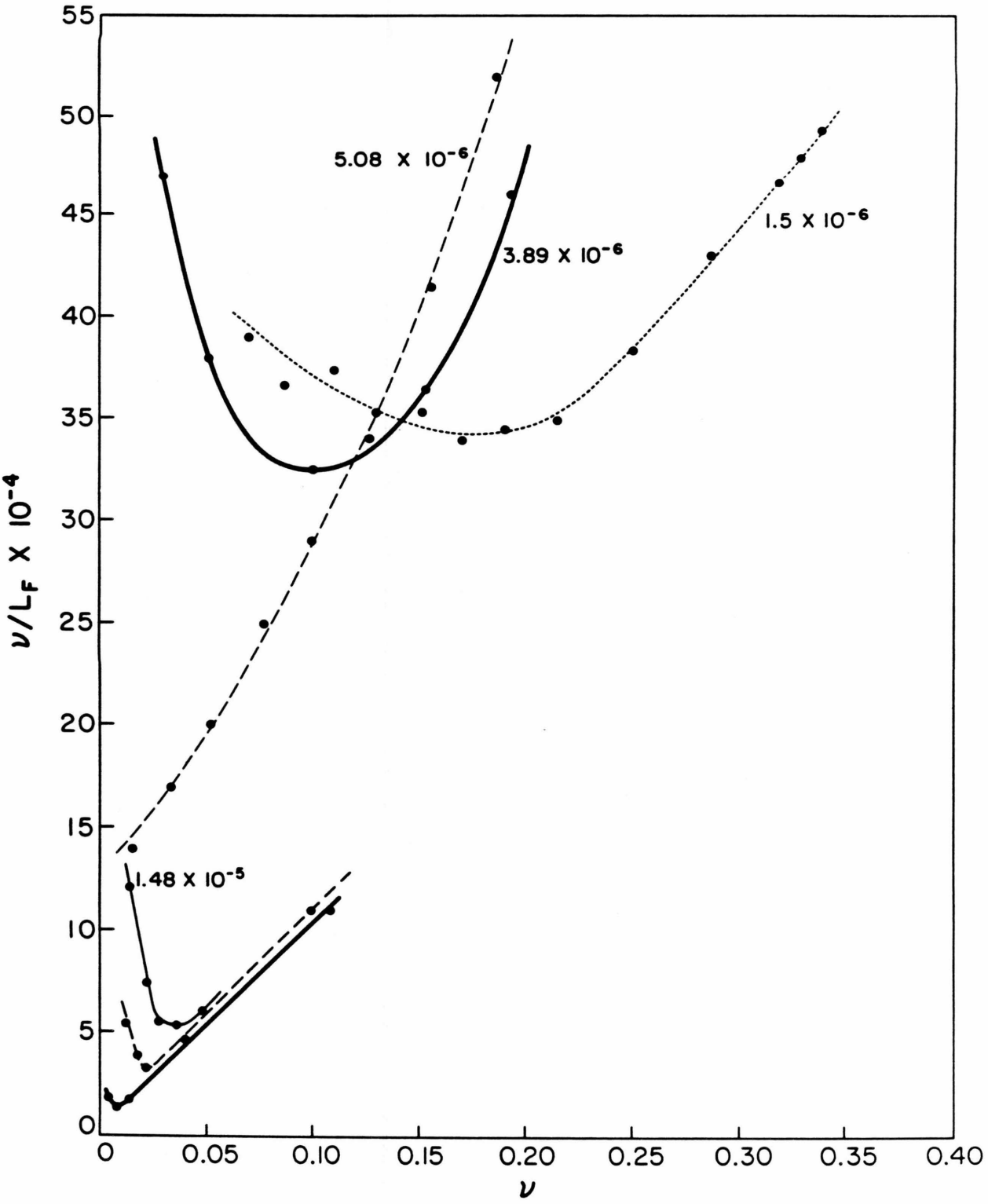
tions for understanding DNA structure and its molecular recognition by other molecules.

Binding of BMSp to PolydAdT

The greatly enhanced affinity of BMSp and EB for the RNA-DNA hybrid form of dAdT must arise from the presence of a 2'-hydroxyl group in the hybrid. Although a 2' hydroxyl group could enhance binding by forming a hydrogen bond to intercalated EB, crystal studies of ethidium intercalated into RNA dinucleotides indicate that it is too far away to do so.[6] Alternatively since a 2'hydroxyl group sterically forces nucleic acid to change structure from a B to an A conformation,[7] BMSp and EB may prefer to bind rAdT because it assumes a more tightly binding A conformation. Since DNA can adopt an (A) conformation in addition to its native B conformation,[8] we reinvestigated the binding of BMSp to polydAdT over a wide BMSp/BP ratio. If BMSp binds A conformations of nucleic acid much more tightly than B conformations, and if this specificity does not depend on the presence of a 2' hydroxyl group, we would expect polydAdT to adopt an A conformation as the concentration of BMSp is increased. Classical indirect spectroscopic techniques have been used to monitor this binding and the results of these measurements are presented in terms of a Scatchard plot in Fig. 1. Because our spectroscopic technique assumes

Figure 1

Scatchard plot of the binding of BMSp to polydAdT at 1 M^+ assuming the formation of one bound complex. Binding was monitored by visible absorption spectroscopy at 550 nm (upper curve) or 485 nm (lower curve).[42] The concentration of polydAdT in base pairs is indicated.

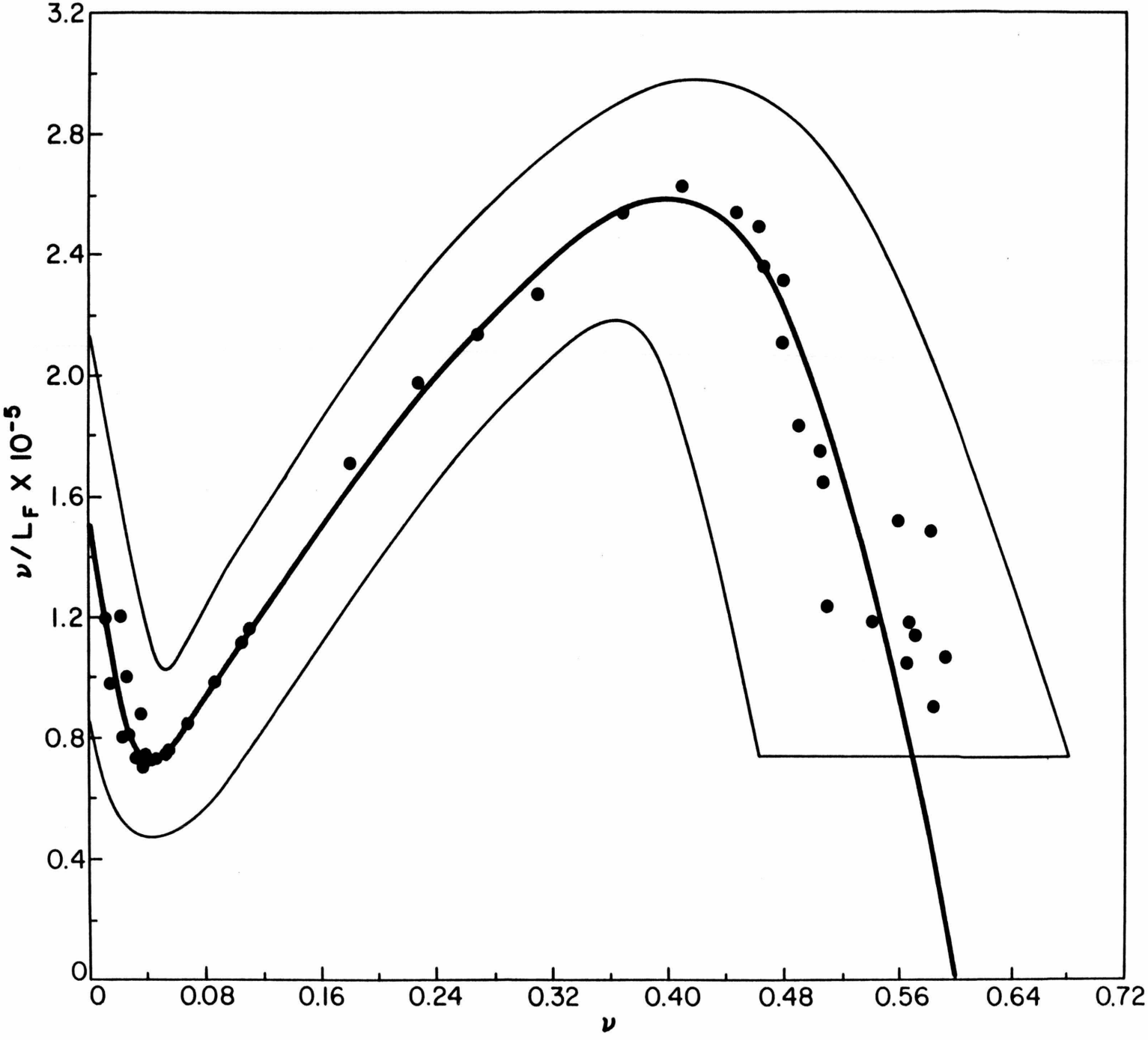

55
50
45
40
35
30
25
20
15
10
5
0
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
$\nu/L_F \times 10^{-4}$
ν
5.08×10^{-6}
3.89×10^{-6}
1.5×10^{-6}
1.48×10^{-5}

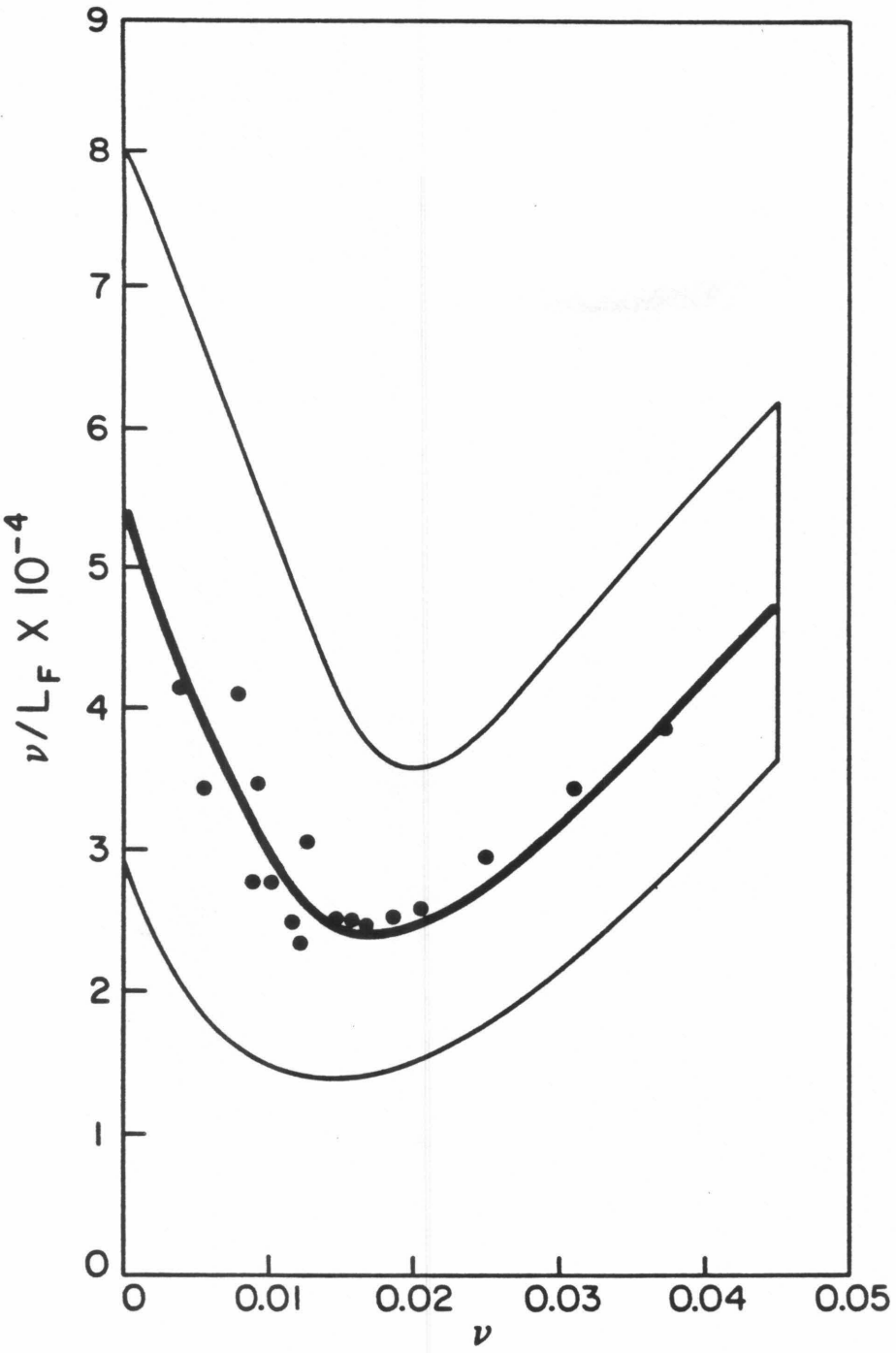
the formation of only one bound complex, the dependence of the Scatchard plot on DNA concentration suggests that BMSp forms at least two spectroscopically distinct complexes with dAdT. Furthermore, the change in slope observed at low binding densities indicates that the binding properties of dAdT are altered as the concentration of BMSp is increased. Although these data suggest that dAdT changes conformation when bound by increasing amounts of BMSp, the assumption of only one bound complex in our spectroscopic analysis precludes quantitative analysis.

Because BMSp appears to form more than one bound complex with dAdT, we have utilized another approach recently developed by us,[3] which can accurately estimate ligand binding in cases where more than one bound complex is formed. Because this technique is also very sensitive, macromolecule binding at very low binding densities can be measured. As shown in Fig. 2, initially BMSp binds dAdT weakly and noncooperatively since at very low binding densities its Scatchard plot exhibits a small vertical intercept and negative slope respectively. However, as the concentration of BMSp is increased, dAdT binds BMSp cooperatively (positive slope). Since bound intercalators do

Figure 2

a) Scatchard plot of the binding of BMSp to polydAdT at 1 M^+. The values of ν and L_F are obtained at constant absorbance[43] by the constant X technique[3] and ν is corrected using ε_B = 5240 and ε_F = 8000 M^{-1} cm^{-1}. That region of the Scatchard plot defined by a $\pm$ 1 standard deviation error limit in the value of ν and ν/L_F is also outlined. b) Same measurements at low binding density. Under these conditions, the extinction coefficient of BMSp bound to polyrAdT or to the B conformation of polydAdT are equal.


3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
$\nu/L_F \times 10^{-5}$
0
0.08
0.16
0.24
0.32
0.40
0.48
0.56
0.64
0.72
ν


9
8
7
6
5
4
3
2
1
0
$\nu/L_F \times 10^{-4}$
0
0.01
0.02
0.03
0.04
0.05
ν

not physical contact one another,[6,9] cooperative binding must arise from changes in the structure of dAdT, induced by BMSp binding, which enhance subsequent BMSp binding. Although a variety of structural changes could conceivably enhance BMSp binding, a Scatchard plot minimum requires that they arise from the induction of an allosteric transition in dAdT.[10]

In the allosteric or two-state model of Monod, macromolecule is assumed to adopt either a tightly binding (T) or weaker binding (R) conformation.[11] In the absence of ligand all subunits adopt the thermodynamically preferred (R) conformation. However, as ligand is added the equilibrium shifts from the (R) to the (T) conformation due to the latter's enhanced affinity for ligand. Ligand binding is cooperative because with increased binding, there is an increased preference for macromolecule to assume a more tightly binding (T) conformation.[11]

An extremely important feature of an allosteric transition is its generality.[11] Rather than distorting macromolecule in some peculiar fashion, ligand induces macromolecule to change conformation by "trapping" preexisting conformational states. Thus <u>any</u> ligand which exhibits the appropriate specificity can cooperatively

induce macromolecule to change conformation uniformly.

Although a minimum in a Scatchard plot demands the induction of an allosteric transition, all allosteric transitions do not exhibit minima. Only when the native (R) conformation of macromolecule is greatly favored thermodynamically, but ligand strongly prefers to bind the unstable (T) form, does a Scatchard plot exhibit a minimum.[10] Since dAdT strongly prefers to adopt a B conformation,[12,13] while BMSp is expected to greatly prefer its A conformation, the induction of a two-state B → A transition in dAdT by BMSp would be expected to exhibit a Scatchard plot minimum. Furthermore, since a B to A transition is cooperative,[14] its induction is expected to exhibit two-state behavior. Thus the reluctance of dAdT to change conformation, the facile ability of BMSp to alter its conformation, and the allosteric character of the induced transition, argue that BMSp binding induces a B to A conformational transition in dAdT by virtue of its enhanced specificity for A conformations of nucleic acid.

Since EB also appears to bind A conformations more tightly than B conformations, large amounts of it should induce a B to A transition in dAdT. Both the equilibrium and kinetic measurements of Bresloff are con-

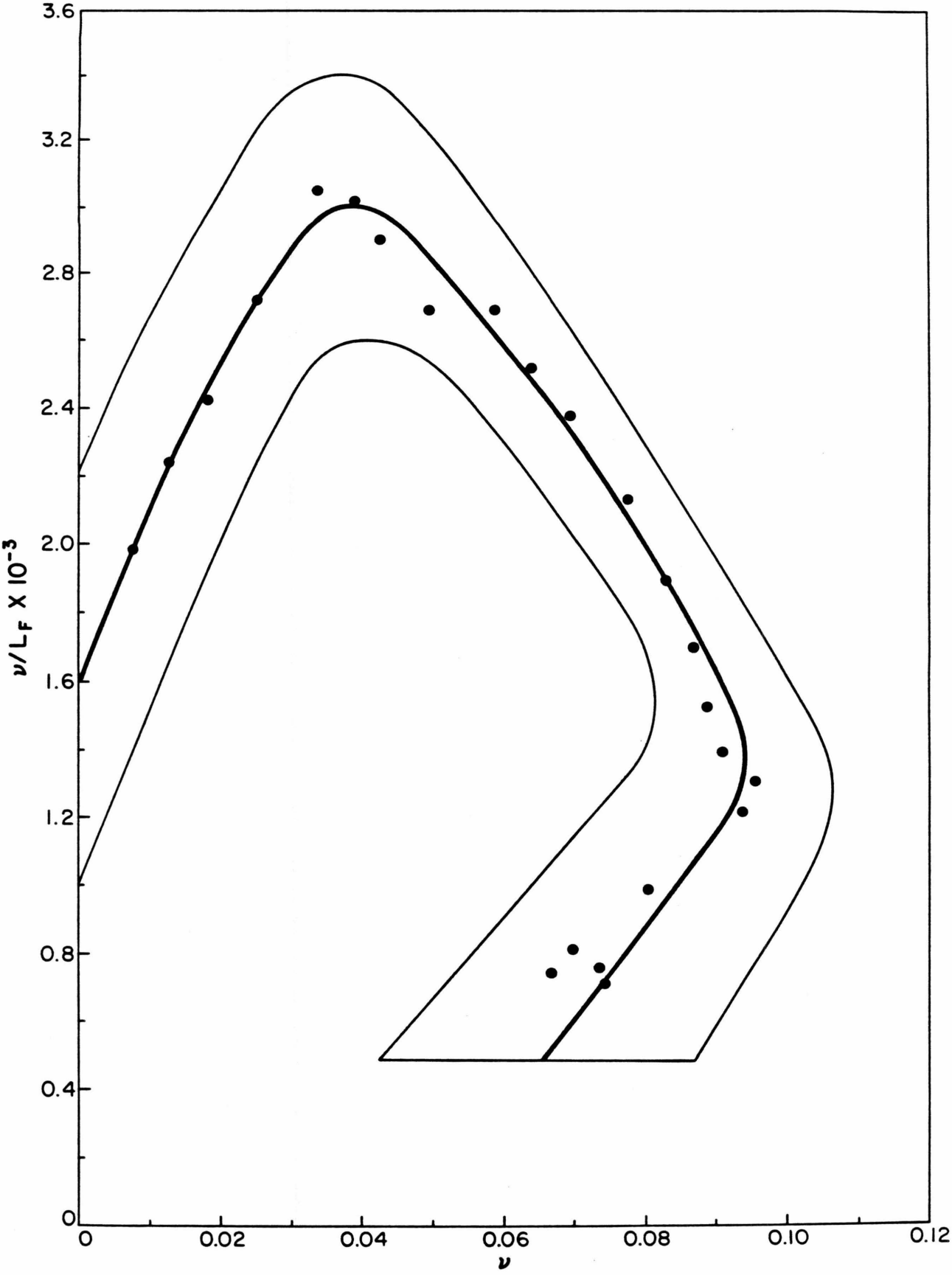
sistent with this hypothesis.[15] These authors found that large amounts of ethidium caused polydAdT to cooperatively change its structure into another form which resembled other intercalated nucleic acids. Although this behavior was attributed to the induction of a local conformational change by closely bound ethidium molecules, it is equally compatible with the induction of an allosteric transition by EB.[11]

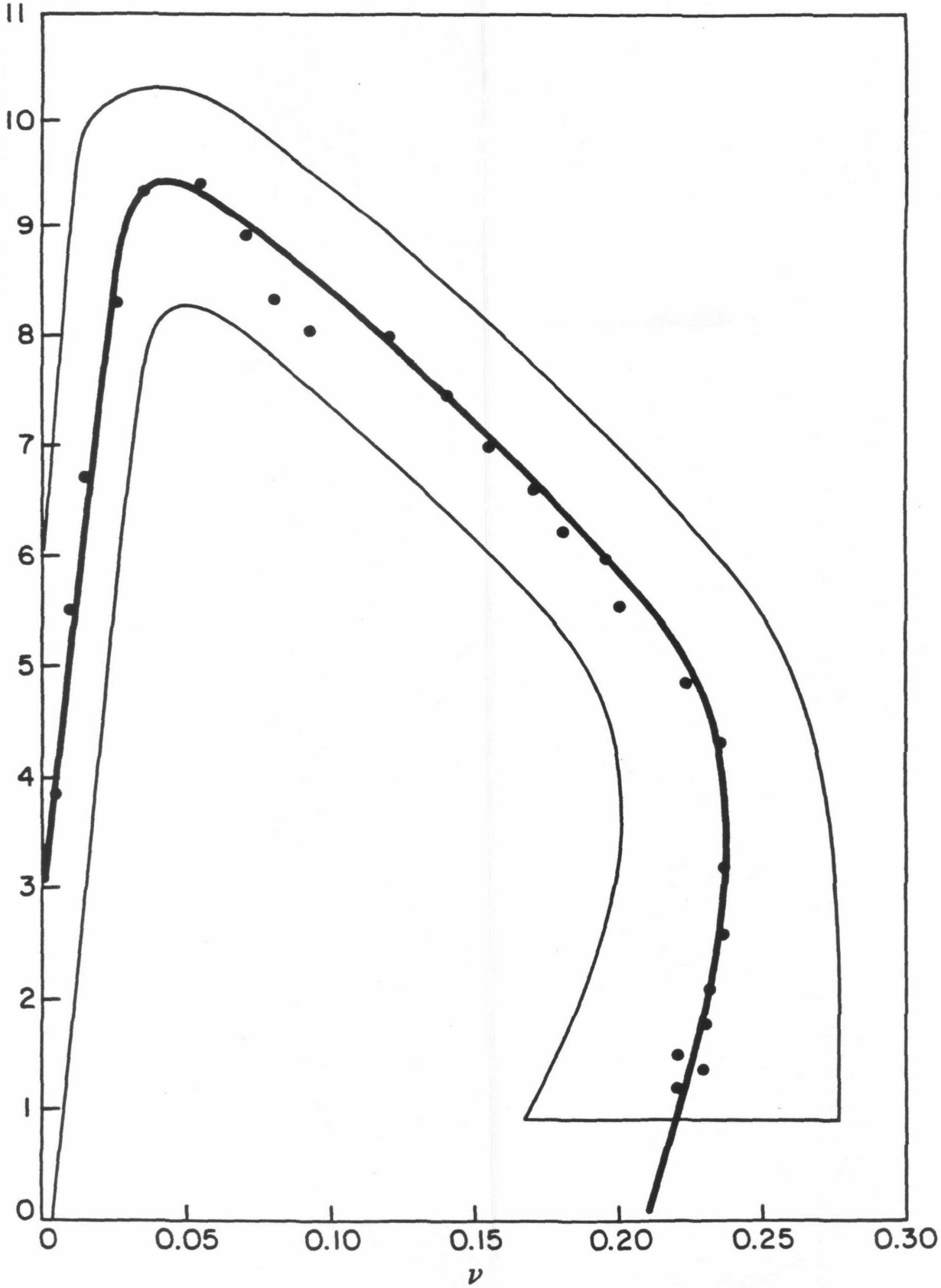
Binding of EB to polyd(C-G) and polydCdG

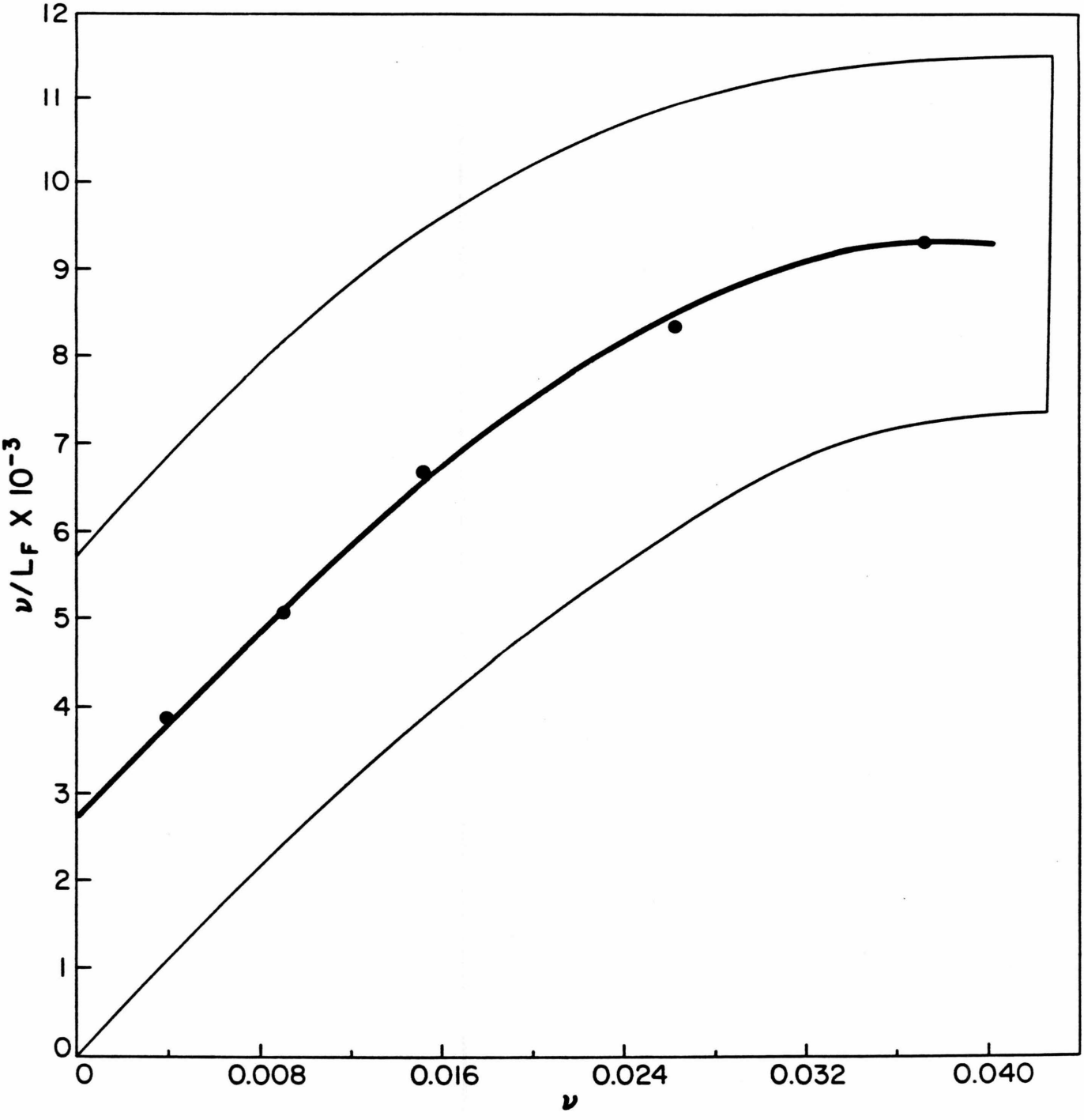
Since a B to A transition is a general property of Watson-Crick DNA,[12,14,16] BMSp and EB have the potential to induce B to A transitions in any DNA base sequence, provided their enhanced affinity for A conformations is not altered by base sequence. We have employed EB rather than BMSp to monitor B → A transitions in other base sequences since it avoids the experimental difficulty of detecting abrupt conformational shifts at very low binding density (Fig. 2). The same sensitive technique used to monitor the dAdT-BMSp allosteric transition has been used to monitor the induction of conformational transitions in polydCdG and polyd(C-G) by EB (Fig. 3).

Figure 3

Scatchard plot of the binding of EB to A) polydCdG and B) polyd(C-G) at 1 M^{+}.[44] The values of ν and L_F are obtained at constant absorbance by the constant X technique.[3] C) Same measurements of polyd(C-G) at low binding density. Under these conditions ε_B = 3100 and ε_F = 600 M^{-1} cm^{-1}. For EB bound to polydCdG and ε_B = 2250 for BMSp bound to polyd(C-G). That region of the Scatchard plot defined by a $\pm$ 1 standard deviation error limit in the value of ν and ν/L_F is also outlined.


3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
$\nu/L_F \times 10^{-3}$
0
0.02
0.04
0.06
0.08
0.10
0.12
ν

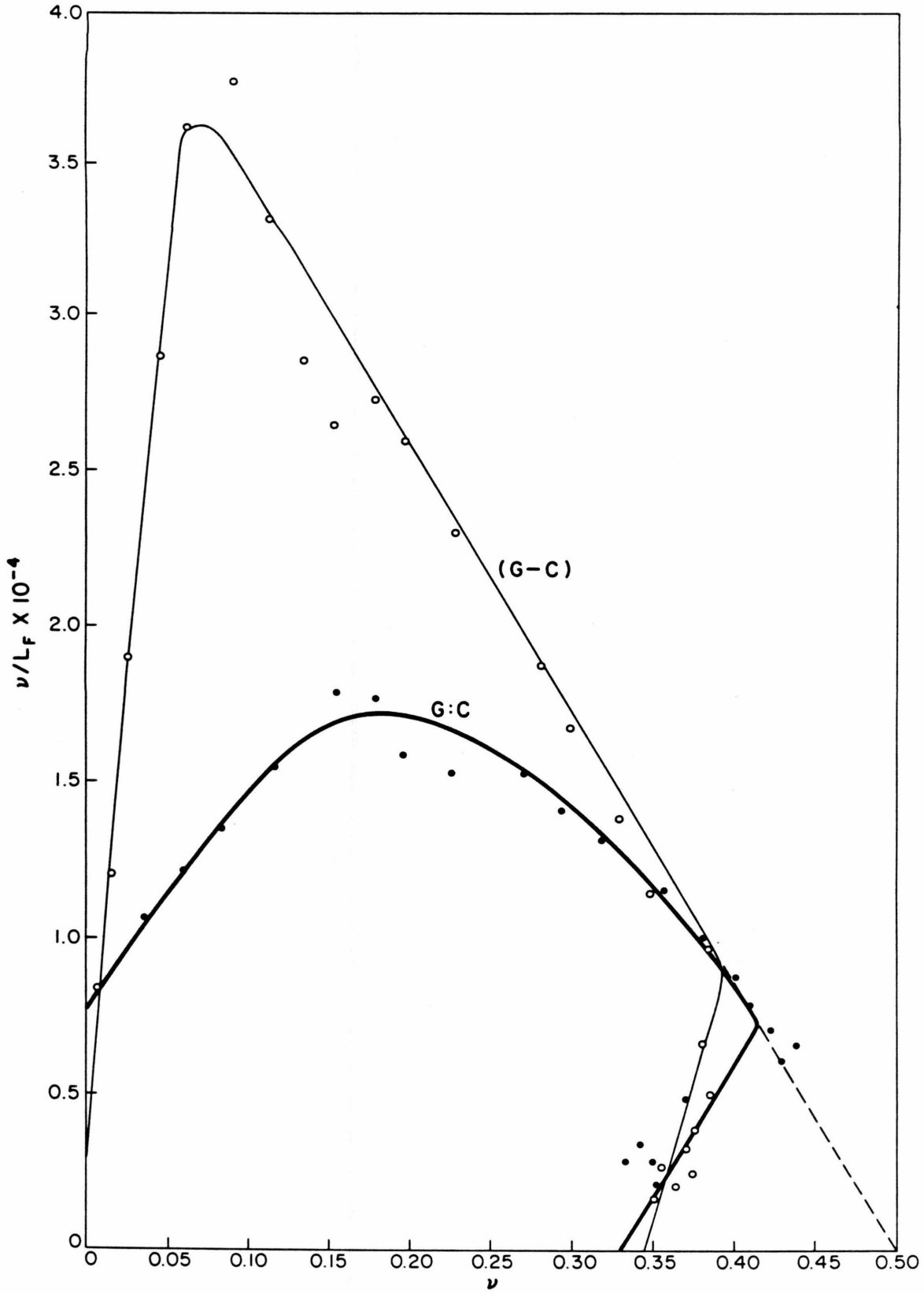

11
10
9
8
7
6
5
4
3
2
1
0
0
0.05
0.10
0.15
0.20
0.25
0.30
ν


12
11
10
9
8
7
6
5
4
3
2
1
0
$\nu/L_F \times 10^{-3}$
0
0.008
0.016
0.024
0.032
0.040
ν

Because measurements were conducted under conditions of low salt concentration (1.0 M), initial ethidium binding is to a B conformation.[17,18] This binding is weak as reflected by a small vertical intercept on Scatchard plot at low binding densities. However, further addition of ethidium induces both nucleic acids to change structure as reflected by the change in slope which occurs in both Scatchard plots. Because this new form binds ethidium with higher affinity than the B conformation, initial binding is cooperative (positive slope). As saturation of this second structure is approached, a third binding form of both nucleic acids is cooperatively induced by ethidium addition. This third form appears to be an A conformation since, like RNA and RNA-DNA duplexes (known A conformations), it binds one ethidium every three base pairs at saturation[15] (Fig. 4). Taken together, these measurements argue that EB can induce B to A transitions in both polydCdG and polyd(C-G) by virtue of its enhanced affinity for A conformations.

Figure 4

Scatchard plot of the binding of EB to polydCdG and polyd(C-G) at 1 M^+. The values of ν and L_F are from Fig. 3 and have been corrected according to the procedure outlined in Chapter II. The correction parameters for dCdG are $\nu = 4.25\ \nu'$ and $L_F = L_F'$ whereas for d(C-G) $\nu = 1.63\ \nu'$ and $M^* = 1 \times 10^{-4}$.


4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
$\nu/L_F \times 10^{-4}$
(G−C)
G:C
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
ν

Mechanism of the B → A Transition

An unexpected feature of both ethidium Scatchard plots is the appearance of an intermediate in the B to A transition. EB apparently traps this intermediate by binding it more tightly than a B conformation but less tightly than an A conformation.

What is the structure of this intermediate? Although a B to A transition alters the structure of DNA in several different ways, previous x-ray studies have shown that these changes are linked to one simple molecular motion: a change in each sugar pucker of B DNA from a 2'endo to 3'endo configuration.[8,19] Because each base pair and phosphate linkage are attached to a sugar ring, a change in its pucker can give rise to a large tilt and twist of the base pairs thus allowing DNA to assume a shorter more compact A conformation. A simple mechanism for the B to A transition which complies with these structural constraints, accounts for the cooperative nature of the transition, and predicts the existence of an intermediate is presented in Fig. 5. Conversion of B DNA to A DNA is postulated to occur in two steps. In the first step, nucleation, one or more base pairs adopt a 2'endo (3' → 5')3' endo sugar pucker alternation. Propogation of this sugar pucker change to adjacent base pairs is cooperative and results in the

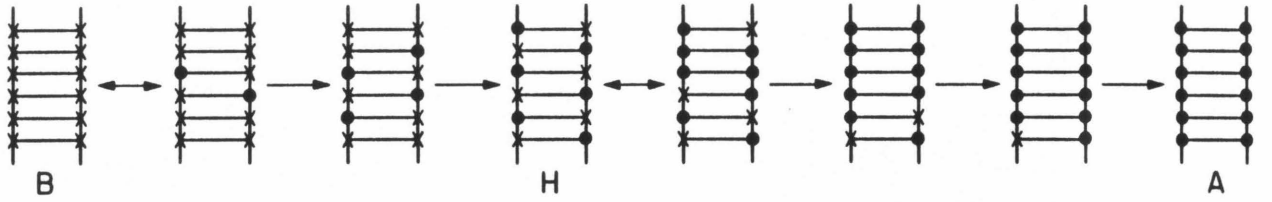


Figure 5. - Proposed mechanism of a B to A transition in DNA. B DNA, which contains all 2'endo sugar puckers (X) is converted to A DNA, which contains all 3'endo sugar puckers (●) in two steps. In the first step, nucleation, one or more base pairs adopt a 2'endo (3' → 5')3'endo sugar pucker alternation. Propogation of this sugar pucker change to adjacent base pairs is cooperative and results in the formation of a species whose sugar pucker alternates every base pair. Completion of the B to A transition proceeds from this species by the same steps used to generate it.

formation of a species whose sugar pucker alternates every base pair. Because completion of the B to A transition can proceed from this species by the same steps used to generate it, this mechanism necessarily predicts the existence of an intermediate in the B to A transition.

Sobell and coworkers have demonstrated by x-ray crystallography that intercalation of ethidium into dinucleotides is accompanied by a change in sugar pucker to a mixed 2'endo (3' → 5')3'endo configuration.[6] Presumably structural changes necessary for DNA to accept an intercalator are easier when a base pair adopts a mixed sugar pucker.[40] An important feature of a mixed sugar pucker at the intercalation site is to sterically prohibit adjacent base pairs from intercalating[11] (Fig. 6). This feature of Sobell's model provides a simple molecular basis for a characteristic feature of ethidium as well as other drug intercalation: at saturation, only one drug is bound every two base pairs (nearest neighbor exclusion).[9] Referring to Fig. 4, the proposed intermediate in the B to A transition also restricts EB intercalation to every other base pair. This observation not only supports the assignment of a 2'endo(3' → 5')3'endo sugar pucker alternation to the intermediate conformation

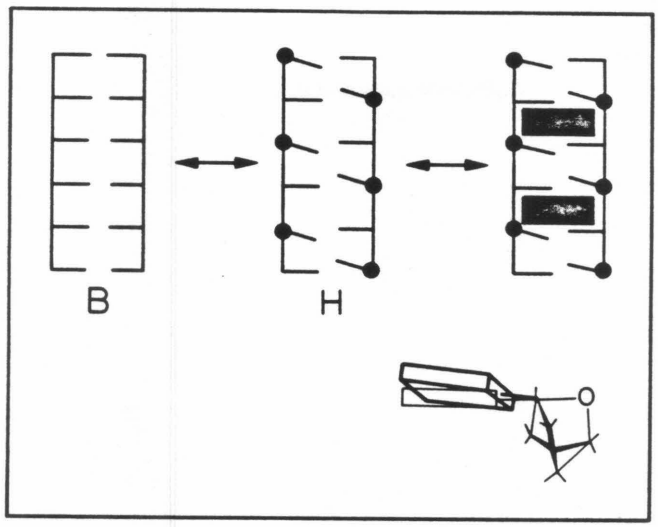


Figure 6
Restricted ethidium binding to an H conformation of DNA. B DNA is shown in equilibrium with H DNA. Because a B → H transition introduces base tilt and twist to every other nucleotide (see lower figure) intercalation is restricted to every other base pair. Structural characteristics are based on the crystal structure of $(ATAT)_2$[45] which appears to adopt an H conformation.

(Fig. 5), but suggests further that a variety of ligands bind DNA in a nearest neighbor exclusion fashion because they induce B DNA to adopt an "intermediate" conformation.

Hybrid DNA

On the basis of the proposed B to A mechanism shown in Fig. 5 and the observed binding properties of the A and intermediate conformation, (Fig. 4) we propose the existence of a new conformational family of right-handed Watson-Crick DNA. Termed H or Hybrid DNA, such DNA is characterized by 2'endo(3' → 5')3'endo sugar pucker alternation every base pair. Since the proposed B → H → A mechanism is independent of base sequence, we further postulate that a hybrid conformation is exhibited by all Watson-Crick DNA as an intermediate in the B to A transition. For RNA and RNA-DNA duplexes,[7] the H conformation, like the B conformation, is expected to be unstable with respect to its native A conformation. As a result, the H conformation will be more accessible to DNA than RNA.

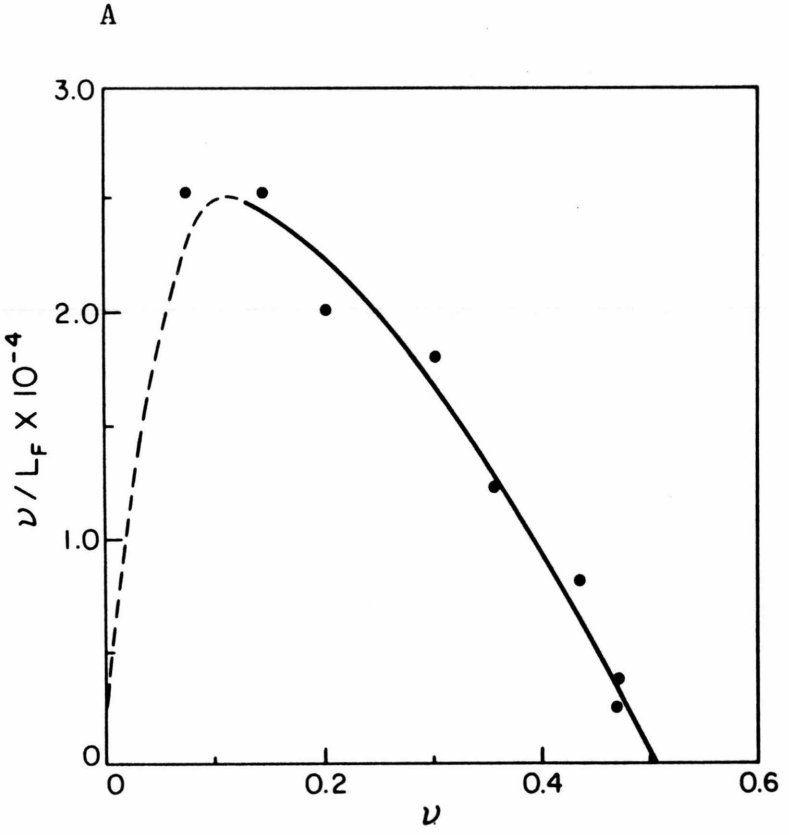
B → H → A Transitions in d(A-T)

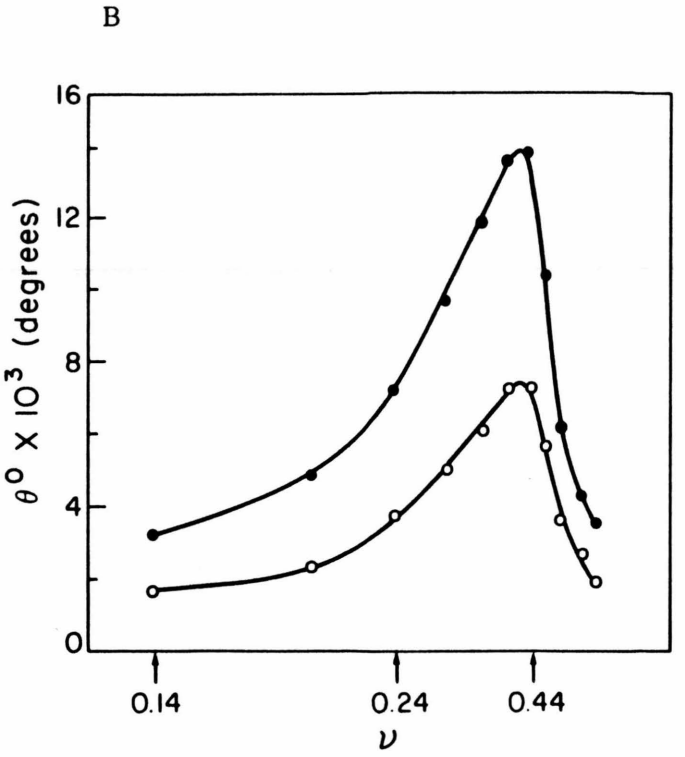
Pilet et. al. have demonstrated that polyd(A-T), like other nucleic acids, can undergo a B to A transition when dehydrated.[12] Since EB is expected to bind the H and A conformations of d(A-T) more tightly than the native B conformation, its binding should progressively induce a B → H → A transition in d(A-T). Bresloff and Crothers have previously investigated the binding of EB to polyd(A-T) (Fig. 7).[15] The curvature of their Scatchard plot, as well as the dependence of d(A-T)'s circular dichroism on EB binding[15] (Fig. 7b), demonstrate that from ν = 0.1 to 0.4, EB binding cooperatively alters the conformation of polyd(A-T). Although neither measurement is sensitive enough to monitor EB binding below ν = 0.1, cooperative binding could arise from the induction of a B → H transition by EB (see dotted line Fig. 7a), since the altered conformation binds one EB every two base pairs at saturation with an affinity characteristic of H DNA (Table I).

Increasing the concentration of EB above ν = 0.4 abruptly induces a second conformational change in d(A-T)[15] (Fig. 7b). This change in conformation may result from the induction of an H to A transition by EB since it occurs

Figure 7

a) Scatchard plot of the binding of EB to polyd(A-T).[15] The solid line represents the best fit through experimental points and the dotted line represents the proposed B → H transition. b) The circular dichroic ellipicity (degress) of polyd(A-T) as a function of the EB binding density (●, 307 nm; ○, 333 nm).[15]


A
3.0
2.0
1.0
0
$\nu / L_F \times 10^{-4}$
0
0.2
0.4
0.6
ν


B
16
12
8
4
0
$\theta^0 \times 10^3$ (degrees)
0.14
0.24
0.44
ν

at the same binding density as EB induced H to A transitions in polyd(C-G) and polydCdG (Fig. 4).

Very recent NMR work by Patel and coworkers[20] may provide direct evidence for the occurrence of a hybrid conformation in polyd(A-T). These workers found that concentrated CsF or Me_3NCl solutions (which presumably dehydrate DNA), as well as a 3α, 17β diprandium steroid, readily induce polyd(A-T) to change structure from a B to a right-handed alternating conformation. Since every glycosidic angle and phosphate linkage adopts a non B conformation, the alternating conformation appears to possess a hybrid DNA sugar pucker alternation.

B → H → A Transitions in Other DNA

Although B to A transitions have been investigated for many years, previous measurements have failed to detect an intermediate.[12,14,16] This failure could be due to the inherent instability of an intermediate in a cooperative transition. For example, the proposed B to A mechanism in Fig. 5 predicts that the free energy of the H intermediate is approximately halfway between the B and A conformations. Since the B to A transition is cooperative, induction of a B to H transition will be just as favorable as induction of an H to A transition.

As a result, H DNA is expected to be an unstable intermediate. EB binding can apparently trap and stabilize the hybrid intermediate because its binding gradually shifts the B → H → A equilibrium.

Another type of experiment which could potentially detect an intermediate in the B to A transition is electron microscopy, since individual molecules can be investigated. Vollenweider et al[21] have demonstrated that when dehydrated, T7 and λ DNA always adopt one of three class lengths. The average axial rise per base pair observed for these classes: 3.2, 2.9, and 2.6 Å are compatible with B, H, and A conformations, respectively. Furthermore, as expected for a B ↔ H ↔ A equilibrium, these three class lengths of DNA appear cooperatively coupled to one another since small differences in sample preparation can greatly alter their relative proportions.

Sequence Specificity of the B → A Transition

The conclusion that BMSp and EB can induce B → H → A allosteric transitions in DNA by virtue of their enhanced affinity for H and A conformations of nucleic acid is quantitatively supported by theoretical analysis[10] of the Scatchard plots in Figs. 2 and 3. The allosteric constants (L) and ligand affinites for the various con-

formations bound by BMSp and EB are summarized in Table I. Due to the complexity of a three state allosteric transition, we have not as yet computed the allosteric constant for the H → A transition or the affinity of ethidium for the A conformations of dCdG and d(C-G). However, the fact that ethidium does induce A conformations in these DNA's requires that its affinity for the A conformation equal or exceed that for an H conformation. Two important points emerge from this analysis. First, the affinity of BMSp and EB for H and A conformations of DNA is significantly greater than for B conformations, the observed affinities being very similar to those found for RNA and RNA-DNA duplexes, known A conformations. Secondly, the allosteric constant, L, which is the equilibrium ratio B/H in the absence of ligand, varies <u>greatly</u> with base sequences.

Table I - Summary of the binding affinity of EB and BMSp to various nucleic acids at 1 M^+. The allosteric constant for the B → H transition as well as the conformation of nucleic acid bound are indicated. Measurements of Bresloff[15] are denoted by (*).

	EB		BMSp		
Nucleic Acid	K	Proposed Conformation	K	Proposed Conformation	L (B/H)
dAdT*	2×10^3	B	$1.6 \pm .7 \times 10^5$	B	10^7-10^{11}
dCdG	5×10^3	B	$2.6 \pm .2 \times 10^8$	H	70-1500
	1×10^5	H			
d(C-G)	3×10^3	B	$7.5 \pm .5 \times 10^7$	H	10- 300
	1×10^5	H			
rAdT*	2×10^5	A	$2.4 \pm .3 \times 10^8$	A	--
Sonicated Calf Thymus	4.5×10^4	H	$1.5 \pm .2 \times 10^7$	H	--
d(A-T)*	5×10^4	H			
rArU*	1×10^5	A			
d(I-C)*	4×10^4	H			
dIdC*	2×10^3	B			

Stability versus Structure

A variety of experiments have demonstrated that changes in the base sequence of DNA alters its secondary structure.[22,23] As a result, local variations in DNA's secondary structure have been postulated to play an important role in base-specific recognition processes, notably regulatory protein-DNA interactions.[24] Although conceptually attractive, structural perturbations arising from base sequence variation[25] are apparently no larger than those induced by torsional (or probably bending) fluctuations.[26] Thus it appears unlikely that local variations in the secondary structure of DNA play an important role in base-specific recognition processes.

A property of DNA which could play an important role in base-specific recognition processes is its ability to undergo B → H → A transition.[52] For a cooperative transition like the B to H, the allosteric constant L can be written as:

1) $$L = \frac{[H]}{[B]} = \beta(\ell)^N$$

where β is the equilibrium constant for nucleation, ℓ is the [H]/[B] ratio for a base pair in the absence of nucleation, and N is the number of base pairs which change conformation in the allosteric transition[27] (refer to Fig. 5).

For the three synthetic DNA's examined, the maximum number of base pairs which could change conformation in the B to H allosteric transition is 700-900 base pairs (see next section). The actual number can be estimated from a Hill plot.[46] If one plots log(sites occupied/sites vacant) versus log(free ligand), at high and low values of free ligand the corresponding slope will equal one, whereas for intermediate values of free ligand the slope will be less than or equal to N_a, the number of base pairs which change conformation in the allosteric transition. When ligand strongly prefers to bind the non-native conformation of macromolecule, but macromolecule strongly prefers to assume its native conformation, the slope of the Hill plot at intermediate values of free ligand equals N_a.[47] For all other conditions, this slope underestimates N_a, the magnitude of the underestimation being a function of the unfavorability of the allosteric transition and the ligands relative specificity for the two conformations of macromolecule.[47] For the BMSp-dAdT Scatchard plot, the Hill plot slope at intermediate free ligand is expected to be very close to the actual number of base pairs which change conformation in the B → H transition, since dAdT strongly prefers the B conformation but BMSp strongly prefers to bind the H conformation (Table One). However, since the experimentally observed binding

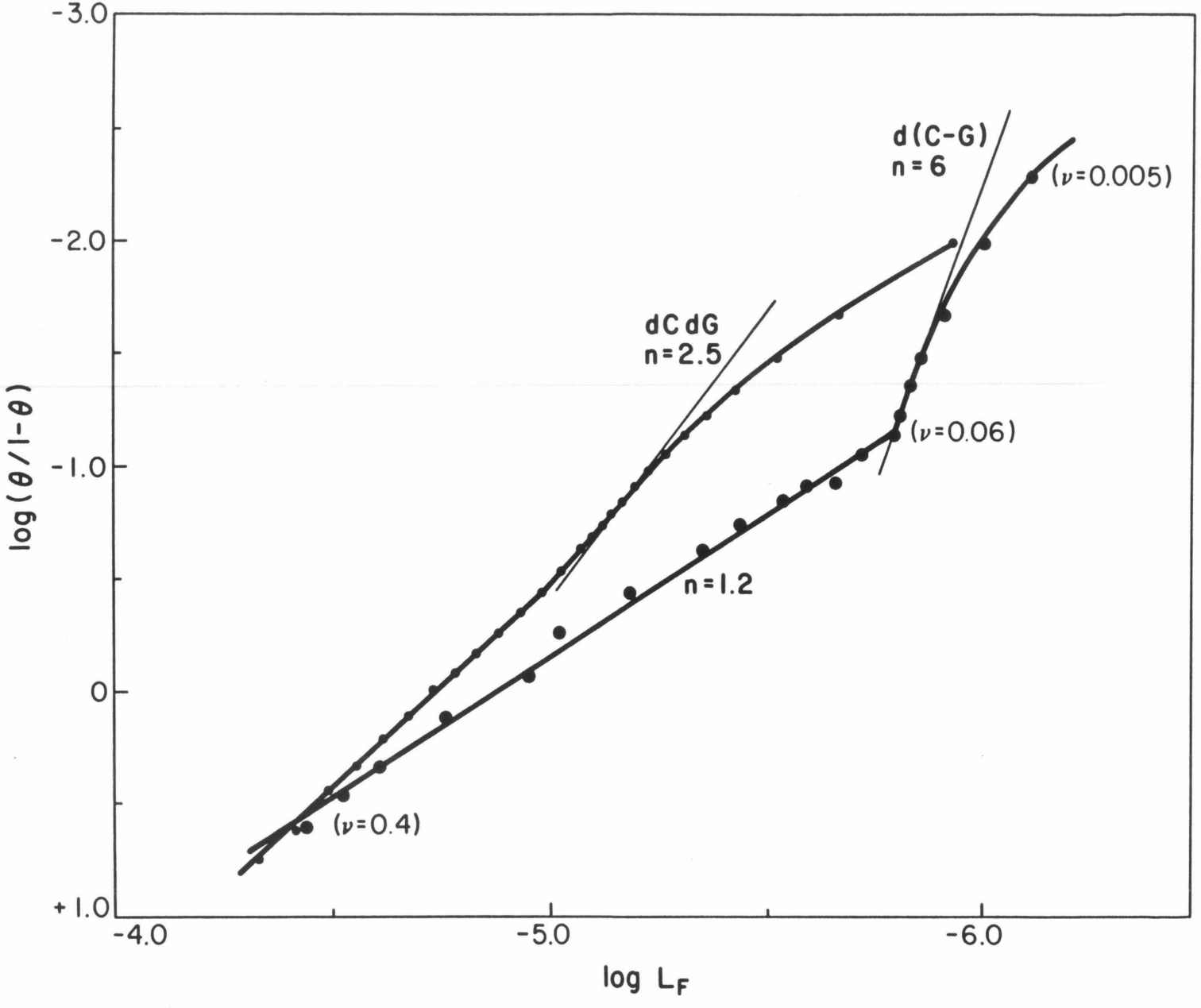
densities (Fig. 2) indicate that BMSp forms multiple bound complexes with dAdT, a meaningful value of Na cannot be calculated from these measurements.

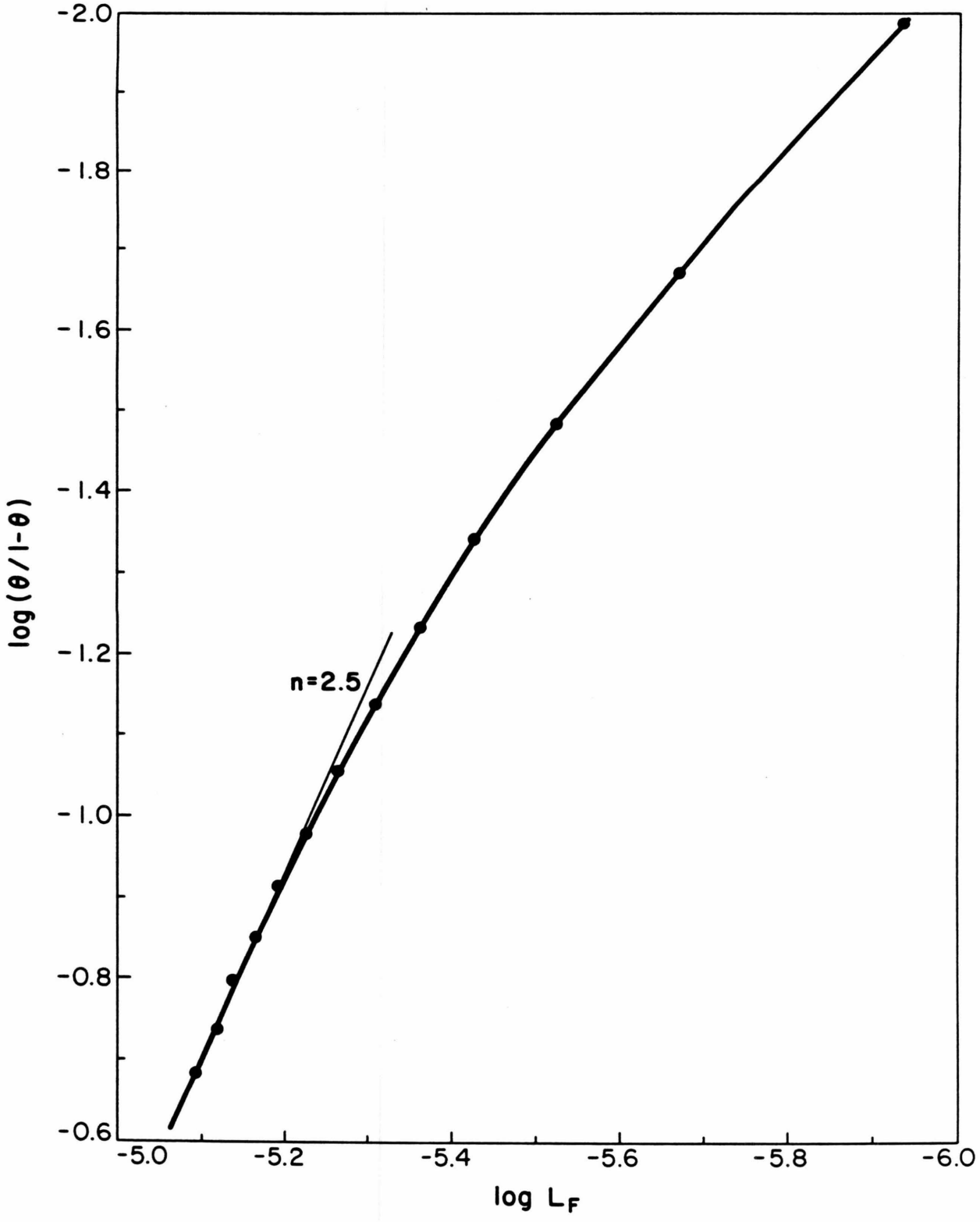
For the B → H transitions in polyd(C-G) and polydCdG, the Hill plot slope at intermediate free ligand is expected to be less than Na, since the preference of both DNA's for the B conformation is not extremely large and ethidium shows a moderate specificity for the H conformation. However, the value of Na calculated from each Hill plot can be corrected using the allosteric constants and relative binding affinities determined for the DNA's (Table One).[47] The values of Na determined for polydCdG (5 base pairs) and polyd(C-G) (10 base pairs) compare favorably with Ivanov's estimate of 10-20 base pairs for the cooperative unit in the B to A transition[14,48] (Fig. 9). Although a larger value of Na for the B to A transition is consistent with the proposal that A DNA is more than H DNA (Fig. 5), additional measurements are necessary to substantiate the generality of this observation.

That Hill plots can be used to estimate the number of base pairs which change conformation in an allosteric transition is demonstrated by the work of Pohl, et al.[49] These workers found that the binding of ethidium to left-handed polyd(C-G) results in an allosteric transition to a right-handed conformation. For oligo $d(C\text{-}G)_{10}$ the

Figure 9

Hill plots of the binding of EB to poly dCdG and polyd(C-G) at 1 M^+. The binding site size of EB is taken as 2 base pairs. Measurements for polydCdG at intermediate concentrations of free ligand are also shown separately.


-3.0
-2.0
-1.0
0
+1.0
-4.0
-5.0
-6.0
log (θ/1-θ)
log L_F
d(C-G)
n = 6
(ν=0.005)
dC dG
n=2.5
(ν=0.06)
n=1.2
(ν=0.4)


-2.0
-1.8
-1.6
-1.4
-1.2
-1.0
-0.8
-0.6
log (θ/1-θ)
n=2.5
-5.0
-5.2
-5.4
-5.6
-5.8
-6.0
log L_F

corrected value of Na is 10 and for oligo $d(C-G)_{200}$ it is too high to measure, but greater than 80.

Although only 5-10 base pairs actually change conformation in a B → H transition, many more base pairs than this will be affected by the transition since adjacent base pairs can change conformation without nucleating. Adjoining DNA is thus "silently activated", its ability to undergo a B to H transition being enhanced without a change in its conformation.

Referring to equation one, the large sequence specificity observed for the B → H → A transition could be due to variations in β, ℓ or both with base sequence. Although the extreme reluctance of polydAdT to adopt an H or A conformation could be attributed to a very unfavorable nucleation event, the measurements of Bresloff[15] suggest that the energy of activation required for the conformational change is low. Thus the large dependence of L on base sequence is probably due to variations in the ability of different base sequences to propogate B to H transitions.

In summary, our measurements suggest that the <u>equilibrium stability</u>, rather than the secondary structure of Watson-Crick DNA, varies greatly with base sequence.

Since the number of base pairs which change conformation in a B → H → A transition is similar to the binding site size of regulatory protein subunits, the extreme base specificity of the B → H → A transition could play an important role in gene regulation.

Alternative Explanations and Summary

Although the allosteric transitions induced in DNA by BMSp and EB binding are consistent with a B → H → A transition, other structural changes in DNA could conceivably exhibit similar behavior. Possible candidates would include single strand ↔ double strand, triple strand ↔ double strand or right-handed ↔ left-handed interconversions. The possibility that the observed allosteric transition arises from a single to double strand conversion is unlikely since under the conditions of our experiments, dCdG, dAdT, and d(C-G) are known to prefer double stranded Watson-Crick configurations.[12,17,22, 28,29] A double to single strand transition is also unlikely since ethidium is known to greatly stabilize the double helix.[30] Although each nucleic acid adopts a right-handed configuration under the conditions of our experiment, if ethidium preferred to bind left-handed forms it could induce a right to left allosteric transition. This possibility is also unlikely since ethidium appears to strongly prefer right-handed forms of DNA.[31]

The ability of dCdG and dAdT to adopt triple stranded structures[29,32] raises the possibility that, in these cases, the observed allosteric transitions may

result from the induction of a triple → double strand interconversion. This hypothesis is made more plausible by the observation that ethidium can induce a triple to double strand allosteric transition in polydArU_2.[33] To ascertain whether the dCdG and dAdT homopolymers used in our binding studies contained triple strands, each DNA preparation was titrated with a single strand known to induce triple strand formation.[28,32] As shown in Figs. 9 and 10, both homopolymers form triple strands at a single strand percentage close to 66%. Since these measurements require that less than 5% of either homopolymer used in our binding study exist as triple strands, the observed allosteric transitions cannot be due to a triple to double strand interconversion.

Although the allosteric constants we have measured are sequence-specific, this conclusion is valid only if the molecular weights of the DNA's are known. We have examined both the single and double stranded length distribution of the dAdT, dCdG, and d(C-G) polymers utilized in our drug binding experiments. The DNA preparations used were chosen because their single-stranded lengths are long and similar: 700-950 bases (Fig. 12). Although single-strand lengths are similar, double-stranded

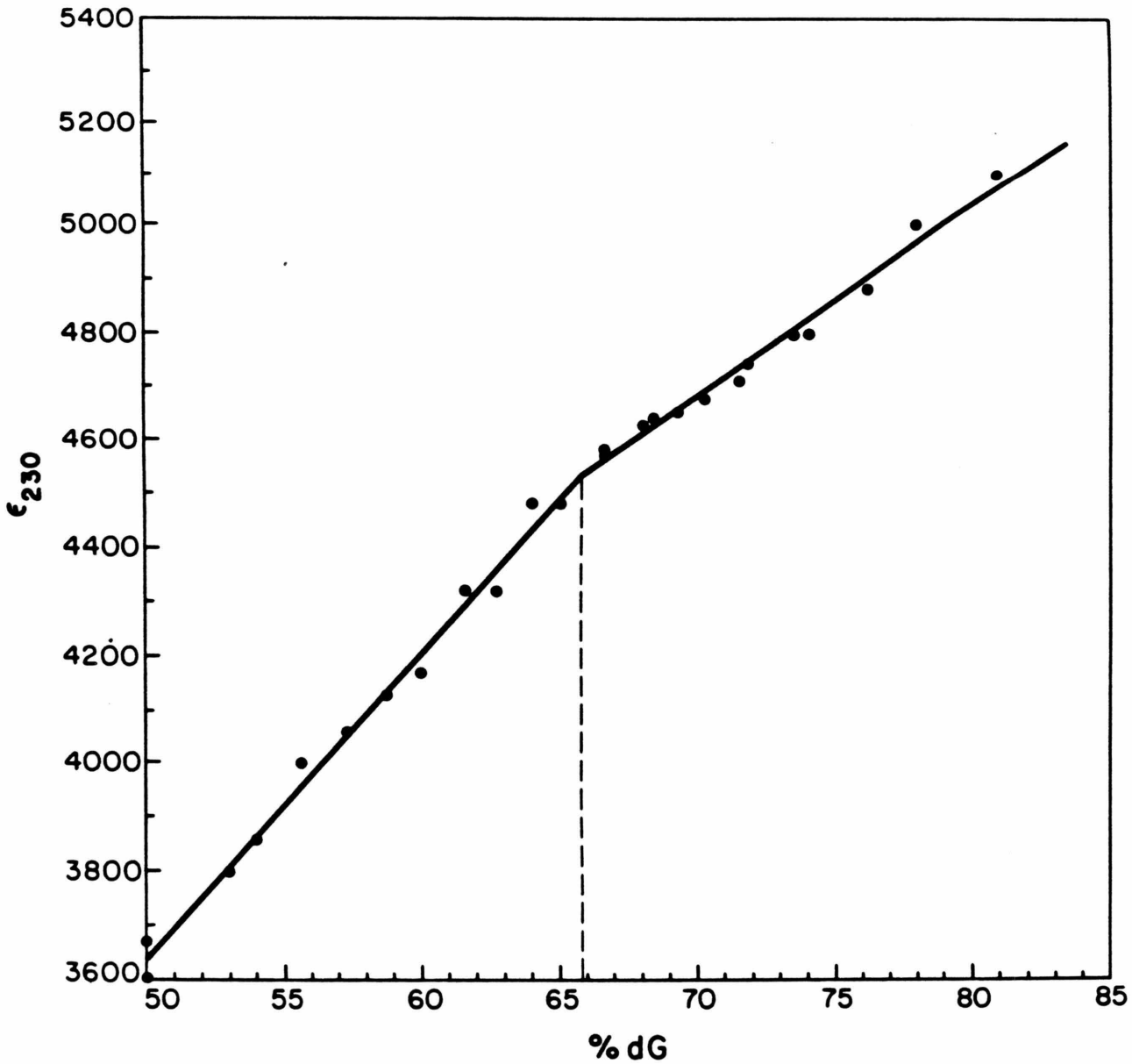
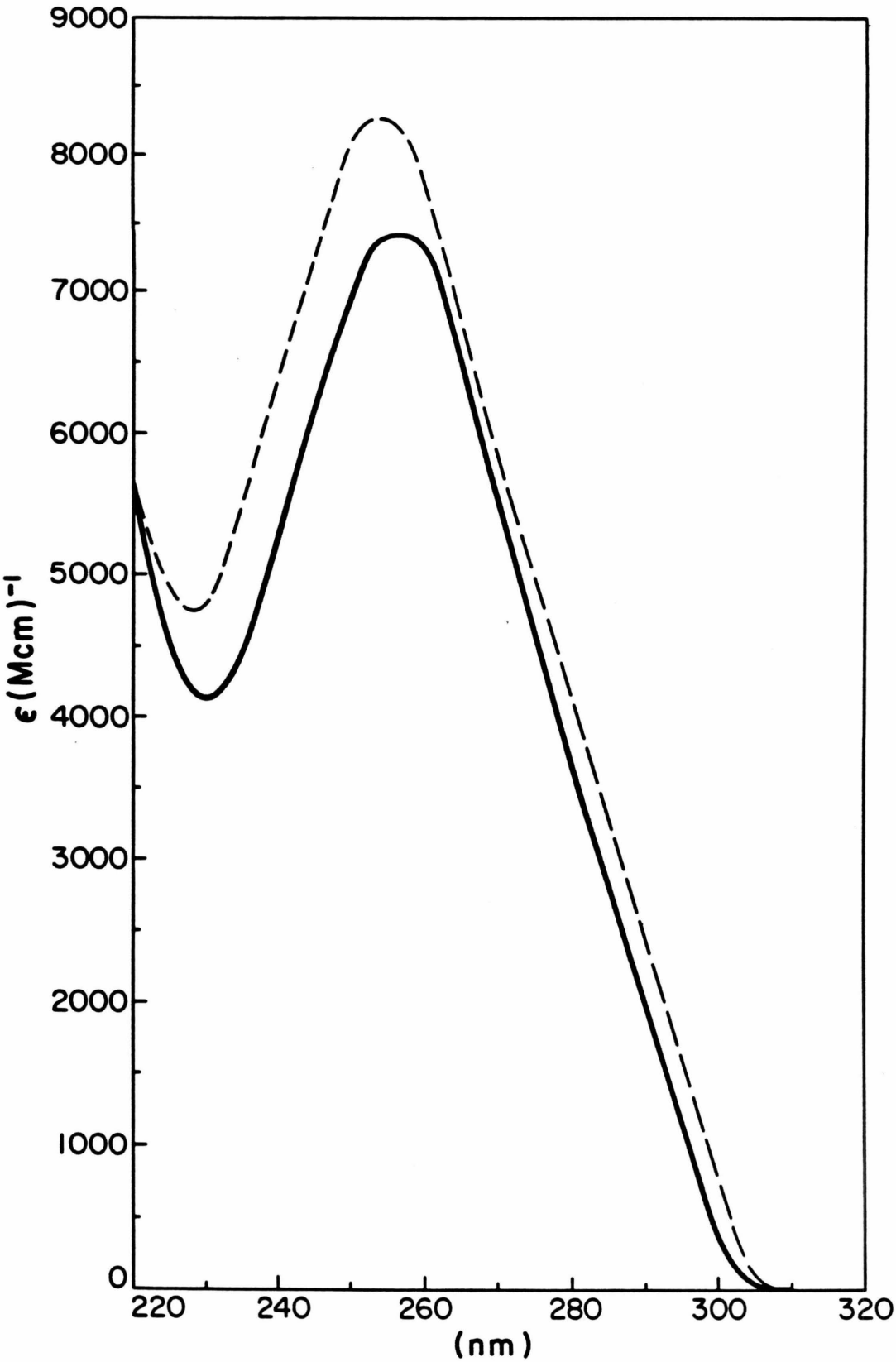


Figure 10
Titration of polydCdG with dG followed by ultraviolet spectroscopy at $[M^+] = 1.0$. The ultraviolet absorption spectrum of polydCdG (solid line) and polydCdG$_2$ (dashed line) are also shown.


9000
8000
7000
6000
5000
4000
3000
2000
1000
0
ε (Mcm)⁻¹
220
240
260
280
300
320
(nm)

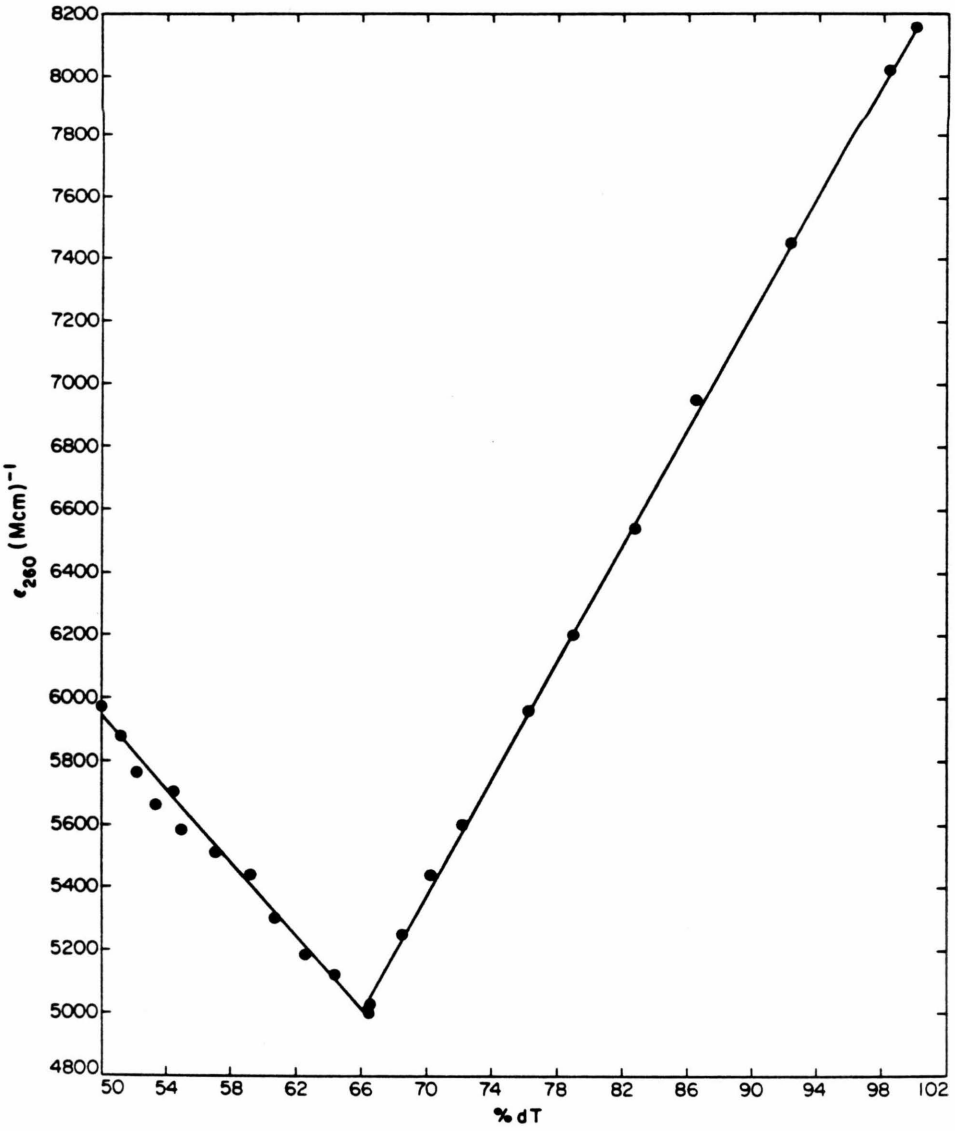


Figure 11
Titration of polydAdT with poly dT followed by ultraviolet spectroscopy at $[M^+] = 1.0$. The ultraviolet absorption spectrum of polydAdT (solid line) and polydAdT_2 (dashed line) are also shown.

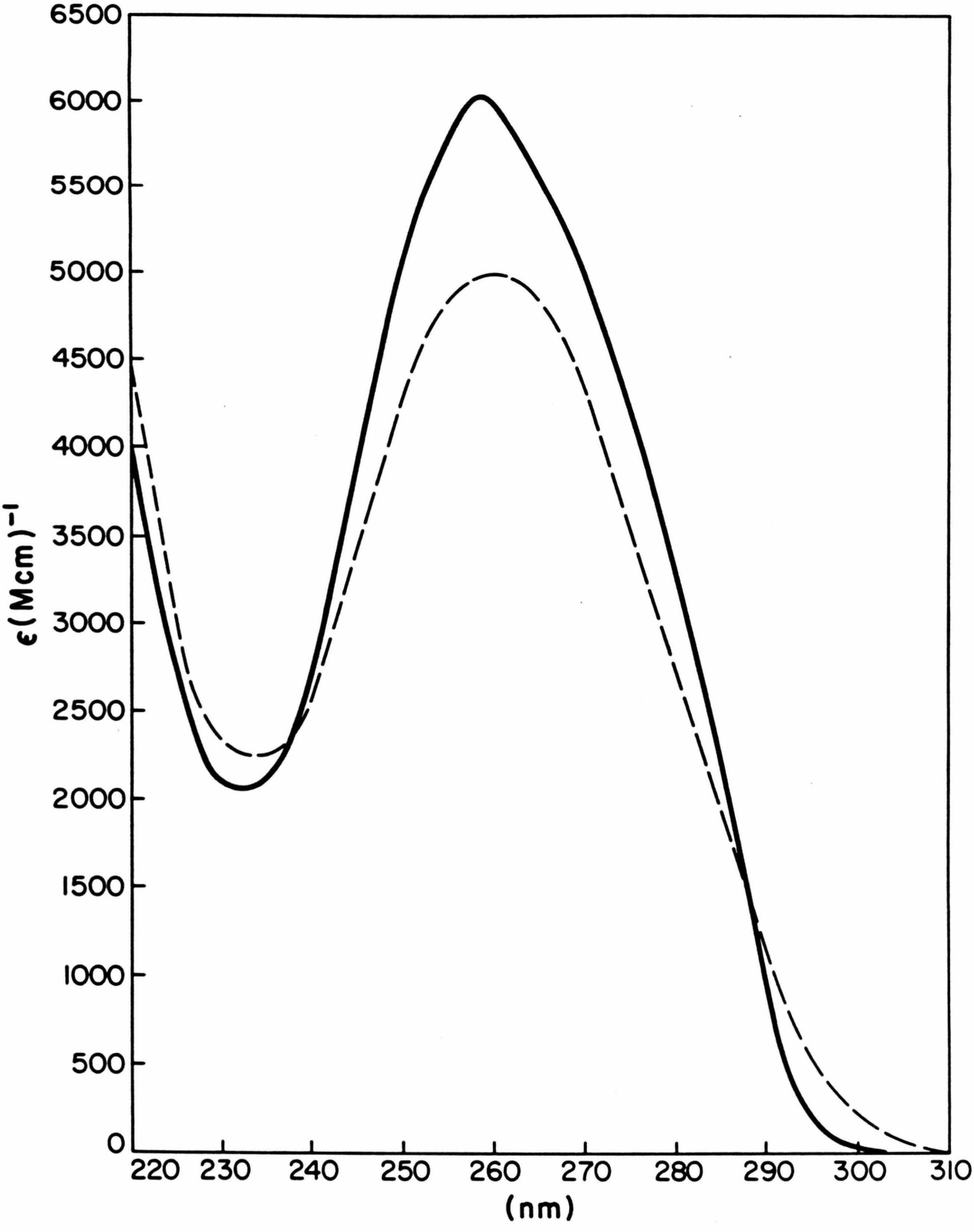
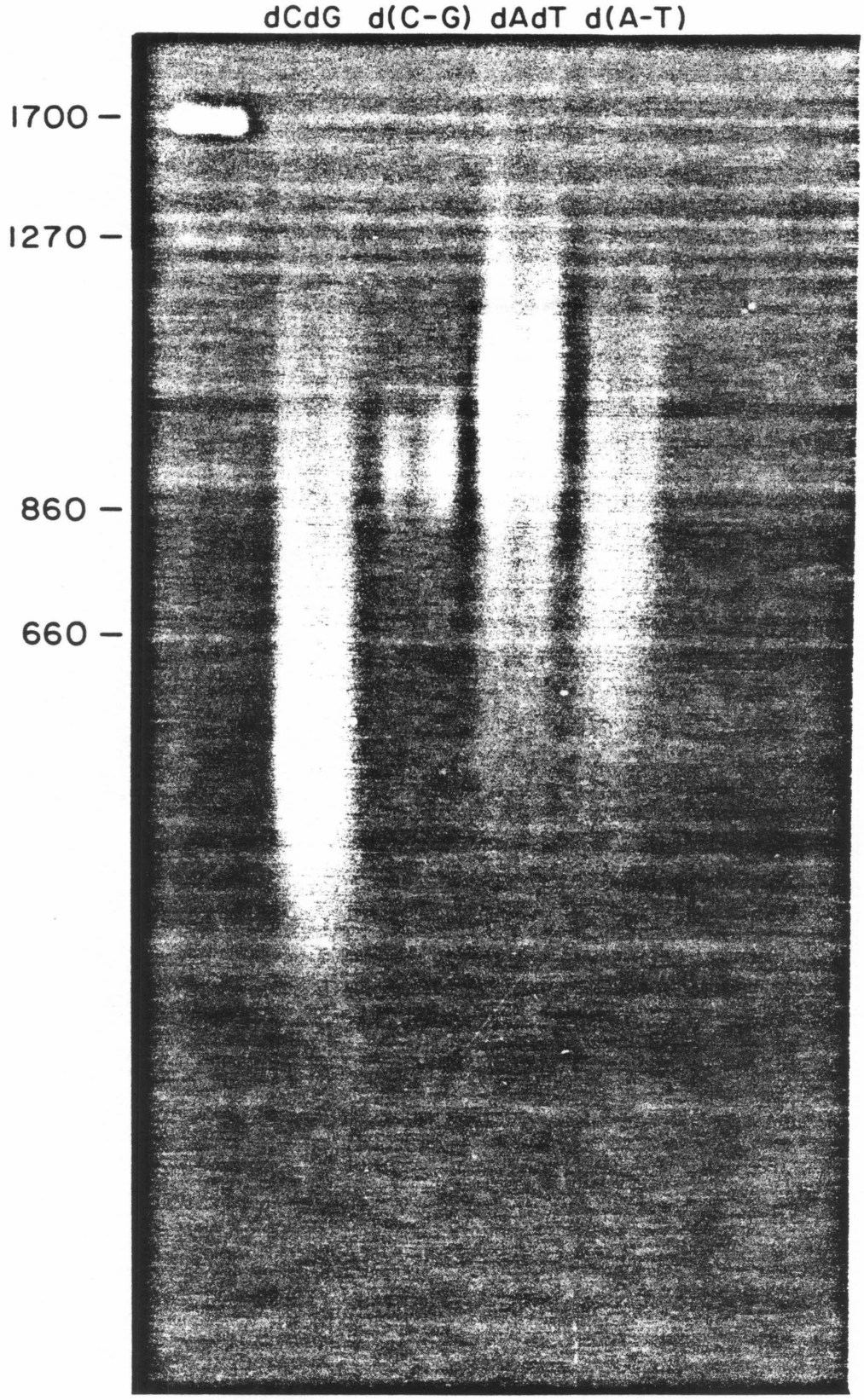
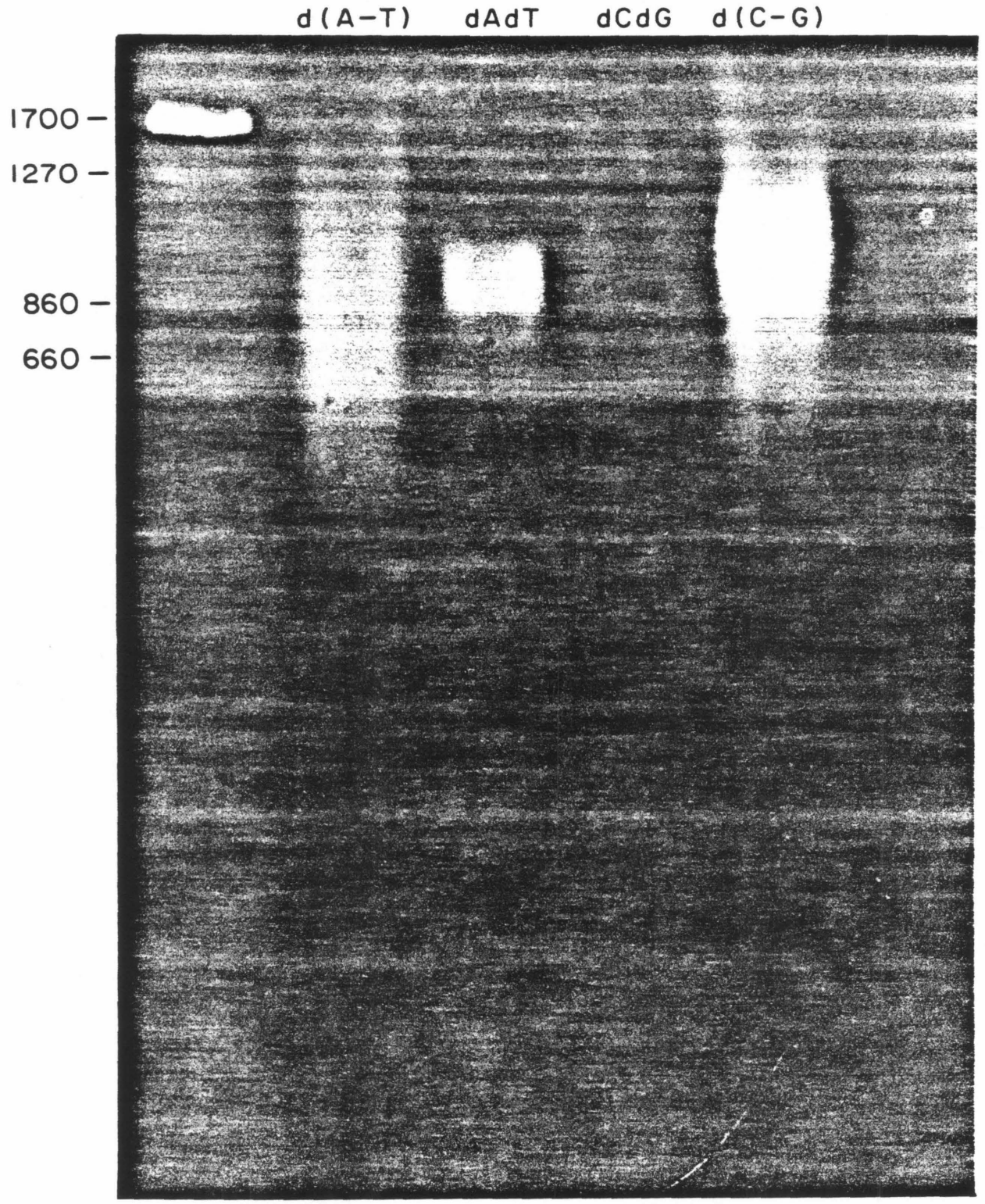

6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
ε(Mcm)⁻¹
220
230
240
250
260
270
280
290
300
310
(nm)

Figure 12

Alkaline gel electrophoresis analysis of polydCdG, polyd(C-G), polydAdT, poly d(A-T), and PM2 DNA restricted by Hae III.


dCdG d(C-G) dAdT d(A-T)
1700
1270
860
660


d(A-T)
dAdT
dCdG
d(C-G)
1700
1270
860
660

lengths differ for each nucleic acid. As shown in Figs. 13-14 and summarized in Table II, each nucleic acid can exist as double helical rods, aggregated rods, or intramolecular hairpins. Although polydAdT and polydCdG form rods and aggregated rods, unlike polyd(C-G), they do not form an intramolecular hairpin since their single strands do not form complementary double helixes.

We have also compared our polyd(C-G) preparation to a sample of polyd(A-T) of similar single strand length to ascertain whether polyd(C-G), like polyd(A-T), forms small "branched" hairpins.[34] As shown in Figs. 13-14 under the conditions of our binding experiments no low molecular weight branched hairpins are formed by polyd(C-G).

The number of base pairs which can undergo a B → H → A allosteric transition, N, depends on both the single and double strand length distribution since gaps created by strand overlap constitute ends of conformational units. For rods and aggregated rods N equals the single-stranded length whereas for intramolecular hairpins N is one-half this length. From the measurements sum-

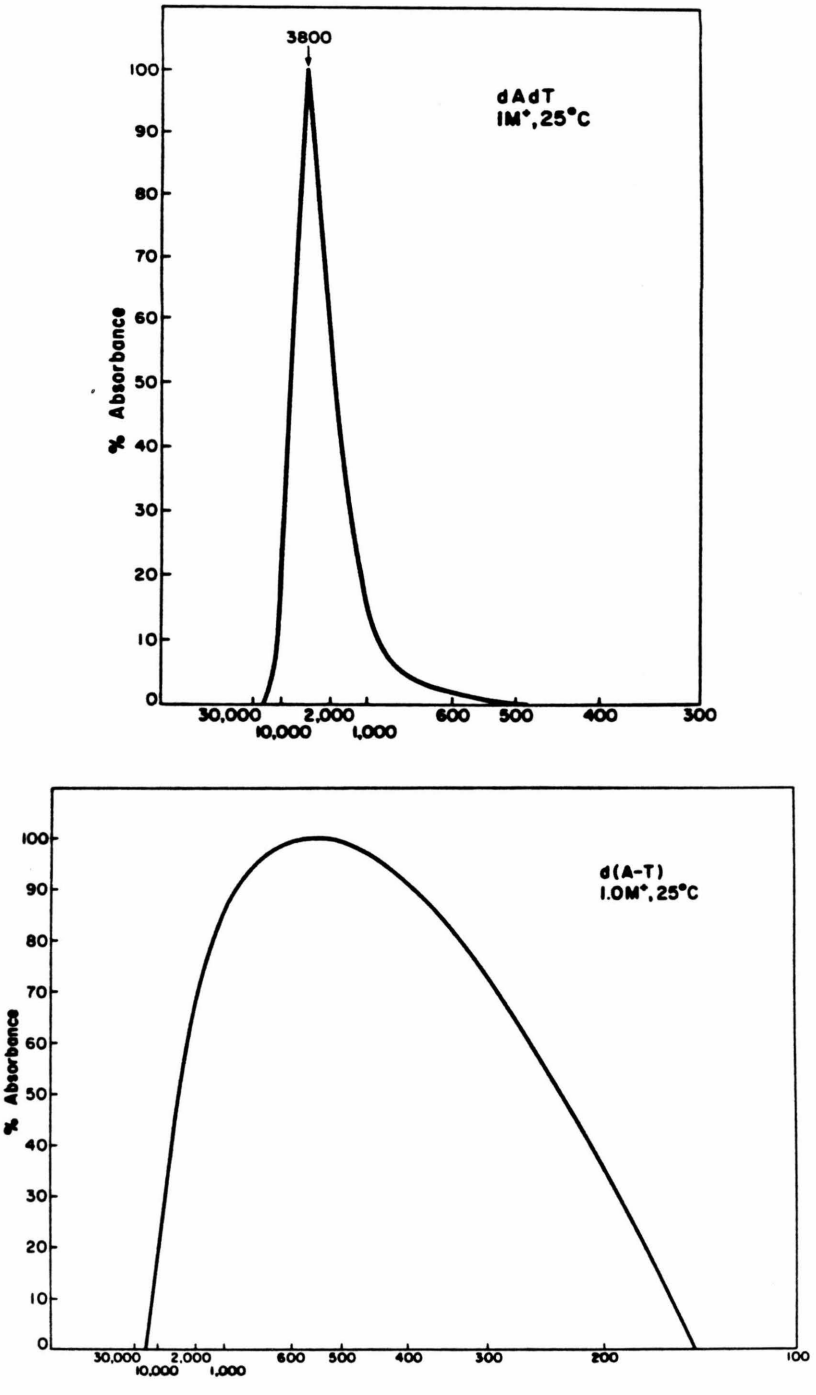
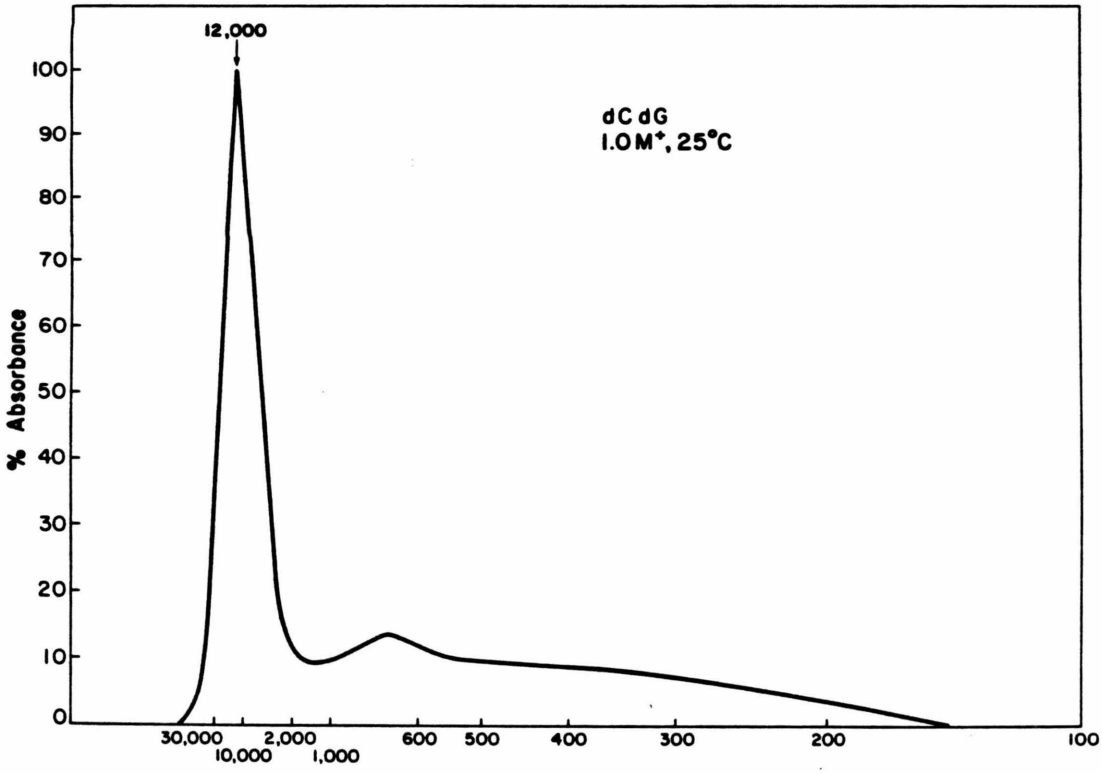


Figure 13
Double stranded molecular weight distribution of polydAdT, polyd(A-T), polydCdG, and polyd(C-G) at $[M^+]$ = 1.0 and 25°C as determined by Sepharose 4B chromatography.


12,000
dC dG
1.0M+, 25°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
2,000
1,000
600
500
400
300
200
100

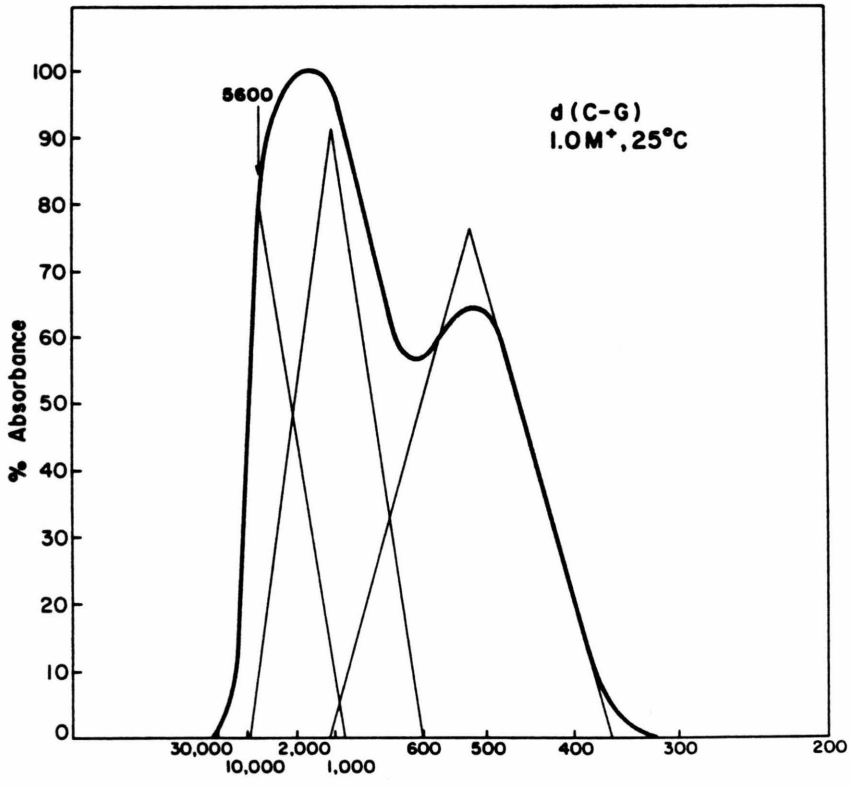
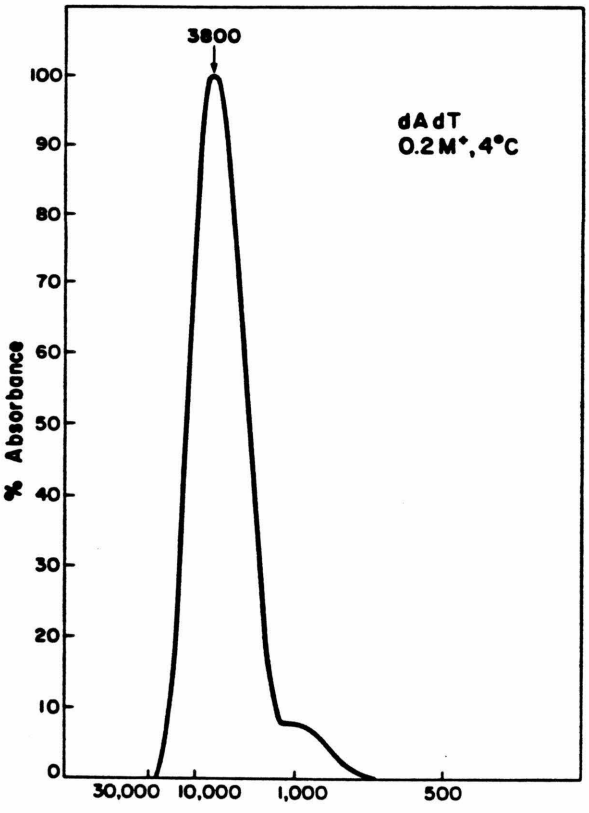
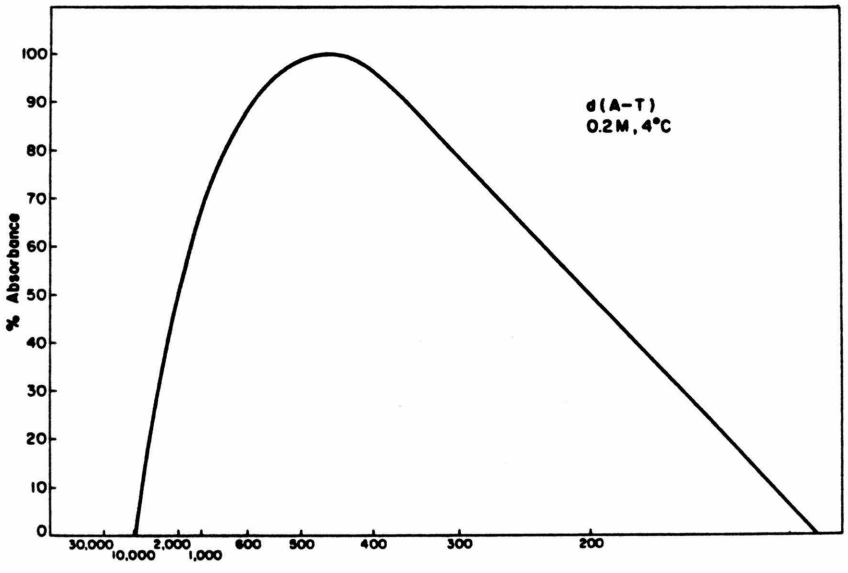
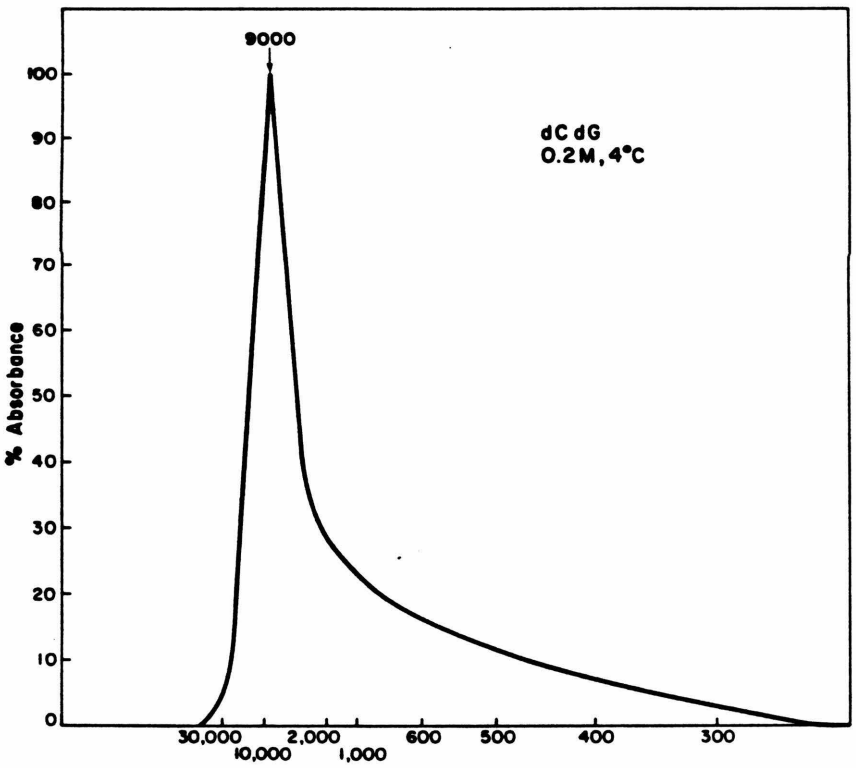

5600
d (C-G)
1.0M+, 25°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
2,000
1,000
600
500
400
300
200

Figure 14

Double stranded molecular weight distribution of polydAdT, polyd(A-T), polydCdG, and polyd(C-G) at $[M^+]$ = 0.2 M and 4^oC as determined by Sepharose 4B chromatography.


3800
dA dT
0.2 M⁺, 4°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
1,000
500


d(A-T)
0.2M, 4°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
2,000
1,000
600
500
400
300
200


9000
dC dG
0.2M, 4°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
2,000
1,000
600
500
400
300

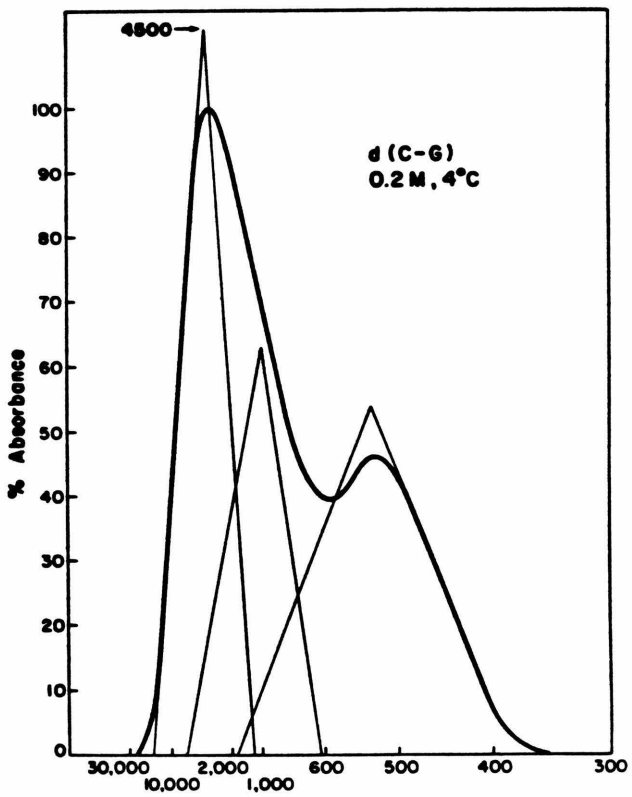

4500
d (C-G)
0.2M, 4°C
% Absorbance
100
90
80
70
60
50
40
30
20
10
0
30,000
10,000
2,000
1,000
600
500
400
300

Table II - Single and double-stranded lengths of synthetic DNA's. The relative distribution of double-stranded forms is summarized below.

			DOUBLE STRAND; 0.2M$^+$, 4°C			DOUBLE STRAND; 1.0m$^+$, 25°C		
NUCLEIC ACID	SINGLE STRAND MAX.	SINGLE STRAND RANGE	AGGREGATE	ROD	HAIRPIN	AGGREGATE	ROD	HAIRPIN
dAdT	950 ±50	1,200 - 600	3,800	1,000	-	3,800	-	-
dCdG	700 ±50	1,100 - 300	9,000	-	-	12,000	700	-
d(C-G)	900 ±100	1,000 - 800	4,500	1,000	500	5,600	1,000	500
d(A-T)	750 ±50	1,100 - 400		450(max)			550(max)	

	0.2M, 4°C			1.0M, 25°C		
NUCLEIC ACID	AGGREGATE	ROD	HAIRPIN	AGGREGATE	ROD	HAIRPIN
dAdT	94%	6%	-	88%	12%	-
dCdG	54%	46%	-	48%	52%	-
d(C-G)	33%	26%	41%	18%	33%	48%

marized in Table II, we conclude that the effective double-stranded lengths of dCdG, dAdT and d(C-G) are similar and equal 700-900 base pairs. Thus the large differences we observe in the ability of these DNA's to undergo B → H → A transitions must arise from differences in base sequence and not molecular weight.

In summary, the binding affinity, binding cooperativity, and binding site size of BMSp and EB to a variety of nucleic acids suggests that these compounds induce sequence specific B → H → A allosteric transitions in DNA by virtue of their enhanced specificity for A and H conformations of nucleic acid.

Possible Involvement of the Hybrid Conformation in Melting

As shown in Fig. 6, when DNA adopts an H conformation vertical stacking interactions and hydrogen bonding between bases are weakened by the alternating propellor twist introduced by a 2'endo(3' → 5')3'endo sugar pucker alternation. Since both forces are known to contribute significantly to the stability of the double helix,[35] an H conformation may also be an intermediate in DNA melting. Loss of structural water,[36] which occurs in a B to A transition,[14,16] may also facilitate melting since its loss is known to destabilize the double helical form of DNA.[12,37]

Because helical stability is a function of several intrinsic as well as extrinsic factors, the reluctance of DNA to undergo a B to H transition may, or may not, inhibit its ability to melt. For example, although our results suggest that alternating DNA undergoes B → H transitions more readily than nonalternating DNA, in 5 to 6 sequences previously examined, the Tm of alternating DNA is 6° higher than nonalternating DNA.[22,38] The one exception to this observation, polyd(A-T), melts 6° lower than the corresponding dAdT homopolymer.[22] Since we have demonstrated that dAdT is very reluctant to undergo a B to H transition, it is possible that its enhanced thermal stability arises from its reluctance to undergo a B to H transition necessary for melting. It should be noted that the unusual instability of d(A-T) compared to dAdT cannot be attributed to hairpin formation in d(A-T) since d(I-C), which forms similar hairpins, is more stable than dIdC.[22]

An additional piece of evidence which implicates a B to H transition in DNA melting comes from the observation of Weischet, et al. that a conformational change whose enthalpy change is zero, preceeds DNA melting.[39] This observation is consistent with the proposal that a B to H transition preceeds melting since the change in

enthalpy for a B to A transition is also zero.[14] Although possible, we have no data concerning the possible involvement of an H conformation in RNA or RNA-DNA hybrid melting.

Relationship of Hybrid DNA to Kinking

We have proposed that the intermediate observed in the B → A transition exhibits a 2'endo(3' → 5')3'endo alternation of sugar pucker on the basis of the proposed B to A mechanism shown in Fig. 4 and the observation that this species excludes intercalation to every other base pair (nearest neighbor exclusion). Sobell first postulated that a 2'endo(3' → 5')3'endo sugar pucker alternation could constrain intercalation to every other base pair and went on further to postulate that this sugar pucker alternation facilitates DNA intercalation and bending by allowing DNA to "kink".[40] If so, H conformations may also be intermediates in DNA kinking.

Hybrid DNA and Previously Proposed Structures

Two groups have previously proposed DNA structures which, like the hybrid conformational family of DNA, apparently exhibit a 2'endo(3' → 5')3'endo sugar pucker alternation. On the basis of x-ray analysis of intercalated

dinucleotide crystals, Sobell et al. postulate a structure for polymer DNA termed "β kinked" DNA which, in addition to possessing a 2'endo(3' → 5')3'endo sugar pucker alternation, is "kinked" every other base pair.[40] Kinking is postulated to arise from partial base unstacking, base pairs forming an angle of about 40° to one another. Klug et al. propose that polyd(A-T) adopts an "alternating B" conformation in which every second phosphate linkage has a conformation different from the B family.[41]

Although it may be possible for DNA to exhibit a "β kinked" or "alternating B" conformation (we do not know), no experimental evidence suggesting that DNA possessing a 2'endo(3' → 5')3'endo sugar pucker alternation is a conformational family, which occurs as an intermediate in a B to A transition and whose equilibrium stability rather than structure varies greatly with base sequence, is presented by these authors. The possibility that "β kinked" or "alternating B" DNA are members of the hybrid conformational family of DNA is not excluded.

BIOLOGICAL IMPLICATIONS

Biological implications of the B → H → A theory are summarized in Propositions 1-5.

MATERIALS AND METHODS

Binding measurements and preparation of DNA's are described in Chapter I. Scatchard plots obtained by the constant X technique are described in Chapter II.

Sepharose 4B Chromatography: Sepharose 4B chromatography was conducted either at 4°C in BPES buffer (0.006 M Na_2HPO_4, 0.002 M NaH_2PO_4, 0.001 M Na_2EDTA, 0.185 M NaCl) or at 25°C in 1[M^+] phosphate buffer (see Chapter I). Approximately 1 O.D. of DNA was dialyzed 24-36 hours and then layered on a column measuring 79 x 2.5 cm (dia). Fractions of 3.4 ml or 1.7 ml were collected at a flow rate of 0.125 ml/min. Increasing the flow rate or the DNA concentration by a factor of two had no effect on the results. Reproducibility was determined to be $\pm$ 0.25 fractions. The elution of linear PM2 (restricted by HPA II) was found to be nearly Gaussian with a half-width of 1.9 fractions in 1 M [M^+] (Fig. 33, Appendix). The dependence of DNA molecular weight on elution volume was determined by chromatography of restricted (Hind III) lambda DNA (Fig. 32, Appendix) and sonicated calf thymus DNA (Fig. 34, Appendix). Molecular weights of sonicated calf thymus DNA were established by electrophoretic comparison to restricted DNA molecular weight standards.

Alkaline Gel Electrophoresis: Electrophoresis was conducted in a 2% agarose slab gel (4 mm x 30 cm) at a constant current of 225 miliampree for 15-26 hours. The gel was prepared in 0.3 M NaCl and 20 mM EDTA; DNA samples were in electrophoresis buffer (0.3 N NaOH, 20 mM EDTA, pH 13) containing 10% glycerol. The buffer was circulated throughout electrophoresis. After electrophoresis the gel was neutralized by soaking in 50 mM Tris acetate - 2.5 mM EDTA, pH 8.3 for four hours, stained with ethidium bromide (4 µg/ml) for one hour, and then destained overnight. The gel was photographically under uv transillumination through a red filter.

CALCULATIONS

Estimation of Allosteric Constants

The following data is used to estimate the values of K_B, K_H, and L for the BMSp-dAdT, EB-d(C-G) and EB-dCdG allosteric transitions:

	EB	BMSp
K_B	2×10^3 dAdT	1.6×10^5 dAdT
K_H	5×10^4 d(A-T)	1.5×10^7 c.thymus
	5×10^4 d(C-G)	2.6×10^8 dCdG
	4.4×10^4 c.thymus	6.5×10^7 d(C-G)
K_A	1×10^5 rArU	2.4×10^8 rAdT
	2×10^5 rAdT	

The corresponding specificities (K_H/K_B) and (K_A/K_B) are given by:

	EB	BMSp
(K_H/K_B)	$5 \times 10^4/2 \times 10^3 = 25$	$6 \times 10^7/1.6 \times 10^5 = 370$
(K_A/K_B)	$1.5 \times 10^5/2 \times 10^3 = 75$	$2.4 \times 10^8/1.6 \times 10^5 = 1500$

A) PolydCdG + EB

The allosteric constant, B/H, can be estimated from the initial slope and initial intercept of the Scatchard plot:[10]

$$(1) \qquad \text{slope} = \frac{1+C^2L^2 - (3C^2+3-8C)}{(1+CL)(1+L)} (K_H)$$

$$(2) \qquad \text{intercept} = K_H\left(\frac{1}{1+L}\right) + K_B\left(\frac{1}{1+L}\right)$$

The experimentally observed slope and intercept are:

$$\text{slope} = 7 \times 10^4 \ (1\text{x}10^4\text{-}1\text{x}10^5)$$

$$\text{intercept} = 8 \pm 3 \times 10^3$$

Typical results, compatible with this intercept and a slope equal to 7×10^4 are shown below.

K_B =	5×10^3	$4\text{x}10^3$	4×10^3	6×10^3	$7\text{x}10^3$	$5\text{x}10^5$
K_H =	$1.5\text{x}10^5$	$2\text{x}10^5$	$1.2\text{x}10^5$	$1.8\text{x}10^5$	$2\text{x}10^5$	$1\text{x}10^5$
K_H/K_B =	30	50	30	30	30	20
L =	150	300	100	150	200	100

Typical results, compatible with the observed intercept and a slope equal to 1×10^4 are shown below:

K_B =	$5\text{x}10^3$	5×10^3	$5\text{x}10^3$	$1\text{x}10^4$	$5\text{x}10^3$
K_H =	$5\text{x}10^4$	$1.5\text{x}10^5$	$8\text{x}10^4$	$1\text{x}10^5$	$2\text{x}10^5$
K_H/K_B =	10	30	15	10	40
L =	70	700	300	100	1500

Typical results, compatible with the observed intercept and a slope equal to 1×10^5 are shown below:

K_B =	5×10^3	$4\text{x}10^3$	$1\text{x}10^4$
K_H =	$1.5\text{x}10^5$	$2\text{x}10^5$	$2\text{x}10^5$
K_H/K_B =	30	50	20
L =	90	200	100

In summary, the allosteric constant, (B/H), for polydCdG is:

$$(B/H) = 70 \leftrightarrow 1500 \text{ for } (C)^{-1} = 10\text{-}40$$

with the best estimate being:

$$(B/H) = 200 \text{ for } (C)^{-1} = 30$$

B) Polyd(C-G) + EB

The intercept and slope at $\nu \to 0$ are:

$$\text{slope} = 3 \times 10^5 \ (6\times10^4\text{-}6\times10^5)$$

$$\text{intercept} \leq 9 \times 10^3$$

Typical results, compatible with this intercept and a slope equal to 6-3 x 10^5 are shown below:

K_B	= 3.3 x 10^3	4 x 10^3
K_H	= 1 x 10^5	2 x 10^3
K_H/K_B	= 30	50
L	= 20	50

Typical results, compatible with the observed intercept and a slope equal to 6 x 10^4 are shown below:

K_B	=	1.7×10^3	5×10^3	3.3×10^3	5×10^3	2×10^3	6.7×10^3
K_H	=	5×10^4	2×10^5	1×10^5	5×10^4	1×10^5	2×10^5
K_H/K_B	=	30	40	30	10	50	30
L	=	30	300	100	10	200	200

In summary, the allosteric constant, (B/H), for polyd(C-G) is:

$$(B/H) = 10\text{-}300 \text{ for } (C)^{-1} = 10\text{-}40$$

with the best estimate being

$$(B/H) = 50 \text{ for } (C)^{-1} = 30$$

C) PolydAdT + BMSp

Since the polydAdT-BMSp Scatchard plot exhibits both a maximum and a minimum, the allosteric constant, (B/H), and ligand specificity constant ($C = K_B/K_H$) are related by the following relationship:[10]

$$(3) \qquad \frac{[(n-1)L]^{1/2}-1}{[(n-1)L]^{1/2}+L} < C < \frac{(n-1)^{1/2}-1}{(n-2)^{1/2}L^{1/4}}$$

where n is the number of subunits which change conformation in the allosteric transition. Assuming c = 1/1500 and n = 4 leads to the following results:

$\frac{[(n-1)L]^{1/2}-1}{[(n-1)L]^{1/2}+L}$	< 0.007 <	$\frac{(n-1)^{1/2}-1}{(n-2)^{1/2}L^{1/4}}$
0.002 ($L=10^6$)	< 0.007 <	0.0025 ($L=10^{10}$)
0.0005 ($L=10^7$)		0.0014 ($L=10^{11}$)
0.0002 ($L=10^8$)		0.0008 ($L=10^{12}$)

Similarly, assuming C = 1/1500 and n = 20 leads to the following results:

$\frac{[(n-1)]^{1/2}-1}{[(n-1)L]^{1/2}+L}$	< 0.007 <	$\frac{(n-1)^{1/2}-1}{(n-2)^{1/2}L^{1/4}}$
0.004 ($L=10^6$)		0.0014 ($L=10^{11}$)
0.001 ($L=10^7$)		0.0008 ($L=10^{12}$)
0.0004 ($L=10^8$)		0.0004 ($L=10^{13}$)

In summary, assuming C = 1/1500 and n = 4-20 the B/H allosteric constant for polydAdT is $1 \times 10^7 - 1 \times 10^{12}$. It should be noted that this estimate is relatively insensitive to the value of C and n .

The value of the (B/H) allosteric constant for polydAdT can also be estimated from the initial slope of the Scatchard plot.[10] Since L is large and C small, the maximum value which the slope can assume is $-(K_B)$ (see later derivation). However, both indirect measurements (Fig. 1, Chapter 3) and constant X measurements (Fig. 15, Chapter 2) yield a value of (-4×10^6) for the initial slope. Since the smallest value for the slope is approximately $-(1 \times 10^6)$ (Fig. ___, Chapter 2), the experimentally observed slope is approximately six times too high. This discrepancy can be attributed to the exclusion of potential binding sites by ligand.[53] For example, since initial BMSp binding is the B conformation of polydAdT, the maximum value of the initial slope is equal to $-K(2n-1)$, where n is the number of base pairs covered by one bound BMSp.[53] Clearly, when n = 4, no discrepancy between theoretical and experimental slopes exists.

A minimum estimate of the (B/H) allosteric constant can be obtained from the initial slope by noting that since this slope is negative, the quantity $\{-1 + C^2L^2 - (3C^2 + 3 - 8C)L\}$ must also be negative (see equation 1). Since the value of this quantity changes from positive to negative at $L = 10^7$ (for $C^{-1} = 1500$), a minimum estimate of L is 1×10^7, in good agreement with estimates from

Scatchard plot extrema (see above).

D) Proof that the maximum value of the initial slope is $-(K_B)$ for large L and small c.

From equation (1) we can write:

$$(4) \quad (K_H)^{-1} \text{ slope} = - \frac{1+C^2L^2-(3C^2+3-8C)L}{(1+CL)(1+L)}$$

Since the initial slope is negative, and the dAdT Scatchard plot exhibits both a maximum and a minimum (e.g. $C < 1$ and $L >> 1$), we can rewrite equation (1) as:

$$(5) \quad (K_H)^{-1} \text{ slope} \simeq \frac{1 + C^2L^2 - 3L}{(1+CL)L}$$

Dividing numerator and denominator by L yields:

$$(6) \quad (K_H)^{-1} \text{ slope} \simeq \frac{L^{-1} + C^2L - 3}{(1+CL)}$$

Since $L >> 1$ and $C < 1$ we can rewrite equation (5) as:

$$(7) \quad (K_H)^{-1} \text{ slope} \simeq C - 3/CL$$

Thus the maximum value which the quantity $(K_H)^{-1}$ slope can assume is C, or equivalently, the maximum value of the slope is K_B.

REFERENCES

1) Dervan, P.B.; Becker, M.M., J. Amer. Chem. Soc., 1978, 100, 1968.

2) Becker, M.M.; Dervan, P.B., J. Amer. Chem. Soc., 1979, 101, 3664.

3) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

4) Bresloff, J.L.; Crothers, D.M., unpublished work.

5) LePecq, J.-B.; Paoletti, C., J. Mol. Biol., 1967, 27, 87.

6) For example, see (a) Tsai, C.-C.; Jain, S.C.; Sobell, H.M., J. Mol. Biol., 1977, 114, 301. (b) Jain, S.C.; Tsai, C.-C.; Sobell, H.M., J. Mol. Biol., 1977, 114, 317.

7) Milman, G.; Langridge, R.; Chamberlin, M.J., Proc. Natl. Acad. Sci. USA, 1967, 57, 1804.

8) For example, see Arnoh, S., Progr. Biophys. Mol. Biol., 1970, 21, 265.

9) Bauer, W.; Vinograd, J., J. Mol. Biol., 1970, 47, 419.

10) Gibson, R.E.; Levin, S.A., Proc. Natl. Acad. Sci. USA, 1977, 74, 139.

11) Monod, J.; Wyman, J.; Changeux, J.-P., J. Mol. Biol., 1965, 12, 88.

12) Pilet, J.; Blicharski, J.; Brahm, J., Biochemistry, 1975, 14, 1869.

13) a) Zimmer, C.H.; Luck, G., Biochim. Biophys. Acta, 1974, 361, 11. b) Arnott, S.; Selsing, E., J. Mol. Biol., 1974, 88, 509.

14) Ivanov, V.I.; Minchenkova, L.E.; Minyat, E.E.; Frank-Kamenetskii, M.D.; Schyolkina, A.K., J. Mol. Biol., 1974, 87, 817.

15) Bresloff, J.L., Ph.D. Dissertation, Yale University, 1974.

16) a) Franklin, R.E.; Gosling, R.G., Acta. Cryst., 1953, 6, 673. b) Minchenkova, L.; Minyat, E.; Schyolkina, A.; Ivanov, V., FEBS Lett., 1975, 51, 38. c) Ivanov, V.I., Minchenkova, L.E.; Schyolkina, A.K.; Poletayev, A.I., Biopolymers, 1973, 12, 89. d) Zimmerman, S.B.; Pheiffer, B.H., J. Mol. Biol., 1979, 135, 1023.

17) For example, see Arnott, S. in "Organization and Expression of Chromosomes," Allfrey, V.G.; Bautz,E.K.F., McCarthy, B.J.; Schimke, R.T.; Tissieres, A., eds., Dahlem Konferenzen, Berlin, 1976, 209.

18) a) Wang, J.C., Proc. Natl. Acad. Sci. USA, 1979, 76, 2000. b) Leslie, A.G.W.; Arnott, S.; Chandrasekaran, R.; Ratliffe, R.L., J. Mol. Biol., 1980, 143, 49. c) For example, see Baase, W.A.; Johnson, W.C., Nucl. Acid Res., 1979, 6, 797.

19) Sundaralingan, M., Biopolymers, 1969, 7, 821.

20) Patel, D.J.; Kozolowski, S.A.; Suggs, J.W.; Cox, S.D., Proc. Natl. Acad. Sci. USA, 1981, in press.

21) Vollenweider, H.J.; James, A.; Szybalski, W. Proc. Natl. Acad. Sci. USA, 1978, 75, 710.

22) For example, see Wells, R.D.; Larson, J.E.; Grant, R.C.; Shontle, B.E.; Cantor, C.R., J. Mol. Biol., 1970, 54, 465.

23) a) Gratzer, W.B.; Hill, L.R.; Owen, R.J., Eur. J. Biochem., 1970, 15, 209. b) Bram, S., Nature New Biology, 1971, 232, 174.

24) Wells, R.D.; Blakesley, R.W.; Hardies, S.C.; Horn, G.T.; Larson, J.E.; Selsing, E.; Burd, J.F.; Chan, H.W.,; Dodgson, J.B.; Jensen, K.F.; Nes, I.F.; Wartel, R.M., CRC Crit. Rev. Biochem. 1977, 304.

25) Dickerson, R.E.; Drew, H.R., J. Mol. Biol., 1981, submitted.

26) Millar, D.P.; Robbins, R.J.; Zewail, A.H., Proc. Natl. Acad. Sci. USA, 1980, 77, 5593.

27) For example see Pohl, F.M.; Jovin, T.M., J. Mol. Biol., 1972, 67, 375.

28) Riley, M.; Maling, B.; Chamberlin, M.J., J. Mol. Biol., 1966, 20, 359.

29) a) Inman, R.B.; Baldwin, R.L., J. Mol. Biol., 1964, 8, 452. b) Patel, D., "Nucleic Acid Geometry and Dynamics," Sarma, R.H., ed., Pergamon, New York, 1980, 185.

30) Aktipis, S.; Martz, W.W., Biochemistry, 1974, 13, 112.

31) Pohl, F.M.; Jovin, T.M.; Baehr, W.; Holbrook, J.J. Proc. Natl. Acad. Sci. USA, 1972, 69, 3805.

32) a) Marck, C.; Thiele, D., Nucl. Acid Res., 1978, 5, 1017. b) Arnott, S.; Selsing, E., J. Mol. Biol., 1974, 88, 509.

33) Lehrman, E.A.; Crothers, D.M., Nucl. Acid. Res., 1977, 4, 1381.

34) Spatz, H.; Baldwin, R.L., J. Mol. Biol., 1965, 11, 213.

35) For example, see Crothers, D.M.; Zimm, B.H., J. Mol. Biol., 1964, 9, 1.

36) Water lost in a B → A transition appears to be structural since $\Delta H = 0$ for the transition.[14]

37) a) Falk, M.; Hartman, K.A.; Lord, R.C., J. Amer. Chem. Soc., 1963, 85, 391. b) Ts'o, P.O.P.; Helmkamp, G., Tetrahedron, 1961, 13, 198.

38) Oliver, A.L.; Wartell, R.M.; Ratliff, R.L., Biopolymers, 1977, 16, 1115.

39) Weischet, W.O.; Tatchell, K.; Van Halde, K.E.; Klump, H., Nucl. Acid Res., 1978, 5, 139.

40) Sobell, H.M.; Tsai, C.-C.; Jain, S.C.; Gilbert, S.G., J. Mol. Biol., 1977, 114, 333.

41) Klug, A.; Jack, A.; Viswamitra, M.A.; Kennard, O.; Shakked, Z.; Steitz, T.A., J. Mol. Biol., 1979, 131, 669.

42) Raw absorbance measurements are shown in Figs. 15-19 and Table V of the Appendix.

43) Raw absorbance measurements are shown in Figs. 20-27 and Table V of the Appendix.

44) Raw absorbance measurements are shown in Figs. 6-14 and Tables III and IV of the Appendix.

45) Viswamitra, M.A.; Kennard, O.; Shakked, Z.; Jones, P.G.; Sheldrick, G.M.; Salisbury, S.; Falvello, L., Nature, 1973, 273, 687.

46) Hill, A.V., J. Physiol. (London), 1910, 40, iv.

47) Rubin, M.M.; Changeux, J.-P., J. Mol. Biol., 1966, 21, 265.

48) Minyat, E.E.; Ivanov, V.I.; Kritzyn, A.M.; Minchenkova, L.E.; Schyolkina, A.K., J. Mol. Biol., 1978, 128, 397.

49) Pohl, F.M.; Jovin, T.M., Baehr, W.; Holbrook, J.J., Proc. Natl. Acad. Sci. USA, 1972, 69, 3805.

50) See Figure 15, Chapter 1.

51) See Table I, Appendix.

52) For example, see propositions 1-5.

53) McGhee, J.D.; von Hippel, P.H., J. Mol. Biol., 1974, 86, 469.

APPENDIX

Table I

Table I - Summary of the visible extinction coefficients (M^{-1} cm^{-1}) of BMSp and EB bound and unbound to nucleic acid. For BMSp bound to C.T. (sonicated calf thymus), polyd(C-G), polydCdG, polydAdT and polyrAdT the monovalent cation concentration was 1.0 M.

	BMSp									EB	
(nm)	H_2O	0.075 M^+	1.0 M^+	4.4 M^+	C.T. (0.075)	C.T.	(dC-dG)	dCdG	dAdT	rAdT	1 M^+	C.T. 1 M^+
640												10
630												18
620		94	138	239	174	257	259	182		109		24
610		307	269	372	336	432	388	243	800	166		26
600	630	532	496	565	683	743	656	471	1290	412	0	58
590	1000	970	885	894	1239	1312	1132	1057	2000	1025	0	142
580	1600	1569	1519	1415	2340	2359	2010	1956	2750	1994	27	329
570	2431	2535	2356	2166	3649	3676	3254	3170	4100	3436	118	729
560	3540	3723	3469	3051	5074	5302	4611	4807	5550	5243	337	1553
550	4820	5081	4707	4330	6360	6683	5980	6479	6900	7029	691	2489
540	6300	6602	6053	5574	7020	7503	6909	7547	7480	8155	1310	3382
530	7730	7948	7285	6732	7124	7797	7235	7933	7860	8597	2074	3874
520	8944	9011	8206	7643	6846	7573	7130	7912	7930	8360	3120	4000
510	9813	9731	8780	8244	6244	7113	6657	7465	7400	7865	4139	3908
500	10160	9825	8852	8372	5421	6266	5839	6648	6750	6929	5095	3600
490	9921	9399	8355	7990	4437	5227	4867	5527	5740	5716	5768	3103
480	9075	8492	7464	7225	3464	4273	3848	4557	4700	4653	6050	2534
470	7924	7197	6322	6200	2653	3374	3009	3652	3500	3538	5931	2020
460	6578	5739	5090	4963	1981	2583	2273	2693	2615	2611	5377	1505
450	5275	4474	3906	3888	1367	1963	1745	1977	1730	1848	4576	1100
440	4125	3411	2937	2971	1019	1514	1383	1455	1090	1341	3630	786
430	3213	2547	2159	2247	753	1286	1166	1211	740	976	2760	563
420	2360	1877	1621	1710	626	1207	1090	962	530	764	1938	411
410	1820	1458	1238	1400	660	1221	1189	1008	610	633	1300	343
400	1485	1252	1035	1201	822	1315	1354	1226	680	779	792	346
390	1350	1264	1005	1194	1193	1822	1711	1710	1060	1325	480	462
380	1500	1489	1286	1374	1981	2575	2239	2395	1830	2357	291	775
370	2100	1940	1980	2062	3012	3704	3119	3337	2950	4122	246	1334

Table II - Summary of selected optical properties of synthetic nucleic acids utilized in binding studies.

λ(nm)/λ(nm)	dAdT	rAdT	$dAdT_2$	(dA-dT)	dCdG	$dCdG_2$	(dC-dG)
290/260	0.164	0.195	0.225	0.116	0.275	0.318	0.284
280/260	0.544	0.547	0.55	0.517	0.498	0.532	0.519
270/260	0.826	0.830	0.857	0.873	0.770	0.750	0.776
250/260	0.868	0.821	0.861	0.758	0.975	1.068	0.962
240/260	0.453	0.411	0.511	0.409	0.744	0.828	0.707
230/260	0.349	0.358	0.468	0.339	0.577	0.621	0.516
220/260	0.671	0.790	0.895	0.746	0.789	0.732	0.556
ε_{max}	6000	6900	5000	6600	7400	8250	7100
λ_{max}	258	258	260	262	254	253	256
λ_{min}	232	233	234	233	229	228	227

Table III - Summary of the binding of EB to polyd(C-G) at 1 M^{+} followed by visible spectroscopy. The cell path length was 1.0 cm and the concentration of nucleic acid is in base pairs (BP).

[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EP/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
0	0	0	0	0	0	0	0	78.76	8.7999	0.89503	0.3205	56.52	2.8773	1.9643	0.2725
1.90^{-6}	9.1312^{-5}	0.02081	0.006	1.367^{-6}	2.9939^{-5}	0.04564	0.005	82.28	8.7845	0.93669	0.338	59.04	2.8714	2.0562	0.285
3.798	9.1229	0.04163	0.012	2.733	2.9909	0.09137	0.0105	85.79	8.7703	0.9782	0.356	61.552	2.8666	2.1472	0.2975
5.691	9.1146	0.06246	0.018	4.0954	2.9885	0.13701	0.0165	0	0	0	0	0	0	0	0
7.582	9.1063	0.08326	0.023	5.4552	2.985	0.1827	0.0215	6.346^{-7}	1.48^{-5}	0.04288	0.003	3.218^{-7}	7.0252^{-6}	0.04581	0.001
9.46	9.0992	0.10396	0.0285	6.8125	2.983	0.2284	0.027	12.68	1.479	0.08571	0.005	6.43	7.0181	0.09162	0.002
11.352	9.0909	0.12484	0.0345	8.167	2.979	0.2742	0.032	18.99	1.477	0.1286	0.007	9.635	7.011	0.1374	0.004
13.233	9.0826	0.14572	0.0415	9.519	2.977	0.3198	0.0375	31.60	1.474	0.2144	0.012	16.02	6.998	0.22897	0.006
16.983	9.0578	0.18747	0.053	12.2144	2.9708	0.4112	0.0495	44.16	1.4706	0.3003	0.018	22.41	6.984	0.3209	0.009
20.72	9.0507	0.22897	0.0655	14.901	2.9648	0.5026	0.0615	56.66	1.4682	0.3859	0.024	28.76	6.9696	0.4126	0.012
23.514	9.0341	0.2602	0.075	16.901	2.9625	0.5705	0.070	69.12	1.4859	0.4716	0.030	35.08	6.955	0.5044	0.016
26.30	9.0259	0.2914	0.0855	18.911	2.957	0.6397	0.0795	81.54	1.4623	0.5576	0.035	41.38	6.9424	0.59606	0.019
29.093	9.0105	0.3229	0.0955	20.898	2.954	0.7074	0.0885	93.90	1.4599	0.6431	0.042	47.66	6.929	0.6877	0.023
31.852	9.0022	0.3538	0.1065	22.898	2.948	0.7767	0.0985	106.21	1.456	0.7293	0.047	53.90	6.9152	0.7795	0.025
35.54	8.9868	0.3954	0.121	25.54	2.9424	0.8681	0.1125	118.48	1.454	0.8148	0.053	60.13	6.9022	0.8712	0.029
39.21	8.9703	0.43707	0.1355	28.18	2.9376	0.9593	0.1245	130.7	1.452	0.90026	0.060	66.265	6.8892	0.9619	0.031
42.87	8.9549	0.4787	0.1505	30.804	2.9317	1.0506	0.1375	142.9	1.448	0.9866	0.065	72.524	6.875	1.0549	0.025
46.51	8.9395	0.52033	0.1665	33.42	2.926	1.1419	0.1505	155.0	1.446	1.0720	0.071	78.672	6.8691	1.1453	0.038
50.15	8.9229	0.562	0.1825	36.03	2.921	1.2332	0.1635	167.1	1.443	1.1575	0.077	84.781	6.856	1.2366	0.041
53.77	8.9075	0.6036	0.1985	38.62	2.9152	1.3245	0.177	179.1	1.4398	1.2442	0.083	90.817	6.835	1.3287	0.044
57.37	8.8922	0.6453	0.214	41.207	2.9092	1.4166	0.1895	191.08	1.4375	1.3293	0.090	96.984	6.8217	1.4217	0.047
60.97	8.8768	0.6868	0.2305	43.78	2.9045	1.5071	0.2035	203	1.4351	1.4147	0.095	103.034	6.8099	1.513	0.050
64.55	8.8602	0.7286	0.2485	46.35	2.8986	1.5992	0.2165	214.9	1.4315	1.5011	0.102	109.043	6.7969	1.6043	0.053
68.122	8.846	0.770096	0.2655	48.91	2.8939	1.6896	0.2305	226.8	1.4292	1.5868	0.108	115.067	6.7898	1.6947	0.056
71.68	8.8306	0.81168	0.2845	51.46	2.8879	1.78176	0.243	238.6	1.4256	1.6735	0.114	121.065	6.7709	1.78802	0.059
75.23	8.8153	0.85335	0.302	53.99	2.882	1.87305	0.2565	250.3	1.4233	1.7589	0.120	126.909	6.7638	1.8763	0.063
								262.05	1.4209	1.8443	0.127	132.95	6.7508	1.9694	0.066

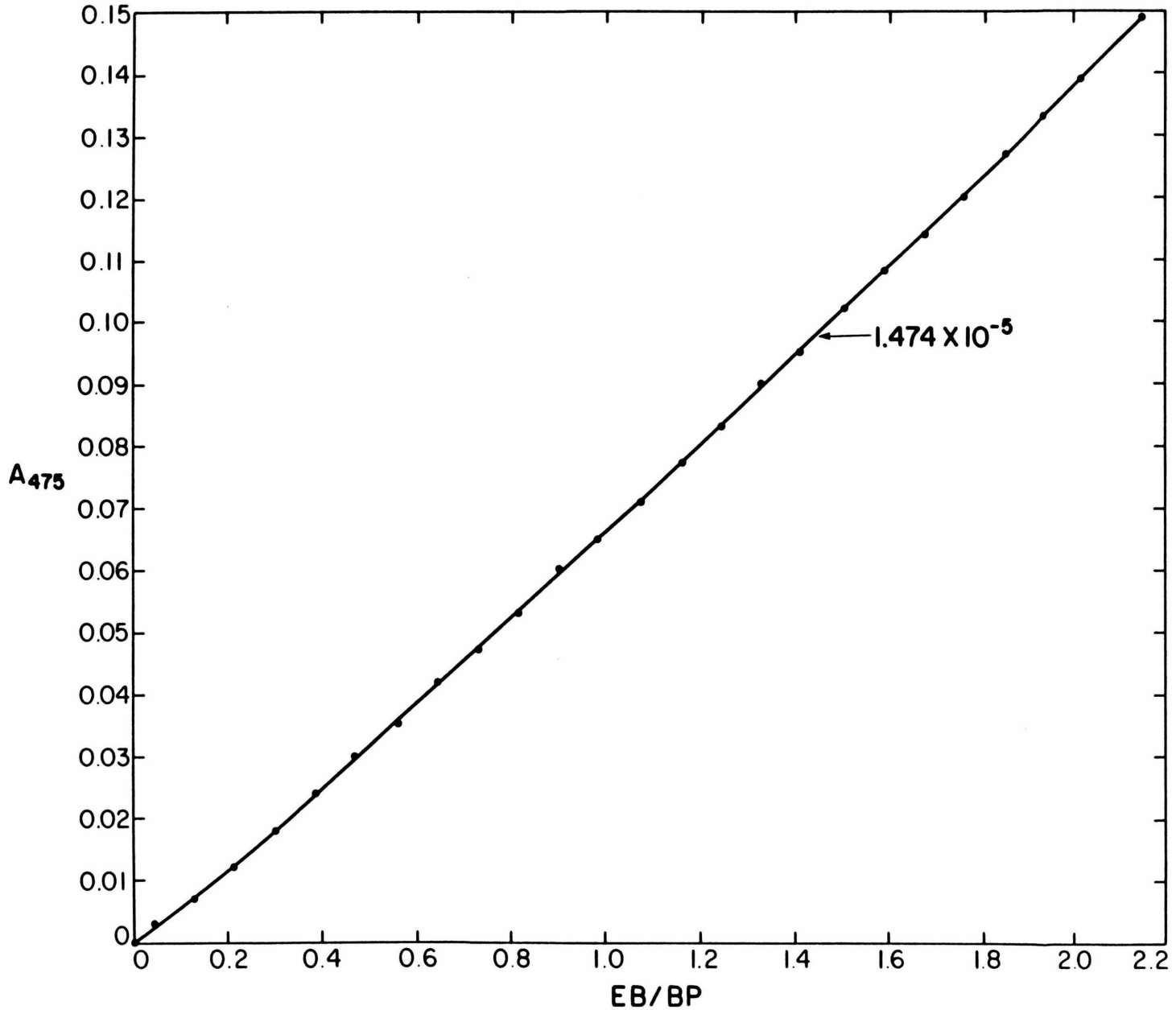
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
22.54	4.058	5.555	1.144	21.15	8.456	25.012	1.124
23.87	4.0462	5.899	1.212	22.75	8.432	26.982	1.205
25.25	4.0391	6.2514	1.282	24.35	8.407	28.9621	1.286
26.78	4.0273	6.6478	1.358	25.94	8.383	30.9408	1.366
28.29	4.0178	7.0417	1.435	27.41	8.36	32.7868	1.443
29.77	4.006	7.4305	1.514	28.863	8.336	34.623	1.514
0	0	0	0	0	0	0	0
3.712^{-5}	7.639^{-5}	0.4859	0.129	2.818^{-5}	2.569^{-5}	1.0971	0.128
7.331	7.591	0.9657	0.302	5.6086	2.557	2.1934	0.274
7.847	7.584	1.0347	0.328	6.1636	2.554	2.413	0.306
8.616	7.571	1.1381	0.369	6.991	2.548	2.743	0.351
9.391	7.564	1.242	0.410	7.822	2.546	3.0724	0.395
10.156	7.553	0.3446	0.451	8.6518	2.542	3.403	0.440
10.927	7.543	1.4486	0.492	9.471	2.539	3.73	0.486
11.946	7.529	1.5866	0.547	10.287	2.535	4.057	0.531
12.962	7.5162	1.7245	0.602	11.222	2.532	4.432	0.580
13.97	7.503	1.8621	0.657	12.198	2.527	4.827	0.634
14.983	7.4902	2.000	0.711	13.282	2.522	5.2657	0.693
15.987	7.476	2.1385	0.765	14.36	2.518	5.7045	0.752
17.239	7.459	2.3109	0.831	15.602	2.509	6.2177	0.817
18.485	7.443	2.4836	0.898	16.82	2.505	6.7167	0.881
19.725	7.426	2.656	0.963	18.053	2.499	7.225	0.946
20.96	7.4109	2.8282	1.029	19.279	2.493	7.734	1.009
22.43	7.391	3.0353	1.108	20.609	2.488	8.283	1.079
23.90	7.372	3.2423	1.185	21.915	2.482	8.8285	1.149
25.36	7.3518	3.449	1.259	23.21	2.476	9.3741	1.217
26.81	7.333	3.6565	1.335	24.53	2.47	9.9282	1.283

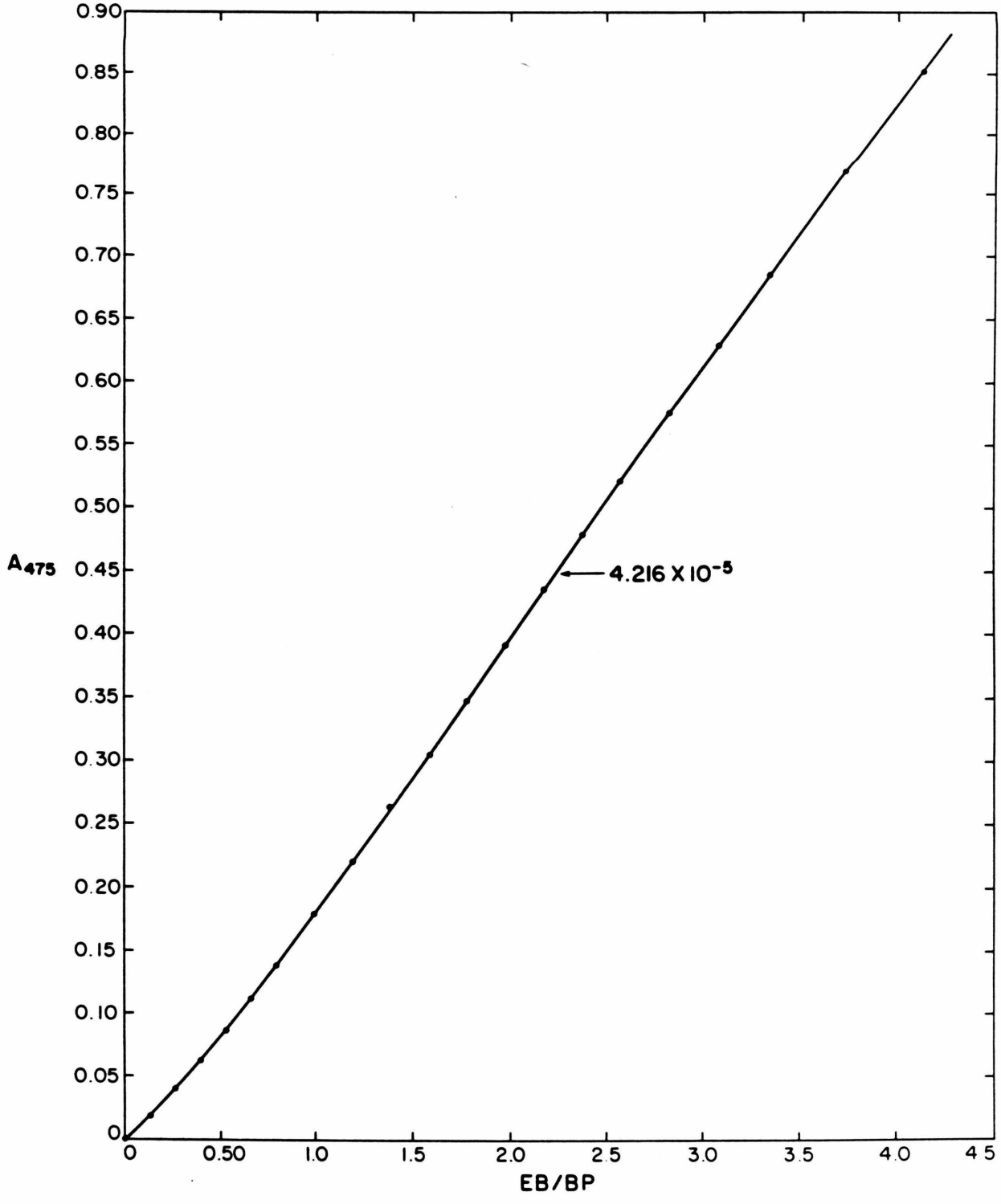
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
28.26	7.314	3.8635	1.409	25.92	2.464	10.5173	1.354
29.69	7.295	4.0704	1.484	27.465	2.458	11.1717	1.434
31.12	7.276	4.2774	1.557	29.004	2.45	11.837	1.514
0	0	0	0	0	0	0	0
3.8267^{-5}	4.983^{-5}	0.7679	0.155	2.8593^{-5}	1.3032^{-5}	2.1942	0.146
5.991	4.964	1.2068	0.265	4.5615	1.299	3.5114	0.239
6.5295	4.9595	1.3165	0.294	5.1267	1.298	3.95	0.270
7.332	4.952	1.4804	0.338	5.9694	1.295	4.6078	0.318
8.139	4.945	1.6458	0.381	6.8155	1.294	5.2658	0.365
9.207	4.937	1.8649	0.44	7.66	1.292	5.9291	0.410
1.0271	4.928	2.0844	0.497	8.608	1.291	6.669	0.463
1.1331	4.918	2.3041	0.555	9.609	1.288	7.4575	0.516
1.2334	4.909	2.512	0.61	10.605	1.286	8.247	0.571
1.333	4.901	2.72	0.664	11.666	1.285	9.0795	0.629
1.4381	4.892	2.9397	0.72	12.712	1.281	9.9214	0.686
1.537	4.884	3.1477	0.774	13.741	1.279	10.7438	0.742
1.636	4.876	3.356	0.828	14.793	1.277	11.5874	0.798
1.736	4.867	3.567	0.880	15.83	1.274	12.4208	0.853
1.834	4.858	3.7757	0.933	16.857	1.272	13.254	0.909
1.937	4.849	3.9946	0.988	17.939	1.269	14.132	0.966
2.0381	4.84	4.2109	1.041	18.999	1.267	14.994	1.023
2.142	4.833	4.432	1.097	20.091	1.265	15.885	1.081
2.244	4.825	4.6511	1.151	21.16	1.262	16.7628	1.137
2.356	4.812	4.8923	1.207	22.23	1.26	17.6398	1.193
2.466	4.803	5.1339	1.265	23.398	1.258	18.605	1.253
2.5772	4.794	5.376	1.321	24.488	1.254	19.5263	1.31
2.686	4.786	5.612	1.387	25.594	1.252	20.448	1.367
2.799	4.777	5.858	1.434	26.73	1.249	21.393	1.425
2.8982	4.768	6.0786	1.486	27.797	1.247	22.291	1.477
2.9974	4.76	6.2977	1.537	28.837	1.245	23.17	1.532

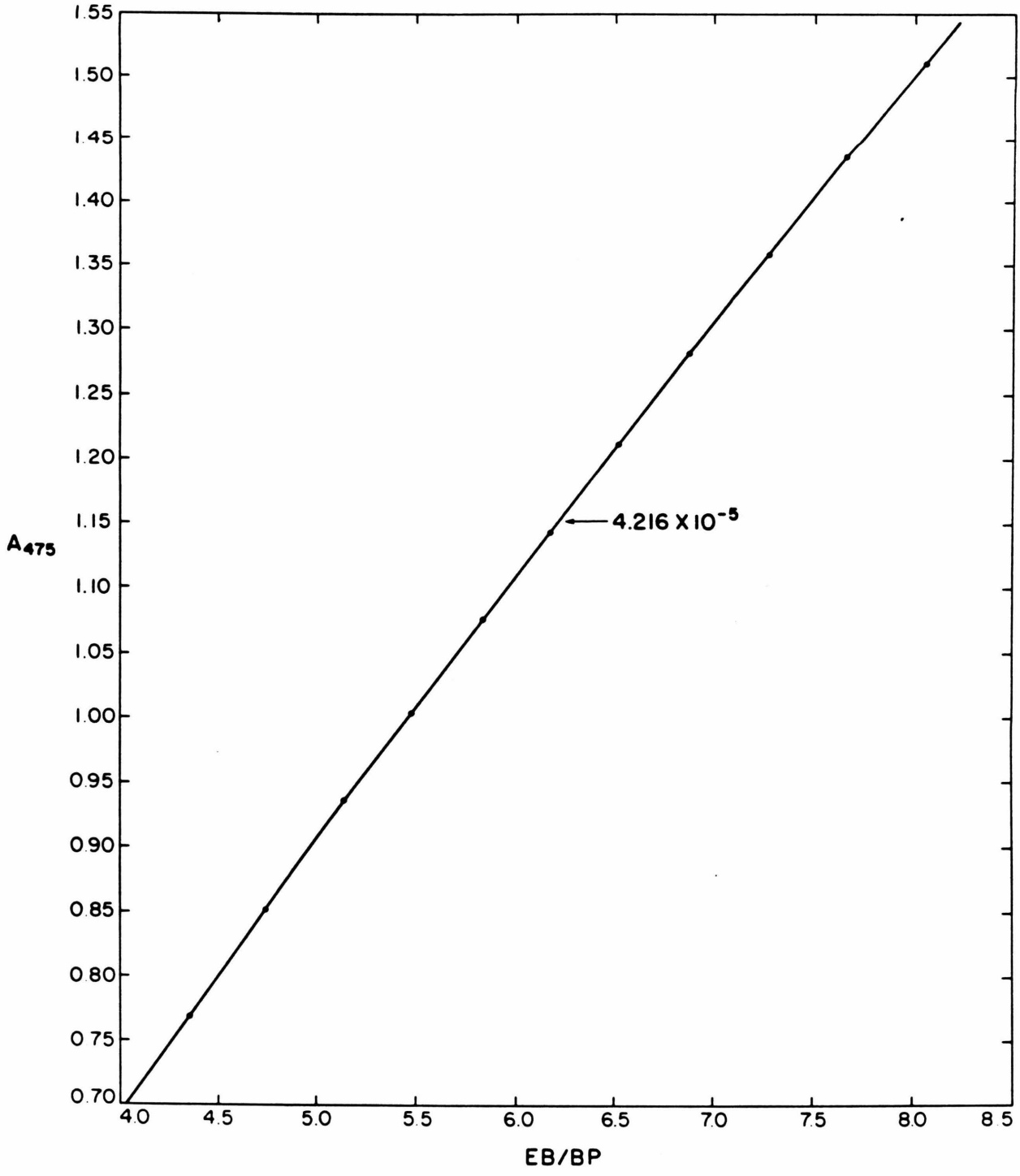
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
273.7	1.4185	1.9297	0.133	138.89	6.7318	2.0632	0.069	253.16	6.2846	4.0284	1.277	242.27	11.6939	2.0717	1.287
285.4	1.4162	2.015	0.139	144.795	6.7259	2.1528	0.072	267.95	6.2681	4.2752	1.351	257.9	12.485	2.0659	1.366
302.7	1.4114	2.1444	0.149	153.561	6.6999	2.2923	0.077	282.66	6.2515	4.5212	1.425	272.38	13.2187	2.0606	1.441
0	0	0	0	0	0	0	0	297.3	6.2349	4.7681	1.50	286.24	13.926	2.0554	1.511
5.396^{-6}	6.5579^{-5}	0.08228	0.018	5.685^{-6}	0.2633	2.15939^{-5}	0.025	0	0	0	0	0	0	0	0
10.78	6.552	0.16456	0.036	11.36	0.5266	2.1573	0.052	5.55^{-5}	4.2166^{-5}	0.1316	0.019	5.77^{-6}	8.7655^{-6}	0.6583	0.029
18.84	6.5402	0.2881	0.066	17.022	0.7898	2.1553	0.080	1.1089^{-5}	4.2118	0.2633	0.040	8.65	8.7561	0.9881	0.043
26.88	6.5335	0.4114	0.098	22.674	1.0532	2.1531	0.110	1.6618	4.2083	0.3949	0.062	1.153^{-5}	8.756	1.3169	0.057
34.90	6.5246	0.5349	0.132	28.316	1.3164	2.151	0.140	2.2137	4.2035	0.5266	0.087	1.44	8.748	1.6461	0.072
42.89	6.515	0.6583	0.169	33.945	1.5796	2.1489	0.170	2.764	4.20	0.6584	0.112	1.727	8.748	1.9745	0.088
50.86	6.5058	0.7817	0.208	39.565	1.843	2.1469	0.202	3.3142	4.1965	0.7898	0.138	2.013	8.7395	2.3033	0.1035
58.81	6.4976	0.90508	0.249	45.173	2.1062	2.1448	0.234	4.137	4.188	0.9877	0.178	2.586	8.73	2.9617	0.134
66.74	6.4893	1.0286	0.290	51.90	2.4226	2.1422	0.271	4.957	4.1846	1.1847	0.220	3.161	8.722	3.6244	0.165
74.645	6.4798	1.152	0.332	58.61	2.739	2.1398	0.308	5.775	4.1763	1.3828	0.263	4.015	8.7135	4.6078	0.211
82.49	6.471	1.2748	0.374	66.94	3.133	2.1367	0.357	6.591	4.173	1.5798	0.305	4.866	8.696	5.595	0.258
90.39	6.4621	1.3988	0.415	75.271	3.528	2.1336	0.400	7.4044	4.1681	1.7763	0.348	5.77	8.687	6.641	0.307
98.23	6.4538	1.5221	0.457	83.57	3.923	2.1303	0.446	8.215	4.1606	1.9745	0.391	6.671	8.671	7.6936	0.356
108.71	6.444	1.6871	0.514	94.604	4.448	2.1265	0.507	9.029	4.1527	2.174	0.436	7.583	8.663	8.7541	0.405
118.99	6.4301	1.8506	0.569	104.48	4.9218	2.1225	0.562	9.831	4.1491	2.369	0.479	8.479	8.645	9.8073	0.454
129.3	6.4183	2.0142	0.625	114.31	5.3941	2.119	0.617	10.63	4.1408	2.567	0.521	9.376	8.628	10.866	0.502
139.6	6.4065	2.179	0.680	124.23	5.8724	2.1155	0.671	11.651	4.1337	2.8189	0.576	10.487	8.612	12.1773	0.561
152.53	6.3922	2.3862	0.749	137.77	6.5279	2.1104	0.743	12.663	4.1255	3.0691	0.63	11.598	8.603	13.4805	0.622
165.38	6.3781	2.5930	0.816	151.37	7.1894	2.1054	0.815	13.724	4.122	3.3294	0.686	12.711	8.579	14.817	0.681
178.03	6.3698	2.7949	0.884	164.91	7.8514	2.1005	0.886	15.31	4.11	3.725	0.769	14.089	8.562	16.455	0.753
193.2	6.352	3.0418	0.963	179.73	8.5792	2.0949	0.964	16.88	4.0994	4.118	0.851	15.525	8.537	18.1853	0.828
208.3	6.335	3.288	1.043	194.61	9.3144	2.0893	1.04	18.452	4.0876	4.5136	0.933	16.935	8.5207	19.875	0.903
223.3	6.3177	3.5356	1.122	210.71	10.1134	2.0834	1.123	19.803	4.07696	4.8576	1.003	18.35	8.496	21.598	0.978
238.3	6.3012	3.7816	1.201	226.53	10.9027	2.0776	1.206	21.2	4.0651	5.2151	1.075	19.75	8.48	23.29	1.050

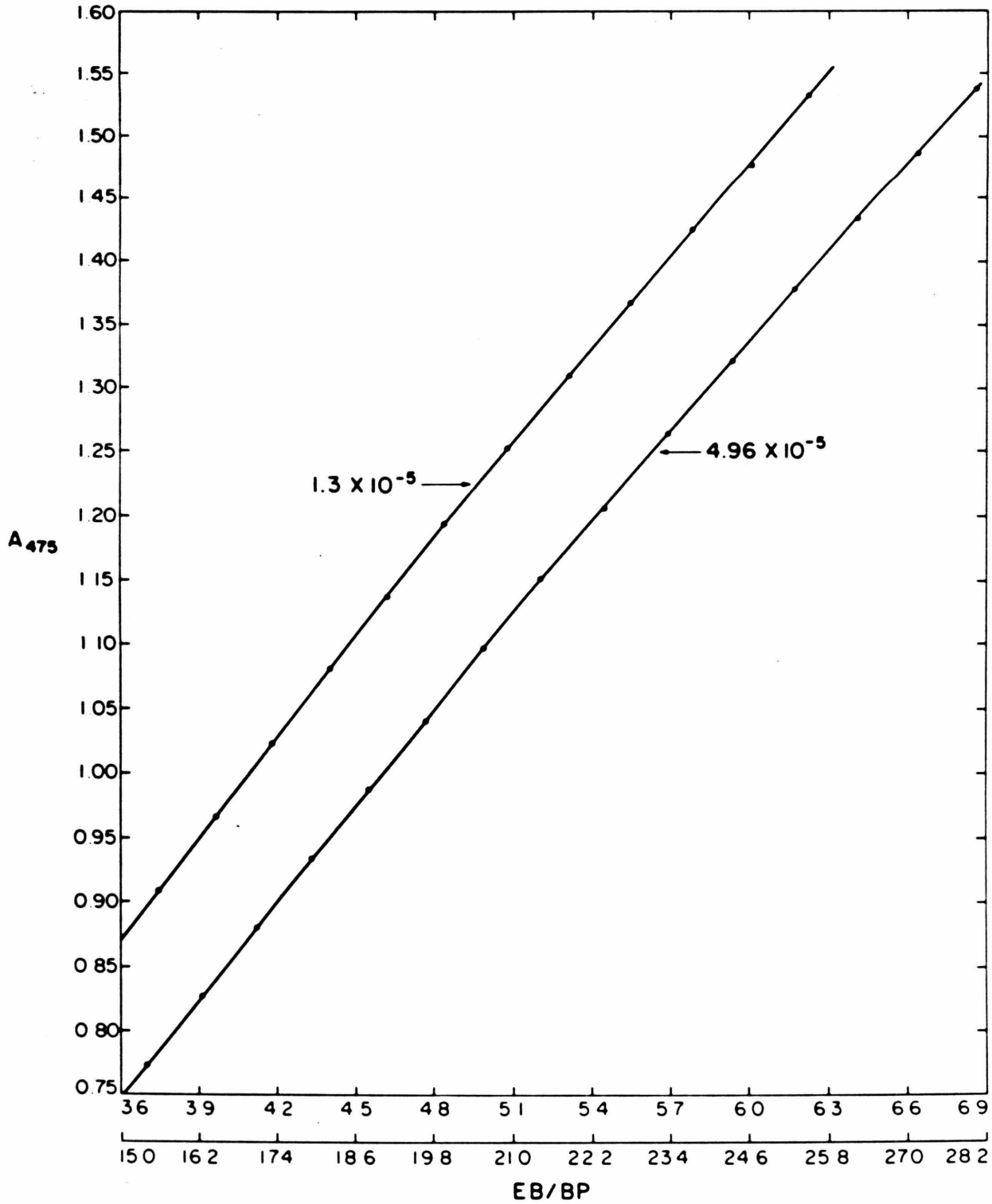
Figures 1-6

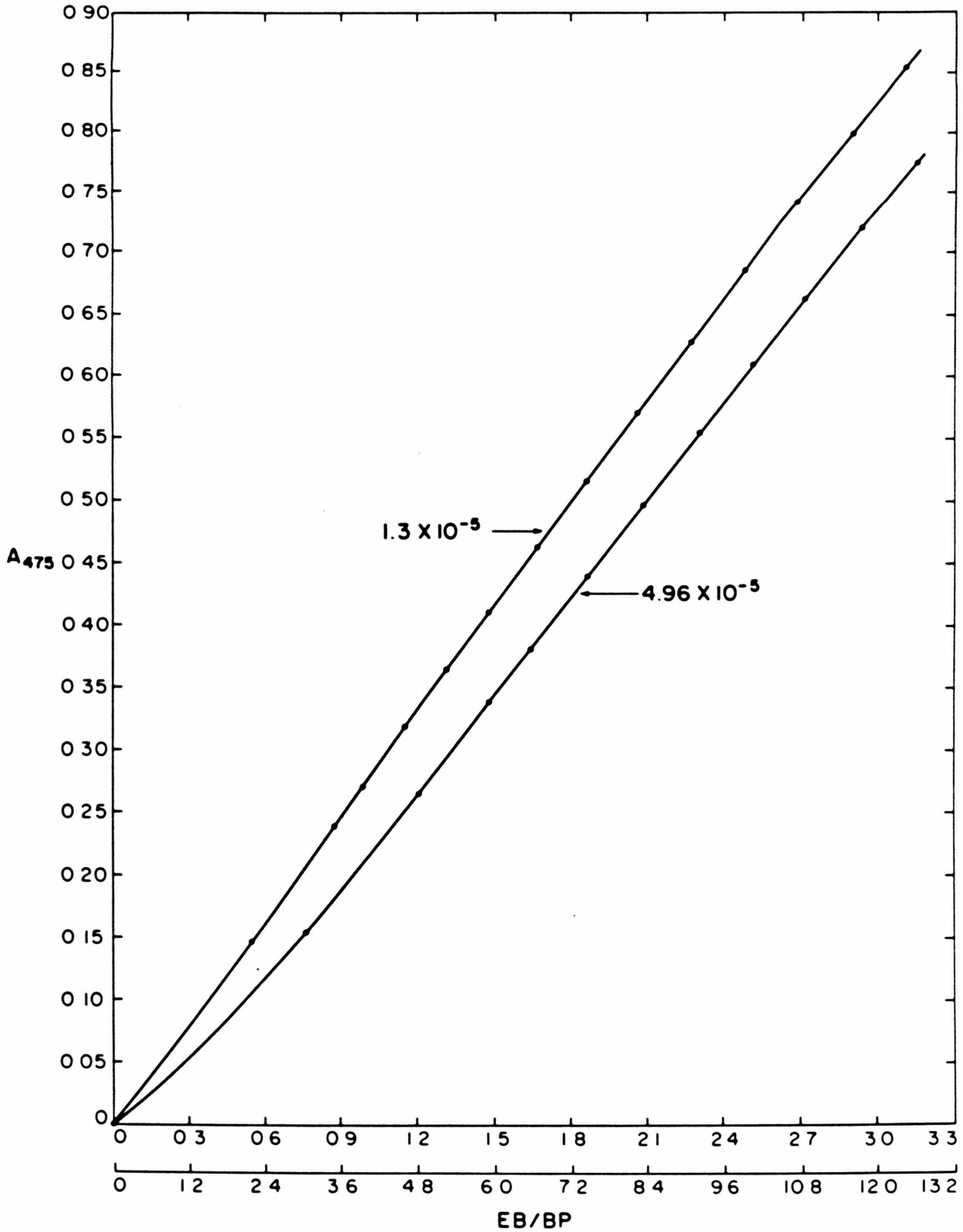
Binding of EB at M^+ = 1.0 to the indicated concentration of poly(dC-dG). The absorbance at 475 nm (path length 1.0 cm) for increasing additions of EB is plotted against the corresponding EB/BP ratio. When more than one EB/BP axis is presented, the lower axis refers to the more dilute d(C-G) concentraion. Five representative examples are shown and all measurements are summarized in Table III of the Appendix.


0.15
0.14
0.13
0.12
0.11
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
A475
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
EB/BP
1.474 X 10-5


0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
A475
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
4.216 X 10-5
0
0.50
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
EB/BP


1.55
1.50
1.45
1.40
1.35
1.30
1.25
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
0.75
0.70
A475
4.216 X 10-5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
EB/BP


1.60
1.55
1.50
1.45
1.40
1.35
1.30
1.25
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
0.75
A475
1.3 X 10-5
4.96 X 10-5
36 39 42 45 48 51 54 57 60 63 66 69
150 162 174 186 198 210 222 234 246 258 270 282
EB/BP


0 90
0 85
0 80
0 75
0 70
0 65
0 60
0 55
0 50
A475 0 45
0 40
0 35
0 30
0 25
0 20
0 15
0 10
0 05
0
1.3 X 10-5
4.96 X 10-5
0 03 06 09 12 15 18 21 24 27 30 33
0 12 24 36 48 60 72 84 96 108 120 132
EB/BP

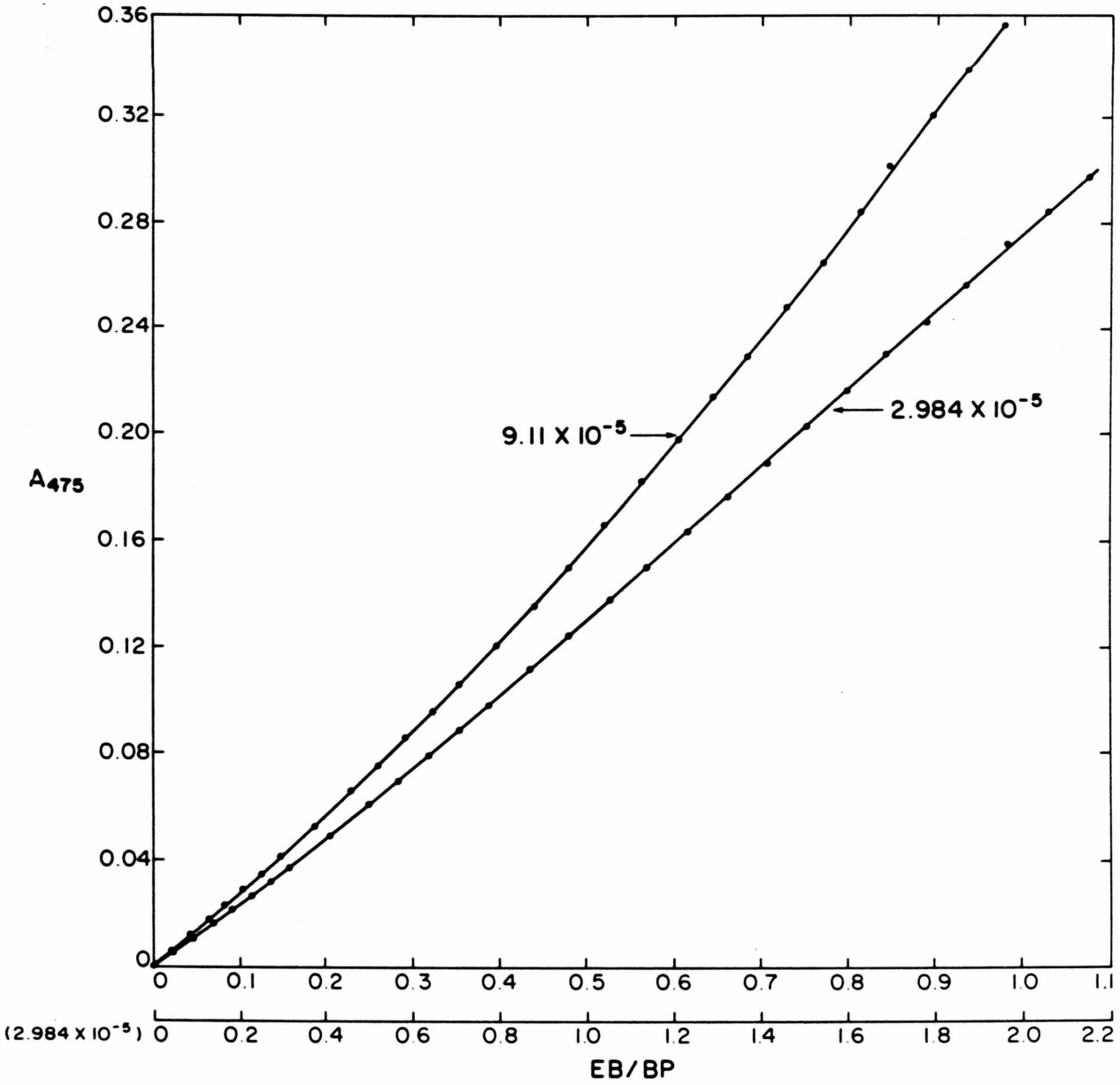

0.36
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0
A475
9.11 X 10-5
2.984 X 10-5
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1
(2.984 X 10-5) 0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2
EB/BP

Table IV - Summary of the binding of EB to polydCdG at 1 M^+ followed by visible spectroscopy. The cell path length was 0.5 cm and the concentration of nucleic acid is in base pairs (BP).

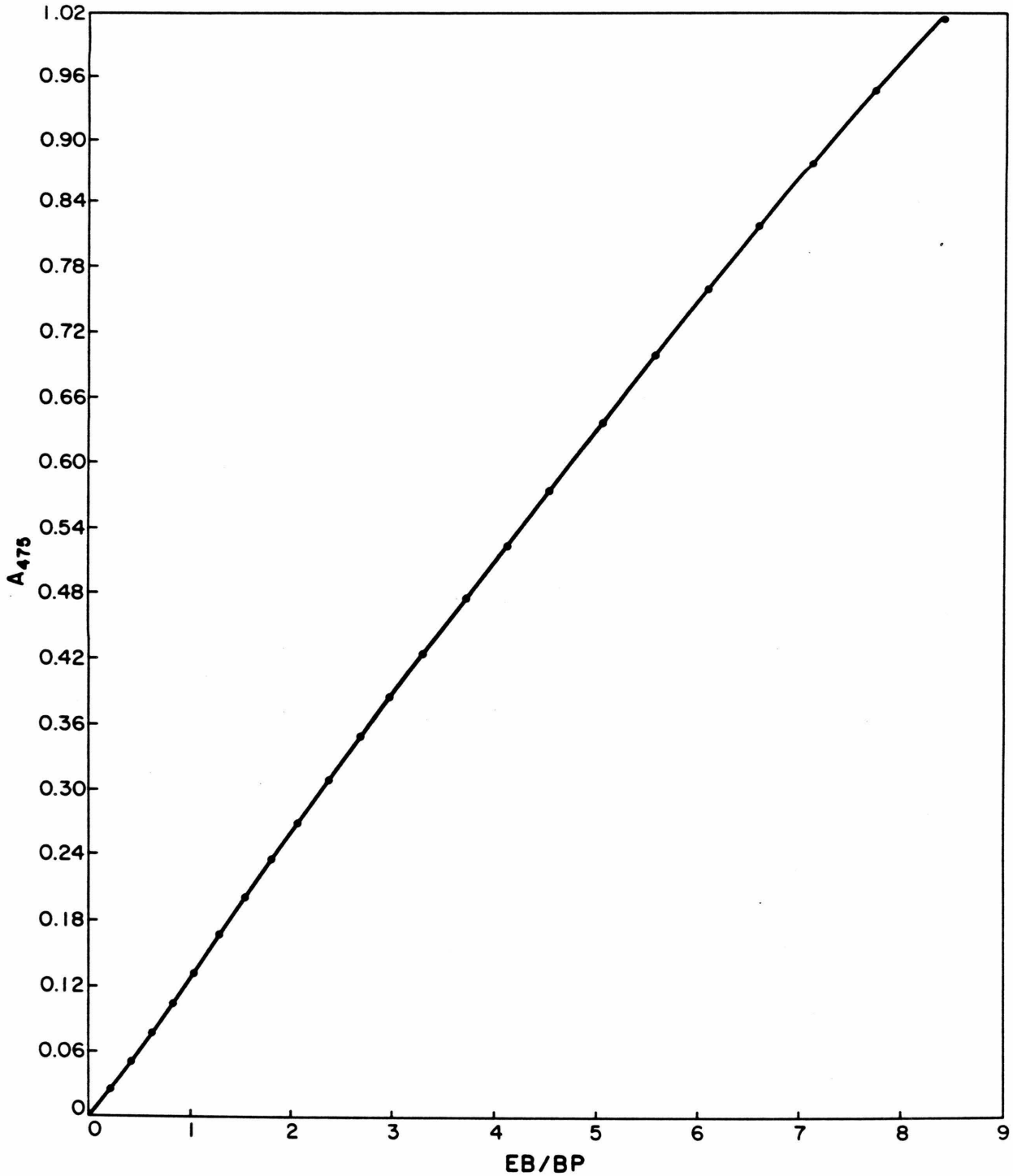
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1.0791^{-5}	3.93^{-5}	0.2746	0.0259	2.2338^{-5}	2.712^{-5}	0.8237	0.0568	1.0696^{-5}	4.2204^{-5}	0.2534	0.02565	1.1347^{-5}	2.239^{-5}	0.5068	0.0281
2.1543	3.923	0.5492	0.529	2.7897	2.709	1.0298	0.0718	2.1353	4.2129	0.5069	0.0522	2.2652	2.2345	1.0137	0.0589
3.2258	3.917	0.8235	0.0816	3.3456	2.707	1.2355	0.0867	3.1974	4.2055	0.7603	0.07985	3.3914	2.2303	1.5206	0.0892
4.2934	3.909	1.0983	0.109	4.4513	2.702	1.6474	0.1159	4.2556	4.1982	1.0137	0.10665	4.5135	2.2262	2.0274	0.119
5.3573	3.903	1.3726	0.1378	5.5539	2.697	2.0593	0.1454	5.3103	4.1908	1.2671	0.1349	5.631	2.222	2.5344	0.1495
6.6847	3.896	1.7158	0.1738	6.923	2.69	2.5651	0.1823	6.6204	4.1835	1.5825	0.17015	7.0261	2.217	3.1692	0.1866
8.000	3.885	2.0593	0.209	8.2927	2.685	3.0885	0.2179	7.9306	4.1725	1.9007	0.2047	8.4079	2.2118	3.8014	0.2231
9.3175	3.875	2.4045	0.2443	9.9241	2.678	3.7058	0.2592	9.2282	4.1616	2.2175	0.2397	9.7821	2.2067	4.4329	0.2585
10.6205	3.868	2.7457	0.277	11.547	2.67	4.3246	0.3017	10.5282	4.1543	2.5343	0.2714	11.1589	2.2016	5.0685	0.2955
12.182	3.858	3.1575	0.3195	13.16	2.663	4.9419	0.3439	12.076	4.1435	2.9144	0.3114	12.7973	2.1955	5.8289	0.3387
13.734	3.848	3.5691	0.3603	14.765	2.656	5.559	0.3849	13.615	4.1329	3.2944	0.3509	14.4268	2.1895	6.5891	0.381
15.279	3.838	3.981	0.3978	16.907	2.646	6.3895	0.4394	15.147	4.1221	3.6746	0.3904	16.0473	2.1835	7.3493	0.421
16.815	3.828	4.3929	0.4383	19.0026	2.638	7.2034	0.4939	16.671	4.1114	4.0549	0.4304	17.6589	2.1775	8.1097	0.4267
18.852	3.815	4.9416	0.4915	21.099	2.627	8.0315	0.5439	18.69	4.0973	4.5617	0.4824	19.7941	2.1696	9.1234	0.5142
20.8748	3.802	5.4905	0.539	23.698	2.616	9.0588	0.6089	20.696	4.0833	5.0686	0.5279	21.9138	2.1618	10.1369	0.5672
22.883	3.789	6.0394	0.589	26.274	2.604	10.0898	0.6739	22.688	4.0694	5.5754	0.5779	24.0183	2.1539	11.1511	0.6192
25.3748	3.772	6.7271	0.6535	28.827	2.593	11.1173	0.7359	25.16	4.0522	6.2089	0.6399	26.627	2.1443	12.4178	0.6852
27.845	3.756	7.4135	0.714	31.358	2.581	12.1496	0.7959	27.61	4.035	6.8425	0.7004	29.2136	2.1347	13.6851	0.7472
30.294	3.740	8.1	0.7745	34.366	2.568	13.3824	0.8689	30.039	4.0181	7.476	0.7594	31.7765	2.1252	14.9522	0.8082
32.723	3.725	8.7846	0.834	37.343	2.554	14.6212	0.9404	32.449	4.0012	8.1097	0.8194	34.3168	2.1158	16.2193	0.8702
35.1308	3.709	9.4718	0.893	40.288	2.541	15.8553	1.0089	34.838	3.9846	8.7431	0.8789	36.835	2.1065	17.4862	0.9307
37.5187	3.693	10.1594	0.949					37.679	3.9647	9.5035	0.9479	39.827	2.0954	19.0068	1.0002
40.358	3.675	10.9818	1.0165					40.024	3.9483	10.1371	1.0029				

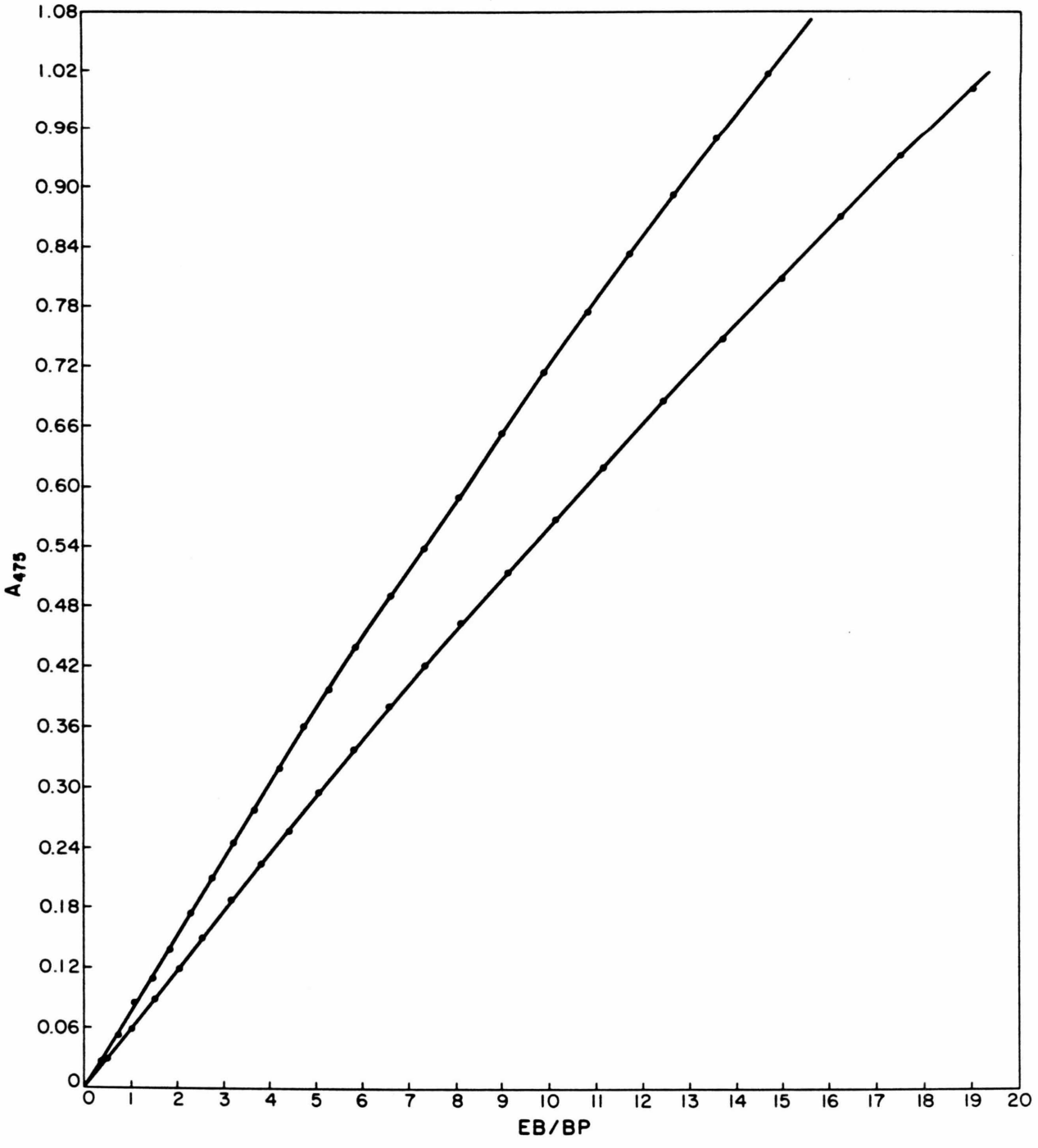
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
0	0	0	0	0	0	0	0	4.1678	5.06	0.8237	0.1038	4.556	1.901	2.3966	0.12
1.0937^{-5}	3.4858^{-5}	0.3138	0.0272	1.1087^{-5}	3.0288^{-5}	0.36606	0.0273	5.2052	5.05	1.0307	0.1315	5.684	1.898	2.9949	0.1503
2.1835	3.4795	0.6275	0.0554	2.2134	2.0233	0.7321	0.0561	6.50	5.05	1.2873	0.1667	7.0196	1.894	3.7062	0.1879
3.2693	3.4732	0.9413	0.0844	3.3141	3.0178	1.0982	0.08545	7.794	5.05	1.5434	0.2012	8.4865	1.889	4.4926	0.2246
4.3513	3.4669	1.2551	0.1123	4.4107	3.0123	1.4642	0.1133	9.0857	5.04	1.8027	0.2352	10.155	1.884	5.3903	0.2675
5.4293	3.4608	1.5688	0.1412	5.5034	3.0068	1.8303	0.1421	10.3748	5.04	2.0585	0.2682	11.815	1.878	6.2912	0.312
6.7745	3.4545	1.9611	0.1769	6.8667	3.0014	2.2878	0.1786	11.9208	5.03	2.3699	0.3089	13.465	1.873	7.1891	0.354
8.1076	3.4453	2.3532	0.2122	8.2177	2.9932	2.7454	0.2143	13.4526	5.03	2.6744	0.3477	15.106	1.868	8.0869	0.397
9.4336	3.4361	2.7454	0.2469	9.5613	2.9852	3.2029	0.2498	14.992	5.02	2.9865	0.3859	17.28	1.861	9.2855	0.453
10.762	3.43	3.1376	0.2797	10.908	2.9797	3.6606	0.2831	16.515	5.01	3.2963	0.4249	19.438	1.854	10.4846	0.508
12.3435	3.4209	3.6083	0.3209	12.5098	2.9717	4.2097	0.3258	18.548	5.00	3.7095	0.4769	21.581	1.848	11.6779	0.5595
13.917	3.4118	4.0789	0.3609	14.104	2.9637	4.7588	0.3663	20.555	4.99	4.1194	0.5244	24.237	1.839	13.1794	0.6265
15.481	3.3998	4.5546	0.4014	15.689	2.9558	5.3078	0.4076	22.573	4.98	4.5328	0.5759	26.869	1.831	14.6744	0.691
17.038	3.3938	5.0202	0.4422	17.25	3.9479	5.8517	0.4498	25.074	4.97	5.0452	0.6379	29.474	1.823	16.1695	0.7565
19.1	3.3819	5.6477	0.4959	19.355	2.9374	6.589	0.5026	27.54	4.95	5.5636	0.6989	32.062	1.814	17.6746	0.8185
21.148	3.37	6.2755	0.5434	21.429	2.927	7.3211	0.5513	30.014	4.94	6.0757	0.7589	34.623	1.806	19.1714	0.880
23.182	3.358	6.9036	0.5944	23.489	2.9167	8.0532	0.6033	32.449	4.93	6.5819	0.8179	37.162	1.798	20.6687	0.9415
25.705	3.344	7.6868	0.6584	26.043	2.9039	8.9683	0.6673	34.866	4.91	7.1012	0.8769	39.679	1.7903	22.1632	1.0015
28.205	3.3294	8.4716	0.7194	28.575	2.8912	9.8835	0.7288	37.836	4.90	7.7216	0.9469				
30.685	3.315	9.2563	0.7794	31.085	2.8786	10.7987	0.7928	40.828	4.87	8.3837	1.014				
33.143	3.3009	10.0405	0.8424	33.573	2.8661	11.7139	0.8513	0	0	0	0	0	0	0	0
35.579	3.2868	10.825	0.9024	36.04	2.8538	12.6288	0.9128	1.12942^{-5}	2.3953^{-5}	0.47151	0.0285	1.1538^{-5}	2.5711^{-5}	0.4488	0.0290
37.996	3.2729	11.6093	0.9599	38.486	2.8415	13.5442	0.9728	2.2546	2.393	0.94218	0.05803	2.3141	2.5783	0.8975	0.0593
39.915	3.2618	12.237	1.0054	39.943	2.8342	14.0932	1.0063	3.3757	2.391	1.4118	0.08853	3.4875	2.5856	1.3462	0.091
0	0	0	0	0	0	0	0	4.4925	2.386	1.8829	0.1174	4.6543	2.5929	1.79506	0.1221
1.0437^{-5}	5.065^{-5}	0.20607	0.0240	1.1454^{-5}	1.9123^{-5}	0.599	0.02925	5.6053	2.382	2.3532	0.1474	5.8343	2.6002	2.2438	0.153
2.0857	5.06	0.4122	0.499	2.2866	1.9086	1.198	0.05985	6.9872	2.375	2.9419	0.1835	7.3066	2.6051	2.8048	0.1914
3.1258	5.06	0.6178	0.0770	3.4234	1.905	1.797	0.0904	8.3692	2.371	3.5298	0.2202	8.7927	2.6126	3.3656	0.2299

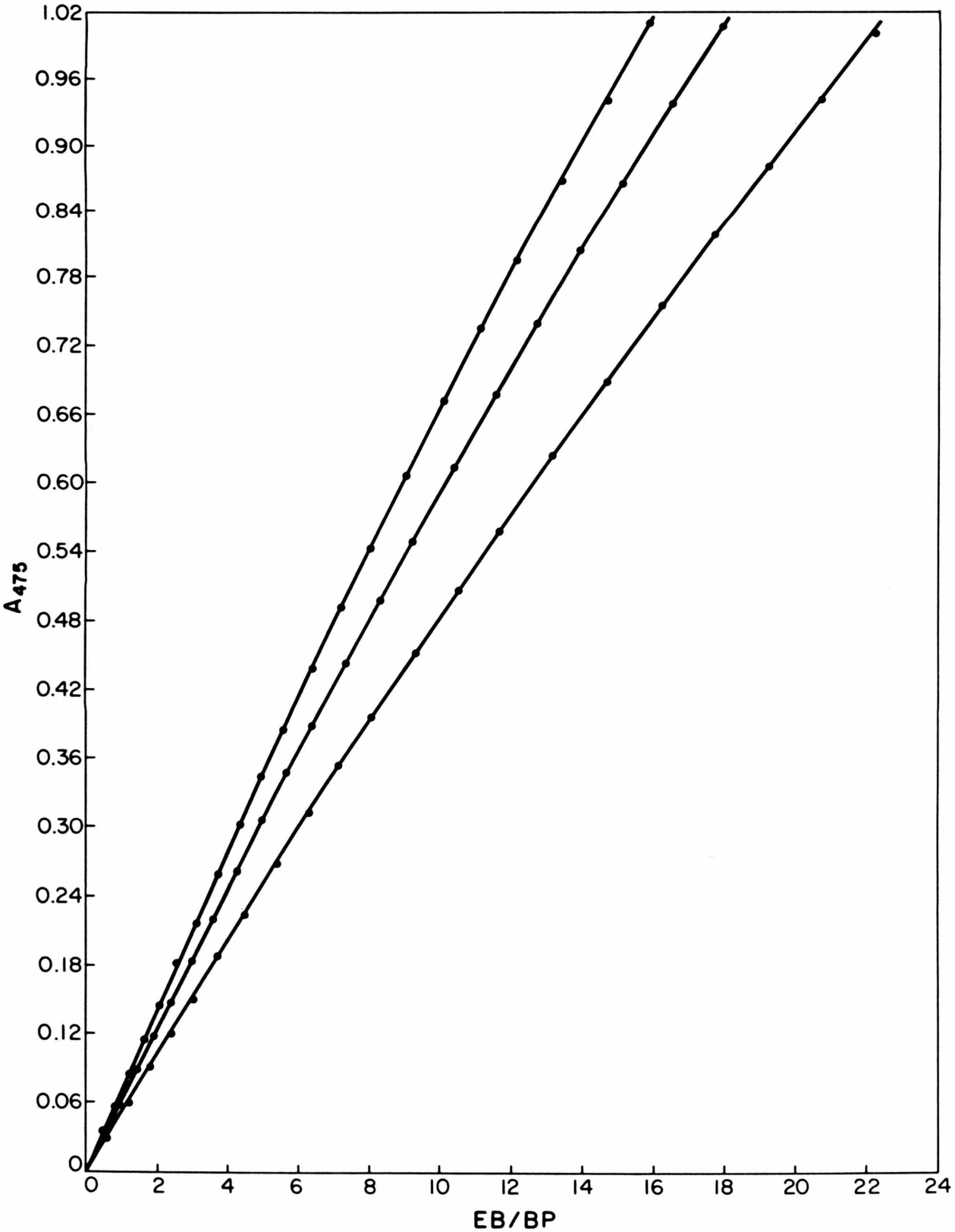
[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}	[EB]	[BP]	EB/BP	A_{475}
10.0153	2.364	4.2366	0.2624	10.2779	2.6175	3.9266	0.2671	16.6849	3.251	5.1327	0.440
11.652	2.358	4.9416	0.3056	11.7239	2.6126	4.4874	0.3039	18.8066	3.239	5.8058	0.4896
13.2805	2.351	5.6489	0.3476	13.1520	2.6051	5.0486	0.3416	20.9134	3.228	6.4789	0.5456
14.8997	2.345	6.3538	0.3895	14.8634	2.5978	5.7216	0.3856	23.0054	3.2166	7.1521	0.5986
17.045	2.337	7.2935	0.44403	16.5653	2.5905	6.3947	0.4291	25.5998	3.2026	7.9934	0.6646
19.175	2.328	8.2365	0.4986	18.258	2.5831	7.0681	0.4721	28.1718	3.1887	8.8348	0.7276
21.289	2.32	9.1763	0.5496	19.9403	2.576	7.7408	0.5121	30.7216	3.1749	9.6763	0.7896
23.91	2.309	10.3553	0.6156	21.6136	2.5687	8.41415	0.5551	33.7523	3.1586	10.6859	0.8648
26.509	2.299	11.5305	0.6791	23.8303	2.5592	9.3116	0.6101	36.7519	3.1425	11.6955	0.9366
29.084	2.289	12.7058	0.7406	26.0304	2.5497	10.2092	0.6671	39.2283	3.129	12.537	0.9966
31.636	2.279	13.8814	0.8046	28.214	2.5403	11.1067	0.7196				
34.166	2.269	15.0708	0.8656	30.3828	2.531	12.0042	0.7741				
37.172	2.55	16.4842	0.9376	32.534	2.5217	12.9018	0.8271				
40.1467	2.243	17.8987	1.0076	34.671	2.5125	13.7993	0.8791				
				37.3196	2.5011	14.9213	0.9406				
0	0	0	0	39.9443	2.4898	16.0432	1.0036				
1.1224	3.355^{-5}	0.3366	0.0276								
2.2407	3.3288	0.6731	0.0566								
3.3549	3.323	1.0097	0.0866								
4.4652	3.317	1.3462	0.1149								
5.5715	3.3108	1.6829	0.1447								
6.6738	3.305	2.01935	0.1736								
8.04604	3.2974	2.44006	0.2111								
9.41218	3.29	2.8608	0.2476								
12.1262	3.2754	3.7022	0.3196								
12.9357	3.271	3.9546	0.3416								
13.743	3.267	4.2070	0.3629								
15.0839	3.259	4.6277	0.3981								

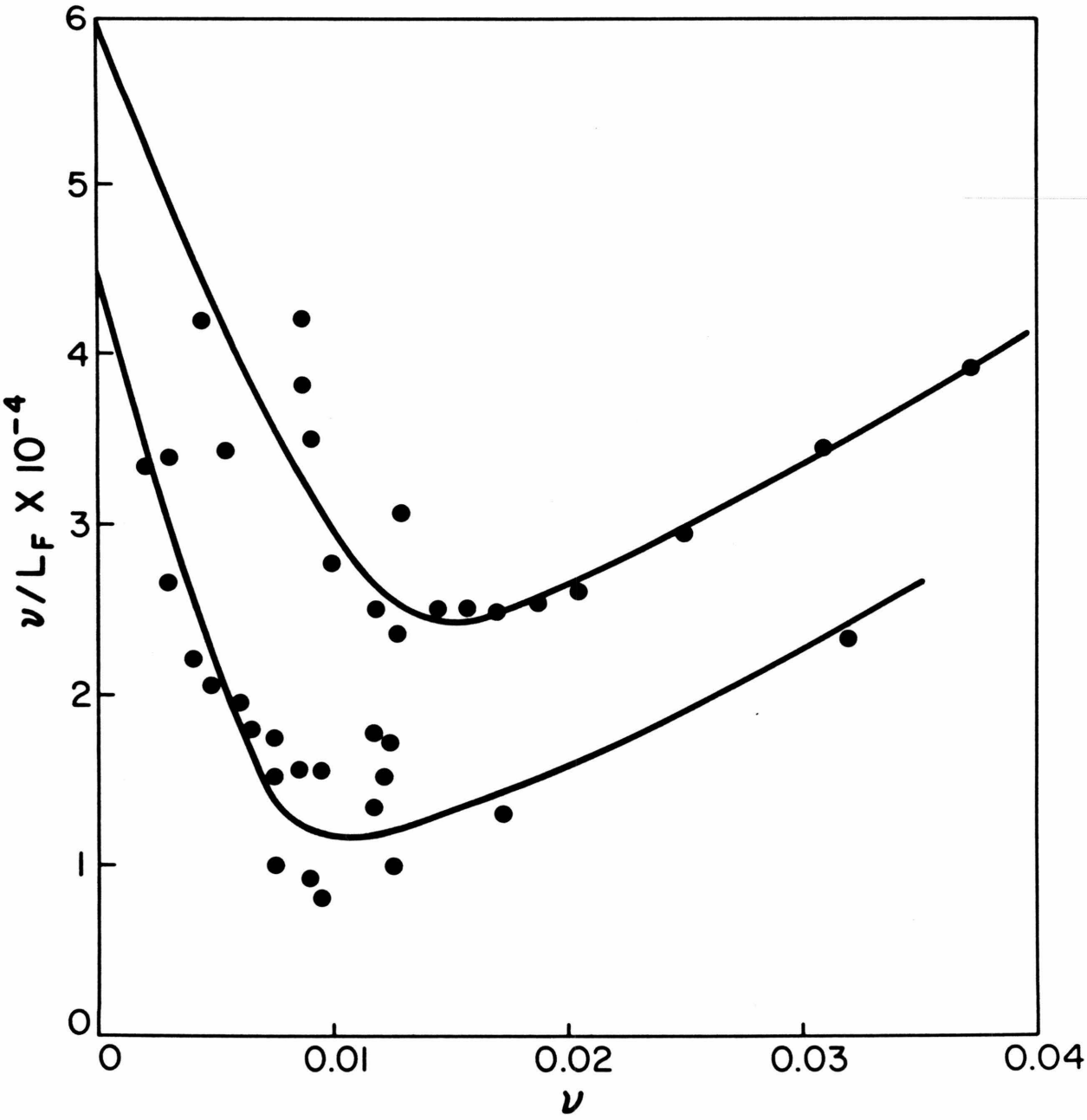
Figures 7-9

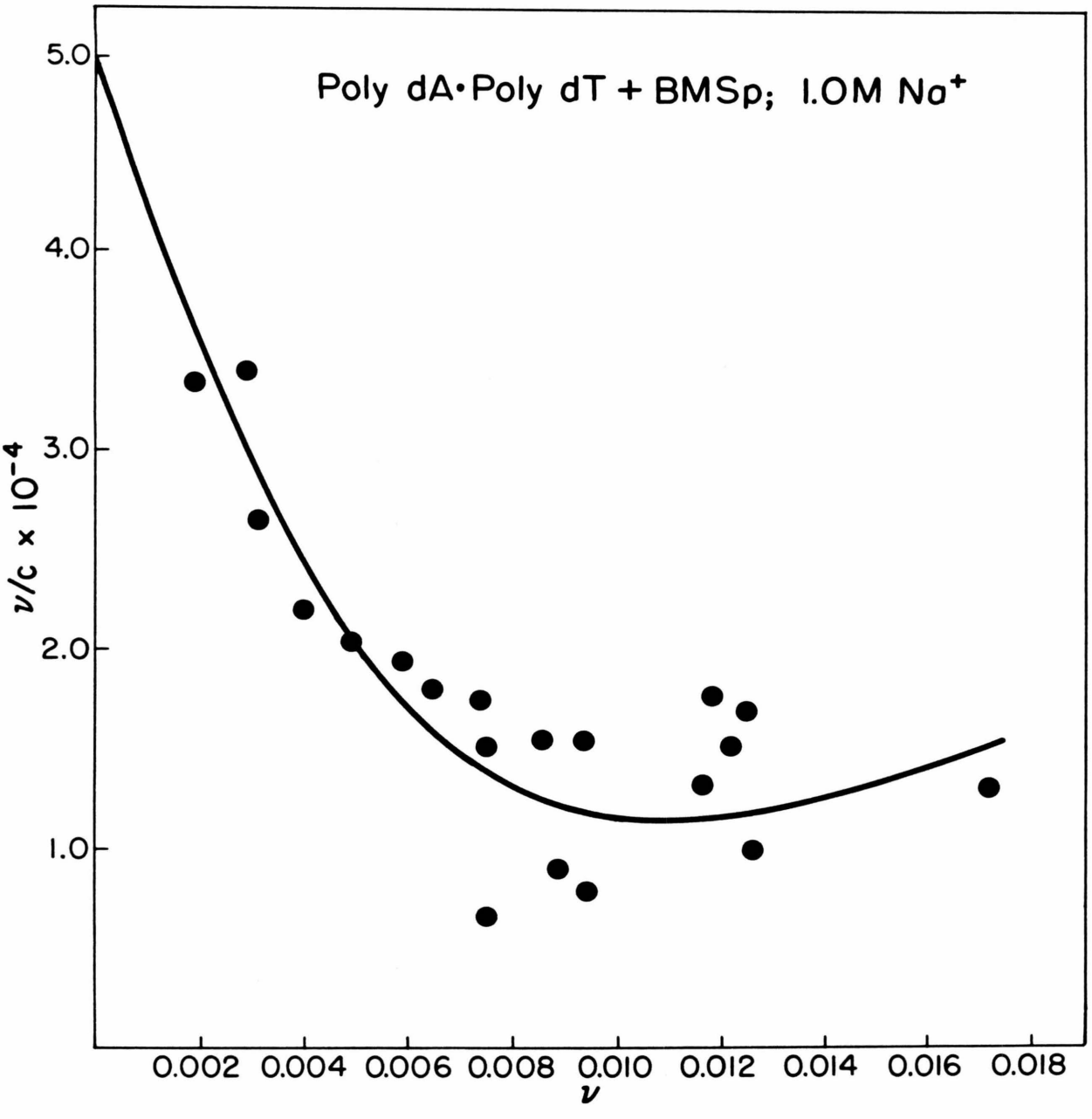
Spectrophotometric titration of polydCdG with EB at 1 M^+. The absorbance at 475 nm (path length 0.5 cm) for increasing additions of EB is plotted against the corresponding EB/BP ratio. Six representative examples are shown and all measurements are summarized in Table III of the Appendix. The concentration of dCdG in base pairs is: Fig. 12, 5.065×10^{-5} M; Fig. 13, upper curve 3.93×10^{-5} M; Fig. 14, left curve 2.712×10^{-5} M, middle curve 2.395×10^{-5}, right curve 1.912×10^{-5} M.

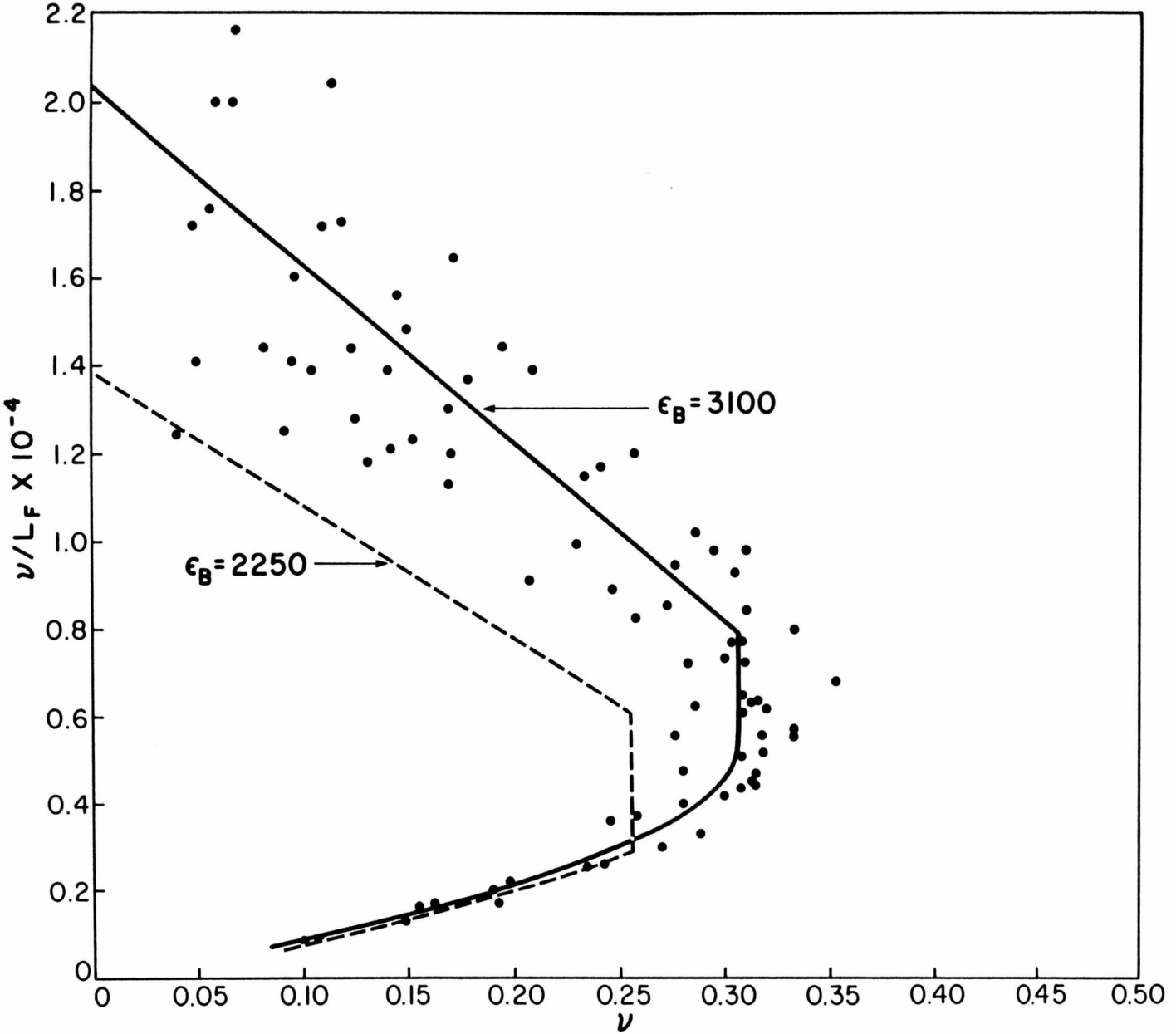

1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A475
0
1
2
3
4
5
6
7
8
9
EB/BP


1.08
1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A475
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
EB/BP


1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A475
0
2
4
6
8
10
12
14
16
18
20
22
24
EB/BP


6
5
4
3
2
1
0
$\nu/L_F \times 10^{-4}$
0
0.01
0.02
0.03
0.04
ν


Poly dA·Poly dT + BMSp; 1.0M Na+
5.0
4.0
3.0
2.0
1.0
ν/c × 10⁻⁴
0.002
0.004
0.006
0.008
0.010
0.012
0.014
0.016
0.018
ν


2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
$\nu/L_F \times 10^{-4}$
$\epsilon_B = 3100$
$\epsilon_B = 2250$
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
ν

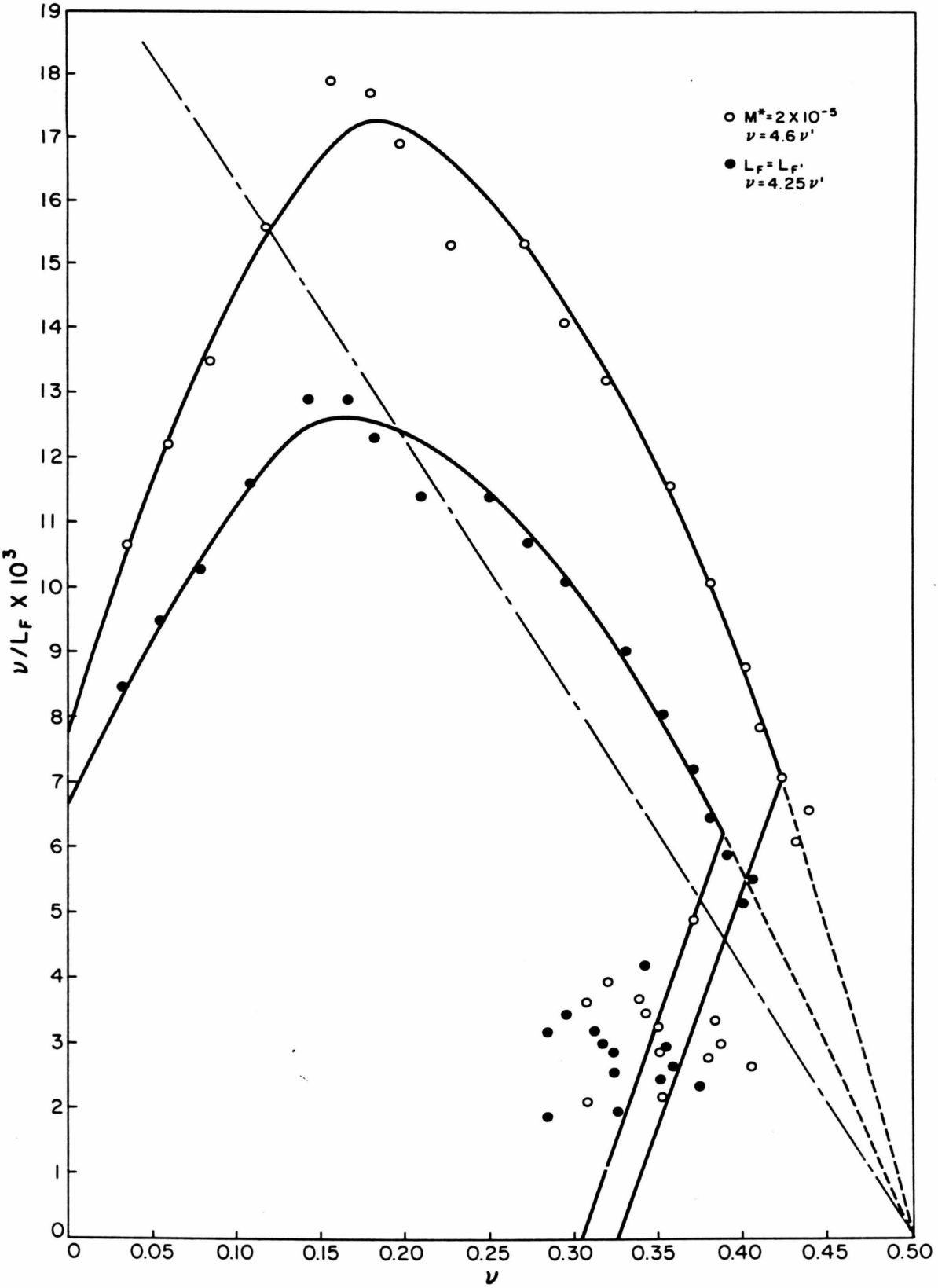
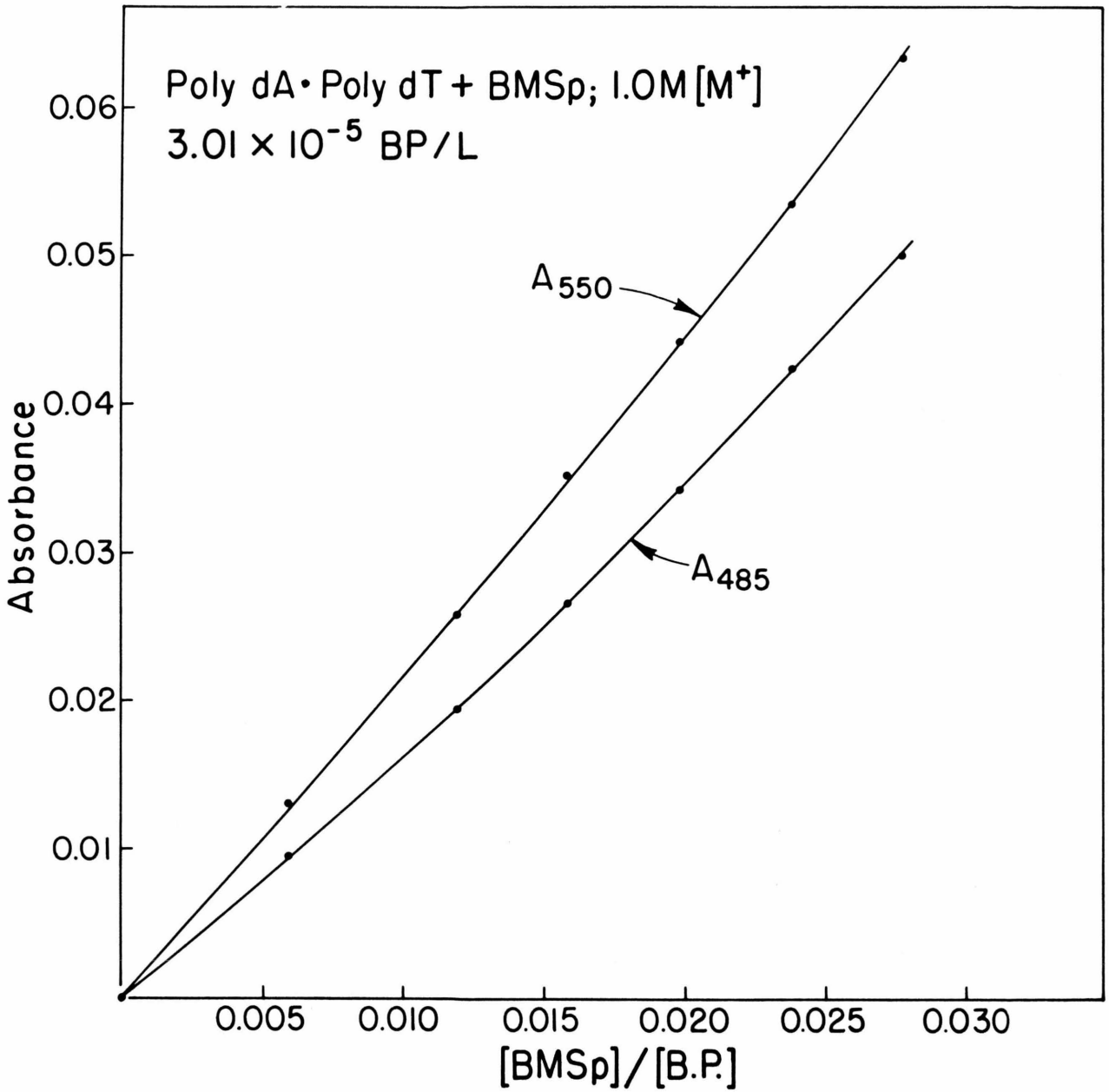

○ M*=2X10⁻⁵
ν=4.6ν'
● L_F=L_F'
ν=4.25ν'
ν/L_F X 10³
ν
0
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

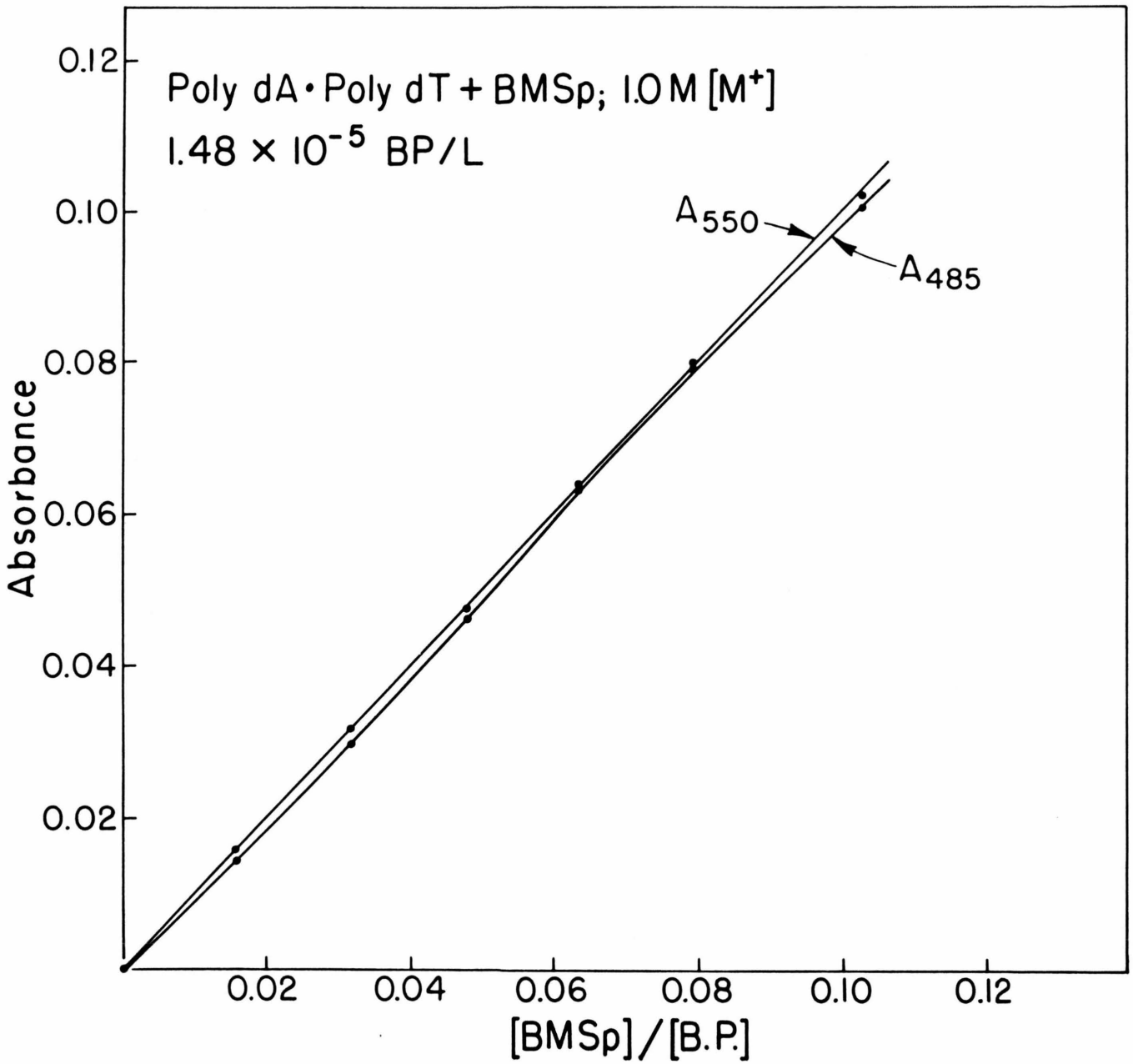
Table V - Summary of the binding of BMSp to polydAdT at 1 M^+ followed by visible spectroscopy. The cell path length was 10 cm and the concentration of nucleic acid is in base pairs.

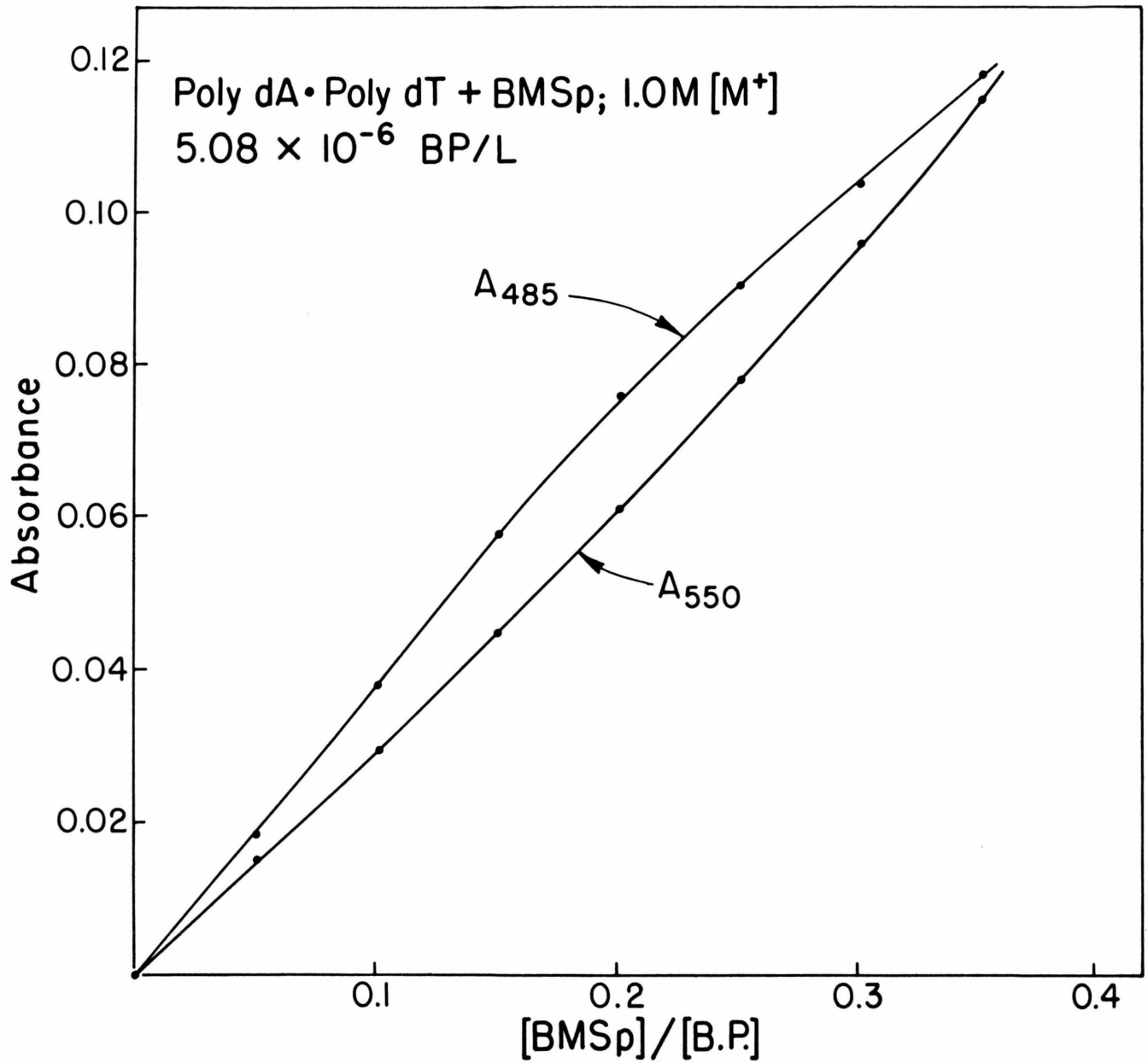
BMSp/BP	A_{550}	A_{485}	BMSp/BP	A_{550}	A_{485}
0	0	0	0	0	0
0.00595	0.0131	0.0095	0.0714	0.006	0.0087
0.1189	0.0259	0.0194	0.1428	0.0117	0.0168
0.0158	0.0351	0.0264	0.2143	0.0175	0.0254
0.0198	0.0442	0.0342	0.2856	0.023	0.0344
0.0237	0.0535	0.0422	0.3571	0.0287	0.0428
0.0277	0.0632	0.0500	0.4283	0.0344	0.0517
	$(3.01x10^{-5})$		0.4999	0.040	0.0597
0.01525	0.0041	0.0047	0.5711	0.0461	0.0672
0.0457	0.011	0.0136	0.6427	0.0519	0.0734
0.0762	0.0176	0.0226	0.7138	0.0584	0.0807
0.1066	0.0244	0.0320	0.7854	0.0654	0.087
0.1371	0.0315	0.0413	0.8566	0.0714	0.0943
0.1674	0.0375	0.050	0.9282	0.0773	0.1024
0.1978	0.0444	0.0597	0.9999	0.0832	0.1118
0.2282	0.0512	0.0677	1.0709	0.0892	0.1207
0.2585	0.0586	0.0756	1.1427	0.0948	0.1303
0.2885	0.0669	0.083	1.2137	0.1002	0.1390
0.319	0.0746	0.0897		$(1.502x10^{-6})$	
0.3493	0.0825	0.0967	0.1009	0.0296	0.0378
0.3795	0.0918	0.1032	0.151	0.0447	0.0577
0.4097	0.1003	0.110	0.2012	0.0609	0.0758
0.455	0.1140	0.1198	0.2513	0.0779	0.0904
	$(3.893x10^{-6})$		0.3012	0.0959	0.1038
0.0159	0.0157	0.0143	0.3507	0.1147	0.118
0.0317	0.0318	0.0299	0.0505	0.0149	0.0181
0.0475	0.0475	0.0460		$(5.082x10^{-6})$	
0.0633	0.0639	0.0630			
0.0791	0.080	0.0790			
0.1026	0.1022	0.1004			
	$(1.48x10^{-5})$				

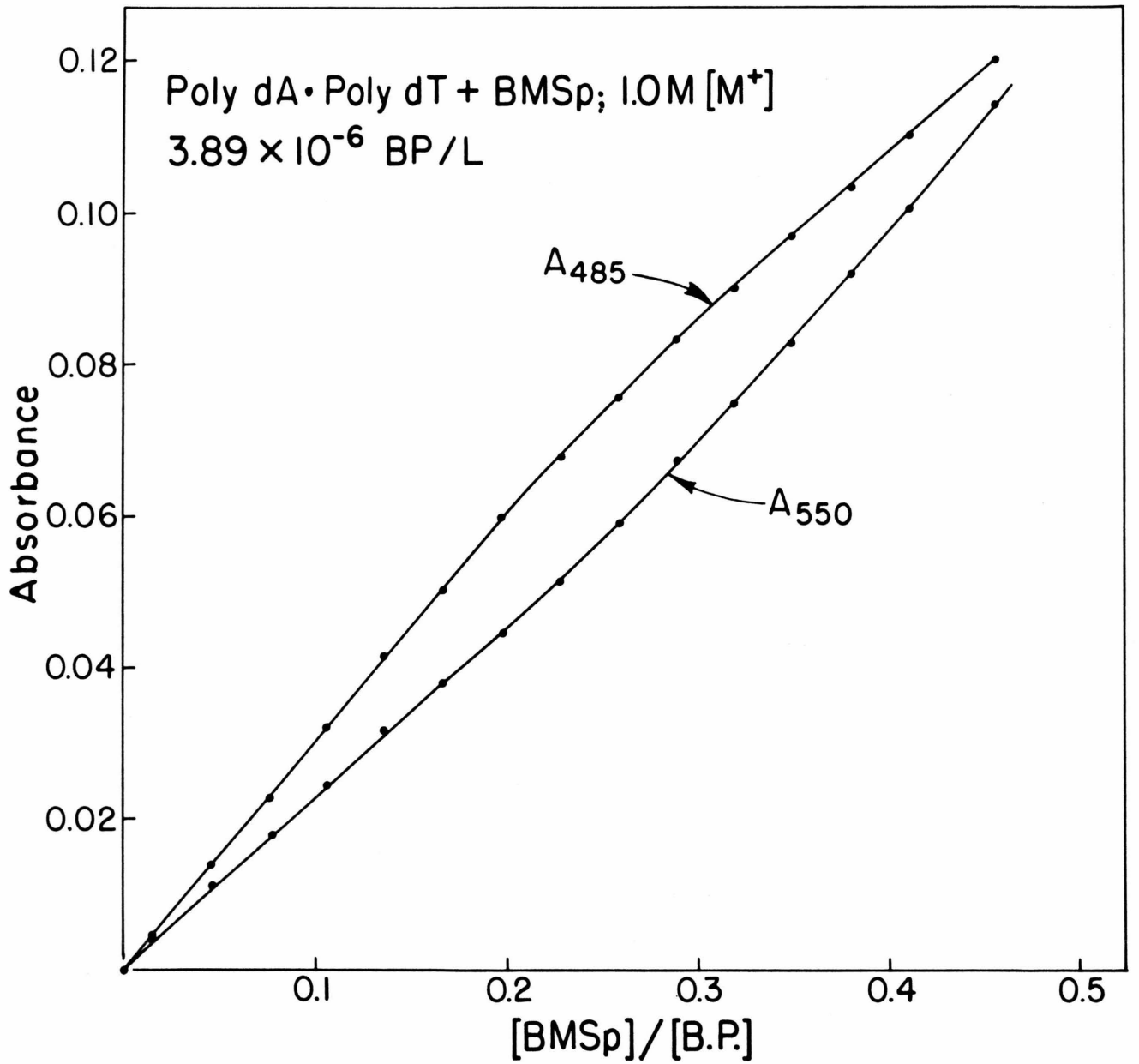
Figures 10-14

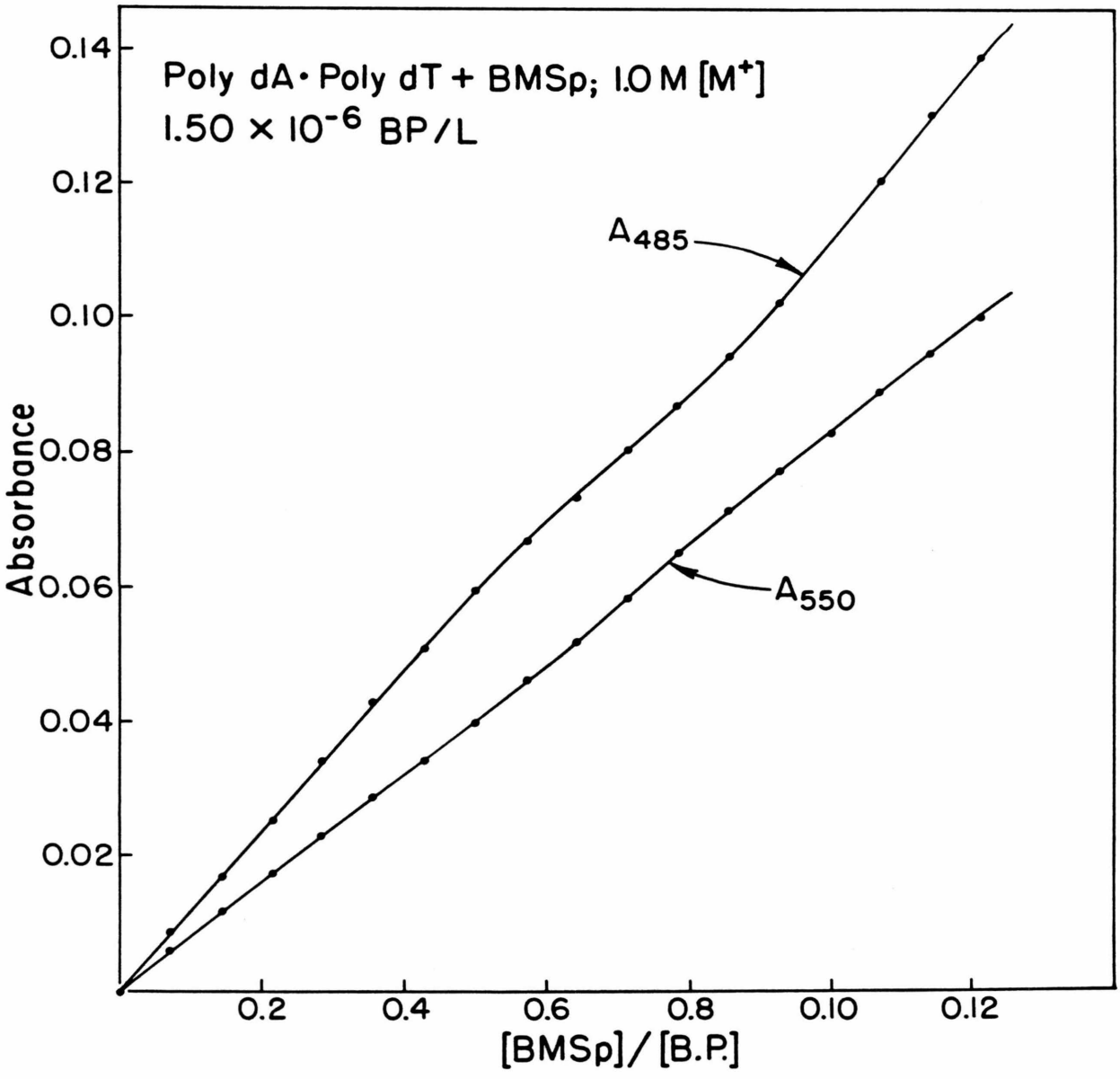
Binding of BMSp at $[M^+]$ = 1.0 to the indicated concentration of poly(dA:dT). The absorbance at 485 and 550 nm (path length 10 cm) for increasing additions of BMSp is plotted against the corresponding BMSp/BP ratio.


Poly dA • Poly dT + BMSp; 1.0M [M+]
3.01 × 10-5 BP/L
A550
A485
Absorbance
0.06
0.05
0.04
0.03
0.02
0.01
0.005
0.010
0.015
0.020
0.025
0.030
[BMSp]/[B.P.]


Poly dA • Poly dT + BMSp; 1.0 M [M+]
1.48 × 10-5 BP/L
A550
A485
Absorbance
0.12
0.10
0.08
0.06
0.04
0.02
0.02
0.04
0.06
0.08
0.10
0.12
[BMSp] / [B.P.]

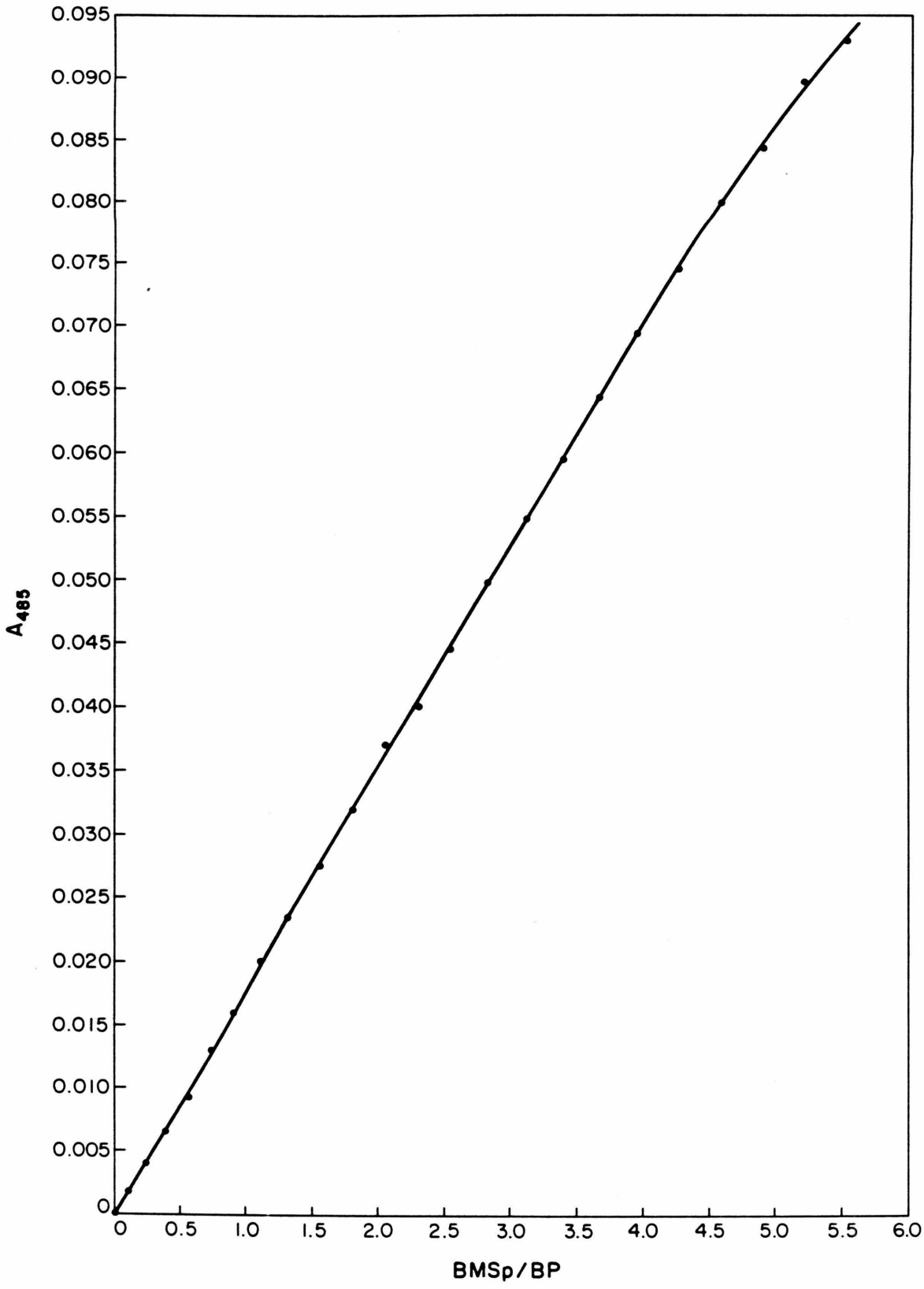

Poly dA• Poly dT + BMSp; 1.0M [M+]
5.08 × 10-6 BP/L
A485
A550
Absorbance
0.12
0.10
0.08
0.06
0.04
0.02
0.1
0.2
0.3
0.4
[BMSp]/[B.P.]

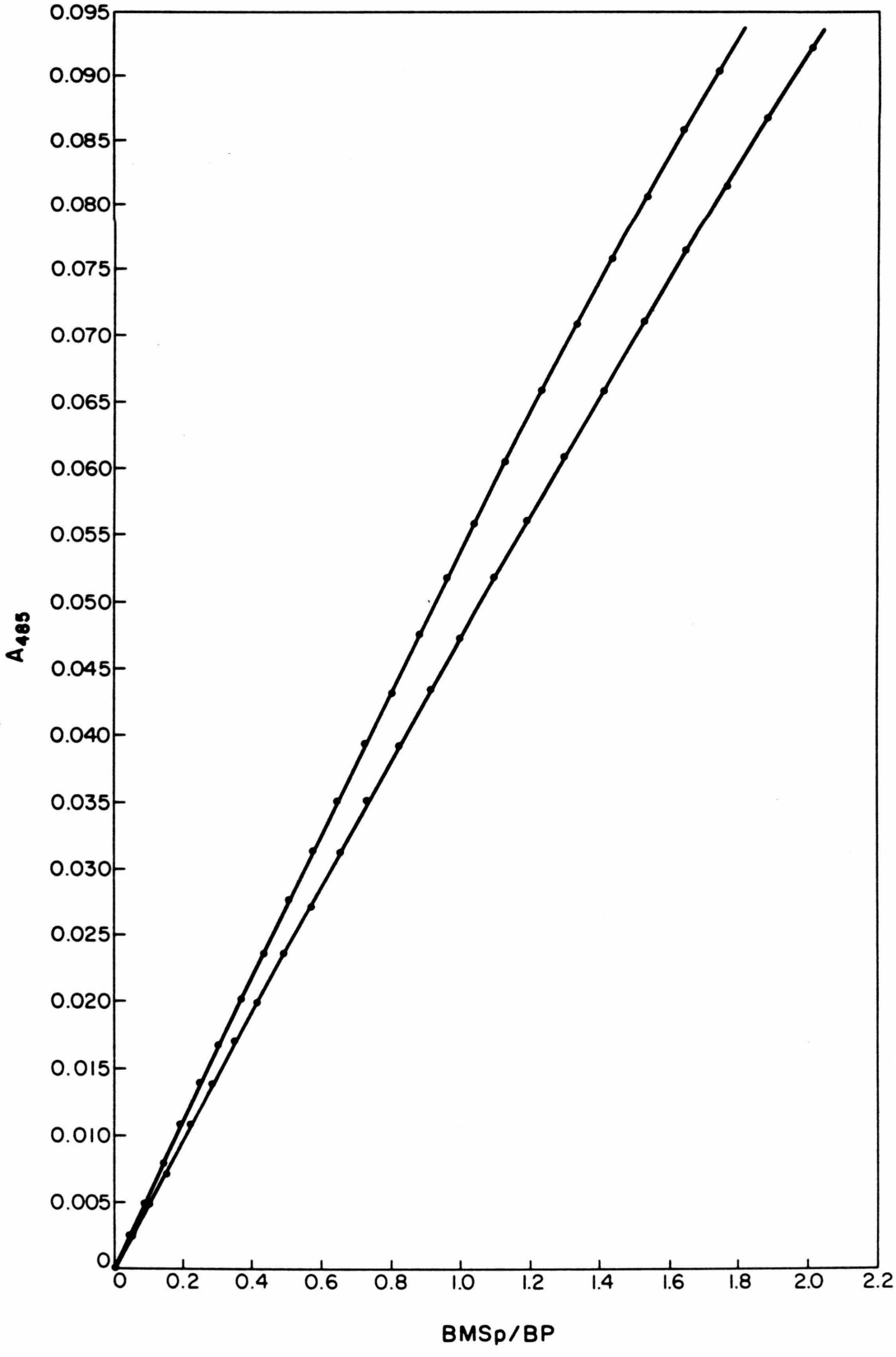

Poly dA• Poly dT + BMSp; 1.0M [M+]
3.89 × 10-6 BP/L
A485
A550
Absorbance
0.12
0.10
0.08
0.06
0.04
0.02
0.1
0.2
0.3
0.4
0.5
[BMSp]/[B.P.]

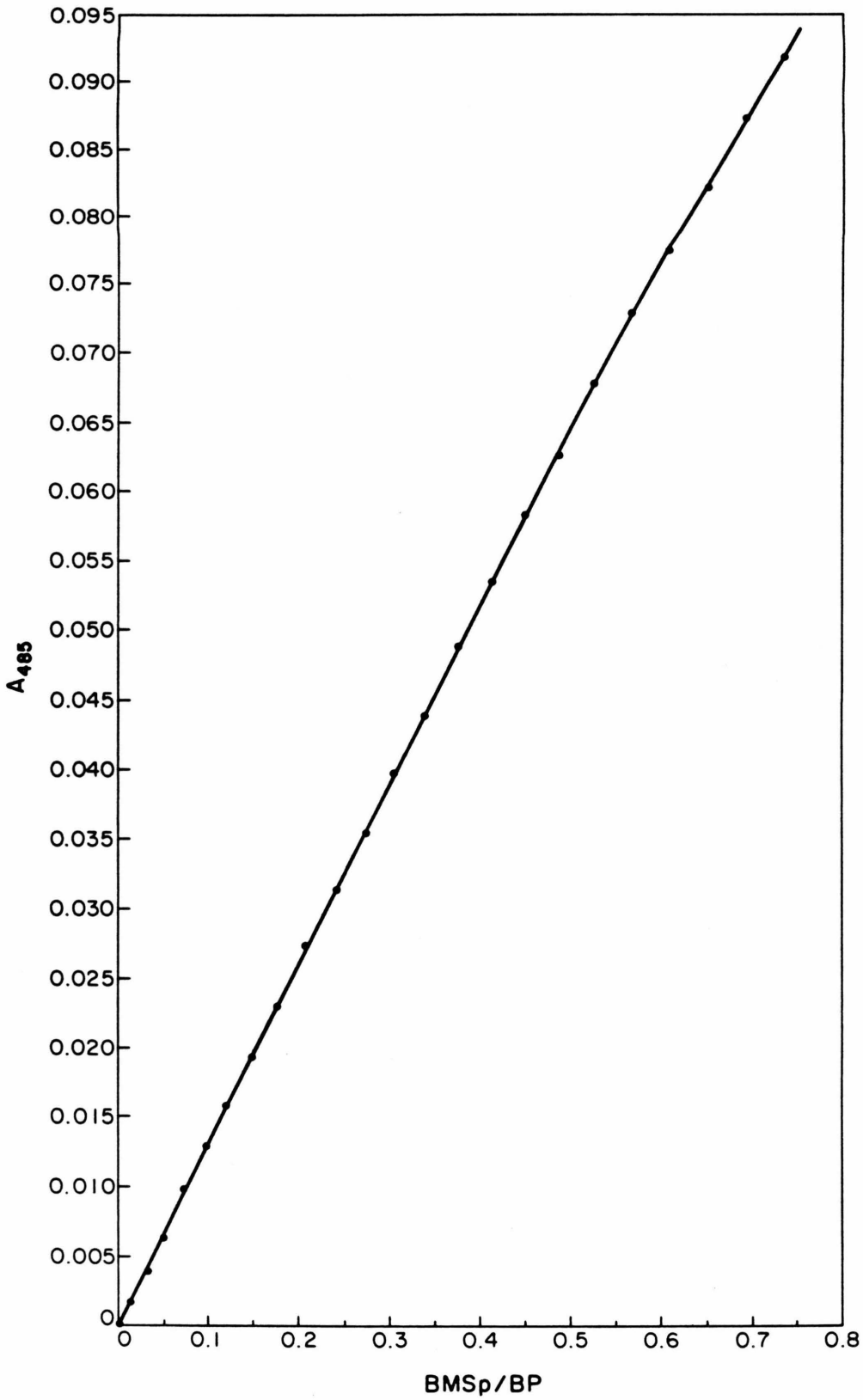

Poly dA • Poly dT + BMSp; 1.0 M [M+]
1.50 × 10-6 BP/L
A485
A550
Absorbance
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.2
0.4
0.6
0.8
0.10
0.12
[BMSp]/[B.P.]

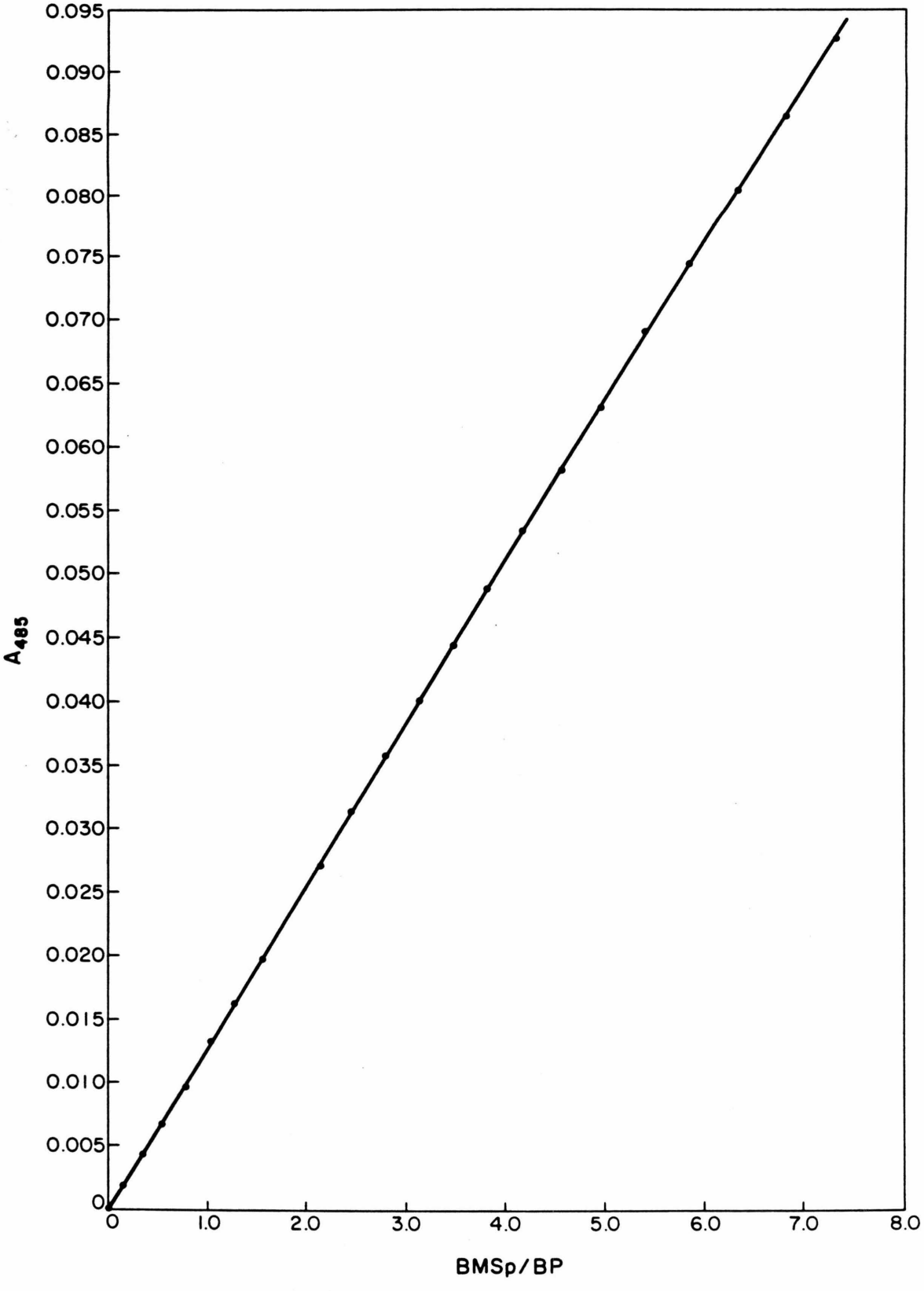
Figures 15-18

Spectrophotometric titration of polydAdT with BMSp at 1 M^+. The absorbance of BMSp at 485 nm (10 cm path length) is plotted against the corresponding BMSp/BP ratio. Five representative examples are shown and all measurements are summarized in Table V of the Appendix. The concentration of dAdT in base pairs is: Fig. 20 upper curve 1.144×10^{-6} M, lower curve 4.955×10^{-7} M; the BMSp/BP axis corresponds to the lower curve; Fig. 21, 1.3941×10^{-7} M; Fig. 22, 1.392×10^{-6} M; Fig. 23, 1.856×10^{-7} M.


0.095
0.090
0.085
0.080
0.075
0.070
0.065
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0
A485
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
BMSp/BP

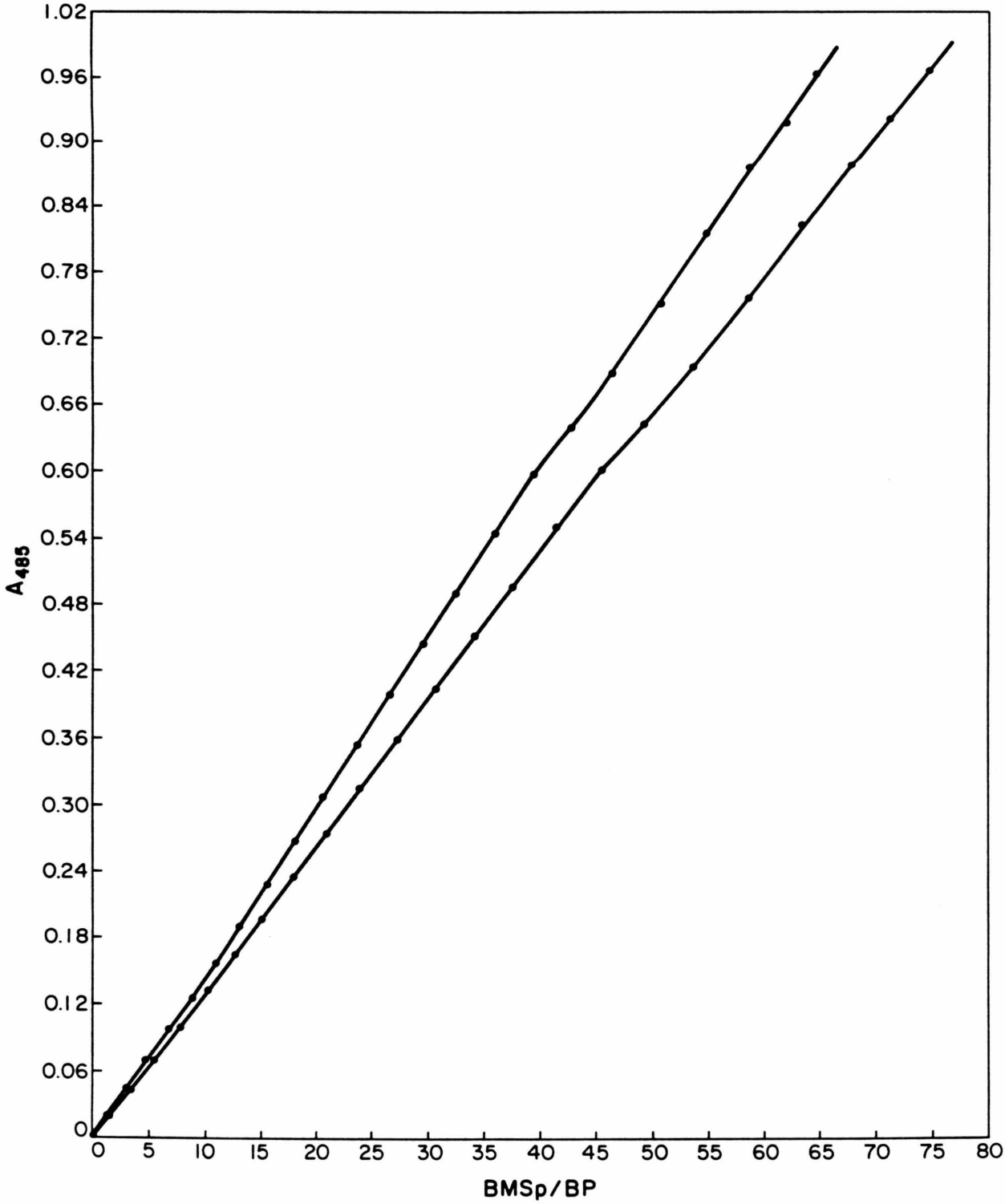

0.095
0.090
0.085
0.080
0.075
0.070
0.065
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0
A465
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
BMSp/BP

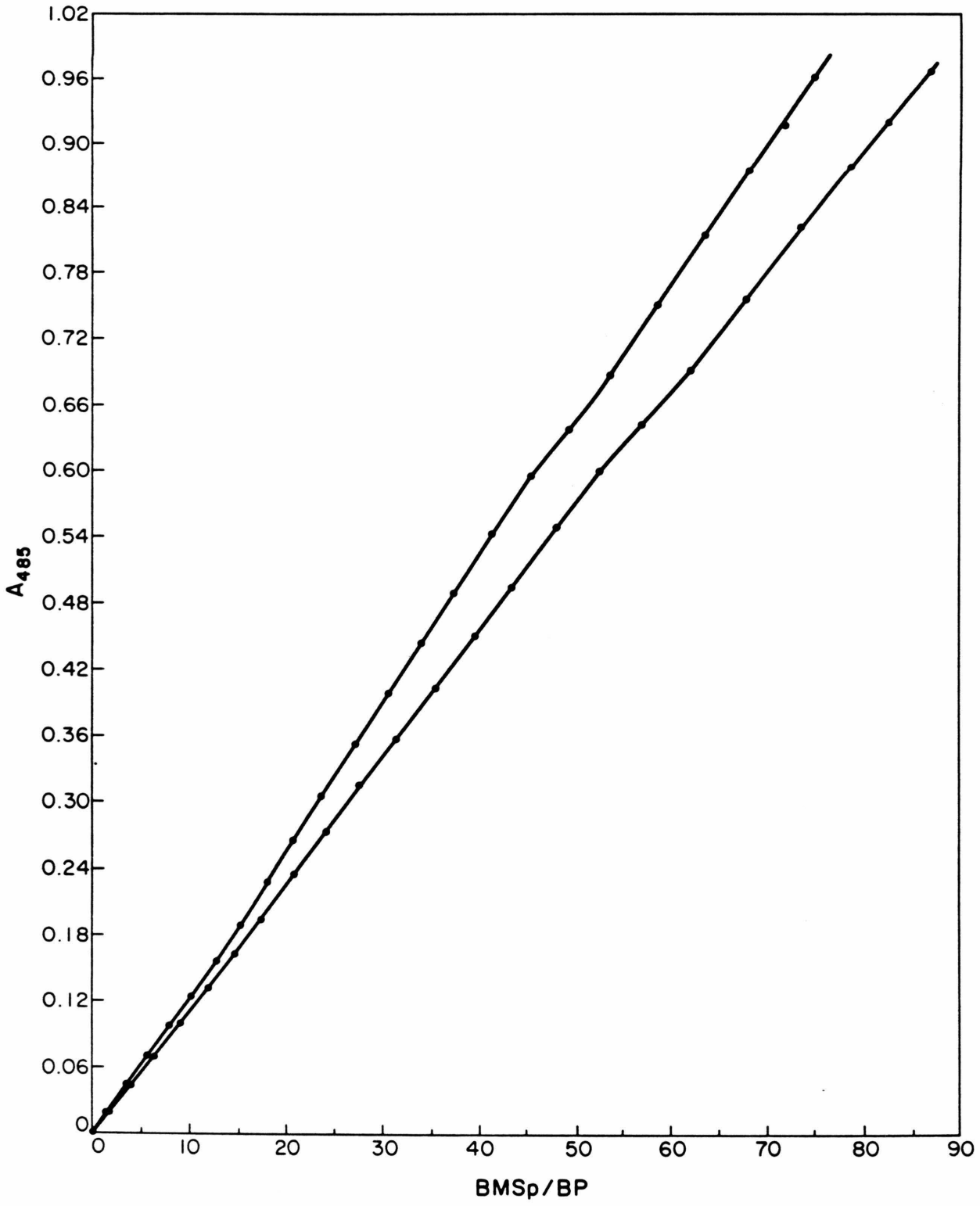

0.095
0.090
0.085
0.080
0.075
0.070
0.065
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0
A485
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
BMSp/BP

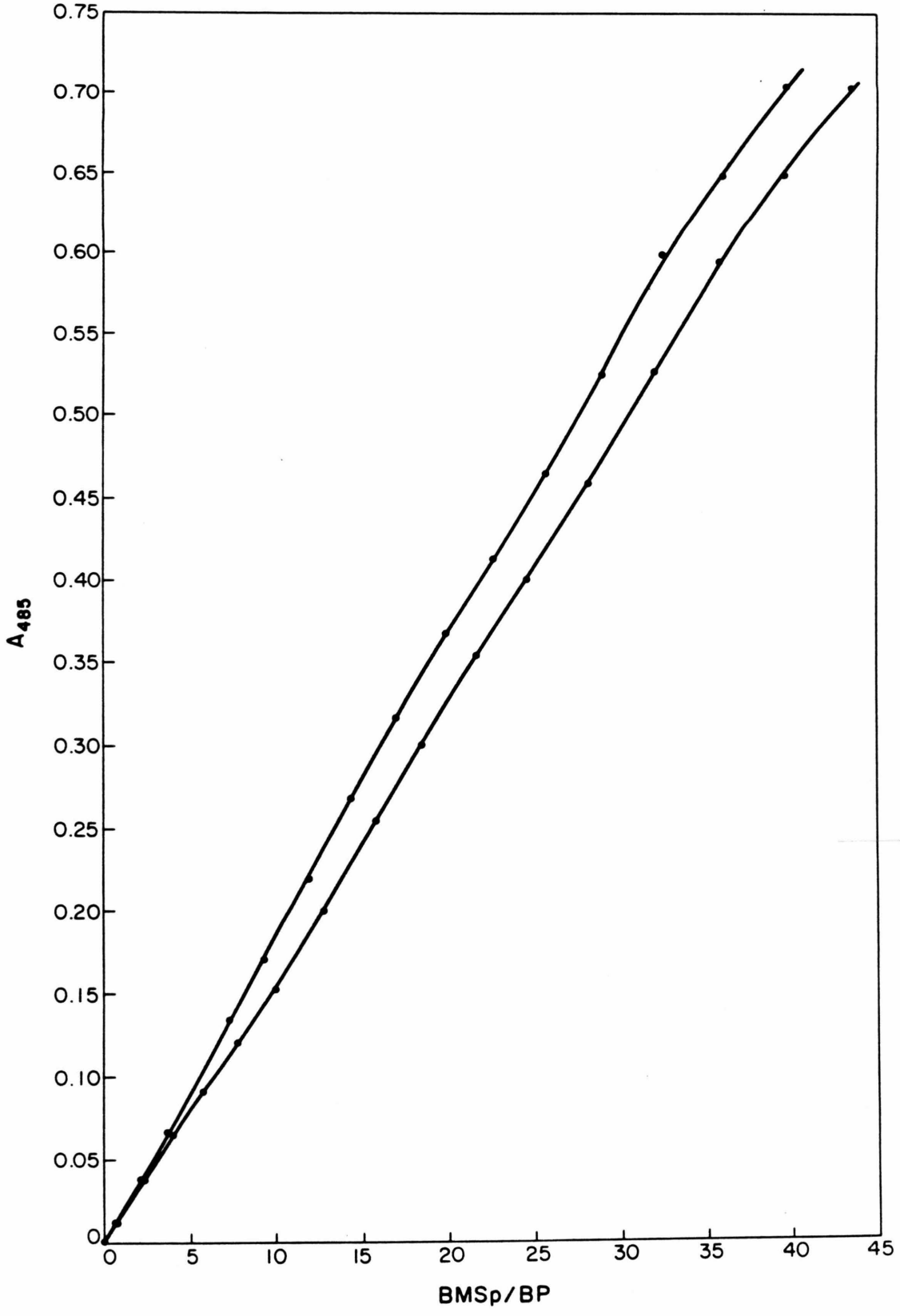

0.095
0.090
0.085
0.080
0.075
0.070
0.065
0.060
0.055
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0
A485
0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
BMSp/BP

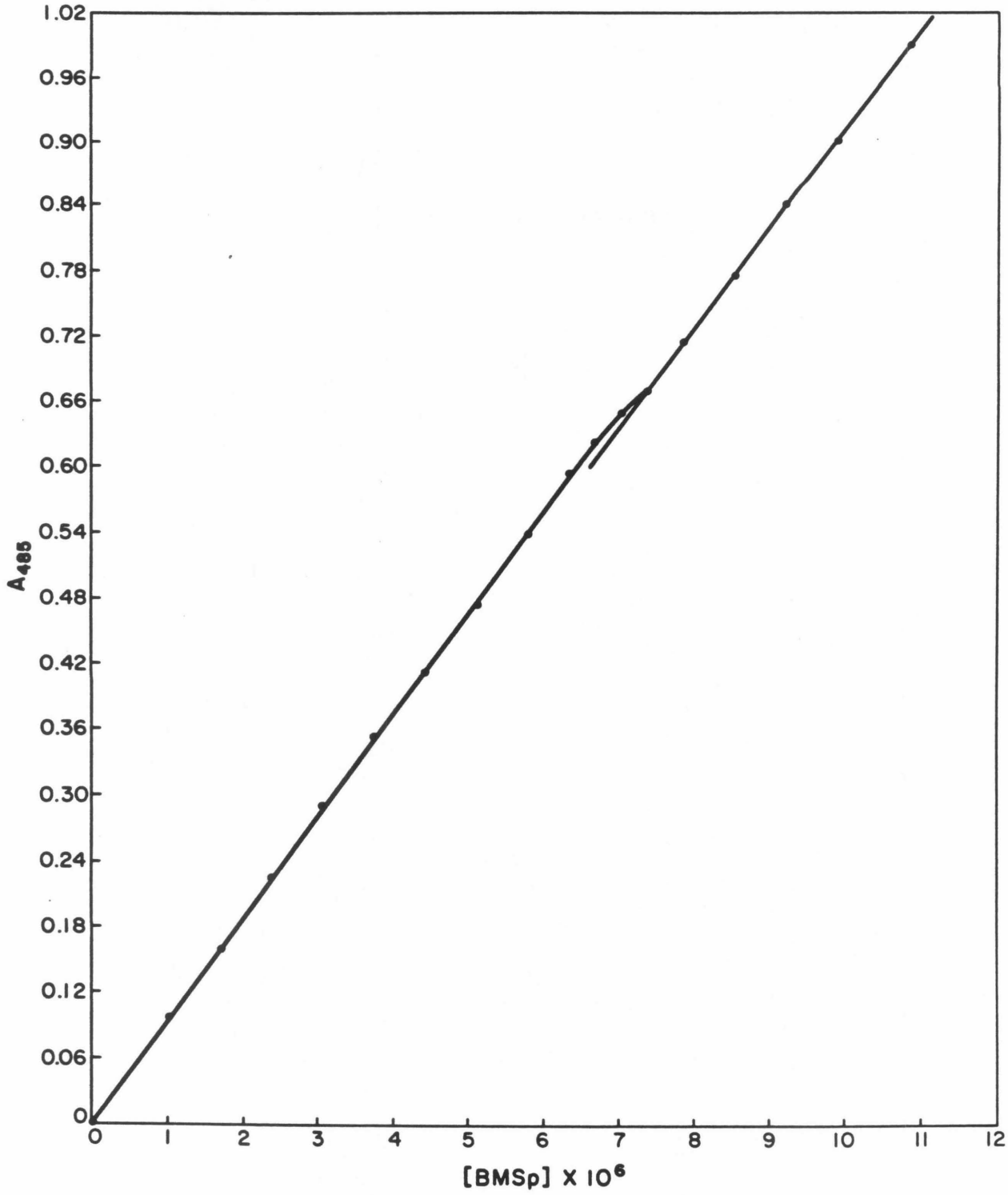
Figures 19-22

Spectrophotometric titration of polydAdT with BMSp at 1 M^{+}. The absorbance of BMSp at 485 nm (10 cm path length) is plotted against the corresponding BMSp/BP ratio. Six representative examples are shown and all measurements are summarized in Table V of the Appendix. The concentration of dAdT in base pairs is: Fig. 24, upper curve 8.049 x 10^{-7} M, lower curve 1.394 x 10^{-7} N, the BMSp/BP axis corresponds to the lower curve; Fig. 25, upper curve 1.559 x 10^{-7} M, lower curve 2.998 x 10^{-7} M, the BMSp/BP axis corresponds to the upper curve; Fig. 26, upper curve 1.903 x 10^{-7} M, lower curve 1.424 x 10^{-6} M, the BMSp/BP axis corresponds to the upper curve; Fig. 27, the unusual dip in the absorbance at 0.65 is due to nonlinearity in the slidewire response of the Cary 14 spectrophotometer.


1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A485
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
BMSp/BP


1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A485
0
10
20
30
40
50
60
70
80
90
BMSp/BP


0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
A485
0
5
10
15
20
25
30
35
40
45
BMSp/BP


1.02
0.96
0.90
0.84
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.12
0.06
0
A485
0
1
2
3
4
5
6
7
8
9
10
11
12
[BMSp] X 10^6

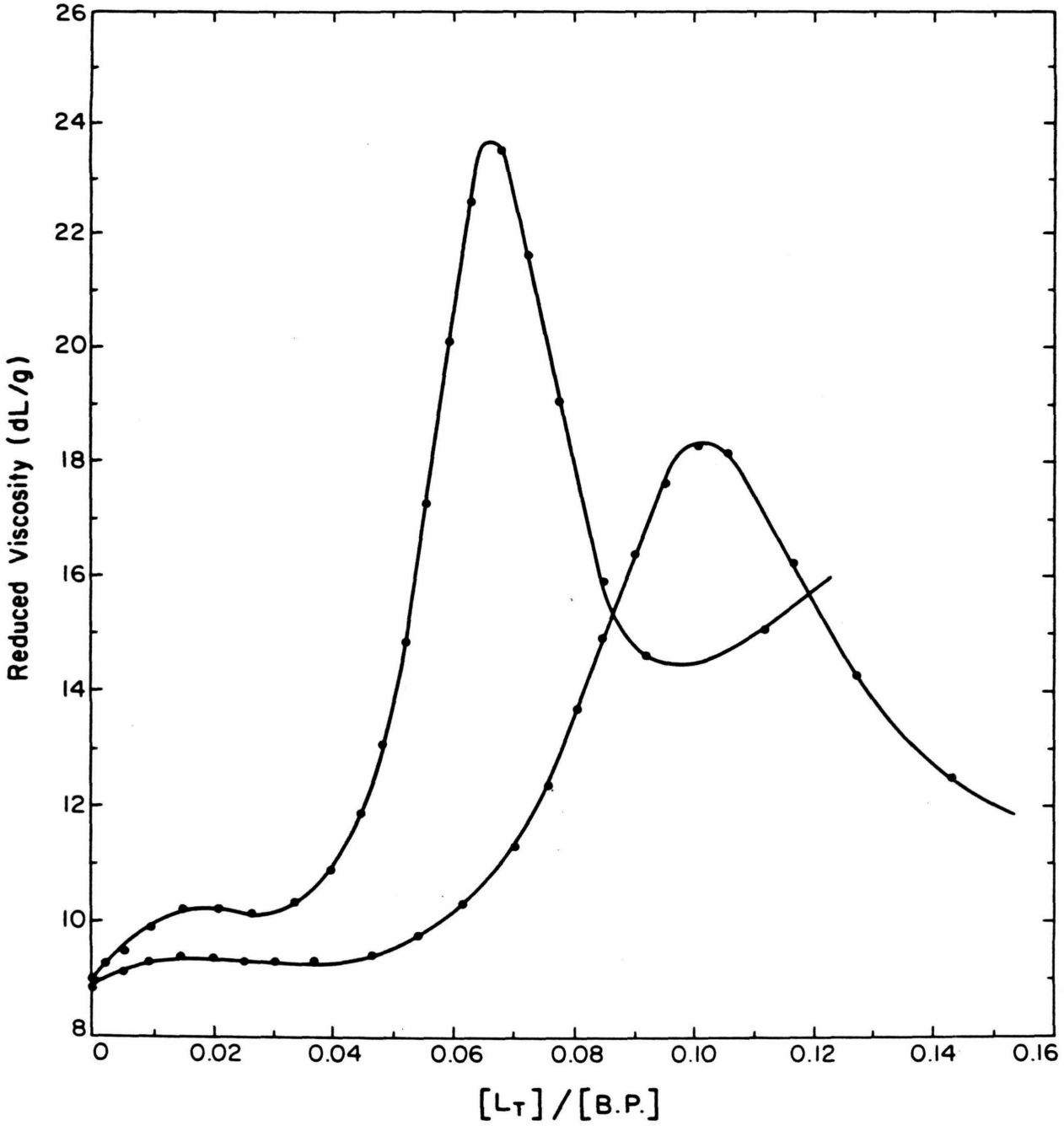
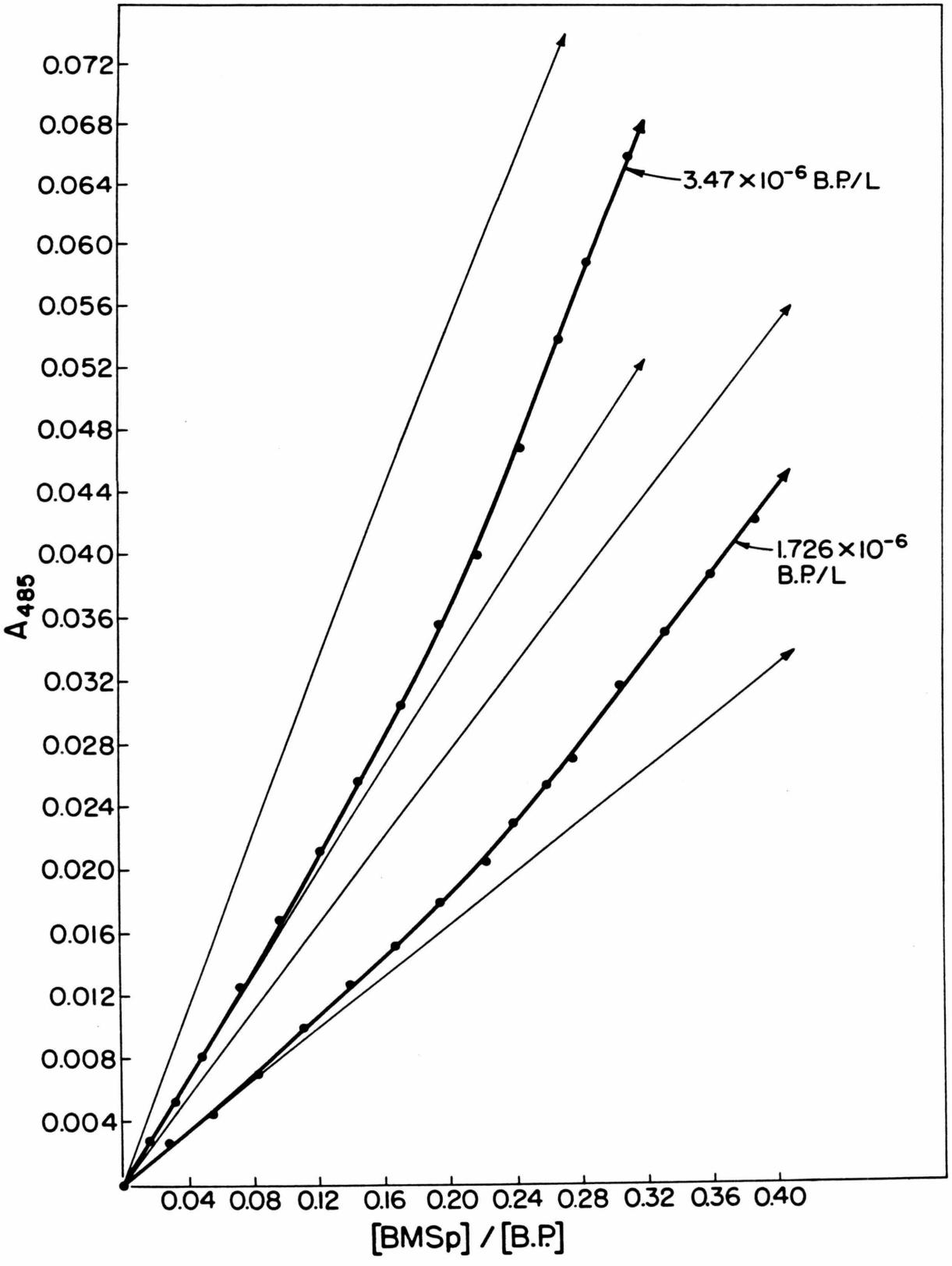


Figure 23
Viscometric titration of closed circular PM2 DNA with BMSp (upper curve) and EB (lower curve) at $[M^+] = 0.075$. The reduced viscosity in deciliter/gram is plotted against the ratio of total ligand added (L_T) and the concentration of DNA in base pairs.

Figure 24

Binding of BMSp to sonicated calf thymus DNA at 1.0 M^+. The absorbance at 485 nm (path length 10 cm) for increasing additions of BMSp is plotted against the corresponding BMSp/BP ratio. The results for two different concentrations of DNA are shown. Thin lines indicate the absorbance which results when all added BMSp is unbound or bound.


0.072
0.068
0.064
0.060
0.056
0.052
0.048
0.044
0.040
0.036
0.032
0.028
0.024
0.020
0.016
0.012
0.008
0.004
A485
3.47×10-6 B.P/L
1.726×10-6
B.P/L
0.04 0.08 0.12 0.16 0.20 0.24 0.28 0.32 0.36 0.40
[BMSp] / [B.P.]

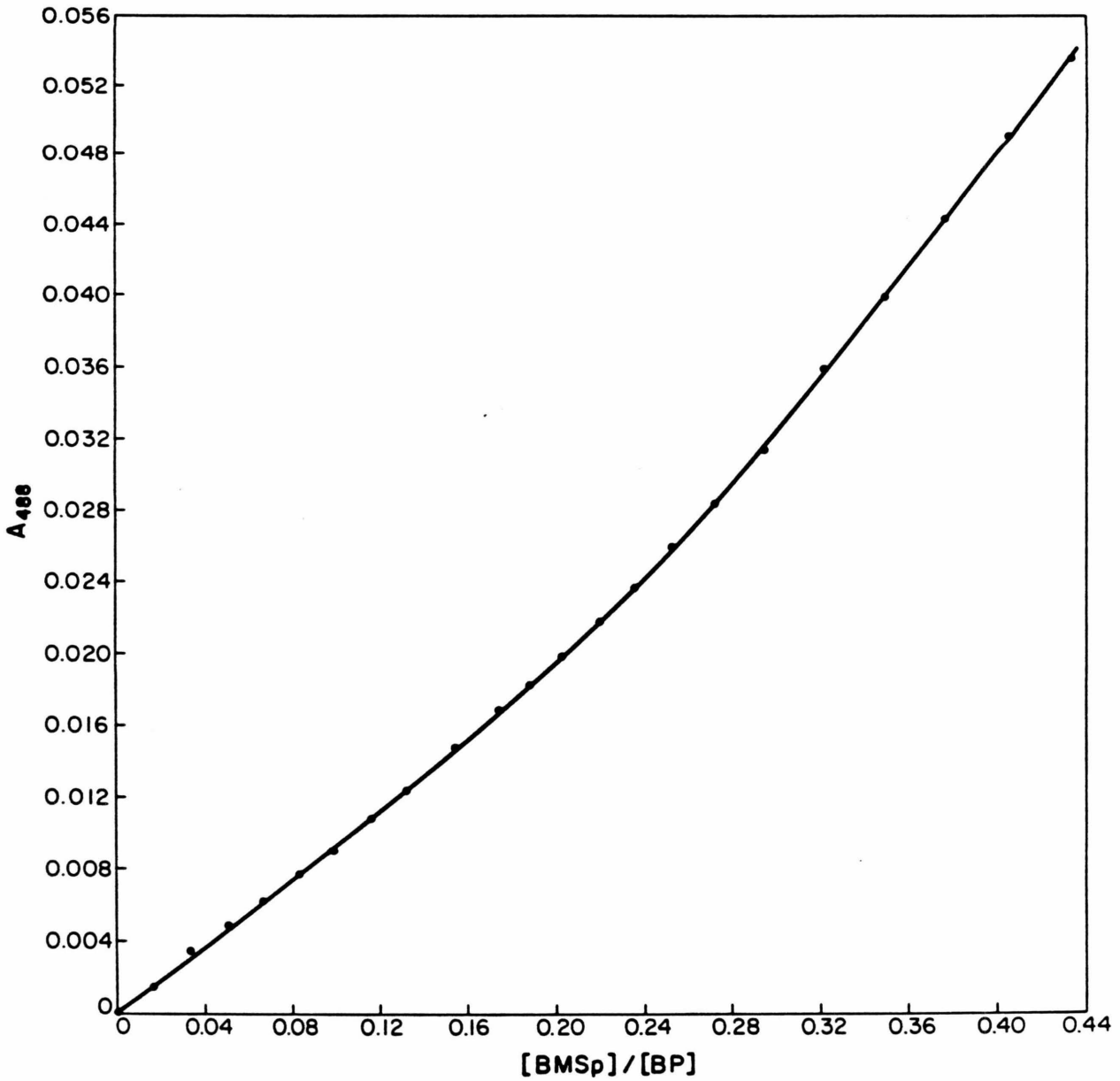
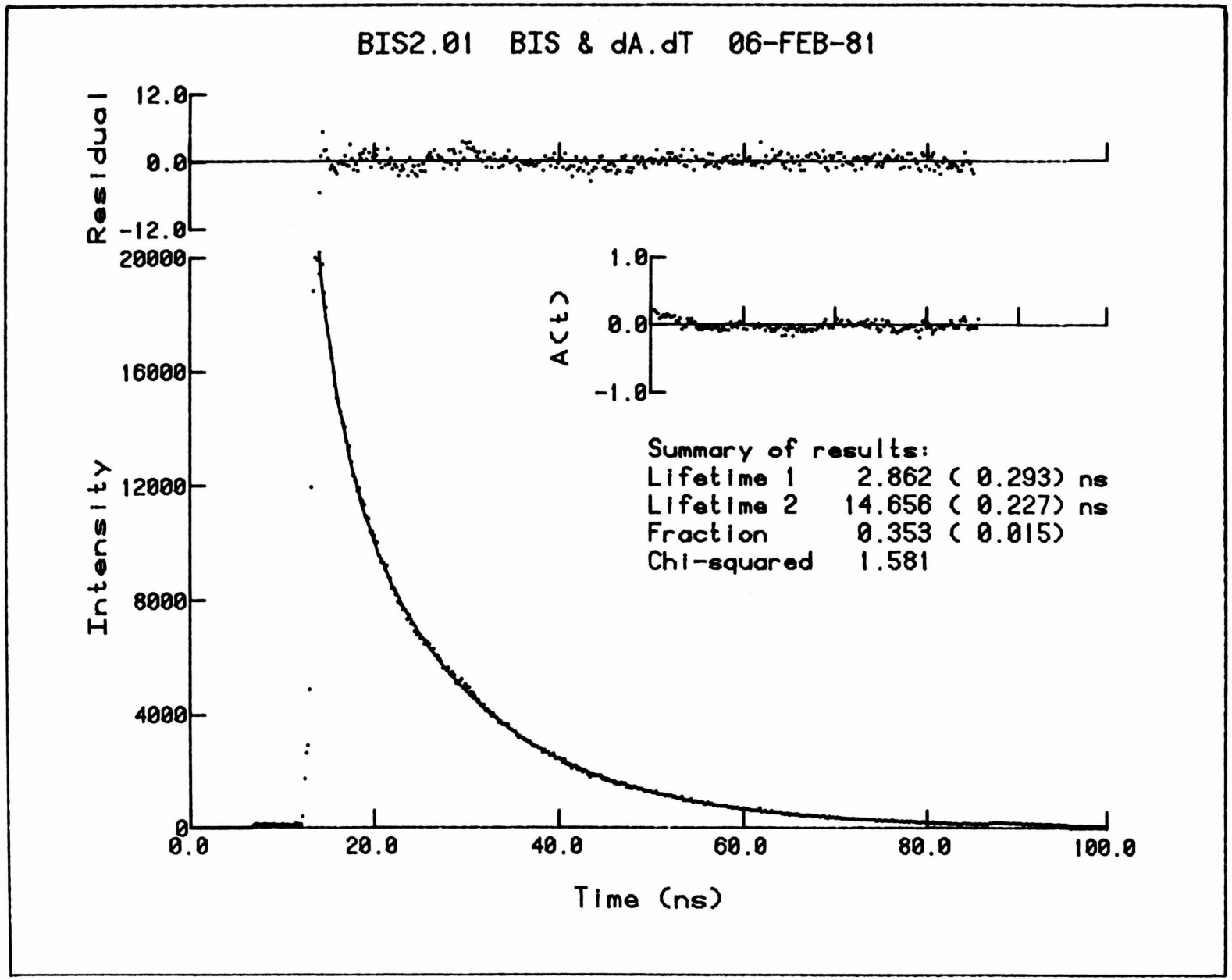


Figure 25
Binding of BMSp to poly(dC-dG) at 1.0 M^+. The absorbance at 488 nm (path length 10 cm) for increasing additions of BMSp is plotted against the corresponding BMSp/BP ratio. The concentration of nucleic acid was 2.042×10^{-6} BP.

Figure 26

Fluorescence lifetime measurement of BMSp bound to A) sonicated calf thymus DNA and B) polydAdT at 1 M^+ in D_2O. The concentration of DNA was .44 x 10^{-4} BP/L and the corresponding BMSp binding density was 2.5 x 10^{-3} BMSp/BP. The observed fluorescence intensity is plotted against time (nano seconds), the best fit to the decay being represented by the solid line. Also shown is the difference between observed and calculated data (residuals) as well as an autocorrelation function of the residuals.


BIS2.01 BIS & dA.dT 06-FEB-81
Residual
12.0
0.0
-12.0
Intensity
20000
16000
12000
8000
4000
0
A(t)
1.0
0.0
-1.0
Summary of results:
Lifetime 1 2.862 (0.293) ns
Lifetime 2 14.656 (0.227) ns
Fraction 0.353 (0.015)
Chi-squared 1.581
0.0
20.0
40.0
60.0
80.0
100.0
Time (ns)

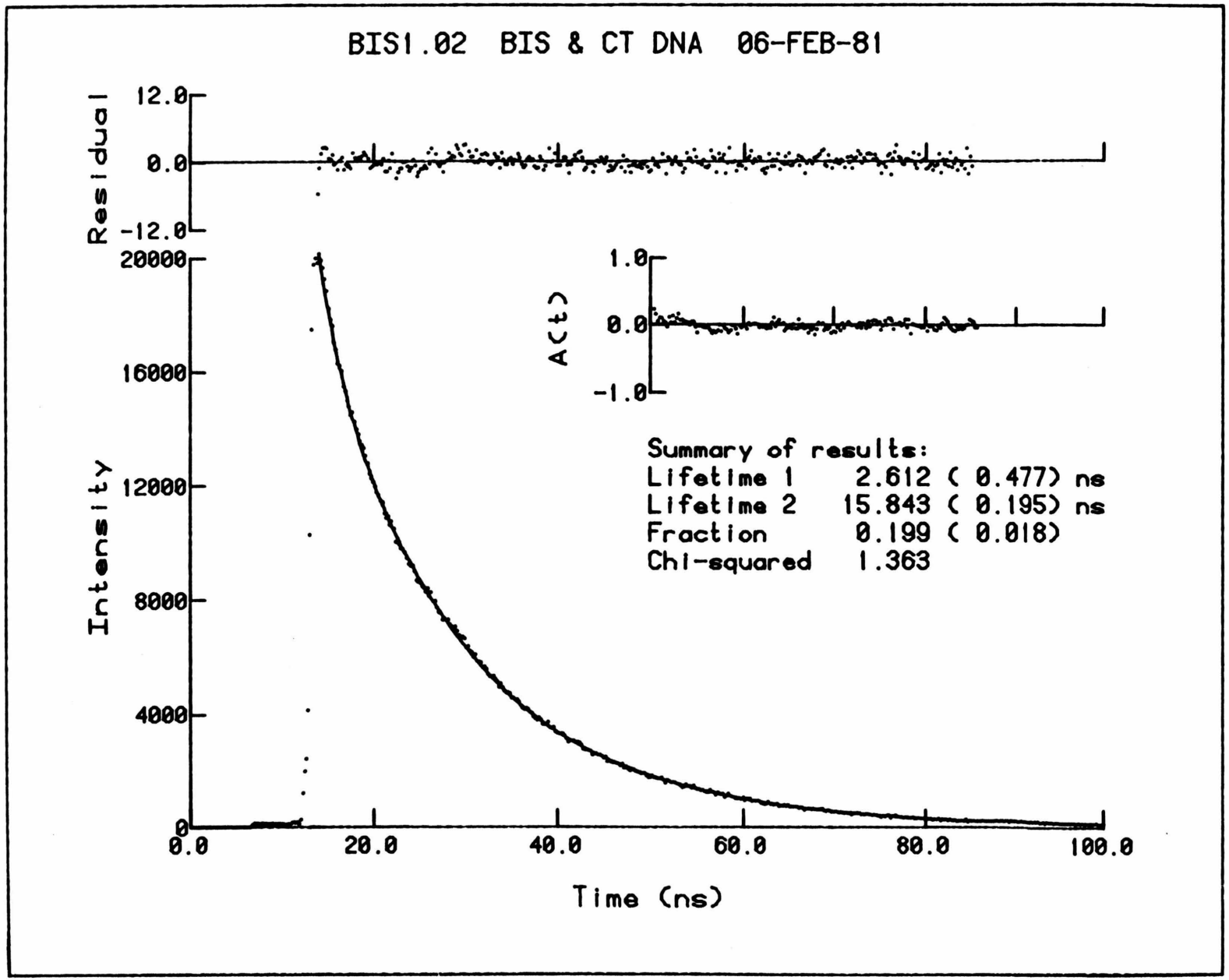

BIS1.02 BIS & CT DNA 06-FEB-81
Residual
12.0
0.0
-12.0
20000
16000
12000
8000
4000
0
Intensity
A(t)
1.0
0.0
-1.0
Summary of results:
Lifetime 1 2.612 (0.477) ns
Lifetime 2 15.843 (0.195) ns
Fraction 0.199 (0.018)
Chi-squared 1.363
0.0
20.0
40.0
60.0
80.0
100.0
Time (ns)

Figure 27

Fluorescence lifetime measurement of BMSp at $1M^{+}$ in D_2O. The concentration of BMSp was 4.1×10^{-6} BP/L. The observed fluorescence intensity is plotted against time, the best fit to the decay being represented by the solid line. Also shown is the difference between observed and calculated data (residuals) as well as an autocorrelation function of the residuals.

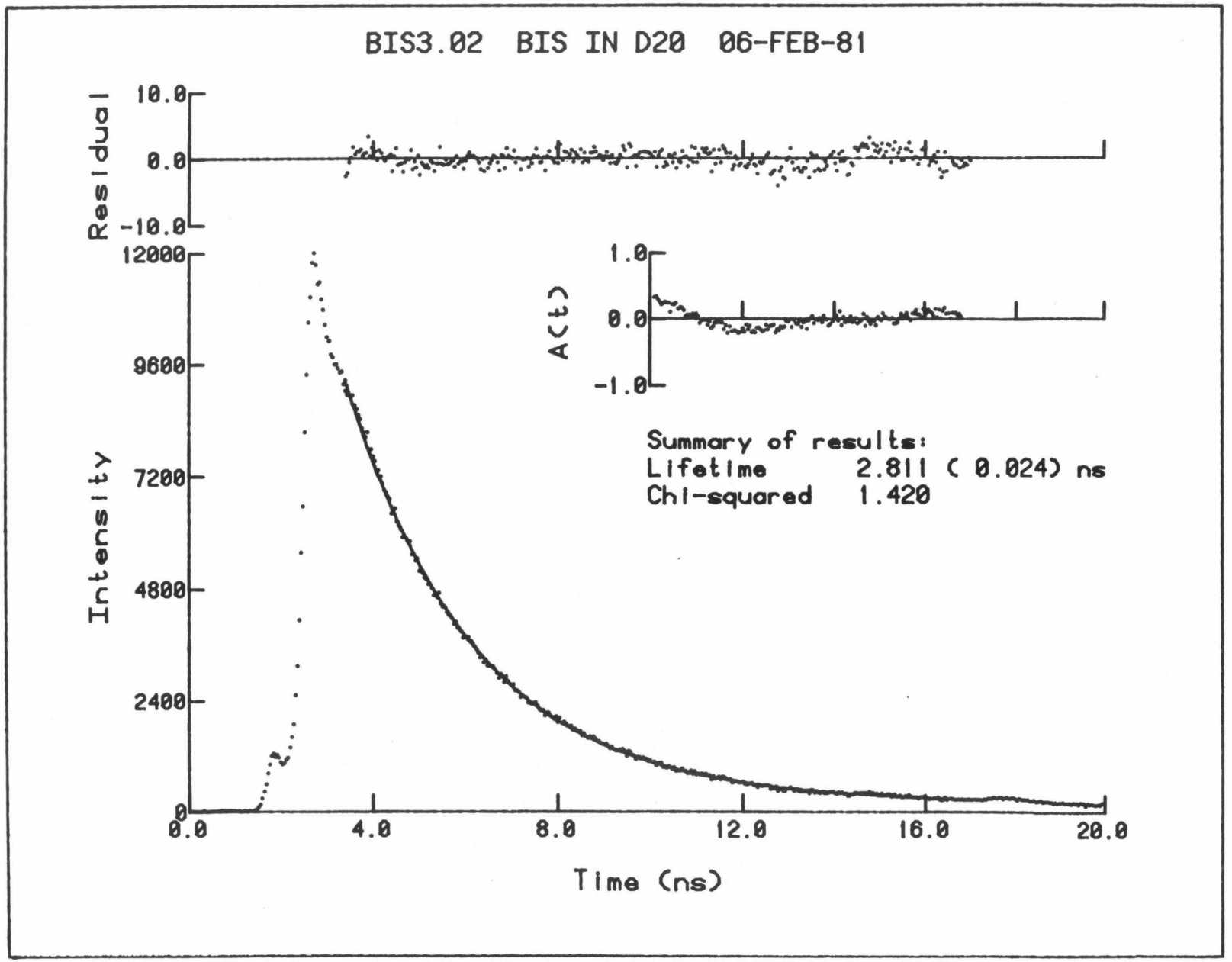
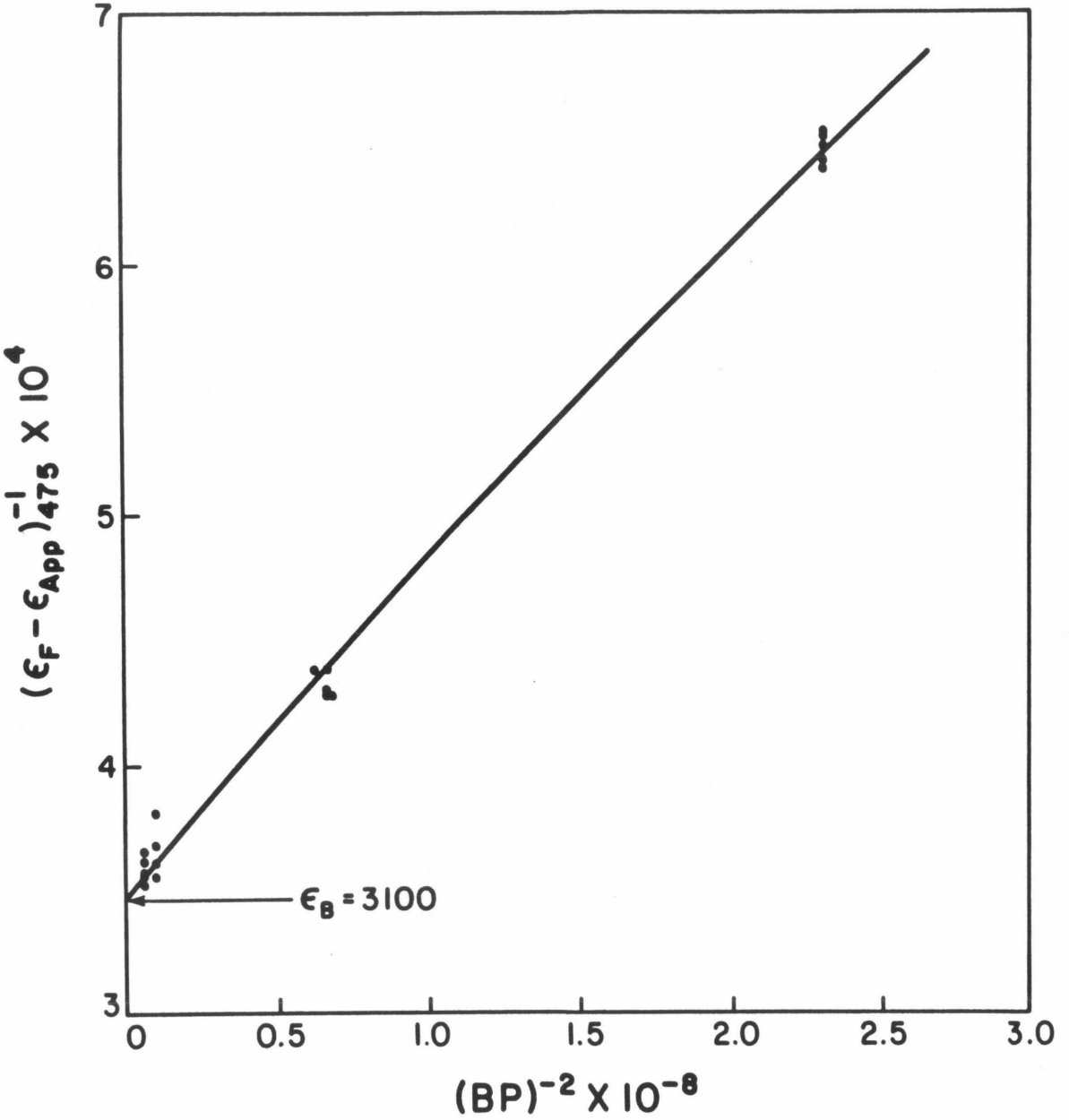

BIS3.02 BIS IN D20 06-FEB-81
Residual
10.0
0.0
-10.0
A(t)
1.0
0.0
-1.0
Summary of results:
Lifetime 2.811 (0.024) ns
Chi-squared 1.420
Intensity
12000
9600
7200
4800
2400
0
0.0
4.0
8.0
12.0
16.0
20.0
Time (ns)

Figure 28

Determination of the extinction coefficient of ethidium bromide (EB) bound to polydCdG at 1 M^+. The inverse of the difference between, ε_F, the extinction coefficient of free drug and the apparent (observed) extinction coefficient ε_{app} is plotted against the square of the inverse of the dCdG concentration in base pairs. Measurements are from Table IV of the Appendix.


7
6
5
4
3
0
0.5
1.0
1.5
2.0
2.5
3.0
$(\epsilon_F - \epsilon_{App})^{-1}_{475} \times 10^4$
$(BP)^{-2} \times 10^{-8}$
$\epsilon_B = 3100$

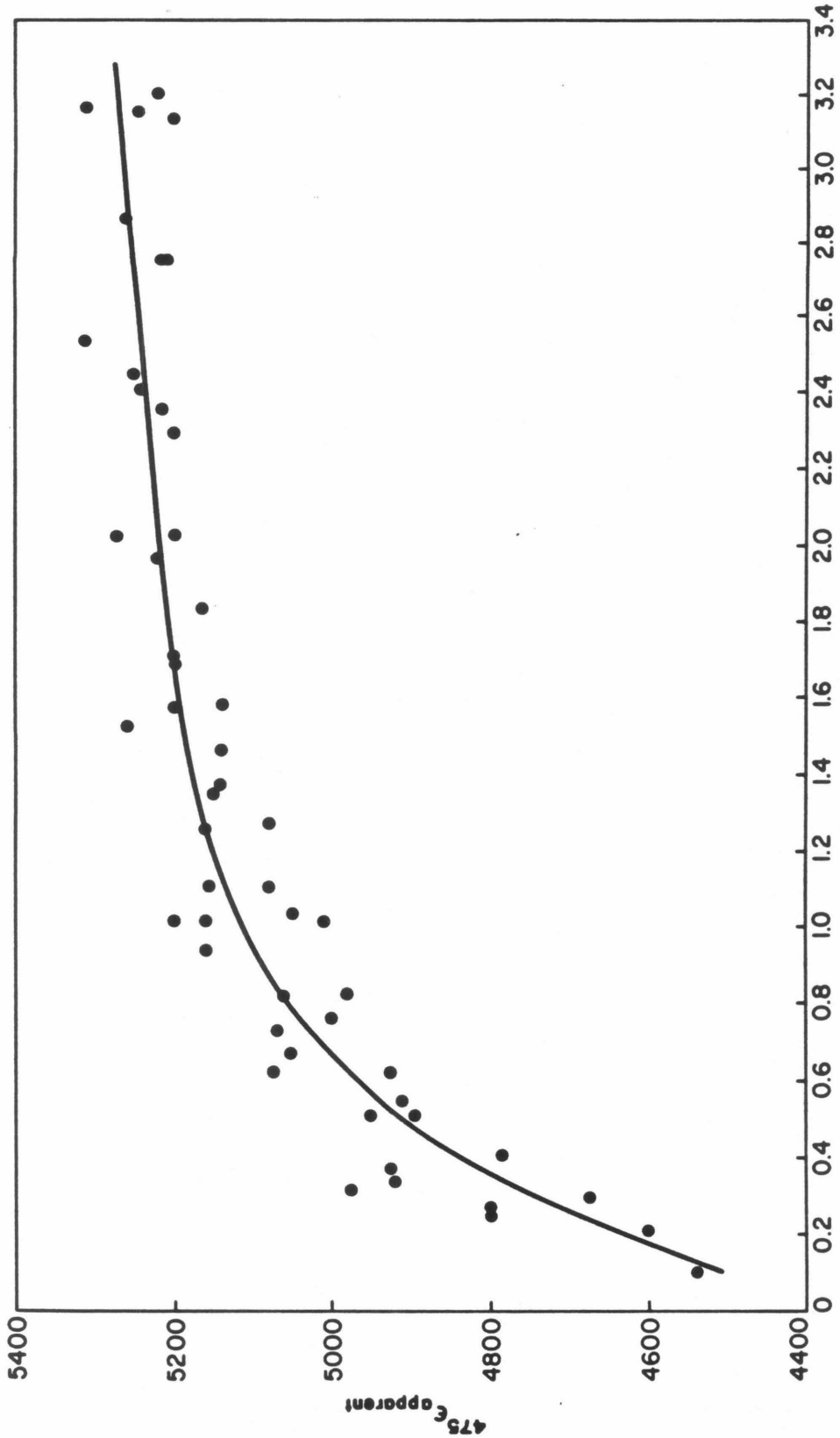


Figure 29 - Determination of the extinction coefficient of EB bound to polydCdG at 1 M^+. Measurements are from Table IV of the Appendix.

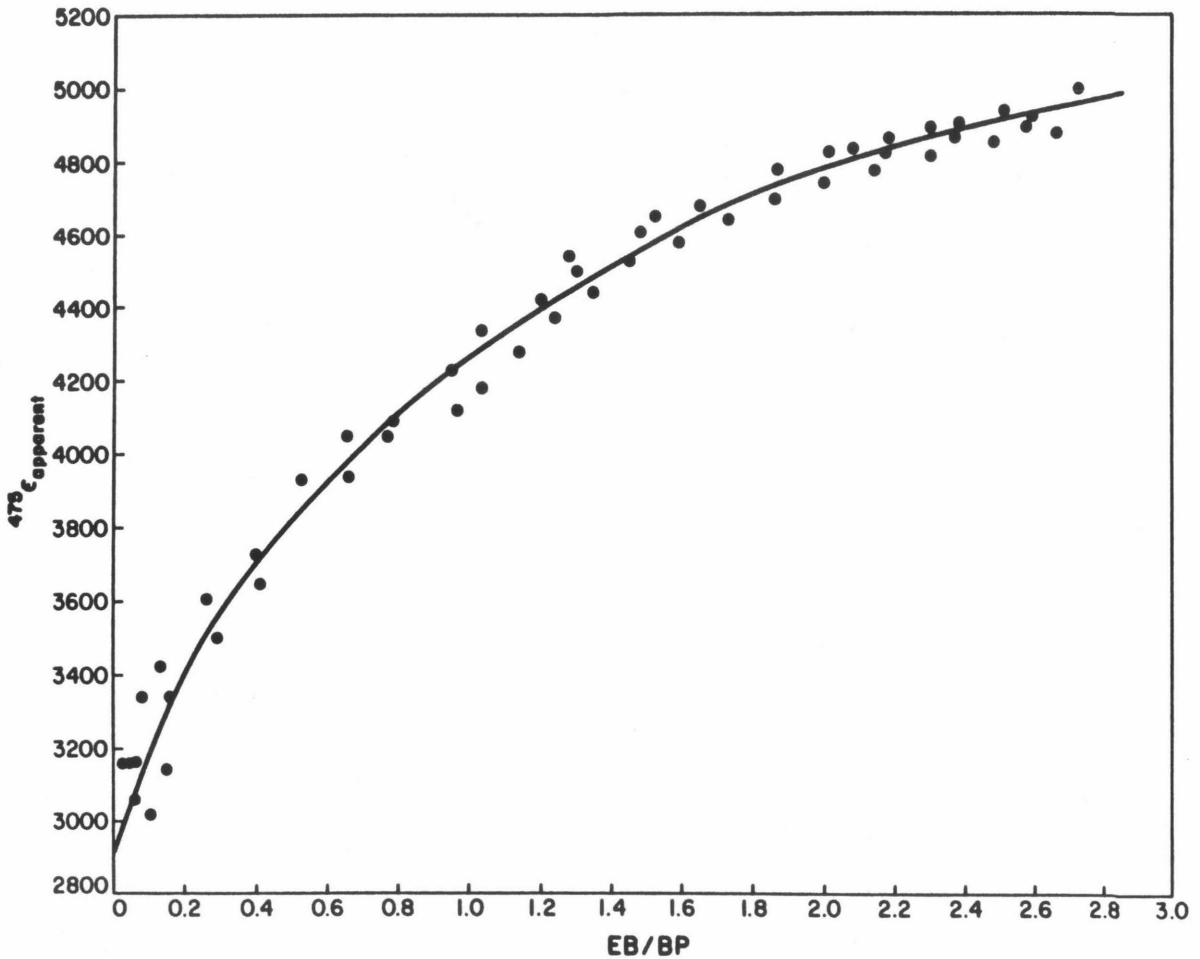


Figure 30 - Determination of the extinction coefficient of EB bound to polyd(C-G) at 1 M^+. Measurements are from Table III of the Appendix.

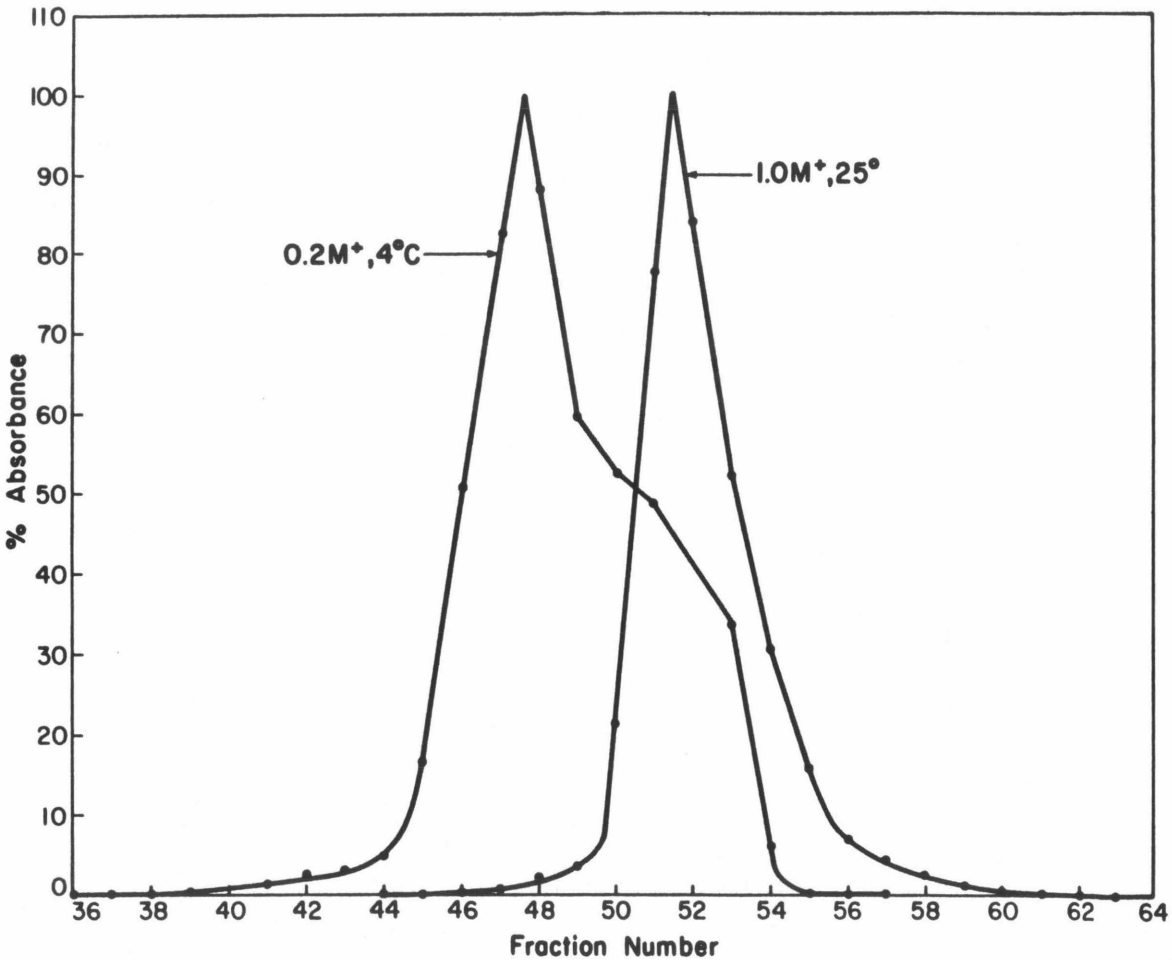


Figure 31 - Comparison of the elution of lambda DNA, restricted by Hind III, from a Sepharose 4B column at the indicated salt and temperature.

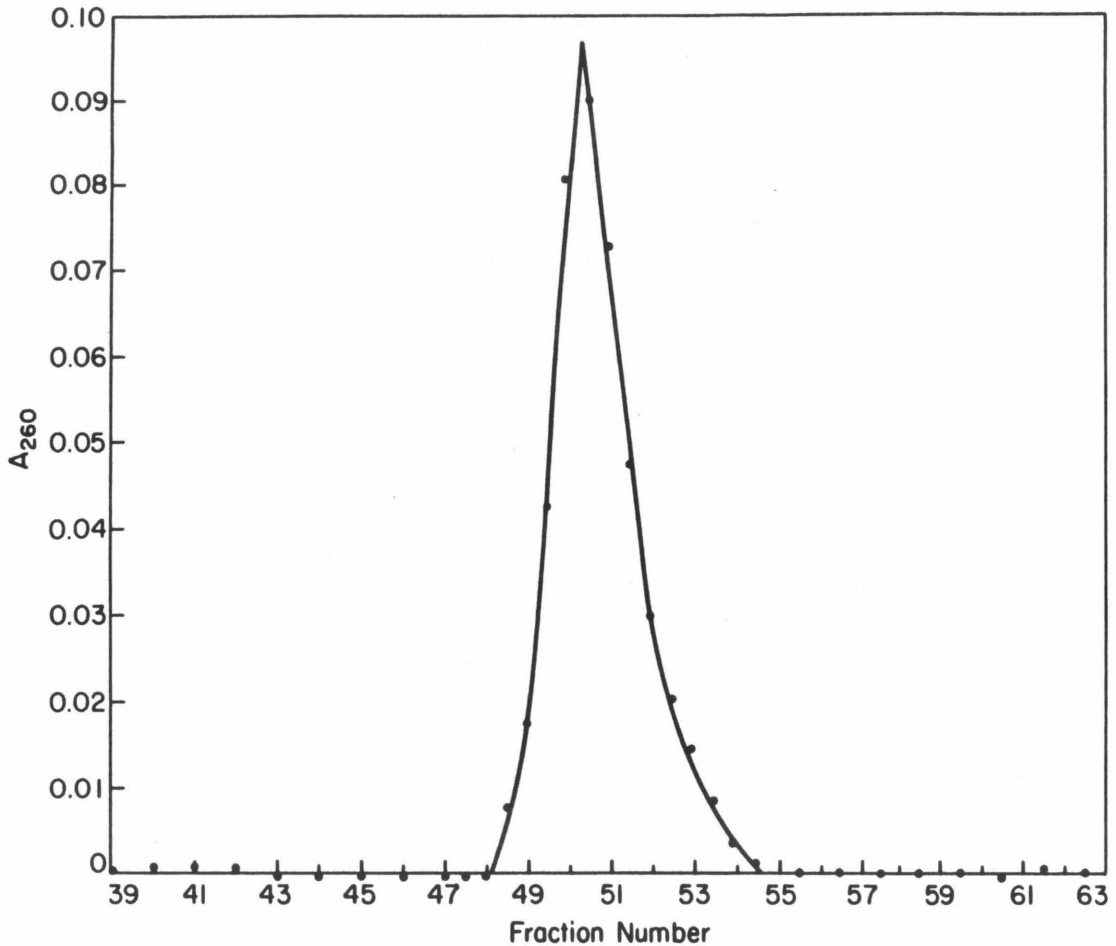


Figure 32 - Elution of PM2 DNA, restricted by HpAII, on Sepharose 4B column at 25°C and M^+ = 1.0.

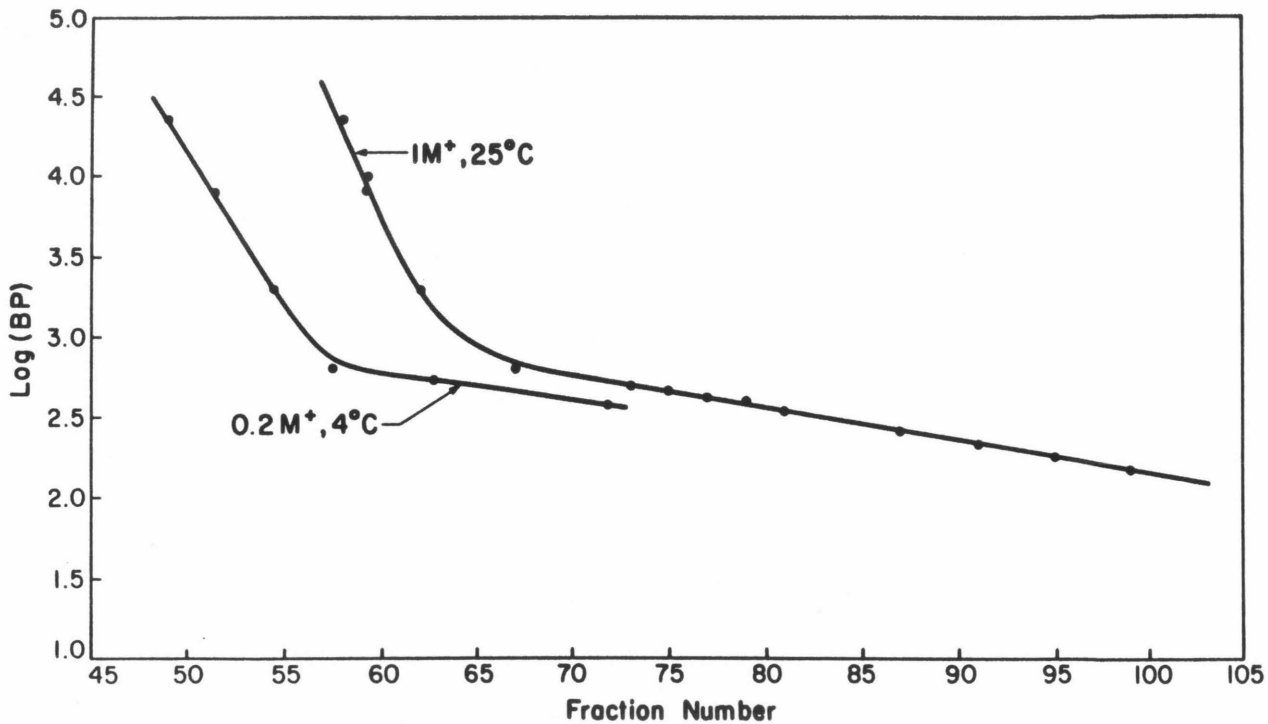


Figure 33 - Dependence of the elution volume of a Sepharose 4B column at 1 M^+: 25°C and 0.2 M^+: 4°C on the log of DNA's molecular weight in base pairs. A combination of restriction fragments and samples of sonicated calf thymust DNA of known molecular weight were utilized to calibrate the column.

PROPOSITIONS

ABSTRACTS OF PROPOSITIONS

Proposition One: A Molecular Basis of Genetic Regulation. The genetic potential of DNA is postulated to be dictated by its conformational state. B DNA is argued to be genetically inactive while its activation is postulated to necessitate at least a B → H conformational transition. Since B ↔ H transitions are cooperative and alter helical unwinding, DNA supercoiling is expected to play an important role in gene regulation.

Proposition Two: Initiation of Gene Expression. A molecular basis for the structure, function, and evolution of procaryotic RNA polymerase and its promoters is presented. Analysis of available data suggests that similar molecular principles may also apply to viral and eucaryotic organisms.

Proposition Three: Genetic Regulation Through Allosteric Control. Reexamination of the binding properties of lambda repressor to linear and supercoiled operator suggest that the cooperativity of repressor binding, the induced unwinding of operator, and the ability of repressor to autoregulate its own synthesis may result from a common mechanism: induction of an allosteric transition in operator by bound repressor. Experiments designed to distinguish this

mechanism from mechanisms invoking protein-protein contacts are described.

Proposition Four: Physiochemical Characterization of a DNA Conformational Code. Experiments designed to directly probe the occurrence of B → H → A transitions in DNA and to characterize their role(s) in gene expression are briefly described.

Proposition Five: Specific Localization of Conformational Hotspots in DNA. Since the ability of DNA to undergo a B → H → A transition varies greatly with its base sequence, DNA whose sequence is heterogeneous should exhibit conformational "hotspots". An experimental method which can specifically localize these hotspots is described.

PROPOSITION 1

A Molecular Basis of Genetic Regulation[1]

Evolution of complex organisms demanded the development of genetic regulatory mechanisms. By regulating when and to what extent a gene is expressed, these mechanisms provided single cells with enhanced ability to adapt to their environment. Eventually, genetic regulation provided a means whereby single cells could differentiate into organisms composed of many different specialized cells thereby enabling higher life forms to evolve.

Conceptually, genetic regulation must occur through transcriptional control, post-transcriptional processing of gene products or some combination of both processes. Control at the level of transcription necessitates base-specific interactions between DNA and effector molecules - principally proteins. The specificity demanded of these interactions requires intimate contacts between effector and DNA, a process which must necessarily disrupt water bound to the grooves and base pairs of the double helix. Any molecular description of gene regulation must therefore consider the consequences of effector induced dehydration on DNA structure.

The double-helical B conformational family is the most stable form of DNA under physiological conditions.[2] When dehydrated, B DNA becomes unstable and cooperatively adopts

an H^3 and then A conformation.[4,5] With additional loss of a small amount of water the double helical form cooperatively denatures.[5-7] Since base-specific recognition necessitates disruption of water bound to DNA surfaces, gene regulation should be associated with destabilization of the native conformation of DNA.

The extent to which B DNA will be destabilized by specific protein recognition can be estimated from the work of Falk, et al. These workers found that hydration of the grooves and base pairs of DNA in fibers begins at a relative humidity of 65%.[8] Since the double helix denatured just below these humidities,[6] protein induced destabilization can be estimated to be approximately equal to the difference in free energy between a native and denatured base pair at 25^{o}, or 1.5 kcal/mole base pair for the heterogeneous DNA examined. A lower limit of 0.7 kcal/mole can be placed on this estimate since states whose free energy is less than this value will be thermally accessible and not perturbed by protein recognition. Because recognition subunits of typical regulatory proteins bind approximately one turn of the helix, the B conformation will be greatly destabilized, destabilization ranging from 6 to 15 kcal/protein subunit. Ivanov has estimated the free energy required to nucleate a B to A transition as 4 kcal/mole,[10] or one-half that

required to nucleate a melting transition.[9] Since the free energy required to propogate a B → A transition is approximately 1.0 kcal or less per base pair,[10] destabilization could be large enough to induce a B to A transition in DNA. Thus base specific recognition of DNA by regulatory proteins is expected to induce <u>at least</u> a B to H transition in DNA since H DNA is an obligatory intermediate in the B to A transition with free energy intermediate between the B and A forms.[3]

These considerations suggest that the native (B) conformation of DNA may be genetically inactive, its function being storage of the genetic code. If so, a B → H transition would represent a <u>minimum</u> and <u>obligatory</u> molecular basis of genetic activation.

Although it could be argued that regulatory proteins have evolved special mechanisms which prevent DNA destabilization destabilization is expected to enhance their function by rendering accessible otherwise hidden or hydrated DNA recognition sites. Furthermore, since a B → H transition is allosteric,[3] protein binding could <u>regulate</u> gene expression by <u>abruptly</u> shifting a B ↔ H equilibrium.[12]

Base-Specific Recognition of DNA by Regulatory Proteins Unwinds the Double Helix.

If base-specific recognition of DNA by regulatory proteins is accompanied by at least a B → H transition in bound DNA, a small but significant unwinding of the double helix should be detected. The unwinding properties of three base-specific regulatory proteins have been examined in some detail. The simplest of these to consider are the lac and lambda repressors, since the unwinding they induce does not appear to arise from denaturation of the double helix or hairpin formation.[13]

Analysis of constitutive mutants[14] as well as methylation protection,[15] crosslinking,[16] and digestion[17] studies indicate that lac repressor binds tightly and specifically to a region of the lac operon termed operator.[18] Base-specific contacts are apparently confined to an 18 base pair region of operator[14-17,19] and unwind operator 90°.[13] If an obligatory B → H transition is limited to the 18 base pair recognition region of operator, the predicted unwinding angle can be estimated to be approximately 18 x 5° or 90°.[3] Although this value is in good agreement with experimental estimates, the close fit is probably fortuitous since uncertainties in both values could be as high as 50%. Uncertainties in theoretical estimates arise from several sources. First, although a B → H

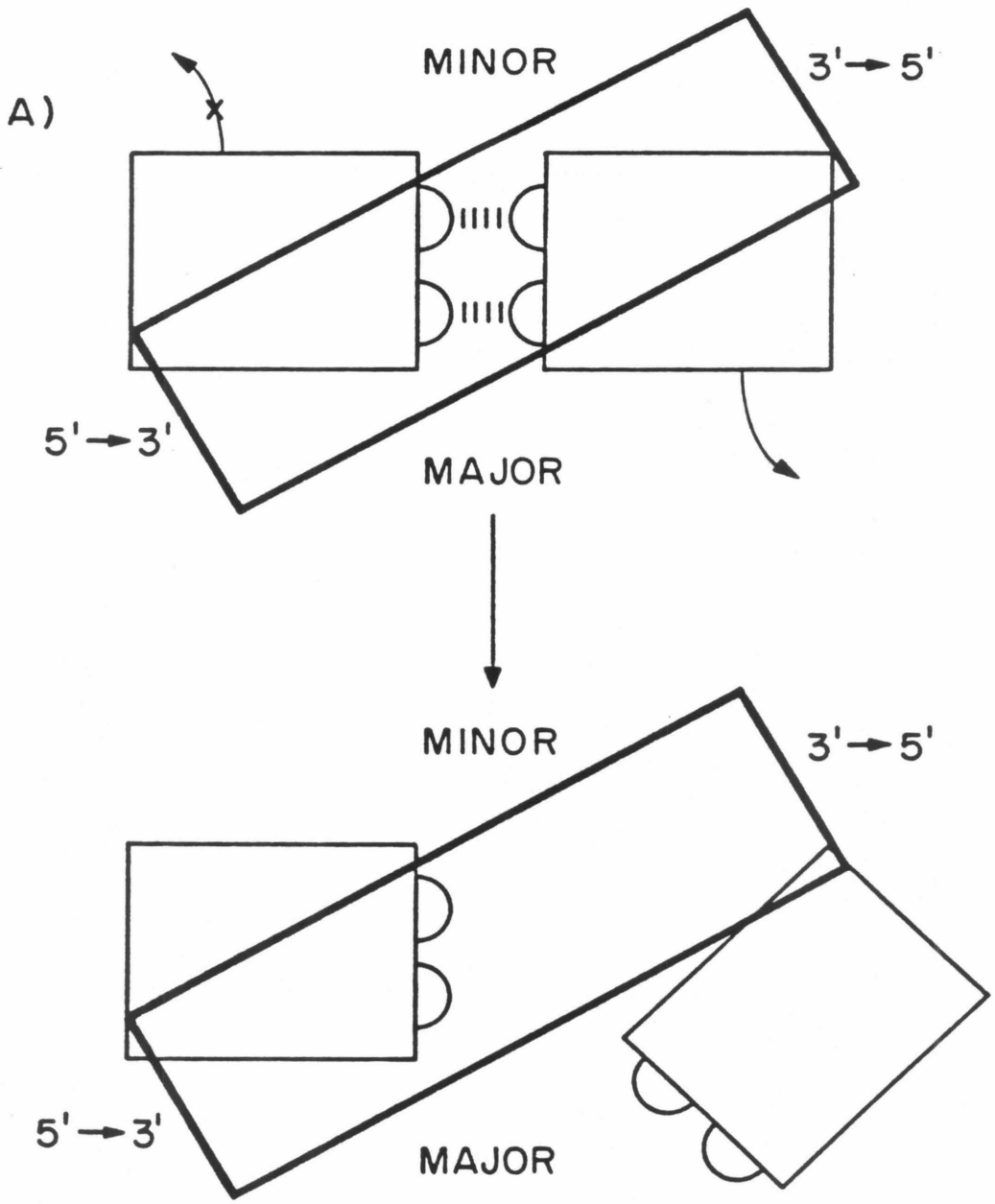
transition unwinds the double helix, the final bound conformation of operator need not be an H conformation. It is only required that operator undergo at least a B → H transition when binding base-specifically to repressor. Secondly, the unwinding angle associated with a B → H transition could be somewhat smaller or larger than 5^{o} owing to small variations in helical unwinding which are possible with the B and H families. Lastly, when bound base-specifically by a regulatory protein, the number of base pairs which undergo a conformational change need not equal exactly the protein binding site size.

Like lac repressor, lambda repressor also appears to unwind its operator since supercoiling can enhance its binding (Fig. 3b). Although the magnitude of unwinding cannot be estimated from these measurements, other features of this data suggest that unwinding is associated with at least a B → H transition in operator (next section).

The unwinding properties of E.coli RNA polymerase have been extensively investigated by Saucier and Wang[20,21] and are of particular interest since gene regulation in all organisms necessitates interactions of similar proteins with the DNA template. E.coli RNA polymerase has been shown by a variety of techniques to bind base-

Figure 1

Proposed mechanism of single-strand template formation induced by RNA polymerase. (A) Top view of two adjacent base pairs of a right-handed DNA double helix in an H conformation. The bold box represents a base pair which is lying on top of an adjacent base pair. This latter base pair is depicted as two nucleotides bonded to one another by Watson-Crick hydrogen bonds. Both major and minor grooves as well as the polarity of each phosphodiester chain are indicated. DNA melting of a base pair is proposed to occur by rotation of each nucleotide with little or no further unwinding of the double helix. For a right-handed double helix rotation is sterically preferred in a right-handed direction as depicted by the arrows. Exposure of only one nucleotide of a base pair will occur because RNA polymerase contacts, which occur in both grooves of the template forming region of promoter, are localized on the 5' → 3' strand.[54]


A)
MINOR
3'→5'
5'→3'
MAJOR
MINOR
3'→5'
5'→3'
MAJOR

specifically 50 ± 10 base pairs of double stranded DNA.[22,23] Like lac and lambda operators, bound DNA also appears to change conformation since polymerase unwinds DNA 240°.[21] The observed unwinding angle agrees with the theoretical value, 250°, predicted to result from a B → H conformational change over a 50 base pair binding site. However, RNA polymerase is also known to break Watson-Crick hydrogen bonds in a 15 base pair region of DNA when it binds.[23] If most of the unwinding results from a B → H conformational change, then it must be assumed that RNA polymerase can expose these 15 bases with little or no further unwinding of the double helix. If correct, this hypothesis provides a simple molecular basis for the observation that only one strand of a gene normally serves as a template in vivo (Fig. 1).

Nonspecific Recognition of DNA

Although base-specific recognition of DNA by regulatory proteins is postulated to necessitate at least a B → H transition, nonspecific recognition process may also induce B → H transitions. For example, the apparent affinity of polyd(A-U) for lac repressor increases by a factor of 20 when the 5-hydrogen or uracil is replaced by a methyl group, and by a factor of 600 when replaced by a bromine.[24] Since both substitutions could facilitate B → H transi-

tions in DNA by disrupting bound water, enhanced binding suggests that these nonspecific complexes may prefer to bind conformations other than the B family. Although contacts between the 5 position of uracil and lac repressor could also explain these data, other experiments have demonstrated that repressor does not contact this site.[25]

Kinetic studies of several base-specific proteins also suggest that B → H transitions may be involved in nonspecific recognition processes. For example, the lac and lambda repressors, as well as E.coli RNA polymerase, all appear to bind their target sequences considerably faster than simple bimolecular collision theory would predict, given the very small percentage of target sequences.[26] Since nonspecific complexes can form at many more sites than specific complexes, these kinetic studies have suggested that nonspecific complexes facilitate specific complex formation.[27] If so, the ability of nonspecifically bound regulatory proteins to induce a B → H transition would play an important role in base-specific recognition processes by directly coupling the formation of nonspecific and specific complexes. Interestingly, removal of flanking sequences next to lac operator, which would prevent nonspecific complexes from forming next to it, decreases operators affinity for repressor by a factor of 100.[28]

Regulatory Proteins May Alter Gene Expression by Enhancing Conformation Transitions in Adjacent DNA

Lambda repressor appears to cooperatively unwind its operators (Fig. 2) since supercoiling can enhance its binding (Fig. 3). However, when more than one repressor is bound at operator, supercoiling no longer enhances repressor binding. This observation indicates that bound repressor can alter the binding properties of adjacent DNA.

The binding cooperativity,[29] operator unwinding, and alteration of DNA binding properties induced by lambda repressor are consistent with the induction of a B → H allosteric[3] transition in operator. In the allosteric model of Moond,[12] operator is postulated to exist in either a strong binding (T) or weaker binding (R) form.[12] In the present discussion (T) represents the final unknown conformation of bound repressor which is derived from an H conformation and (R) represents the B conformation of operator. In the absence of repressor, operator assumes the more thermodynamically favored B conformation. However since repressor prefers to bind to the (T) form, its binding will shift the equilibrium from the B to the (T) form. Repressor binds cooperatively because with increased repressor binding there is an increased preference for DNA to assume a more tightly binding (T) conforma-

Figure 2

Sequence organization of the right and left lambda operators.[52] The 17 base-pair repressor binding sites are boxed within each operator. Transcription initiation sites for genes N, cro, and cI (repressor) are indicated by arrows. Alternating sequence clusters, which enhance B → H transitions, are depicted by bold [py(3' → 5')pu] or thin [pu(3' → 5')py] vertical lines. Dashed Pribnow boxes represent sites within operator where RNA polymerase must bind in order to transcribe the subsequent gene.[53]

λ Operator with Pribnow Boxes

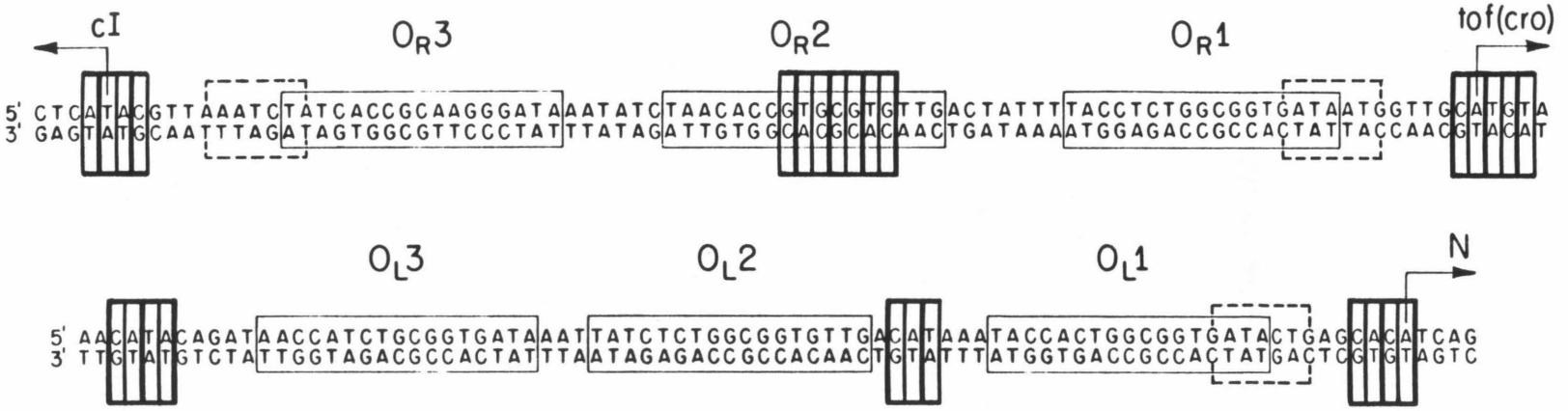
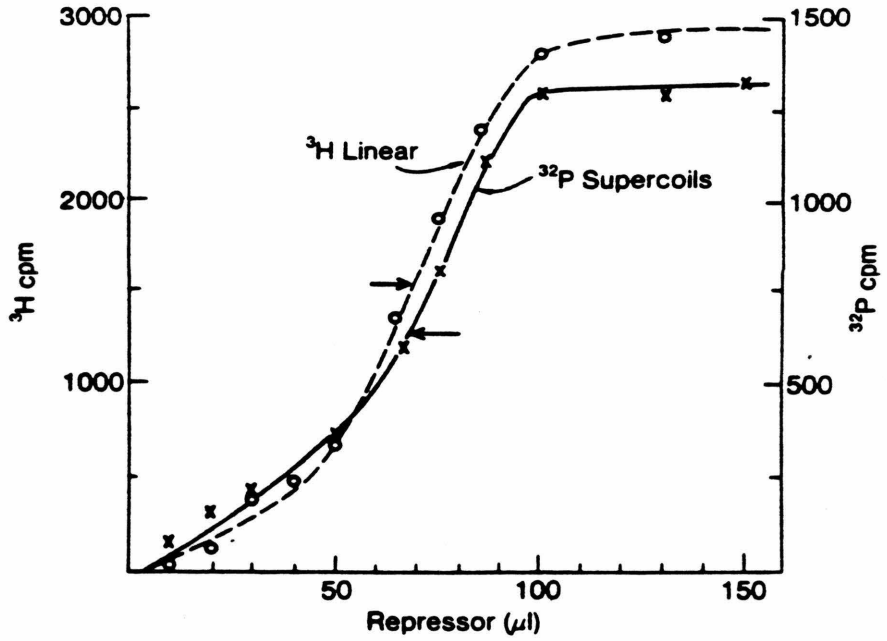
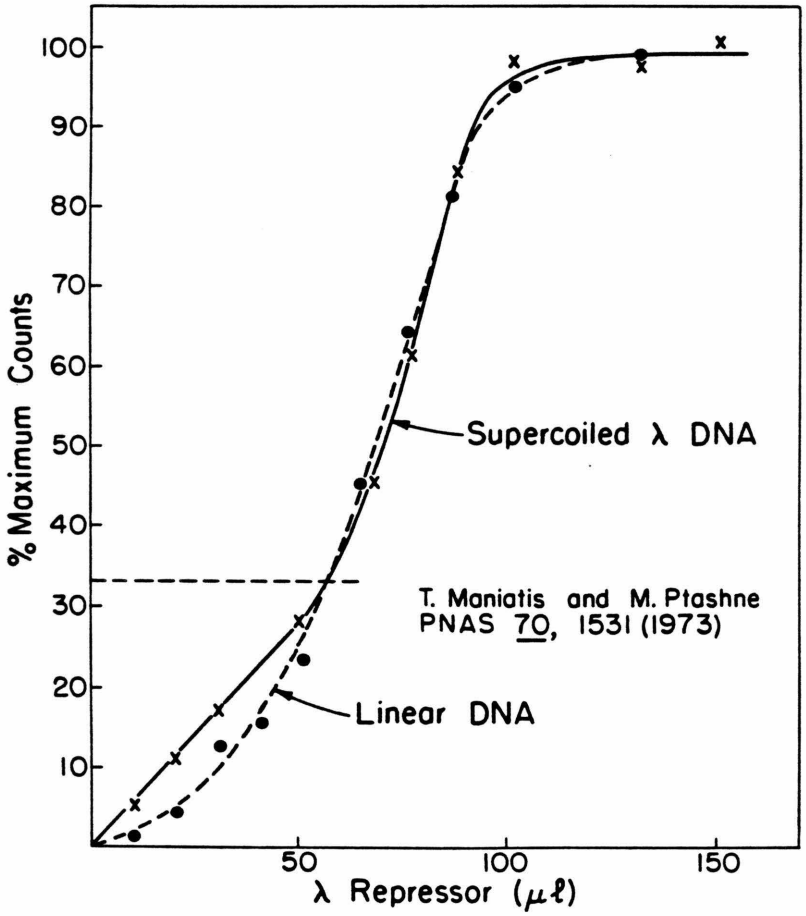


Figure 3

A) Binding of lambda repressor to supercoiled and linear lambda DNA as determined by Maniatis and Ptashne.[29] Arrows indicate the points of half-maximal binding. B) Data in Fig. 3a has been re-expressed in terms of percent maximum counts retained by the binding filter. The horizontal dashed line at one-third saturation (one bound repressor per operator) denotes the point where both linear and supercoiled lambda DNA exhibit the same affinity for repressor.


3000
2000
1000
1500
1000
500
^{3}H cpm
^{32}P cpm
^{3}H Linear
^{32}P Supercoils
50
100
150
Repressor (μl)

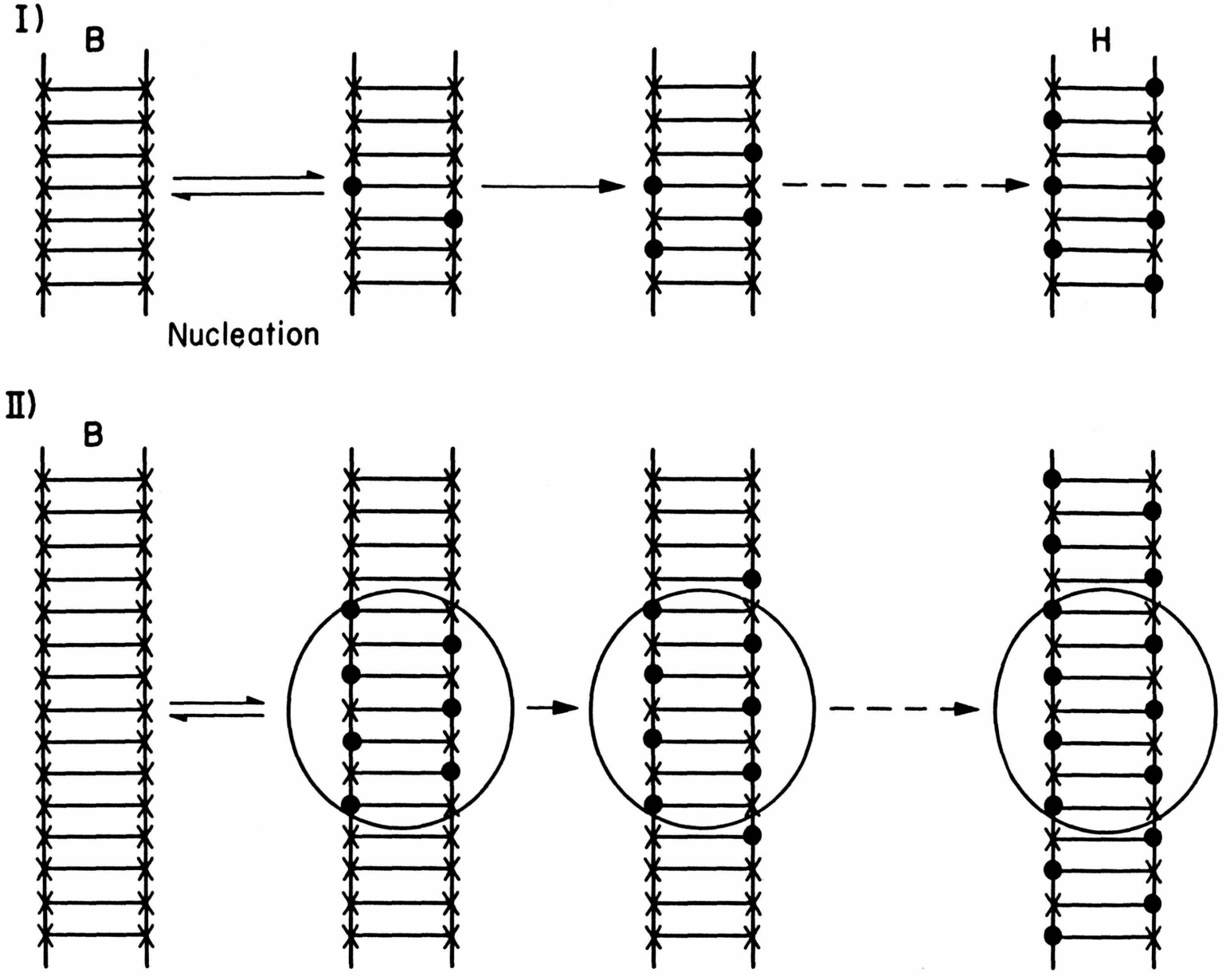

100
90
80
70
60
50
40
30
20
10
% Maximum Counts
Supercoiled λ DNA
T. Maniatis and M. Ptashne
PNAS 70, 1531 (1973)
Linear DNA
50
100
150
λ Repressor (μℓ)

tion[12] (cooperative repressor binding arising from repressor-repressor contact[30] is not excluded; it is distinguished from the present hypothesis, however by its relative insensitivity to DNA supercoiling; see proposition 3). Unwinding of operator by repressor is also consistent with the induction of a B → H transition in operator since recognition involves base-specific contacts. Lastly, the observed alteration of operator binding properties induced by repressor binding would result from the formation of B/T conformational interfaces in operator near each each of bound repressor (Fig. 4). Since nucleation of unwinding B → T conformational transitions in adjacent DNA would be facilitated by these interfaces, bound repressor could enhance subsequent repressor binding more than does supercoiling as required experimentally (see dashed line, Fig. 3b).

The apparent ability of lambda repressor to alter the conformational potential of adjacent DNA may provide a simple molecular basis for the observation that lambda repressor can activate as well as deactivate its own transcription.[31] For example, when lambda repressor binds O_{R3}, it both <u>inhibits</u> transcription from its own Ic gene and facilitates transcription from the distant <u>cro</u> gene (refer to Fig. 2). Similarly, lambda repressor when bound at O_{R1} inhibits transcription from the <u>cro</u> promoter and <u>facilitates</u> transcription from its distant IC gene. These results are readily interpreted by the allosteric B → H model presented above.

Figure 4

A) Mechanism of a B → A conformational transition in DNA.[3] B DNA, which contains all 2' endo sugar puckers (X), is converted to A DNA which contains all 3'endo (●) sugar puckers in two steps. The first step involves nucleation of a sugar pucker change in two or more adjacent and diagonal base pairs. Propogation of this sugar pucker change to adjacent base pairs is cooperative and results in an H conformation. Completion of the B→A transition proceeds from the H conformation by the same steps used to generate it. B) Induction of a B → H transition by protein binding. In the first step, protein (shaded circle) binds to a stretch of H DNA. Since DNA near each end of protein has a nucleation configuration, propogation of a B → H transition to adjacent base pairs is facilitated. Therefore, at equilibrium the probability that sequences adjacent to protein will assume an H conformation is enhanced.


I)
B
H
Nucleation
II)
B

When bound to O_{R1}/O_{R3} repressor sterically prevents polymerase from binding the cro/Ic promoter. However, since bound repressor enhances nearby B → H transitions, bound repressor will also increase the probability that its distant Ic/cro promoter will assume an activated (H) conformation necessary for transcription. It should be noted that activation may result in no apparent change in the physical properties of adjacent DNA (e.g. unwinding) since the probability of undergoing a B → H transition, although enhanced by protein binding, may still be less than one.

Other proteins which modulate transcription, like catabolite activating protein[32] and T antigen,[33] may also exert all or part of their effect by enhancing B → H transitions in associated DNA.

Genes Can Adopt Activated As Well As Active States

Although postulated as an obligatory step in gene expression, a B → H transition may only serve to activate a gene since additional conformational changes, such as DNA melting may also be required for full activity. Of particular interest in this regard is the initial finding by Weintraub and Gourdine that an alteration in nucleosome structure, as revealed by enhanced sensitivity of DNA to DNAse I digestion, is also a necessary but insufficient criterion

for genetic activity.[34] For example, globin genes in erythrocytes show enhanced sensitivity to DNAse I digestion whereas the same genes in fibroblasts or brain, tissues which do not express globin, do not. Since transcription of globin genes in adult erythrocytes eventually ceases while maintaining enhanced DNAse I sensitivity, DNAse I probes activated rather than active genes.

The adoption of an H conformation by an activated gene may enhance its sensitivity to DNAse I digestion by rendering DNA in a conformation which is more readily bound or digested by DNAse I, by disrupting pre-existing DNA-protein interactions which otherwise retard enzymatic digestion, or by some combination of both effects. The observation that alternating DNA is digested five times more readily by DNAse I than nonalternating DNA[35] suggests that DNAse I digests H conformations more readily than B conformations, since alternating DNA also undergoes B → H transitions more readily than nonalternating DNA.[2] Of the two possible sites in alternating DNA, pu(3' → 5')py sites are more readily digested than py(3' → 5')pu sites.[35] Since DNAse I cuts those base pairs directly adjacent to the one which it binds,[36] DNAse I prefers to bind py(3' → 5')pu sites more readily than pu(3' → 5')py sites.

This binding specificity exhibited by DNAse I also suggests that it binds H conformations more readily than B conformations, since ethidium bromide, which prefers to intercalate H rather than B DNA,[3] also exhibits the same binding preference?[57] Finally, the proposed conformational specificity of DNAse I is consistent with the observation that base specific recognition of DNA by regulatory proteins enhances DNA's sensitivity to DNAse I digestion,[37] since regulatory protein binding is postulated to induce at least a B → H transition in DNA.

The adoption of an H conformation by an activated gene may also enhance its sensitivity to DNAse I digestion by disruption pre-existing nucleosomal protein-DNA interactions. For example, if nucleosomal DNA is topologically constrained so that free rotation of DNA is prevented, then a B → H transition throughout nucleosomal DNA would necessarily disrupt the native nucleosome structure by uncoiling DNA from its surface. As a result, protein-DNA contacts would be reduced thereby enhancing DNA's sensitivity to DNAse I digestion. An interacting feature of this model is its ability to reconcile apparently contradictory evidence regarding the association of nucleosome with active genes. For example, a variety of techniques have demonstrated that DNA is stabilized in a B conformation when complexed in a nucleosome structure.[38] If B DNA is genetically inactive,

nucleosomal DNA should be more refractory to genetic activation processes then naked DNA. Indeed, numerous experiments support this prediction.[38,39] However, previous measurements have demonstrated that although active genes are depleted of nucleosomes, all sequences, regardless of their genetic activity, can be isolated as nucleosomes by nuclease digestion of chromatin.[40] Since native (inactive) and disrupted (activated) nucleosomes are proposed to be in equilibrium with one another, association of active genes with native nucleosomes is not excluded.

Torsional or Bending Stresses Exerted on DNA May Regulate Gene Expression

Since the discovery of supercoiled DNA by Vinograd and coworkers,[41] supercoiling has been increasingly implicated in DNA function. Proteins which supercoil DNA occur throughout nature, their enzymatic activity varying with both the growth and development of the organism.[42] In addition, increasing evidence has demonstrated that DNA supercoiling can play an essential role in DNA replication,[43] RNA transcription,[44] DNA transposition,[45] genetic recombination,[46] chromosome condensation,[47,48] nucleosome formation,[49] and virus encapsidation.[48]

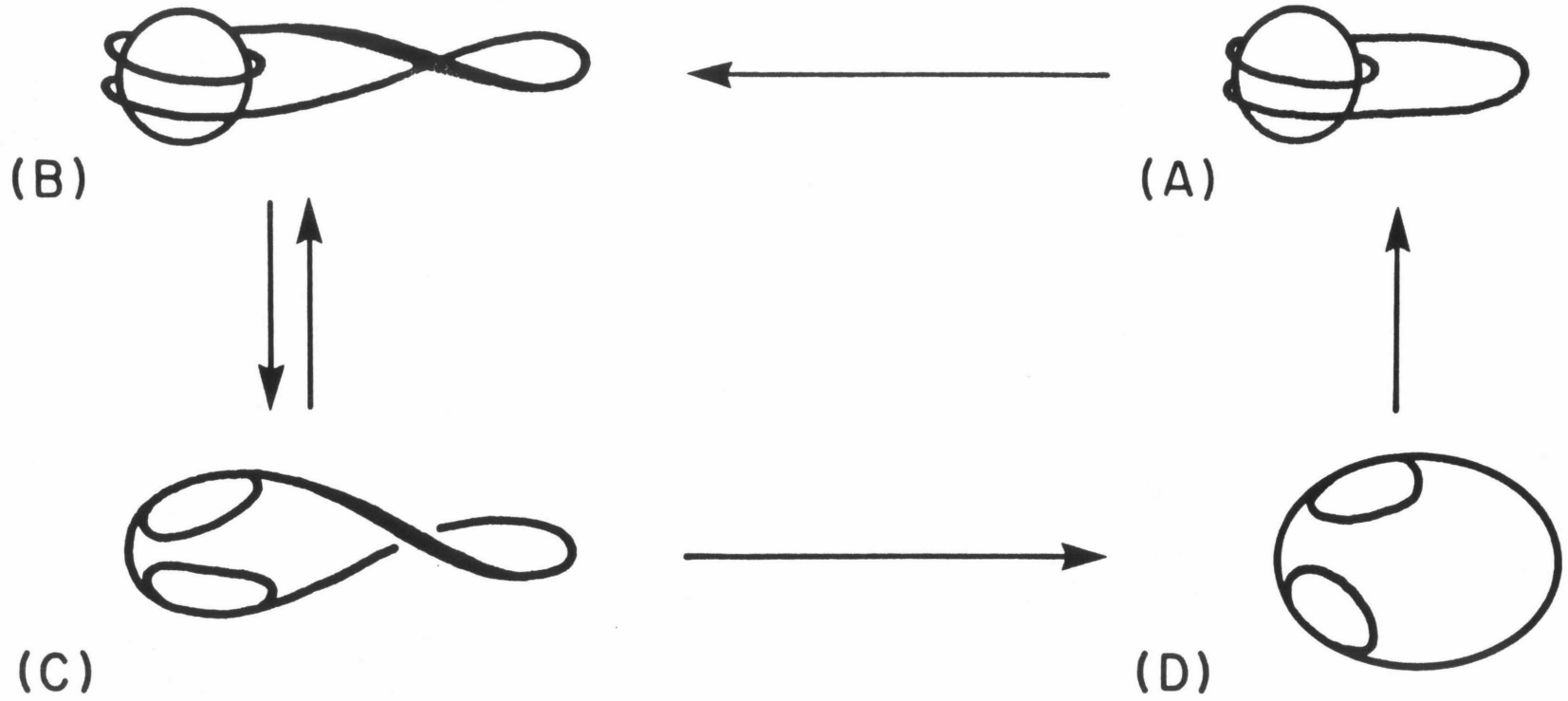
The postulated role of B ↔ H transitions in gene regulation allows one to make specific predictions regarding the role of supercoiling in these processes. For example, since B → H transitions are cooperative and unwind DNA, small torsional (or bending) stresses, arising from supercoiling or protein binding, could activate/deactivate gene expression by abruptly shifting the B ↔ H equilibrium to the right/left. This behavior is a necessary consequence of the cooperative nature of a B ↔ H transition and is absolutely essential if supercoiling is to play an important role in regulating gene expression.[12] That the enzymatic activity of a variety of DNA unwinding proteins can be abruptly controlled by supercoiling has been clearly demonstrated by Wang.[50] Included in his study were regulatory proteins which bind base-specifically as well as those which induce DNA melting (like a B → H transition, melting also cooperatively unwinds DNA).

Whether a B → H transition is inhibited or initiated by torsional (and/or bending) stresses will depend on the type of DNA supercoiling. When DNA is supercoiled as a left-handed torroid, as in a nucleosome, B to H transitions are enhanced. In contrast, when DNA is supercoiled as a left-handed interwound structure (as in a hairpin), B → H transition will be inhibited. Since a left-handed interwound structure is equivalent (topologically) to a right-

handed torroid,[51] supercoiling chirality is also expected to play an important role in gene regulation. In Fig. 5 a simple model relating supercoiling stresses to changes in gene regulation is presented.

Figure 5

Hypothetical model of genetic regulation at the level of transcription in chromatin. (A) One nucleosome, consisting of 1-3/4 supercoiled turns of left-handed torroidal B DNA[55] wrapped around an octameric protein core is shown associated with additional DNA. (B) Activation of a nucleosome. Nucleosomal DNA is subject to torsional (or bending) stress, as for example by the introduction of right-handed interwound turns (or equivalently, left-handed torroidal turns) in neighboring DNA. Although the resulting nucleosome is genetically inactive, supercoiling stress may be sufficient[56] so that (C) induction of a B → H transition will occur throughout nucleosomal DNA upon the introduction of further small torsional or bending stresses, disrupting the nucelosome[58] and activating it towards transcription. These additional torsional or beding stresses could arise, for example, from the enhancement of B → H transitions by regulatory protein binding. (D) Relaxation of torsional and bending stress by nicking-closing enzyme generates an inactive nucleosome.


(B)
(A)
(C)
(D)

References

1) Becker, M.M., submitted to Nature, 1979.

2) a) Watson, J.D.; Crick, F.H.C., Nature, 1953, 171, 737. b) Hamilton, L.D.; Barclay, R.K.; Wilkins, M.H. F.; Brown, G.L.; Wilson, H.R.; Marvin, D.A.; Taylor, H.E.; Simmons, N.S., J. Biophysic. Biochem. Cytol., 1959, 5, 397. c) Arnott, S., Progr. Biophys. Mol. Biol., 1970, 21, 265. d) Bram, S., J. Mol. Biol., 1971, 58, 277. e) Wang, J.C., Proc. Natl. Acad. Sci. USA, 1979, 76, 200.

3) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

4) Franklin, R.E.; Gosling, R.G., Acta. Cryst., 1953, 6, 673.

5) a) Pilet, J.; Blicharski, J.; Brahm, J. Biochemistry, 1975, 14, 1869. b) Ivanov, V.I.; Minhenkova, L.E.; Minyat, E.E.; Frank-Kamenetskii, M.D.; Schyolkina, A.K., J. Mol. Biol., 1974, 87, 817. c) Minchenkova, L.; Minyat, E.; Schyokina, A.; Ivanov, V., FEBS Lett., 1975, 51, 38. d) Ivanov, V.I.; Minchenkova, L.E.; Schyolkina, A.K.; Poletayev, A.I., Biopolymers, 1973, 12, 89. e) Zimmerman, S.B.; Pheiffer, B.H. J. Mol. Biol., 1979, 135, 1023.

6) Falk, M.; Hartman, K.A.; Lord, R.C., J. Amer. Chem. Soc., 1963, 85, 391.

7) Ts'o, P.O.P.; Helmkamp, G., Tetrahedron, 1961, 13, 198.

8) Falk, M.; Hartman, K.A.; Lord, R.C., J. Amer. Chem. Soc., 1963, 85, 387.

9) Frank-Kamenetskii, M.D.; Zaxurkin, Y.S., Ann. Rev. Biophys. Bioeng., 1974, 3, 127.

10) See reference 5b.

11) a) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981. b) Levitt, M.; Warshel, A., J. Amer. Chem. Soc., 1978, 100, 2607.

12) Monod, J.; Wyman, J.; Changeux, J.-P., J. Mol. Biol., 1965, 12, 88.

13) Wang, J.C.; Barkley, M.D.; Bourgeois, S., Nature, 1974, 251, 247.

14) a) Gilbert, W.; Gralla, J.; Majors, J.; Maxam, A. in "Protein Ligand Interactions," Sund, H.; Blaver, G., eds., De Gruyter, Berlin, 1975, 193. b) Bourgeois, S.; Barkley, M.D.; Hobe, A.; Sadler, J.R.; Wang, J.C., ibid, 251. c) Jobe, A.; Sadler, J.R.; Bourgeois, S., J. Mol. Biol., 1974, 85, 231.

15) Ogata, R.T.; Gilbert, W., J. Mol. Biol., 1979, 132, 709.

16) Ogata, R.T.; Gilbert, W., Proc. Natl. Acad. Sci. USA, 1977, 74, 4873.

17) Gilbert, W.; Maxam, A., Proc. Natl. Acad. Sci. USA, 1973, 70, 3581.

18) Jacob, F.; Monod, J., J. Mol. Biol., 1961, 3, 318.

19) a) Bahl, C.P.; Wu, R.; Stawinsky, J.; Narang, S., Proc. Natl. Acad. Sci. USA, 1977, 74, 966.
b) Goeddel, D.V., Yonsura, D.G.; Caruthers, M.H., Proc. Natl. Acad. Sci. USA, 1978, 75, 3578

20) Saucier, J.M.; Wang, J.C., Nature, New Biol., 1972, 239, 167.

21) Wang, J.C.; Jacobsen, J.H.; Saucier, J.M., Nucl. Acid Res., 1977, 4, 1225.

22) a) Walz, A.; Pirrotta, V., Nature, 1975, 254, 118.
b) Pribnow, D., Proc. Natl. Acad. Sci. USA, 1975, 72, 784. c) Schaller, H.; Gray, C.; Herrmann, K., Proc. Natl. Acad. Sci. USA, 1975, 72, 737.
d) Hirsh, J.; Schleif, R., J. Mol. Biol., 1976, 108, 471.

23) Melnikova, A.F.; Baebealashvilli, R.; Mirzabekov, A.D., Eur. J. Biochem., 1978, 84, 301.

24) Lin, S.; Riggs, A.D., Biochem. Biophys. Res. Commun., 1971, 45, 1542.

25) Richmond, T.; Steitz, T.A., J. Mol. Biol., 1976, 103, 25.

26) a) Hinkle, D.C.; Chamberlin, M.J., J. Mol. Biol., 1972, 70, 187. b) Riggs, A.D.; Bourgeois, S.; Cohn, M., J. Mol. Biol., 1970, 53, 401. c) Chadwick, P.; Pirrotta, V.; Steinberg, R., Hopkins, N.; Ptashne, M. Cold Spring Harbor Symp. Quant. Biol., 1970, 35, 283.

27) a) von Hippel, P.H.; McGhee, J.D., Ann. Rev. Biochem., 1972, 41, 231. b) Richter, P.H.; Eigen, M., Biophys. Chem., 1974, 2, 255.

28) Heyneker, H.L.; Shine, J.; Goodman, H.M.; Boyer, H.; Rosenberg, J.; Dickerson, R.E., Narang, S.A.S; Itakura, K.; Lin, S.Y.; Riggs, A.D., Nature, 1976, 263, 748.

29) Maniatis, T.; Ptashne, M., Proc. Natl. Acad. Sci. USA, 1973, 70, 1531.

30) Johnson, A.D.; Meyer, B.J.; Ptashne, M., Proc. Natl. Acad. Sci. USA, 1979, 76, 5061.

31) For example see Ptashne, M.; Jeffery, A.; Johnson, A.D.; Maurer, R.; Meyer, B.J.: Pabo, C.O.; Roberts, T.M.; Saucier, R.T., Cell, 1980, 19, 1.

32) a) Zubay, G.; Schwartz, D.; Beckwith, J., Proc. Natl. Acad. Sci. USA, 1970, 66, 104. b) Eron, L.; Block, R., Proc. Natl. Acad. Sci. USA, 1971, 68, 1828.

33) Reed, S.I.; Stark, G.R.; Alwine, J.C., Proc. Natl. Acad. Sci. USA 1976, 73, 3, 83.

34) Weintraub, H.; Gourdine, M., Science, 1976, 193, 848.

35) Klug, A.; Jack, A.; Viswamitra, M.A.; Kennard, O.; Shakked, Z.; Steiz, T.A., J. Mol. Biol., 1979, 1313, 669.

36) Sollner-Webb, B.; Melchoir, W.; Felsenfeld, G., Cell, 1978, 14, 611.

37) a) Schmitz, A.; Galas, D.J., Nucl. Acid. Res., 1979, 6, 111. b) Taniguchi, T.; O'Neil, M.; DeCrombrugghe, B., Proc. Natl. Acad. Sci. USA, 1979, 76, 5090.

38) a) Bram, S., J. Mol. Biol., 1971, 58, 277. b) Herskovitis, T.T.; Brahm, J., Biopolymers, 1976, 15, 687. c) Liquier, J.; Taboury, J.; Taillandier, E.; Brahm, J. Biochemistry, 1977, 16, 3262. d) Goodwin, D.C.; Brahm, J., Nucl. Acid Res., 1978, 5, 835. d) Liquier, J.; Gadenne, M.C.; Taillander, E., Nucl. Acid. Res., 1979, 4, 1479.

39) a) Bonner, J.; Dahmus, M.E., Fambrough, D.; Huang, R.C.; Marushigue, K.; Tuan, D.Y.H., Science, 1968, 159, 47. b) Cedar, H.; Felsenfeld, G., J. Mol. Biol., 1973, 77, 237. c) Williamson, P.; Felsenfeld, G., Biochemistry, 1978, 17, 5695. d) Lilley, D.M.J.; Jacobs, M.F.; Houghton, M., Nucl. Acid Res., 1979, 7, 377.

40) a) Lacy, E.; Axel, R., Proc. Natl. Acad. Sci. USA, 1975, 72, 3978. b) Reeves, R., Science, 1976, 194, 529. c) Kuo, M.T.; Sahasrabuddhe, C.G.; Sanders, G.F., Proc. Natl. Acad. Sci. USA, 1976, 73, 1572.

41) Vinograd, J.; Lebowitz, J.; Radloff, R., Watson, R.; Laipis, P., Proc. Natl. Acad. Sci. USA, 1965, 53, 1104.

42) For example see Wang, J.C. and Liu, L.L., in "Molecular Genetics, Part III," Taylor, J.H., ed., Academic Press, New York, 1979, 65.

43) a) Wang, J.C., J. Mol. Biol., 1971, 55, 523. b) Champoux, J.J.; Dulbecco, R., Proc. Natl. Acad. Sci. USA, 1972, 69, 143. c) Gellert, M.; O'Dea, M.H.; Itoh, T.; Tomizawa, J., Proc. Natl. Acad. Sci. USA, 1976, 73, 4474. d) Sumida-Yasumoto, C.; Hurwitz, J., Proc. Natl. Acad. Sci. USA, 1977, 74, 4195.

44) a) Wang, J.C. in "DNA Synthesis in Vitro," Wells, R.D.; Inman, R.B., eds., University Park Press, Baltimore, 1973, 163. b) Smith, C.L.; Kubo, K. Imamoto, F., Nature, 1978, 275, 420.

45) Shapiro, J.A., Proc. Natl. Acad. Sci. USA, 1979, 76, 1933.

46) a) Champoux, J.J., Proc. Natl. Acad. Sci. USA, 1979, 74, 3800. b) Kirkegaard, K.; Wang, J.C., Nucl. Acid Res., 1978, 5, 3811. c) Kikuchi, Y.; Nash, H.A., Proc. Natl. Acad. Sci. USA, 1979, 76, 3760.

47) Baase, W.A.; Wang, J.C., Biochemistry, 1974, 13, 4299.

48) Bauer, W.R.; Ressner, E.C.; Kates, J.; Patzke, J.V., Proc. Natl. Acad. Sci. USA, 1977, 74, 1841.

49) Germond, J.E.; Rouviere-Yaniv, J.; Yaniv, M.; Brutlag, D., Proc. Natl. Acad. Sci. USA, 1979, 76, 3779.

50) Wang, J.C., J. Mol. Biol., 1974, 87, 797.

51) Bauer, W.; Vinograd, J., J. Mol. Biol., 1968, 33, 141.

52) For example see, Humayun, Z.; Jeffery, A.; Ptashne, M., J. Mol. Biol., 1977, 112, 265.

53) Pribnow, D., J. Mol. Biol., 1975, 99, 419.

54) a) Johnsrud, L., Proc. Natl. Acad. Sci. USA, 1978, 75, 5314. b) Siebenlist, U.; Gilbert, W., Proc. Natl. Acad. Sci. USA, 1980, 77, 122.

55) For example see Lutter, L.C., J. Mol. Biol., 1978, 124, 391.

56) Vologodskii, A.V.; Lukashin, A.V.; Anshelevich, V.V.; Frank-Kamenetskii, M.D., Nucl. Acid. Res., 1979, 5, 967.

57) For example see Patel, D.J.; Canuel, L.L., Proc. Natl. Acad. Sci. USA, 1976, 73, 3343.

58) Weintraub, H.; Worcel, A.; Alberts, B., Cell, 1976, 9, 409.

PROPOSITION 2

Initiation of Gene Expression[1]

Certain sequences along the DNA chromosome, when mutated, alter gene expression. A large number of these sites have been isolated and sequenced in the hope that new insights into the molecular basis of gene expression and its regulation might emerge. One example which has necessarily received much attention are DNA sequences which initiate gene expression.[2,3] Termed promoters, such sites, by virtue of their special sequence, are recognized, bound, and melted by RNA polymerase.[4] Subsequent uptake of ribonucleotide triphosphates then initiates specific transcription of promoter associated gene(s).

The most extensively characterized promoters are those of procaryotes. They are of special interest since their sequence variability is an important element in distinguishing not only their transcriptional efficiency but their control by regulatory factors as well.[3,5] In this paper recent postulates concerning DNA conformation[6] and its role in gene expression[6] are used to examine the sequence organization of procaryotic promoters. This analysis suggests a simple molecular basis for the structure, function, and evolution of procaryotic RNA polymerase and its promoters. Analysis of available data suggests that similar molecular principles may also apply to viral and eucaryotic organisms.

Pribnow's Box

One of the most thoroughly characterized promoters whose structure is representative of procaryotic promoters in general is the lac promoter shown in Fig. 1.[7] Three regions of sequence homology have been observed in procaryotic promoters: a recognition region, Pribnow's box,[9] and the initiating nucleotide triphosphate (NTP) binding site(s).[9,10] The close proximity of Pribnow's box to the site where RNA transcription starts suggests that melting of the presumptive DNA template initiates at Pribnow's box. Indeed, since RNA polymerase breaks Watson-Crick hydrogen bonds in a 15 base pair region of DNA,[11] melting initiated at Pribnow's box would just extend to the NTP incorporation site. Direct evidence for these expectations has been presented by Sibenlist,[12] who trapped melted base pairs in promoter by methylating the one position of adenine in the presence of bound polymerase. After removal of polymerase, DNA was allowed to reanneal and mismatched base pairs were detected by S_1 digestion. Melting was found to cover at least an 11 base pair region extending from the later half of Pribnow's box to the NTP incorporation site.

How does Pribnow's box initiate promoter melting? One clue is suggested by experiments which probe close contacts between promoter and polymerase.[13] These experiments have detected base-specific contacts between RNA

Figure 1
Lac promoter sequence. Three important regions of sequence homology are denoted. From left to right these are the recognition region, Pribnow's box, and the nucleotide triphosphate incorporation site. The startpoint of transcription is denoted by an arrow.

polymerase and Pribnow's box suggesting that polymerase could induce melting by binding to sites normally involved in Watson-Crick hydrogen bonding. The failure of trapping experiments to detect the formation of 1-methyladenine in the first four base pairs of Pribnow's box may reflect binding of polymerase to this Watson-Crick hydrogen bonding site.[12]

A second feature of Pribnow's box which appears to play an important role in promoter melting is its sequence organization. For example, in procaryotes, the most efficient Pribnow box, the lac UV-5 mutant, has the sequence $\frac{\text{TATAATG}}{\text{ATATTAC}}$.[14] Three features of this sequence are expected to contribute significantly to its efficient role in initiating transcription. First, since AT base pairs melt more readily than GC pairs, its high (A+T) content would enhance melting. Secondly, since the sequence is largely alternating it can easily adopt transcriptionally active conformations. For example, it has been proposed that transcription demands a B → H → melt conformational transition in DNA because the necessity of base-specific recognition as well as template melting will greatly destabilize the native B conformation of DNA.[6] Since alternating DNA undergoes B → H → melt transitions more readily than nonalternating sequences,[6] sequence alternation in Pribnow's box should

facilitate its transcription. Lastly, since the UV-5 Pribnow box begins with an AT base pair and is alternating in the rightward direction whereas in the opposite direction alternation is interrupted and begins with a GC base pair, initiation of a B → H → melt transition would be preferred in the rightward direction. As a result, polymerase would engage promoter in only one direction thereby insuring specific transcription of associated gene(s).

In addition to initiating promoter melting, Pribnow's box may also be utilized in other DNA melting processes. For example, integration of lambda DNA into the E.coli chromosome involves recombination between specific sites on lambda (POP') and E.coli (BOB').[15] DNA supercoiling is essential for recombination since at least the POP' site must be located on a negatively supercoiled DNA molecule.[16] In addition to supercoiling, a phage protein, INT (integrase), as well as accessory proteins are required.[17] A 15 base pair core sequence, the crossover region for site-specific recombination, is common to both the POP' and BOB' sites and binds INT base specifically.[18] As shown in Fig. 2, the POP' site, which overlaps the common core sequence, consists of two Pribnow boxes directed towards one another. Since recombination necessitates strand separation, it may be suggested that when supercoiled, each Pribnow box induces a B → H → melt transition

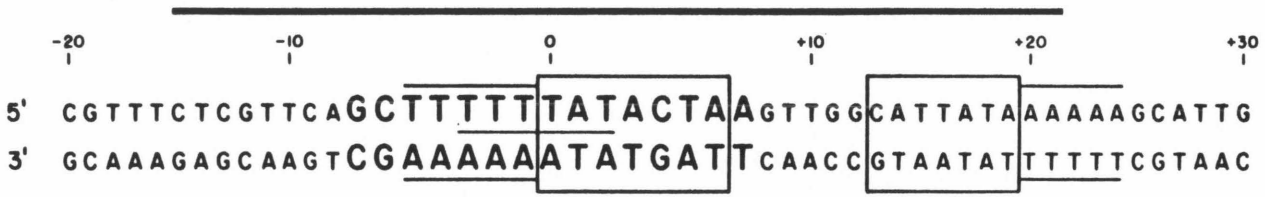


Figure 2
Nucleotide sequence of the POP' integration site. Letters in bold type are the 15 base pairs of the common core sequence. Each Pribnow box as well as adjoining dAdT tracts are denoted. A bold line denotes that region of the sequence protected by integrase from pancreatic and neocarzinostatin digestion.

over the common core sequence which is trapped by the INT protein. Subsequent steps, catalyzed by INT and accessory proteins, then exchange reciprocal strands.[19]

RNA Polymerase-Promoter Evolution: Pribnow's Box

The suggestion that the sequence of Pribnow's box is designed to initiate a B → H → melt transition in the direction of transcription when bound base specifically by RNA polymerase suggests a plausible scheme for the evolution of RNA polymerase and its promoters. For example, since alternating sequences readily undergo B → H transitions[6] and AT rich DNA readily melts, the first efficient promoter may have been a stretch of alternating AT DNA (eg. d(A-T)). At this early stage in evolution RNA polymerase could have been a protein which bound electrostatically and disrupted structural water bound to the double helix. Loss of structural water would not only induce a B → H transition in promoter thereby enabling the eventual development of base specific contacts, but it would also simultaneously enhance DNA melting. Melting would be enhanced because vertical stacking interactions and hydrogen bonding between base pairs will be weakened by the propellor twist which occurs when a base pair adopts an H conformation.[6]

Although integration of an alternating d(A-T) sequence into a segment of DNA would facilitate transcription of associated DNA, transcription of <u>both</u> flanking sequences would occur owing to the two-fold symmetry of this primitive promoter. Since the most efficient, as well as the most conserved Pribnow box[3] has the UV-5 sequence ($^{TATAATG}_{ATATTAC}$), it may be suggested that transcriptional polarity was established in a primordial alternating d(A-T) promoter by introduction of a mirror plane to generate the sequence $^{-TATAATA-}_{-ATATTAT-}$ followed by substitution of the last AT base pair with a GC pair. Introduction of polarity in this manner would preserve those sequence features (alternation and high AT content) which are essential for its efficient transcription.

<u>RNA Polymerase-Promoter Evolution: Duplication</u>

In addition to Pribnow's box, procaryotic promoters exhibit a second region of sequence homology. This region has been termed the recognition region because although it is not necessary for template melting, it is necessary for functional promoter complex formation.[8-10] For example, if transcriptionally active promoter-RNA polymerase complexes are digested with DNAse I, the undigested DNA, which now lacks the recognition region, no longer rebinds RNA polymerase in a productive complex.[8-10] A number of promoter mutants involving single base changes have been isolated

and located within the recognition region. Analysis of a number of these mutations strongly suggests that the recognition region is actually an inverted Pribnow box (Fig. 3). For example, mutations in this region which depress transcription can be interpreted to result from the formation of an inverted Pribnow box whose sequence deviates more from the inverted UV-5 sequence than wild type. (Promoters whose recognition region overlaps the binding sites of bound by additional regulatory factors such as CAP may exhibit more complicated behavior[3].)

These observations suggest that when bound to promoter, RNA polymerase makes base-specific contacts within the recognition region which are maximized for the UV-5 sequence. As a result, mutations which render the recognition region more like an inverted UV-5 Pribnow box should enhance transcription. Although this appears to be true for the Iq mutant[20], the λP_L sexl mutation[21] suggests that possession of a more efficient Pribnow box in the recognition region than in the melting region actually depresses transcription from Pribnow's box. This behavior may result from transcription in the wrong direction as the recognition region rather than Pribnow's box would melt.

The observation that procaryotic promoters possess two Pribnow boxes suggests that following evolution of a

Figure 3

Promoter mutations in the recognition region are shown for the lac gene, the lac repressor I gene, and the lambda P_L promoter. In each case the wild type sequence is listed first followed by mutant sequence(s). Nucleotides in the vicinity of the recognition region are shown along with each associated Pribnow box and RNA transcript initiation site. Underlined sequences represent proposed inverted Pribnow boxes in the recognition region. Arrows above each sequence delineates the site of mutation. All mutations shown impair promoter function expect the Iq mutant which enhances promoter function.

Promoter Mutants in the "Recognition Region"

	-35	-10	0
Lac wild type	CAGGCTTTACACTTT GTCCGAAATGTGAAA	TATGTTG ATACAAC	AA
ΔL305	CAGG TTTACACTTT GTCC AAATGTGAAA		
L241	CAGGCTTAACACTTT GTCCGAATTGTGAAA		
L157	CAGGCTTTAAACTTT GTCCGAAATTTGAAA		
I	GAATGGCGCAAAACC CTTACCGCGTTTTGG	CATGATA GTACTAT	GG
Iq	GAATGGTGCAAAACC CTTACCACGTTTTGG		
λp_L	GCGGTGTTGACATAAA CGCCACAACTGTATTT	GATACTG CTATGAC	A
λp_L Sex 1	GCGGTGTTGATATAAA CGCCACAACTATATTT		

Pribnow box duplication of promoter and RNA polymerase occurred. Through duplication, the complexity of the transcriptional process could be enhanced thereby facilitating further control of transcription as demanded by more complex organisms (eg. see discussion on sigma factor). Although two Pribnow boxes could be combined in several ways, combination of a poorly melting inverted Pribnow box with an efficient Pribnow box, assures that promoter is capable of both recognition and unidirectional transcription (Fig. 4a).

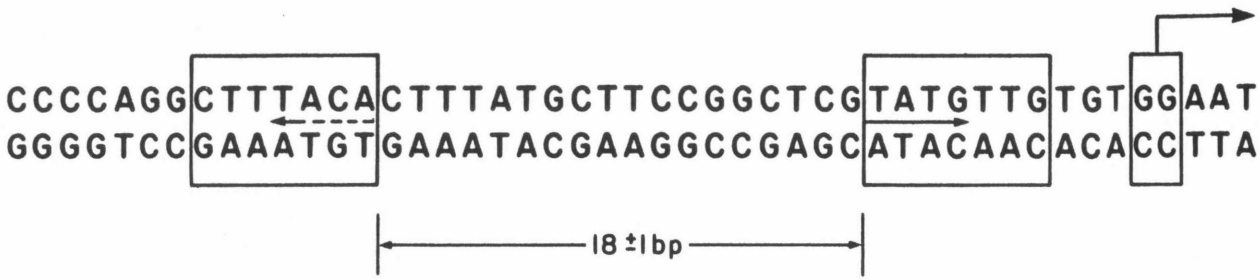
In order that a duplicated promoter be bound with enhanced affinity and specificity, each Pribnow box must be bound simultaneously by polymerase. RNA synthesis by modern day procaryotic RNA polymerase is carried out by four subunits: β,β', and the α_2 dimer.[22] Since the α_2 dimer binds the β and β' subunits rather than DNA,[23] it may be suggested that it binds two DNA binding monomers into a functional dimer (Fig. 4b). The observation that the proposed DNA binding monomers of polymerase differ in their amino acid constitution[24] suggests that one subunit may have evolved mechanisms which prevent its melting the recognition region when bound.

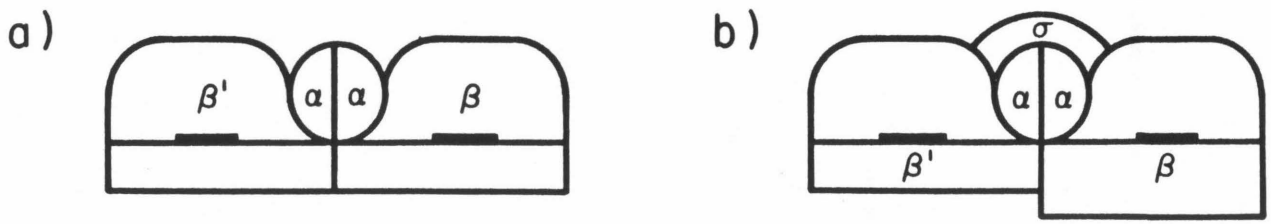
RNA-Polymerase-Promoter Evolution: The Sigma Factor

The two α subunits together with the β and β' subunits constitute the core of procaryotic RNA polymerase. Al-

Figure 4

I) Sequence organization of a procaryotic promoter. Promoter is postulated to consist of two Pribnow boxes separated by 18 ± 1 base pairs (the exact distance was deduced from an analysis of 60 promoters[3]). The startpoint and direction of transcription are denoted by a bent arrow (⌐→). Arrows within each Pribnow box denote the direction of the B → H (dotted arrow) or B → H → melt transition (solid arrow) induced by RNA polymerase binding. The Pribnow box farthest from the startpoint of transcription (lying within the so-called recognition region) binds polymerase base-specifically but does not melt. Its sequence, relative to the other Pribnow box, is inverted and exhibits less homology with the lac UV-5 Pribnow box. The second Pribnow box, begins 12-15 base pairs from the startpoint of transcription, binds polymerase base specifically, and initiates template melting. The example considered is the lac promoter.
II) Proposed subunit organization of E.coli RNA polymerase (A) core polymerase (B) holoenzyme. The β and β' subunits are postulated to possess a fixed region which binds Pribnow's box base specifically. These regions are denoted by heavy vertical lines in each subunit. The proposed structures are consistent with previous crosslinking studies.[51]


CCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAAT
GGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAACACACCTTA
18 ±1 bp


a)
β'
α
α
β
b)
σ
α
α
β'
β

though core polymerase can elongate an RNA transcript, it does not bind or transcribe promoter specifically.[25] Specific transcription of promoter is induced by the binding of a fifth subunit, sigma, to core polymerase resulting in a "holoenzyme".[25] Since sigma is lost from core after RNA synthesis is initiated, its binding regulates transcription by reversibly switching polymerase from a promoter-specific initiating structure (holoenzyme) to a nonspecific chain elongating structure (core polymerase).[26] Sigma apparently regulates transcription by inducing a conformational change in RNA polymerase. For example, when bound to nonpromoter DNA, polymerase unwinds DNA by 240°.[27,28] As previously discussed, unwinding can be interpreted to result from the induction of an obligatory H conformation in DNA over the 50 base pair binding site of polymerase.[6] However, when bound to promoter, polymerase only unwinds 150°[28] suggesting that sigma restricts polymerase contact to approximately 30 base pairs of promoter. These observations suggest that both DNA binding monomers of polymerase undergo a conformational change in the presence of sigma which reduces their binding capacity from approximately 25 to 15 base pairs.

Two additional observations support this hypothesis. First, indirect binding measurements have shown that one-half of the

electrostatic interactions between polymerase and double-stranded DNA are lost when sigma binds core.[29] Secondly, in addition to protecting 12 base pairs of the RNA transcript,[30] holoenzyme-promoter complexes also protect double-stranded fragments of DNA 15 base pairs long from extensive nuclease digestion.[31]

Since polymerase-promoter contact is apparently reduced when core polymerase binds sigma, the affinity of polymerase for DNA in general should decrease. Indeed sigma is known to decrease the affinity of polymerase for nonpromoter DNA by a factor of 10,000.[25] However, simultaneously, sigma enhances the affinity of polymerase for promoter.[25] This apparent contradiction may be reconciled by suggesting that in the absence of sigma, polymerase assumes a conformation which prevents its DNA binding monomers from simultaneously engaging both Pribnow boxes base specifically. With incorporation of sigma, polymerase then assumes a conformation which not only reduces nonspecific DNA contacts, but simultaneously orients both DNA binding monomers in such a way as to allow formation of simultaneous base-specific contacts to each Pribnow box. The observation that binding of sigma to polymerase results in a more compact core enzyme[32] suggests that sigma may "pull" the two DNA binding monomers of polymerase into a position where simultaneous contacts to each Pribnow box are possible.

Initiation of Transcription

Previous measurements have implicated the formation of two base specific complexes in the initiation of transcription.[4] The first complex, termed the closed complex, presumably binds promoter base-specifically but does not melt DNA. In terms of the proposed structure of RNA polymerase and promoter, this complex would involve simultaneous binding of the β and β' subunits of holoenzyme to both Pribnow boxes of promoter. Since this complex does not melt DNA, recognition would be restricted to sites accessible on the outside of the helix. Direct evidence for simultaneous base-specific contacts in both Pribnow boxes by polymerase has been obtained by a variety of chemical probes.[13]

In equilibrium with the closed complex is an open complex which can rapidly initiate RNA synthesis. Open complex formation exhibits a pronounced sigmodial dependence on temperature which varies with the sequence of promoter.[33] This, as well as other observations, indicates that transition from a closed to an open complex cooperatively melts promoter.[4] However, since induction of a native conformation in unbound RNA polymerase by sigma exhibits a similar temperature dependence,[34] open complex formation may involve simultaneous conformational changes in both promoter and polymerase. If so, sigma may induce the closed complex to initiate promoter melting by "pulling" one or more bases of Pribnow's box into a melted configuration.

In most cases, transition from closed to open complex is the rate-limiting step in transcription.[4] However, a notable exception is the lac UV-5 mutant whose rate of transcription is governed by a step subsequent to open complex formation.[35] As previously discussed, the sequence organization of the lac UV-5 Pribnow box appears to greatly enhance its ability to undergo a B → H → melt transition when bound by RNA polymerase. Since, initiation of melting by the UV-5 Pribnow box may not require polymerase to pull bases into a melted configuration, it may be suggested that a slow rate-limiting conformational change in polymerase is avoided.

Promoter Strength

The experimental observation that (in general) transition of promoter from a closed to an open complex is the rate-limiting step in transcription allows one to identify those features of promoter sequence which are expected to play an important role in its transcriptional efficiency. For example, mutations which render the sequence of Pribnow's box more homologous to the UV-5 sequence are expected to increase promoter efficiency for two reasons. First, these mutations would enhance the ability of Pribnow's box to form both closed and open complexes by facilitating initiation of a B → H →

melt transition in the direction of transcription when bound by polymerase. Secondly, since base-specific contacts between Pribnow's box and polymerase are utilized in the formation of both closed and open complexes, single base pair mutations could lead to incorrect or unproductive contacts thereby impairing promoter recognition. Approximately 15 promoter mutants in Pribnow's box have been sequenced and those mutations which increase the homology between Pribnow's box and the UV-5 sequence enhance promoter efficiency whereas decreases in homology impair its efficiency.[3] This observation indicates that those base changes in Pribnow's box which enhance its ability to undergo obligatory conformational transitions also enhance its ability to contact polymerase. Since the UV-5 sequence is the most conserved as well as the most efficient Pribnow box[3], polymerase presumably has evolved base-specific contacts which are maximized for this sequence.

A second feature of a promoter sequence which is expected to play an important role in its transcriptional efficiency is the (G+C) content of the 15 base pair melting region. Whereas a low (G+C) content would enhance transcription, a high (G+C) content would not necessarily prevent transcription since formation of a tight closed polymerase-promoter complex

may still be possible. An interesting example in this respect is the lac promoter. Transcription of wild type lac promoter requires catabolite activating protein (CAP).[36] CAP apparently stimulates transcription by binding base-specifically to a site 60 base pairs upstream from the startpoint of lac transcription.[37] The location of the CAP binding site has suggested two possible mechanisms for its stimulation of lac transcription: CAP could directly contact polymerase thereby stabilizing its interaction with promoter[37,38] or it could induce a transcriptionally active conformation in promoter.[6,39] Although direct evidence for or against these hypotheses is lacking, the proposed sequence organization of procaryotic promoters suggests a third possible mechanism. For example, overlapping the lac promoter, and adjacent to the CAP binding site, one can locate a second potential promoter (which will be termed the pseudo lac promoter) (Fig. 5). Since the Pribnow box in the melting and recognition region of both promoters are very similar, closed polymerase complexes could form equally well at either promoter. However, the (G+C) content in the melting region of the pseudo lac promoter is twice that found in the lac promoter. As a result, the efficiency of transcription from the pseudo lac promoter would be much lower than from the lac promoter. Since the pseudo lac promoter overlaps the CAP binding site, it may be suggested that

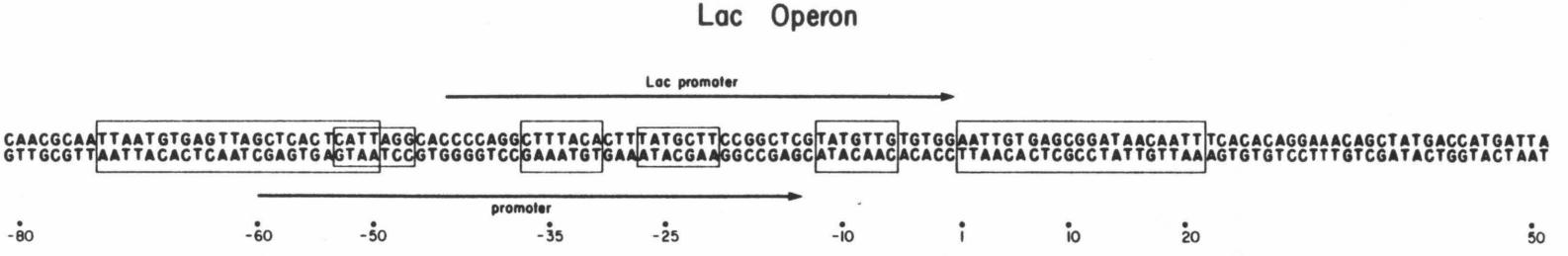


Figure 5
Nucleotide sequence of the lac operon. The catabolite activating protein binding site at (-60) and the lac repressor binding site at (+10) are denoted by large boxes. The Pribnow box and recognition of the pseudo lac promoter are shown at positions -2 and -50, respectively. Similar regions of the lac promoter are indicated at positions -10 and -35, respectively.

one way in which CAP may stimulate lac transcription is to prevent polymerase from being "trapped" by the non-productive pseudo lac promoter. Interestingly, replacement of the wild type Pribnow box with the UV-5 sequence removes the dependency of the lac promoter on CAP stimulation.[40] Since the UV-5 lac promoter would bind polymerase much more tightly than the pseudo lac promoter, CAP binding would not be required to block access of polymerase to the pseudo lac promoter. Thus the lac operon may consist of two promoters, one repressed by CAP and the other repressed by lac repressor.

As previously discussed, analysis of mutations in the recognition region indicates that polymerase makes base-specific contacts within this region which are maximized for an inverted UV-5 Pribnow box. Since similar homology is demanded within Pribnow's box, transcriptional efficiency should be governed by the relative strength of each Pribnow box. Comparison of 60 procaryotic promoter sequences[3] indicates that the Pribnow box nearest the startpoint of transcription exhibits greater homology to the UV-5 sequence than the inverted Pribnow box in the recognition region.

Eucaryotic and Viral Promoters

The following observations suggest that in some, but perhaps not all cases, the sequence organization of procaryotic and eucaryotic promoters are similar.

1) Eucaryotic RNA polymerase II and E.coli RNA polymerase have been shown by electron microscopy to form stable initiation complexes at identical sites on eucaryotic DNA (SV-40 and polyoma DNA).[41]

2) The fibroin promoter,[42] which is transcribed by eucaryotic RNA polymerase II, exhibits a procaryotic like sequence organization (Fig. 6).

Figure 6
Fibroin promoter. Numbers denote the distance from the 5' end of the mRNA coding part of the gene.

3) The adenovirus VA1 RNA promoter,[43] the yeast,[44] X borealis oocyte[45] and X borealis somatic 5S promoters[45] as well as the $tRNA^{phe}$ and $tRNA^{tyr}$ promoters of X laevis,[46] all of which are transcribed by eucaryotic RNA polymerase III, exhibit a procaryotic like sequence organization (Figs. 7 and 8).

Figure
Adeno virus VA1 promoter. Arrows mark the 5' ends of the VA1(A) and VA1(G) RNA's and the two base pairs missing in the 309 promoter mutant are denoted (Λ).

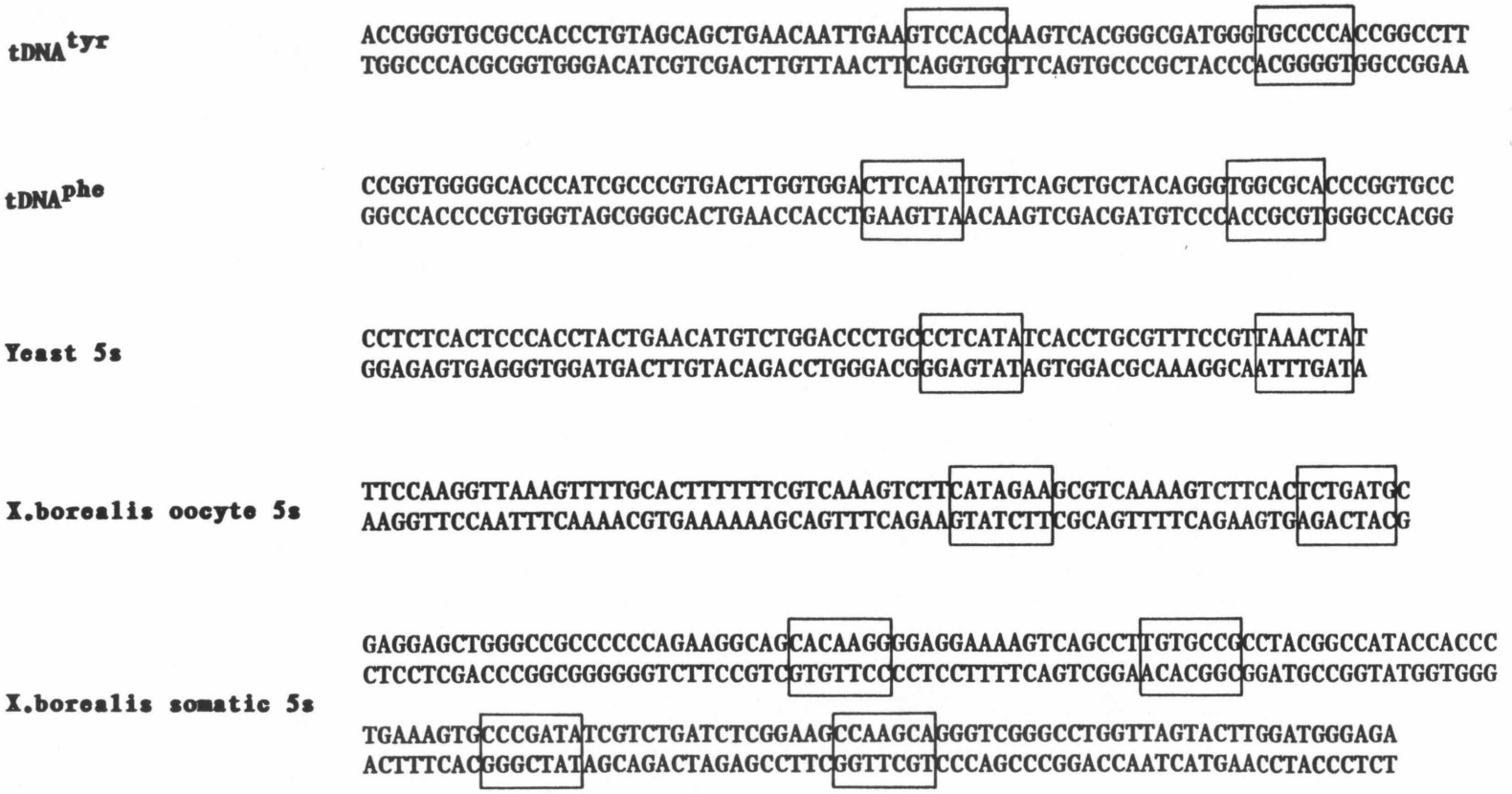


Figure 7
Numbers denote the distance from the 5' end of the mRNA coding part of the gene and arrow indicate transcription start points.

a) The observed variance in transcription startpoints may arise from subunits differences within the RNA polymerase III class of eucaryotic polymerases.

b) The suggestion that the Xbl gene contains two promoters, one of which lies within the "control region,"[47] suggests that transcriptional control may be mediated by RNA polymerase-RNA polymerase as well as RNA polymerase-control factor protein interactions.

4) The ribosomal gene unit of X laevis,[48] which is transcribed by RNA polymerase I, exhibits procaryotic-like promoters (Fig. 9). The observation of low but detectable transcription from the BAM islands[49] may be attributable to the presence of the proposed promoters at positions 2270 and 3380.

T7 Promoters

Comparison of 13 class II and III T7 promoters[50] suggests that these phage specific promoters contain a Pribnow box separated 10-11 base pairs from an inverted Pribnow box and thus exhibit a sequence organization similar to procaryotic promoters (Fig. 10).

Since the T7 viral polymerase contains only one very large polypeptide chain (molecular weight 107,000),[52] fusion

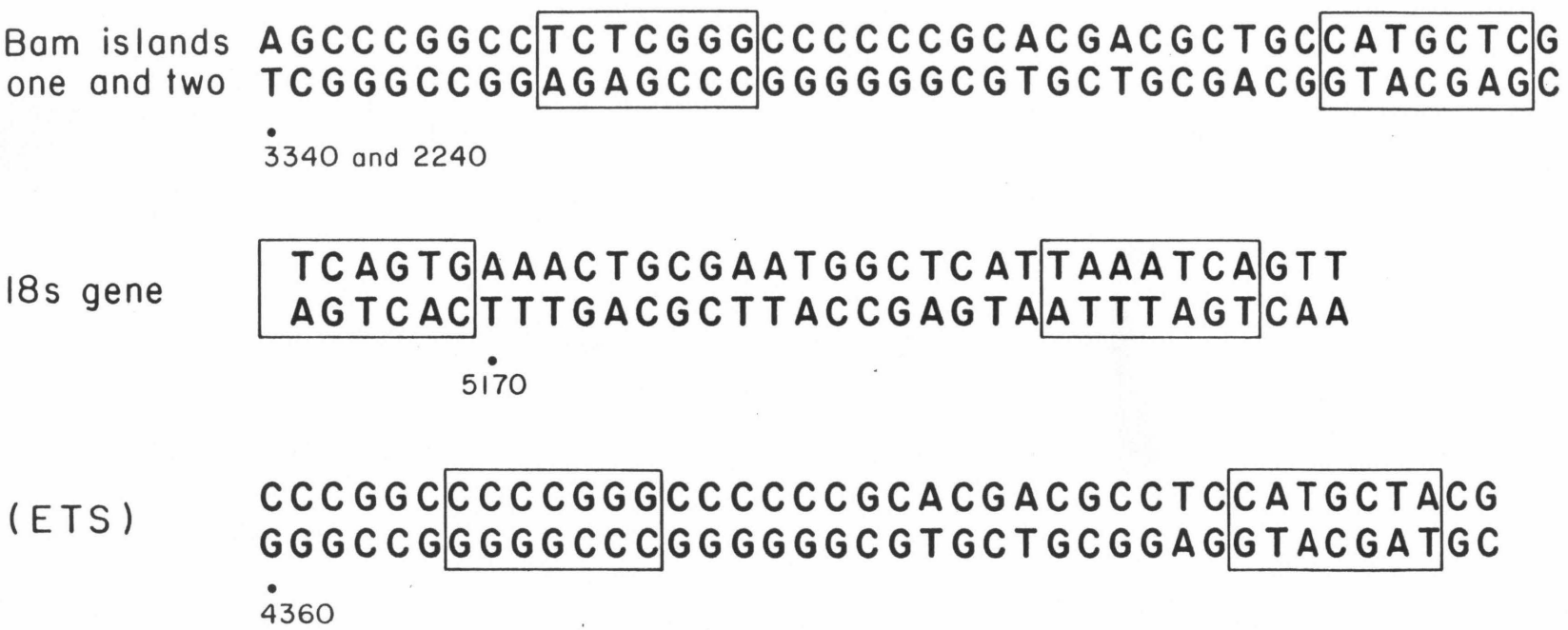


Figure 9

X. laevis ribosomal gene unit.

of ancestral procaryotic polymerase monomers may have occurred later in its evolution. In contrast, the large number of subunits characteristic of eucaryotic polymerases[53] suggest that they may have evolved from procaryotic polymerases by incorporating additional subunits into polymerase.

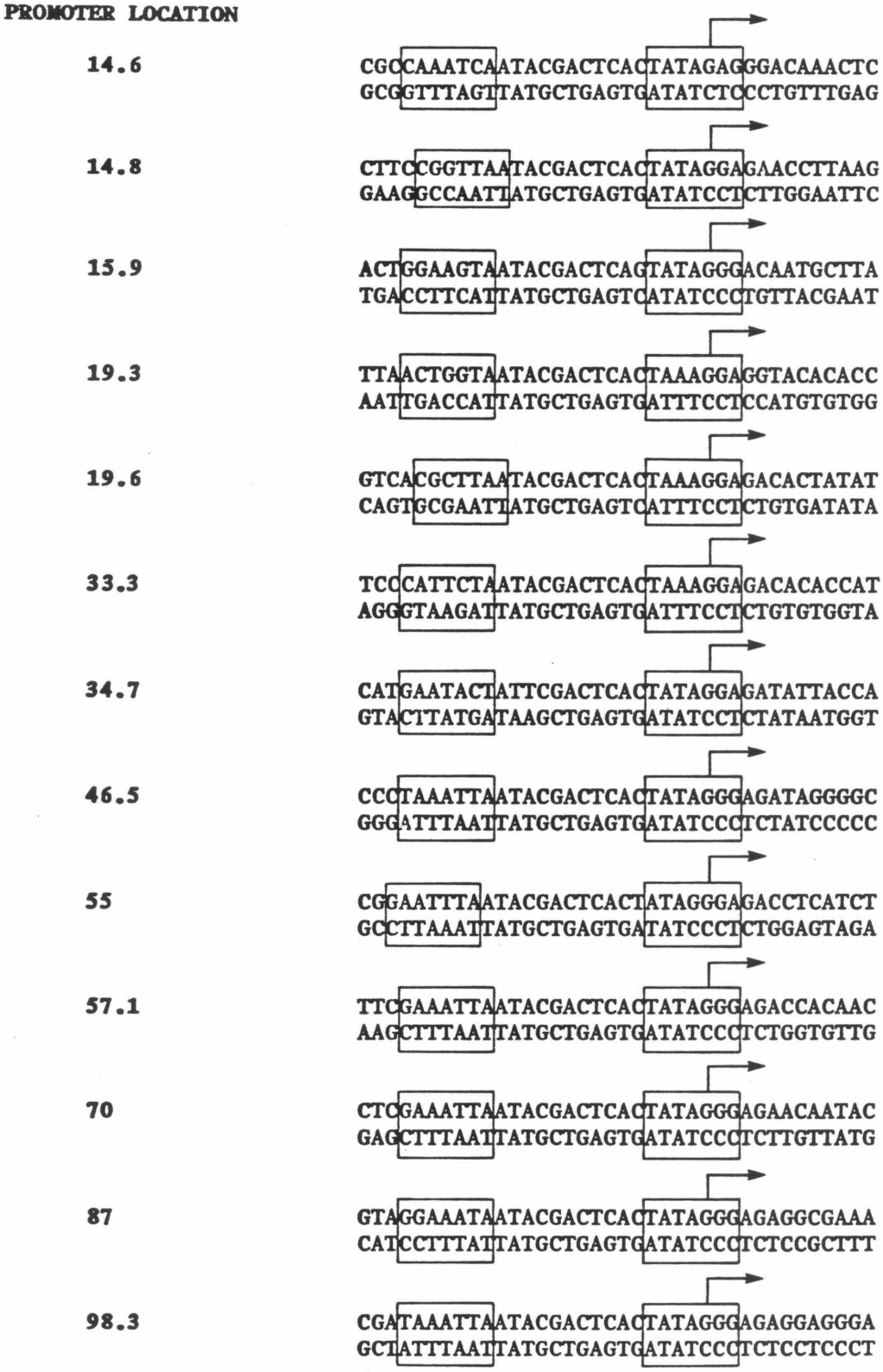


Figure 10
T7 class II and III promoters. Arrows denote start points of transcription.

References

1) Becker, M.M., submitted to Nature, 1979.

2) Jacob, F.; Ullman, A.; Monod, J., C.R. Acad. Sci. Paris, 1964, 258, 3125.

3) For a recent review see Rosenberg, M.; Court, D., Ann. Rev. Gen., 1979, 13, 319.

4) For example, see Chamberlin, M.J., Ann. Rev. Biochem., 1974, 43, 721.

5) a) Silverstone, A.E.; Arditti, R.R.; Magasanik, B., Proc. Natl. Acad. Sci. USA, 1970, 66, 773. b) Eron, L.A.; Block, R., ibid, 68, 1828.

6) Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

7) Dickson, R.; Abelson, J.; Barnes, W.; Reznikoff, W., Science, 1975, 187, 27.

8) a) Pribnow, D., Proc. Natl. Acad. Sci. USA, 1975, 3, 784. b) Takanami, M.; Sugimoto, K.; Sugisaki, H.; Okamoto, T., Nature, 1976, 260, 297. c) Seeburg, P.; Nusselelin, C.; Schaller, H., Eur. J. Biochem., 1977, 74, 107.

9) Pribnow, D., J. Mol. Biol., 1975, 99, 419.

10) Schaller, H.; Gray, C.; Herrman, K., Proc. Natl. Acad. Sci. USA, 1975, 72, 737.

11) Melnikova, A.F.; Beabealashvill, R.; Mirzabekov, A.D., Eur. J. Biochem., 1978, 84, 301.

12) Siebenlist, U., Nature, 1979, 279, 651.

13) For example see Siebenlist, U.; Simpson, R.T.; Gilbert, W., Cell, 1980, 20, 269.

14) Gilbert, W.; Gralla, J.; Majors, J.; Maxan, A., "Protein-Ligand Interactions," Sand, H.; Blauer, G., eds., Walter de Gruyter and Co., Berlin, 1975, p. 193.

15) For an example see Nash, H.A., Curr. Top. Microbiol. Immunol., 1977, 78, 171.

16) a) Mizuuchi, K.; Gellert, M.; Nash, H.A., J. Mol. Biol., 1978, 121, 375. b) Mizuuchi, K.; Mizuuchi, M., Cold Spring Harbor Symp. Quant. Biol., 1979, 43, 1111.

17) a) Nash, H.A.; Kikuchi, Y., J. Biol. Chem., 1978, 253, 7149. b) Kikuchi, Y.; Nash, H.A., Cold Spring Harbor Symp. Quant. Biol., 1979, 43, 1099.

18) a) Landy, A.; Ross, W., Science, 1977, 16, 1147. b) Ross, W.; Landy, A.; Kikuchi, Y.; Nash, H., Cell, 1979, 18, 297.

19) Kikuchi, Y.; Nash, H.A., Proc. Natl. Acad. Sci. USA, 1979, 8, 3760.

20) Caolos, M.P., Nature, 1978, 274, 762.

21) Kleid, D.; Humayun, Z.; Jeffery, A.; Ptashne, M., Proc. Natl. Acad. Sci. USA, 1976, 73, 293.

22) Burgess, R.R., J. Biol. Chem. 1969, 244, 6168.

23) a) Fukuda, R.; Ishihama, A., J. Mol. Biol., 1974, 87, 523. b) Ishihama, A.; Taketo, M.; Saltoh, T.; Fukuda, R., "RNA Polymerase," Losick R.,; Chamberlin, M.J., eds., Cold Spring Harbor Laboratory, 1976, 485.

24) For example, see Burgess, R.R., ibid, 69.

25) For example, see Hinkle, D.C.; Chamberlin, M.J., J. Mol. Biol., 1972, 70, 157.

26) Travers, A.A.; Burgess, R.R., Nature, 1969, 222, 537.

27) Saucier, J.M.; Wang, J.C., Nature, New Biology, 1972, 238, 167.

28) Wang, J.C., Jacobsen, J.H.; Saucier, J.M., Nucl. Acid Res., 1977, 4, 1225.

29) deHaseth, P.L.; Lohman, T.M.; Burgess, R.P.; Record, M.T., Biochemistry, 1978, 17, 1612.

30) Kumar, S.A.; Krakow, J.S., J. Biol. Chem., 1975, 250, 2878.

31) Giacomoni, P.U.; LeTalaer, J.Y.; LePecq, J.-B., Proc. Natl. Acad. Sci. USA, 1974, 8, 3091.

32) Pilz, I.; Ratky, O.; Rabussay, D., Eur. J. Biochem., 1972, 28, 205.

33) a) Mangel, W.F.; Chamberlin, M.J., J. Biol. Chem., 1974, 249, 3007. b) Seeburg, P.H.; Nusstein, C.; Schaller, H., Eur. J. Biochem., 1977, 74, 107.

34) Wu, F.Y.-H.; Yarbrough, L.R.; Wu, C.-W., Biochemistry, 1976, 15, 3254.

35) Stefano, J.E.; Gralla, J., Biochemistry, 18, 1063.

36) For example, see Majors, J., Proc. Natl. Acad. Sci. USA, 1975, 11, 4394.

37) a) Simpson, R.B., Nucl. Acid Res., 1980, 8, 759. b) Majors, J., J. Biol. Chem., in press.

38) a) Majors, J., Nature, 1975, 256, 672. b) Gilbert, W., "RNA Polymerase", Losick, R.; Chamberlin, M., eds., Cold Spring Harbor Laboratory New York, 1976, 193.

39) Dickson, R.C.; Abelson, J.; Barnes, W.M.; Reznikoff, W.S., Science, 1975, 187, 27.

40) Eron, L.; Block, R., Proc. Natl. Acad. Sci. USA, 1971, 68, 1828.

41) a) Lescure, B.; Dauguet, C.; Yaniv, M., J. Mol. Biol., 1978, 124, 87. b) Saragosti, S.; Croissant, O.; Yaniv, M., Eur. J. Biochem., 1980, 106, 25.

42) a) Tsujimoto, Y.; Suzuki, Y., Cell, 1979, 16, 425. b) Tsuda, M.; Ohshima, Y.; Suzuki, Y., Proc. Natl. Acad. Sci. USA, 1979, 76, 4872.

43) Thimmappaya, B.; Jones, N.; Shenk, T., Cell, 1979, 18, 947.

44) a) Maxam, A.; Tizard, R.; Skryabin, K.G.; Gilbert, W., Nature, 1977, 267, 643. b) Valenzuela, P.; Bell, G.I.; Masiarz, F.R.; DeGennaro, L.J.; Rutter, W.J., Nature, 1977, 267, 641.

45) Korn, J.L.; Brown, D.D., Cell, 1978, 15, 1145.

46) Muller, F.; Clarkson, S.G., Cell, 1980, 19, 345.

47) a) Sakonju, S.; Bogenhagen, D.F.; Brown, D.D., Cell, 1980, 19, 13. b) Engelke, D.R.; Ng, S.-Y.; Shastry, B.S.; Roeder, R.G., Cell, 1980, 19, 717.

48) Boseley, P.; Moss, T.; Machler, M., Portman, R.; Birnstiel, M., Cell, 1979, 17, 19.

49) For example see Rungger, D.; Drippa, M.; Trendelenburg, M.F.; Scheer, U.; Franke, W.W., Exp. Cell Res., 1978, 116, 481.

50) For example, see McAllister, W.T.; Carter, A.D., Nucl. Acid Res., 1980, 8, 4821.

51) Hillel, Z.; Wu, C.-W., Biochemistry, 1977, 16, 3334.

52) Chamberlin, M.J.; Ring, J., J. Mol. Biol., 1972, 70, 221.

53) See for example Roeder, R., "RNA Polymerase," Losick, R.; Chamberlin, M.J., eds., Cold Spring Harbor Laboratory, 1976, 285.

PROPOSITION 3

Genetic Regulation Through Allosteric Control

Macromolecules, composed of subunits, can exhibit different conformations whose relative proportions may be greatly altered by small changes in the concentration of effector molecules.[1] If effector binding induces all macromolecule subunits to change conformation, the transition is termed allosteric.[1] Since different biological activities are often associated with different conformational states, the abrupt response of an allosteric transition to an effector molecule represents a simple mechanism for controlling biological activity.[1] When the free energy difference between different conformational states is large, an allosteric transition may occur throughout many subunits enabling control of biological activity from a distance.

Although allosteric transitions play a critical role in the function of many regulatory proteins[1,2] little is known about their function in nucleic acids. However, the postulated obligatory involvement of DNA conformation in genetic regulation[3] suggests that allosteric transitions could play an important role in regulating the genetical potential of DNA. Consider, for example, the interaction of lambda repressor with its operon. Lambda repressor has been shown to bind two operators, O_R and O_L, which

contain multiple repressor binding sites 17 base pairs long (Fig. 2, Proposition 1).[4] Binding of repressor to operator is cooperative since binding isotherms (Fig. 3a, Proposition 1) are sigmoidial.[5] Reexamination of these measurements (Fig. 3b, Proposition 1) also indicates that lambda repressor unwinds operator since supercoiling enhances repressor binding. However, when more than one repressor is bound at operator, supercoiling no longer enhances repressor binding. These results indicate that bound repressor also alters the binding properties of adjacent DNA.

The binding cooperativity, operator unwinding and alteration of DNA binding properties by repressor may result from a common mechanism: an induced allosteric transition in operator by repressor. For example, the two state or allosteric model of Monod[1] would postulate that operator exists in a strong binding (T) and weaker binding (R) form. In the absence of repressor, operator is in the more thermodynamically favored (R) form. However, as repressor is added the equilibrium shift from the (R) to the tightly binding (T) form due to the latter's enhanced affinity for ligand. This model accounts for the cooperativity of repressor binding because with increased repressor binding there is an increased preference for DNA to assume a more tightly binding (T) conformation.[1] Operator unwinding is also compatible with

the allosteric model since repressor is postulated to induce a conformational change, R → T, in operator. A likely candidate would be a B → H conformational transition[5,18] since base-specific recognition of DNA by regulatory protein binding has been postulated to necessitate at least a B → H allosteric transition in DNA (see Proposition 2). Lastly, the observed alteration of operator binding properties induced by repressor binding would represent the enhancement R → T conformational transitions in operator which result from the creatment of conformational interfaces by bound repressor (Fig. 4, Proposition 1).

In addition to providing a molecular basis for the binding properties of lambda repressor, the allosteric model also provides a simple molecular basis for the observation that lambda repressor can activate as well as deactivate its own transcription.[4,7] For example, when lambda repressor binds O_{R3}, it both <u>inhibits</u> transcription from its structural gene Ic and facilitates transcription from the distant <u>cro</u> gene (see Fig. 2, Proposition 1). Similarly, lambda repressor when bound at O_{R1} inhibits transcription from the <u>cro</u> promoter and <u>facilitates</u> transcription from its distant structural Ic gene. These results can readily be understood in terms of the allosteric model presented above. When bound to O_{R1}/O_{R3} repressor sterically prevents polymerase from binding to the <u>cro</u>/Ic promoter. However, since repressor enhances nearby

R → T transitions, bound repressor will also increase the probability that its distant Ic promoter will assume an activated (T) conformation. Other proteins which modulate transcription, like catabolite activating protein,[19] T antigen,[16] and the cro repressor[8] may also exert all or part of their effects by enhancing B → H transitions in DNA.

Although an allosteric transition in operator appears to be induced by lambda repressor binding, alternative models must be considered. For example, physical contact between lambda repressors could account for the cooperativity of repressor-operator binding whereas repressor activated transcription may arise from contacts between repressor and RNA polymerase. These two models, allosteric and contact, may be distinguished from one another by monitoring the effect of supercoiling on repressor binding and repressor activated transcription. For example, if the contact model is correct, little change in the binding cooperativity of repressor should result when the supercoiled density of operator is altered. In contrast, as the negative supercoiled density of operator is increased/decreased the allosteric model predicts that repressor binding will become less/more cooperative. For positively supercoiled operator, binding is predicted to be weaker and more cooperative than

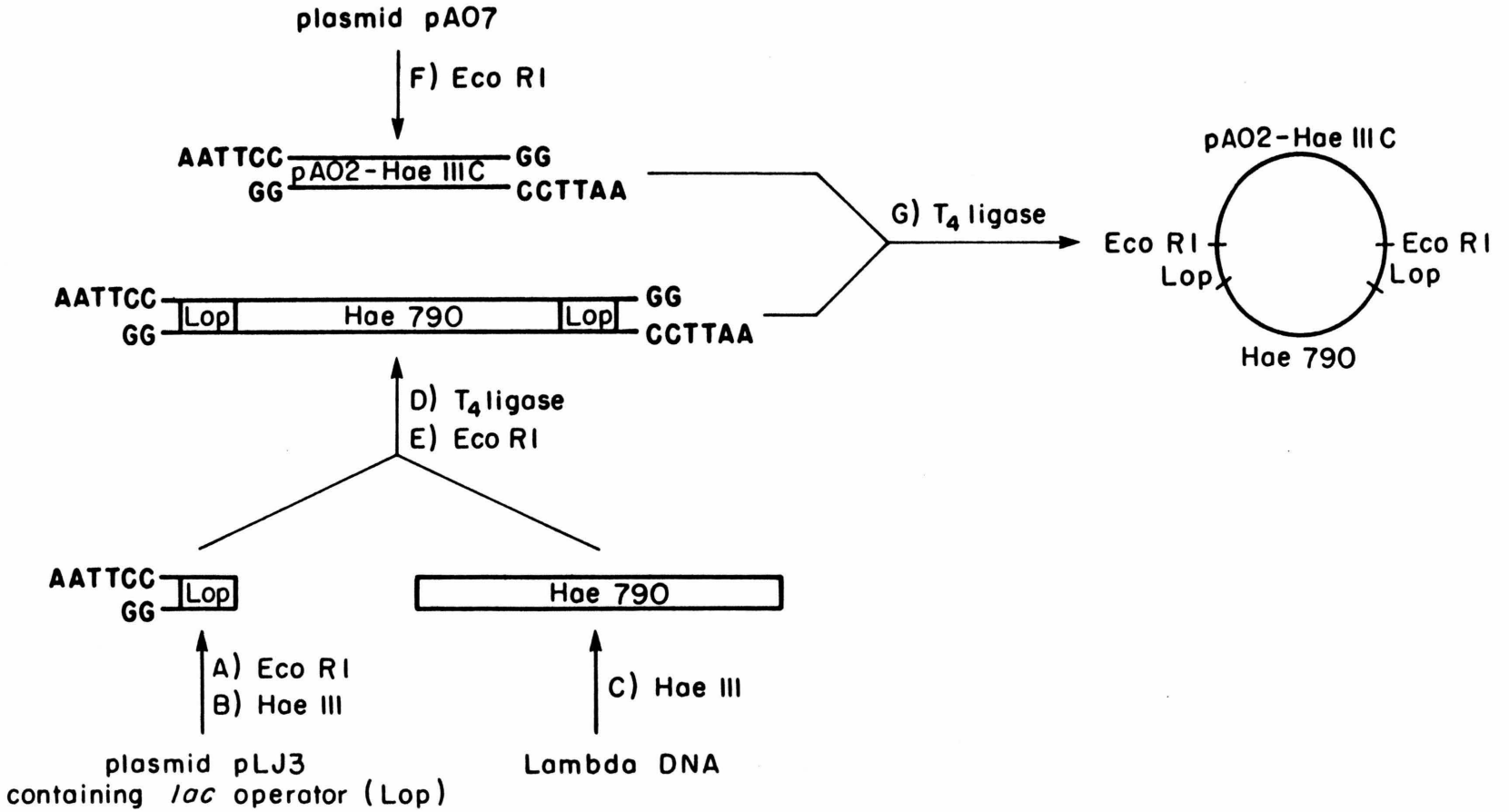
to linear operator. The allosteric model also predicts that as the negative supercoiled density of operator is increased/decreased, transcriptional activation by repressor should decrease/increase relative to the case where no repressor is present. In contrast, supercoiling should have little effect on repressor stimulated transcription if activation requires repressor-polymerase contact. It should be noted that both the allosteric and direct contact models, either acting separately or in concert, could modulate repressor-repressor and/or repressor-polymerase interactions.

The effect of supercoiling on lambda repressor binding cooperativity and transcriptional activation may be assayed using a plasmid whose construction is outlined in Figure 1. This plasmid contains the three O_R operators necessary for monitoring lambda (and/or cro) repressor binding. In addition, the P_R and P_{RM} promoters necessary for monitoring lambda (and/or cro) repressor stimulated transcription are present and adjacent to lac operators. Transcription in vitro of this plasmid in the presence of lac repressor should yield only two transcripts: a 140 bp P_R and 330 bp P_{RM} mRNA. Subjecting purified plasmid to N/C enzyme in the presence of varying amounts of ethidium bromide followed by removal of the drug should generate plasmids of varying negative supercoiled density.

Positively supercoiled plasmid may be obtained by the action of N/C enzyme in the presence of high concentrations of cations followed by their removal.[9] *In vitro* assays for monitoring repressor binding cooperativity[5] and transcriptional activation[8,10] have been previously described.

Figure

Construction of a plasmid for investigating the effect of supercoiling on lambda repressor-operator binding and lambda repressor activated transcription. A 117 base pair (bp) fragment containing the lac operator is isolated (A,B) from the pLJ3 plasmid[11] and blunt end ligated[12] (D) to the 790 bp H_{AE}III fragment[10] (C) of lambda DNA. Following E_{CO} RI restriction (E) the resulting fragment is ligated[13] (G) to the pAO2-H_{AE}III C fragment (F) of the CO1 EI pAO2 plasmid.[14] The final plasmid contains the entire right operator of lambda and bears viable P_R and P_{RM} promoters flanked on each side by a lac operator (LOP). The lac operators allow selection (β galactosidase production)[15,17] of the final desired plasmid and assure that transcription of the P_R and P_{RM} promoters will yield reproducible and discrete RNA transcripts. Transcription in vitro of the final plasmid should yield a 140 bp P_R and 330 bp P_{RM} mRNA. No other transcripts will be produced since the pAO2-H_{AE}III C fragment cuts in the middle of the pAO2 RNA-I promoter.[14] Incorporation of the pAO2-H_{AE}III C fragment enables the final plasmid to autonomously replicate.[14]


plasmid pAO7
F) Eco RI
AATTCC
GG
pAO2-Hae III C
GG
CCTTAA
AATTCC
GG
Lop
Hae 790
Lop
GG
CCTTAA
G) T4 ligase
pAO2-Hae III C
Eco RI
Lop
Eco RI
Lop
Hae 790
D) T4 ligase
E) Eco RI
AATTCC
GG
Lop
Hae 790
A) Eco RI
B) Hae III
C) Hae III
plasmid pLJ3
containing lac operator (Lop)
Lambda DNA

References

1. Monod, J., Wyman, J. & Changeux, J.P., J. Mol. Biol. 12, 88-118 (1965).
2. For example see Hames, G.C. & Wu, C-W., Ann. Rev. Biophys. 3, 1-33 (1974).
3. Becker, M.M., Nature, submitted.
4. See for example, Ptashne, M., Backman, K., Humayun, M.Z., Jeffry, A., Maurer, R., Meyer, B. & Sauer, R.T., Science 194, 151-161 (1976).
5. Maniatis, T. & Ptashne, M., Proc. Natl. Acad. Sci. USA 70, 1531-1535 (1973).
6. Humayun, Z., Jeffery, A. & Ptashne, M., J. Mol. Biol. 112, 265-277 (1977).
7. Walz, A., Pirrotta, V. & Ineichen, K., Nature 262, 665-669 (1976).
8. Johnson, A., Meyer, B.J. & Ptashne, M., Proc. Natl. Acad. Sci. USA 75 1783-1787 (1978).
9. Anderson, P. & Bauer, W., Biochemistry 17, 594-601 (1978).
10. Meyer, B.M., Kleid, D.G. & Ptashne, M., Proc. Natl. Acad. Sci. USA 72, 4785-4789 (1975).
11. Johnsrud, L., Proc. Natl. Acad. Sci. USA 74, 5314-5318 (1978).

12. Sgaramella, V., van de Sande, J. & Khorana, H.G., Proc. Natl. Acad. Sci. USA 67, 1468-1475 (1970).

13. Sgaramella, V. & Khorana, H.G., J. Mol. Biol., 72, 427-444 (1972).

14. Oka, A., Nomura, N., Morita, M., Sugisaki, H., Sugimoto, K. & Takanami, M., Molec. Gen. Gent. 172, 151-159 (1979).

15. Bolivar, F., Betlach, M., Heyneker, H., Shine, J., Rodriquez, R. & Boyer, H.W., Proc. Natl. Acad. Sci. USA 74, 5265-5269 (1977).

16. Reed, S.I., Stark, G.R. & Alwine, J.C., Proc. Natl. Acad. Sci. USA 73, 3083-3087 (1976).

17. Heyneker, H.L., Shine, J., Goodman, H.M., Boyer, H.W., Rosenberg, J., Dickerson, R.E., Narang, S.A., Itakura, K., Lin, S., & Riggs, A., Nature 263, 748-752 (1976).

18. Becker, M.M. & Dervan, P.B., Nature, submitted.

19. Zubay, G., Schwartz, D. & Beckwith, J., Proc. Natl. Acad. Sci. USA 66, 104-110 (1970).

PROPOSITION 4

Physiochemical Characterization of a DNA Conformational Code

In Chapter III evidence for a new conformational family of Watson-Crick DNA was presented. Termed H or hybrid DNA, such DNA was postulated to be an obligatory intermediate in the interfamily B → A transition. The ease with which DNA undergoes B → H → A transitions was also found to vary greatly with its sequence. On the basis of these results, the equilibrium stability, rather than the structure of Watson-Crick DNA was postulated to vary greatly with base sequence.

Since the ability of DNA to undergo B → H transitions varies greatly with base sequence, expression of DNA functions which necessitate such transitions is expected to be influenced by base sequence. Thus, in addition to a genetic code, the base sequence of DNA may harbor a conformational code which specifies the ease with which a given sequence of DNA assumes conformations other than the B family. If, as proposed in Proposition 1, B DNA is genetically inactive, such a conformational code would be expected to play an extremely important role in gene expression. The following experiments are designed to probe, both directly and indirectly, the occurrence of B ↔ H ↔ A transitions in DNA and their possible role(s) in genetic regulation.

Since the free energy of H DNA is intermediate between B and A conformations, and its transition to B or A forms cooperative, it will (in general) be an unstable intermediate in the B ↔ A transition. Thus classical techniques normally used to induce B → A transitions, such as the addition of small amounts of ethanol, are likely to miss B → H or H → A transitions. However since intercalators such as EB and BMSp strongly prefer A and H conformations over B conformations, their binding could be used to progressively shift the B ↔ H ↔ A equilibrium allowing detection of each conformation.[1] For DNA which is 1000 base pairs long, the amount of BMSp needed to induce a B → H transition in the most conformationally inert DNA (dAdT) is approximately one BMSp per 25 base pairs; other base sequences would require significantly less BMSp.[1] Since the concentration of BMSp required to induce a B → A transition is so small, changes in the physical properties of DNA could be used to directly monitor the transition without interference from bound drug. A variety of physical techniques, such as circular dichroism, NMR spectroscopy, fluoresence depolarization measurements[2] or electric dicrhoism[3] could be used to simultaneously monitor shifts in the B ↔ H ↔ A equilibrium upon BMSp or EB addition. Once a convenient technique was established, the indiction of a B → H → A

transition by other drugs or proteins could be investigated.

It has been proposed that base specific recognition of DNA by regulatory proteins necessitates at least a B → H transition in DNA. Since the ability of DNA to undergo a B → H transition varies greatly with base sequence, sequences known to inhibit B → H transitions, when ligated to a regulatory protein target sequence, would be expected to impair protein expression. Presumably if the binding of a regulatory protein to its target necessitates a B → H transition, its association rate and/or affinity will progressively decrease when the linker is changed from d(C-G) to dCdG and then dAdT (ligation linkers must be of a similar length since the ease of a B → H → A transition is dependent on molecular weight).[1]

Another approach which could be used to monitor the effect of B → H transitions on regulatory protein binding would involve insertion[4] of d(C-G), dCdG and dAdT fragments, of variable but known length, into restriction sites nearby protein target sequences and monitoring subsequent protein expression.

Lastly, since B → H transitions unwind DNA,[1] the effect of supercoiling on the proposed ligation and insertion experiments could also be investigated. Negative supercoiling is expected to enhance B → H transitions whereas positive supercoiling is expected to inhibit their occurrence.

References

1. Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.

2. a) Barkely, M.D. & Zimm, B.H., J. Chem. Phys. 1979, 70, 2991. b) Millar, D.P., Robbins, R.J. & Zewail, A.H., Proc. Natl. Acad. Sci. USA 1980, 77, 5593.

3. Hogan, M., Dattagupta, N. & Crothers, D.M., Biochemistry 1979, 18, 280.

4. Wang, J.C., Proc. Natl. Acad. Sci. USA 1979, 76, 2000

PROPOSITION 5

Specific Localization of Conformational Hotspots in DNA

The polyintercalator BMSp has been shown to exhibit greatly enhanced specificity for A and H conformations of DNA.[1] Since the ease with which DNA undergoes a B → H → A transition varies greatly with base sequence, DNA whose sequence is heterogeneous should exhibit conformational hotspots.[1] We wish to understand how base sequence, supercoiling, drug or protein binding, etc. can influence the reactivity and distribution of conformational hotspots. Since BMSp nicks DNA photochemically (personal observation) specific localization of conformational hotspots could be achieved by first binding BMSp to DNA and then photochemically inducing a nick at its binding site. Those sequences or regions of the DNA which have been nicked can be identified by either subsequently treating with S_1 nuclease to introduce a specific double strand break[2] or if a free 3'-hydroxyl group can be generated at the nick, carrying out a nick translation reaction with DNA polymerase I.[3] Restriction analysis of BMSp-S_1 induced double strand breaks or hybridization of nick translated material to known sequences localizes the conformational hotspots.[4] Alternatively, if BMSp binds DNA covalently when it nicks, one could incorporate a radioactively labeled

BMSp at the nick. Subsequent restriction and hybridization of labeled single strand specifically localizes the hotspot.

References

1. Becker, M.M., Ph.D. Dissertation, California Institute of Technology, 1981.
2. Beard, P., Morrow, J.F. & Berg, P., J. Virol. 1973 12, 1303.
3. Levitt, A., Axel, R. & Cedar, H., Devel. Biol. 1979, 69, 496.
4. Zasloff, M. & Camerini-Otero, D.R., Proc. Natl. Acad. Sci. USA 1980, 77, 1907.