The Thermodynamics of Oligonucleotide-Directed Triple Helix Formation at Single DNA Sites

Thesis by

Scott F. Singleton

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1995

(Submitted June 24, 1994)

For my family…

To my parents, Earl and Beverly:
One of the pleasures of maturing is developing a richer understanding of your unswerving dedication to your family and all the blessings that has brought me.

To Greg and Suz:
You taught me the importance of laughter, and
you remind me that life cannot be adequately described by statistics.

To my wife, Jennifer:
Your encouragement, patience, and loving support have helped make this possible.
We travel together now. This I promise with joy, today and always.
143∞

ACKNOWLEDGMENTS

I would like to thank my research advisor, Professor Peter Dervan, for his support and encouragement throughout my thesis work. He consistently gave me the space and the respect for my own independent thought, even when that meant the research progressed at a slow and deliberate pace. Much of whatever maturity I have gained as a scientist I owe directly to him.

Warren Wade was just ending a biophysical project when I joined the research group. His example and early guidance helped to get me started. Jumi Shin developed a protocol, similar to that described here, for measuring binding constants by quantitative titration of protein binding domains. We endured many of the same trials, and her experience helped me tremendously.

Mark Distefano, Andrea Staubli, Ken Turnbull, Pete Beal, Hogyu Han, George Best, Natalia Colocci, Michelle Parks, Scott Priestley, Will Greenberg, and David Liberles were all group members who have been involved in measuring equilibrium triplex formation constants. Their collective insight into problems and limitations inherent to such measurements have helped guide the development of the experimental protocol described here. Their efforts have also increased our appreciation of the scope and limitations of this method.

Laura Kiessling and Dave Horne were both post-docs in my lab when I joined the group. That was a particularly difficult time for new beginnings because Peter was traveling much of my first year. Dave and Laura were always there with support, encouragement, and practical guidance when I needed them.

David Mack, Tom Povsic, and Kevin Leubke were older students in the group who are all my good friends. All three had unique perspectives on science, and each served as an example during my own development. Martha

Oakley joined the group at the same time I did, and we shared a laboratory for five years. I can't think of anyone I'd have rather shared with, nor anyone else who would have put up with me for that long. Matt Taylor and Scott Priestley have been good friends and colleagues, especially over the last year and a half when many of my classmates were graduating and moving on.

Bryan Coughlin was a great friend and roommate for more than four years. I'm thankful for all the televised sporting events and bike rides to Long Beach we shared. Rick Muller has also been a truly good friend during my stay at Caltech. We shared a lot of fun times and many conversations that enhanced the breadth of my understanding and appreciation of chemistry. Nat Finney has been my best friend for many years now. I can't think of anyone who exemplifies the notion of "chemist" any better than Nat. I would like to thank him for all the times he was there to help smash things when they needed breaking, to help drink things when they needed drinking, and to do the victory dance when it was time for dancing. It's difficult to imagine coming through this without the Finney.

Finally, I'd like to thank Jennifer Wales Singleton. Jen has been my best friend and constant companion for the past two years. She was a terrific proofreader and editor when I needed her, and I don't think I could have done all this thesis writing without her help. Of all the things I accomplished in the lab, our relationship is the single most important development in my life. I am proud and happy to be her husband. I look forward to our days ahead. Together...

ABSTRACT

The local triple-helical complexes formed upon binding of pyrimidine oligonucleotides to double-helical DNA offer a versatile structural motif for designing sequence specific duplex DNA-binding molecules. The experimental determination of oligonucleotide association constants as a function of solution conditions is necessary to characterize the noncovalent forces which contribute to the affinity and specificity of DNA recognition by triple helix formation. Such information will be important for the rational application of oligonucleotide-directed triple helix formation to such tasks as the manipulation of large DNAs, and the artificial modulation of biological events. We are interested in controlling the relative amount of a local triple-helical complex formed at equilibrium under conditions which approximate those found within cells. Thus, we chose to measure equilibrium constants for the formation of triple-helical complexes at single sites within larger duplex DNAs in mixed valence salt solutions near neutral pH.

Concepts borrowed from existing footprint titration methods were combined with the powerful affinity cleaving strategy in the development of the affinity cleavage titration method for the measurement of single-site DNA-binding isotherms (Chapters One and Two). The amount of site-specific cleavage of a radiolabeled plasmid DNA fragment produced by a 15mer oligonucleotide-EDTA·Fe is measured over an oligonucleotide concentration range covering four orders of magnitude, allowing a binding curve to be constructed. Curve-fitting a Langmuir titration isotherm to the data points affords an equilibrium binding constant, $K_T = 4 \pm 1 \times 10^6\ M^{-1}$ ($\Delta G_T =$ –9 kcal/mol), that is identical within experimental uncertainty to those obtained

from quantitative DNase I footprint titrations of the same oligonucleotide with and without EDTA·In . Further affinity cleavage titrations demonstrate that the energetic stabilization of this local triple-helical structure depends on the length of the oligonucleotide, the presence of base triplet mismatches, and the solution conditions, including the pH, the concentration and valence of cations, and the temperature.

Decreasing the length of the oligonucleotide from 15 to 11 nts reduces the stability of the corresponding triplexes by ~0.3 kcal/mol per nucleotide removed (Chapter Two). Single internal base triplet mismatches result in a destabilization of the local triple-helical structure by ~ 3 kcal/mol. The equilibrium association constants for the 15mer increased by 10-fold as the pH was decreased from pH 7.6 to pH 5.8, indicating that the corresponding triple-helical complex was stabilized by 1.4 kcal·mol^{-1} at the lower pH (Chapter Three). Equilibrium association constants for the 15mer in the presence of various concentrations of KCl, $MgCl_2$, and spermine tetrahydrochloride ($SpmCl_4$) are reported in Chapter Four. Varying the K^+ concentration by a factor of 28 in the presence NaCl, $MgCl_2$, and $SpmCl_4$ resulted in an overall 100-fold *decrease* in the binding affinity from the lowest to the highest concentration. In contrast, measured binding constants *increased* 500-fold as the Spm^{4+} concentration was increased 10-fold. There was a modest effect on the binding constant (a 3-fold decrease) upon increasing the Mg^{2+} concentration by a factor of 100. The influence of temperature on the energetics of oligonucleotide-directed triple helix formation in mixed valence salt solutions at pH 7.0 are reported in Chapter Five. In three solutions differing in their salt compositions, the equilibrium association constants decreased at least 100-fold (from $>10^7$ M^{-1} to ~ 10^5 M^{-1}) as the temperature was increased from 8°C to 37°C. Least squares analysis of van't

Hoff plots (ln*K versus* 1/*T*) of the data revealed that in each solution the triplex is enthalpically stabilized by ~ 2 kcal per mole of base triplets.

This value for the enthalpic contribution to triplex stability, which reflects contributions from both T·AT and C+GC base triplets, is identical within experimental error to that measured previously using differential scanning calorimetry for the 15mer binding to 15 bp within a 21-bp duplex. We collaborated with Professor Kenneth J. Breslauer of Rutgers University to make the calorimetric measurement, and the complete thermodynamic characterization of the stability and the melting behavior of the triplex are described in Chapter Six.

Taken together, the equilibrium association constants, the thermal denaturation data, and the calorimetric data reported here provide a quantitative measure of the influence of third strand length, base triplet mismatches, pH, salt, and temperature on the stability and the melting behavior of a DNA triplex. Such information should prove useful in designing triplex-forming oligonucleotides and in defining solution conditions for the effective use of triple-helical structure formation as a tool for modulating biochemical events.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Introduction.

Figures.

Tables.

Chapter 1.

Figures.

Chapter 2.

Figures.

Chapter 5.

Figures.

Tables.

Chapter 6.

Figures.

Tables.

INTRODUCTION

Deoxyribonucleic acid is the molecular agent at the core of life processes, providing the primary code that is translated to yield all of the proteins and RNA necessary for cellular function. Because the genetic code is composed of the sequence of base pairs in the DNA molecule, the ability of native proteins to recognize specific control sequences of genomic DNA is essential for the regulation of cellular processes such as gene replication, transcription, and cell division. The central importance of the DNA molecule to these functions makes it an attractive target for molecular recognition studies. An analysis of the structure of DNA provides insights for understanding the foundations for sequence specific recognition.

The gross structure of Watson-Crick double-helical DNA is well known: two polydeoxyribonucleotide strands interact to form a right-handed helical structure, stabilized by vertical stacking interactions between pairs of hydrogen-bonded planar heterocycles at its core (Figure 1).[1] In this intertwined structure, the deoxyribose phosphate diester backbones are placed away from the helix long-axis, forming grooves along the cylindrical morphology of the double helix (Figure 2). The edges of the base pairs are located at the bottom of the grooves. These edges present distinct surfaces capable of interaction with potential ligands, with the properties of the surface depending on the identity of the base pair at the site. Because the base pairs are not truly symmetric with respect to C_2 rotation about an axis in the plane of each base pair, there are formally four different base pairs, which each present a unique combination of atoms at distinct location within the groove (Figure 3). Thus, endogenous DNA-binding

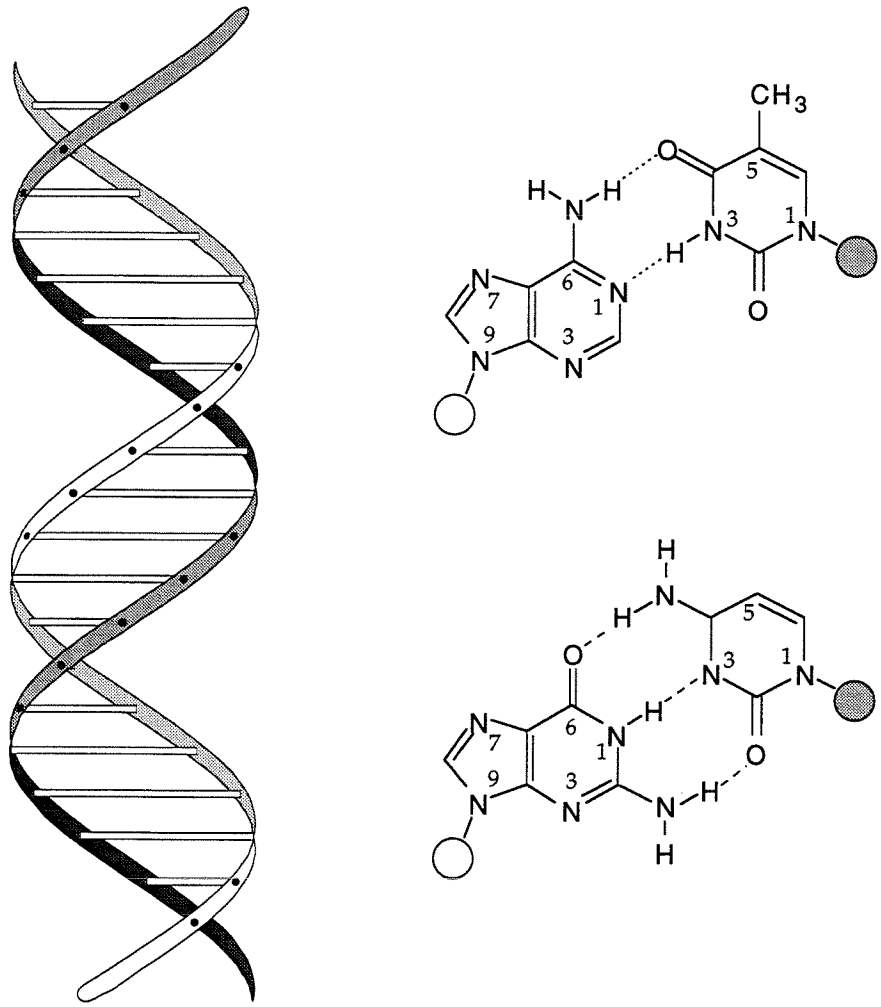


Figure 1. *Left,* Ribbon model of double-helical DNA. The deoxyribose phosphodiester backbone is represented by the ribbon strands, and base pairs are represented by the rods joining the ribbons. *Right,* The structures of the Watson-Crick base pairs A·T (top) and G·C (bottom). The circles replace the deoxyribose to which the bases are attached, with the shading of the circle corresponding to the shading of one of the ribbons. The numerals indicate the IUPAC numbering scheme.

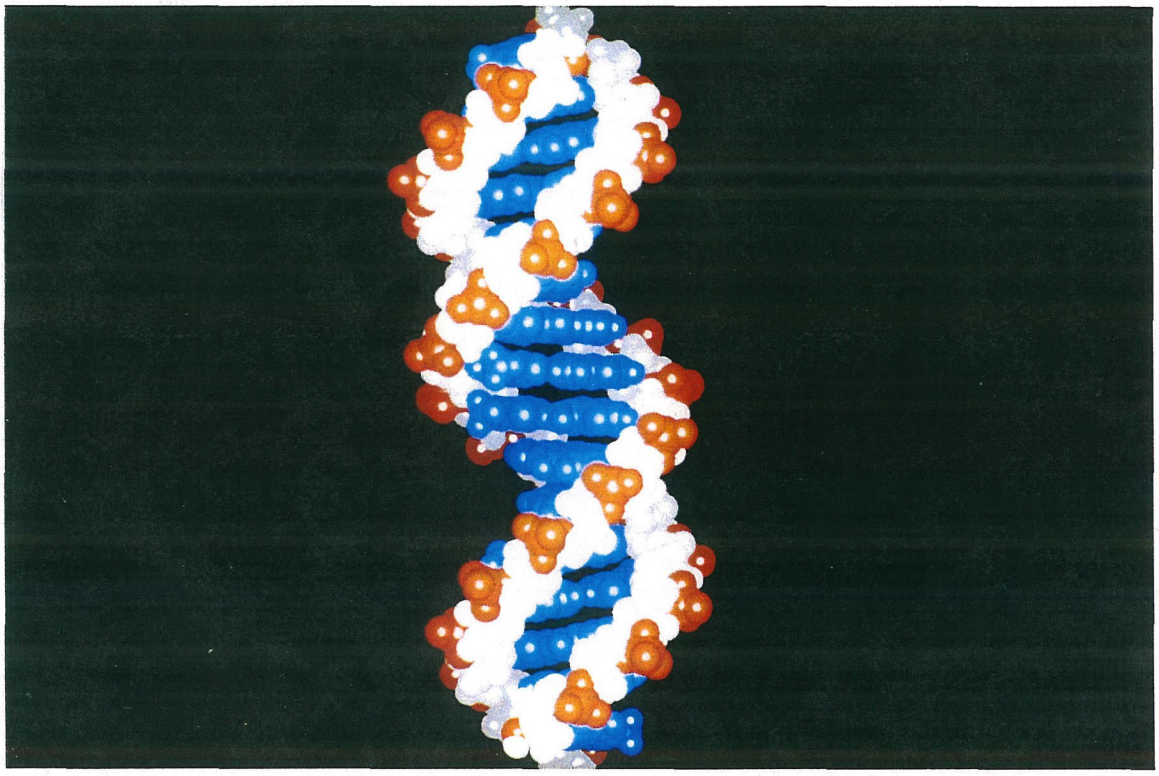


Figure 2. Computer-rendered CPK model of the structure of canonical B-form double-helical DNA. The atoms of the bases are colored blue, those of deoxyribose sugars are white, and phosphate atoms are red. Note that the edges of the base pairs are visible on the floor of the grooves.

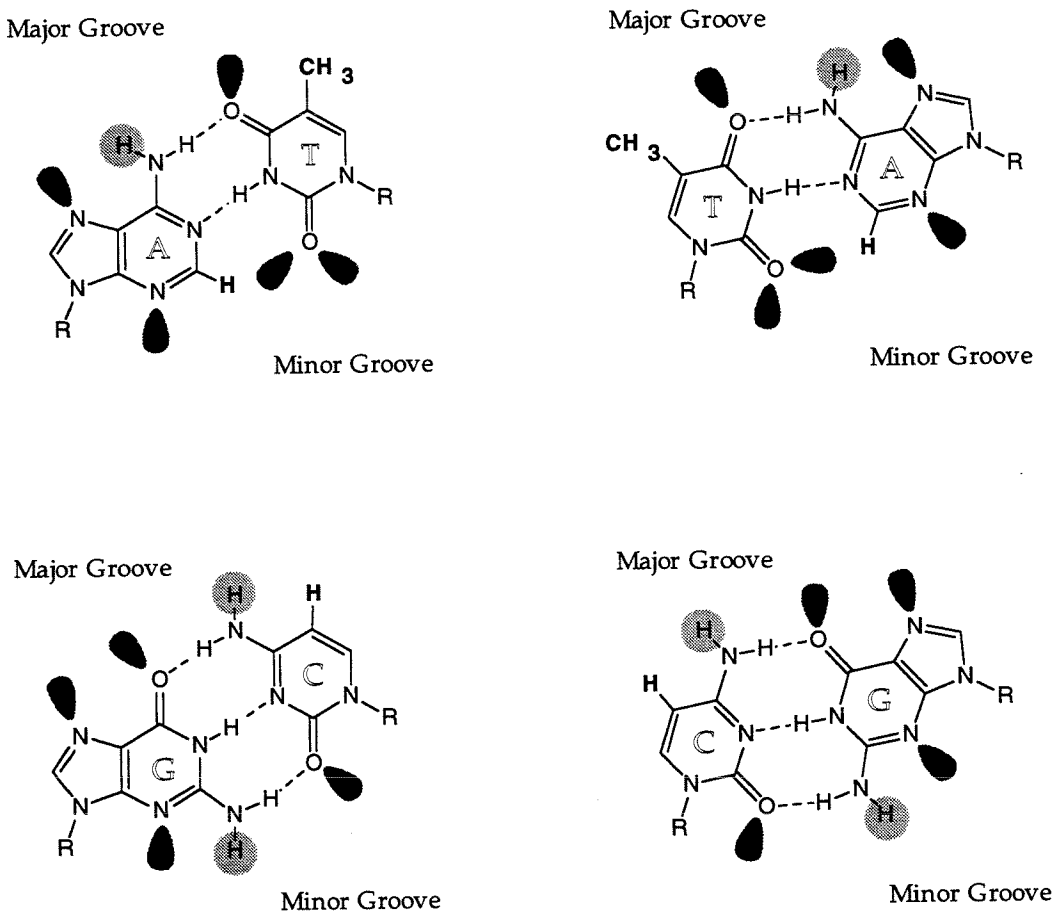


Figure 3. The structures of the four possible base pairs of DNA, (clockwise from top left) A·T, T·A, G·C, and C·G. Lone pairs are explicitly represented and amino hydrogens capable of forming hydrogen bonds are shaded. The locations of the major and minor grooves relative to the edges of each base pair are noted.

ligands, such as restriction endonucleases, can distinguish sequences based on the array of functional groups found on the floor of the grooves.

A Chemical Approach to DNA Recognition

Inspired early on by the remarkable ability of restriction enzymes to discriminate among DNA sequences, our laboratory undertook a chemical design-synthesis approach to the construction of DNA-binding ligands. The approach relies on the analysis of structure-function relationships, with a primary goal of identifying structural motifs that allow for the generalizable design of molecules capable of binding desired DNA sequences being a primary goal. By characterizing DNA complexes in solution, principles relating ligand covalent structure to the non-covalent interactions required for stabilizing the corresponding complex can be derived. A long-range functional goal is the binding of 15 – 18 contiguous base pairs (bp) of DNA, a size which is sufficient to statistically define a single site in human genomic DNA, which contains ~ 3×10^9 bp (Table 1).[2] The use of a ligand with such sequence specificity could, for example, complement the implementation of restriction endonucleases, which recognize fewer base pairs and thus more sites in genomic DNA, for the physical manipulation of chromosomes.

The first step in initiating a chemical investigation of a particular family of DNA complexes is the design of the first generation ligand. Although it might be possible to design the ligand *de novo*, an alternative approach is to begin the process by the analysis of previously characterized DNA complexes. For example, the complexes formed by small molecular weight natural products, such as the antibiotics netropsin and distamycin, have been used to design DNA-binding ligands.[3-19] A second area of investigation was initiated by the isolation of DNA-binding domains from native proteins.[20-28] Most recently, oligonucleotides capa-

ble of forming triple-helical complexes on duplex DNA have been studied. The characterization of the thermodynamic foundations of the stabilities of these triple-helical structures has been the goal of the studies described in this dissertation.

Table 1. The relation between duplex DNA binding site size and the number of unique sequences.

Binding Site Size[a]	Unique Sequences[b]	Binding Site Size	Unique Sequences
4	136	12	8,390,656
6	2080	15	536,870,912
8	32,896	18	68,719,607,808
10	524,800	21	2,199,023,300,000

[a] The binding site size is given in number of base pairs. [b] The number of unique sequences is determined from $4^n/2$ when n (the binding site size) is odd, and $(4^n + 4^{n/2})/2$ when n is even.[29,30]

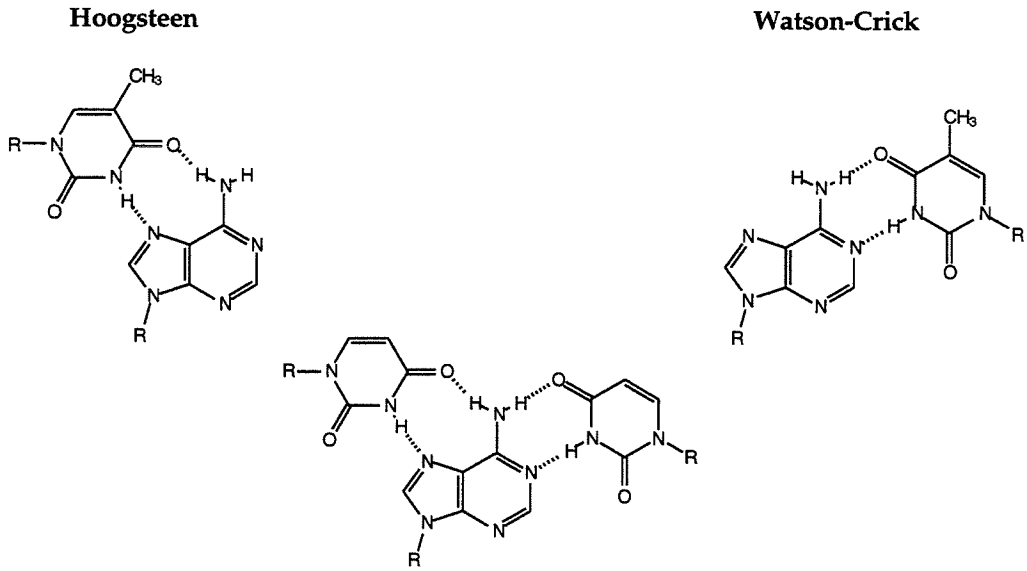
Triple-Helical Nucleic Acids

In 1957, only four years after the elucidation of the structure of double-helical DNA, Felsenfeld, Davies, and Rich first observed nucleic acids comprised of three strands.[31] In solutions containing millimolar concentrations of Mg^{2+} ions, the formation of polyribonucleotide complexes containing three RNA strands in the stoichiometry of 2 poly(rU) to 1 poly(rA) were observed. It was suggested that the stability of the complex was derived from hydrogen bonding between a second U and the A of a typical Watson-Crick A·U base pair, to form a U·A·U

base triplet (Figure 4). Hydrogen bonds between *N*3 and *O*4 of U and *N*7 and *N*6 of A, respectively, were observed by Hoogsteen in the crystals of 1-methylthymine and 9-methyladenine monomers (Figure 4). Because of the importance of this report, the interaction between a second pyrimidine and the purine of a Watson-Crick base pair is often referred to as Hoogsteen hydrogen bonding, and the third strand is referred to as the Hoogsteen strand.

Soon thereafter, poly(rC) was shown to form complexes with oligoriboguanidinylates [oligo(rG)] at acidic pH that contained two moles of cytosine bases per mole of guanine.[32,33] It was reasoned that these complexes were comprised of C^+·G·C base triplets which would be isomorphous with U·A·U triplets and stabilized by two Hoogsteen-type hydrogen bonds between *N*3-protonated C and the G of a G·C base pair (Figure 4). The requirement for protonation of one-half of the cytosines resulted in the observed pH sensitivity of the complexes. The association of poly[r(U-C)] with a poly[d(G-A)]·poly[d(C-T)] duplex to form triple-helical complexes at low pH suggested that triple helix formation could be generalized to any poly$(Y)_2$·poly(R) sequence, where Y is any pyrimidine and R is any purine.[34,35]

Subsequent to the observation of triple-helical nucleic acids in solution, x-ray diffraction studies of poly$(rU)_2$·poly(rA)[36,37] and poly$(dT)_2$·poly(dA)[37,38] fibers provided a model for the triple helix structural family. In this structure, one of the pyrimidine strands binds in the major groove of a duplex joined by Watson-Crick base pairs (Figure 5). The third strand is bound with a 5' $\rightarrow$ 3' orientation parallel to that of the poly(A) strand by Hoogsteen hydrogen bonds between thymine (or uracil) and adenine heterocycles. Although a high resolution x-ray crystal structure of a triple-helical DNA or RNA has not yet been



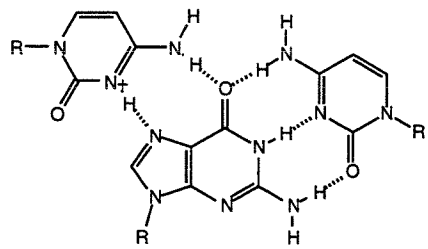


Figure 4. The structures of two possible pairing interactions between uracil and adenine are indicated above. The Watson-Crick base pair is that formed in double-helical nucleic acids, while the Hoogsteen base pair was formed between monomers in the crystallline state. Both types of pairing are present in the U·A·U base triplet. A C^+·G·C base triplet is shown below, in which the protonation at cytosine *N*3 allows the formation of two hydrogen bonds to the GC base pair. Note that the structures of the two base triplets are isomorphous.

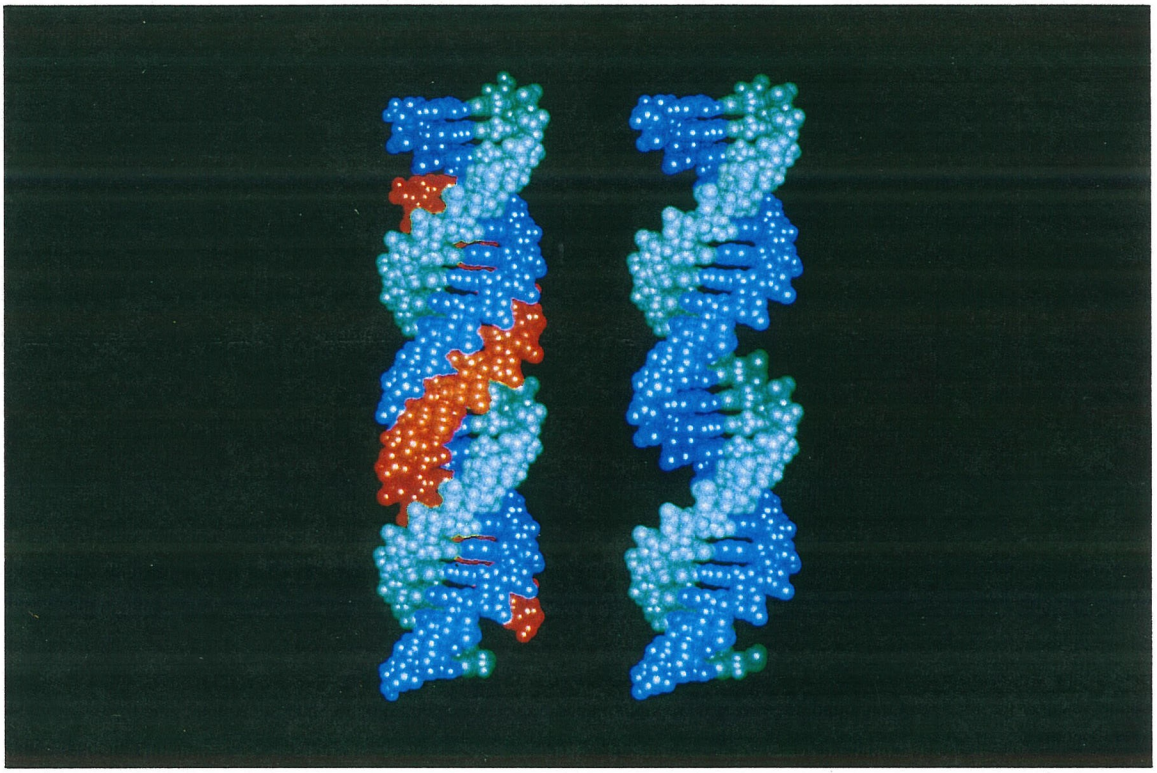


Figure 5. A computer-rendered CPK model showing the general structural features of the poly(dT)·poly(dA) double helix (left) and thepoly(dT)$_2$·poly(dA) triple helix (right). The poly(dT) strand bound in the major groove of the duplex is colored red.

reported, recent NMR,[39-48] circular dichroism,[49-54] infrared,[55,56] and Raman[57] spectroscopic studies have elaborated the general features of this structural model, providing a more detailed view of the rich conformational diversity of triple-helical complexes.

Oligonucleotide-Directed Triple Helix Formation

The length of the observed polynucleotide complexes suggested that single-stranded nucleic acid molecules could, in principle, be designed to bind large sequences of double-helical DNA by triple helix formation. The more recent development of chemical methods for the high-yield automated synthesis of relatively short nucleic acid strands allowed the preparation of oligonucleotides of any desired sequence.[58-60] Interestingly then, the application of triplex formation toward the goal of binding 15 contiguous base pairs of duplex DNA raised the question of whether a short triple-helical complex would be stable. This represented a diametrically different problem from that presented by the use of small peptides or DNA-binding protein domains that typically recognize 4 to 8 bp.

In 1987, Moser and Dervan provided direct evidence that a chemically synthesized 15mer oligonucleotide binds in the major groove parallel to a single homopurine sequence, 15 nucleotides (nt) in length, within a 4.5-kbp plasmid DNA fragment (Figure 6).[61] Subsequent reports from other laboratories confirmed the generality of this result.[62-65] Moser and Dervan also found that changing the sequence of the oligonucleotide at a single position and creating a base triplet "mismatch" (Figure 7) dramatically reduced the apparent stability of the triple-helical complex. The triple helix involving the 15mer oligonucleotide therefore contained 30 sequence-specific hydrogen bonds which were sufficient

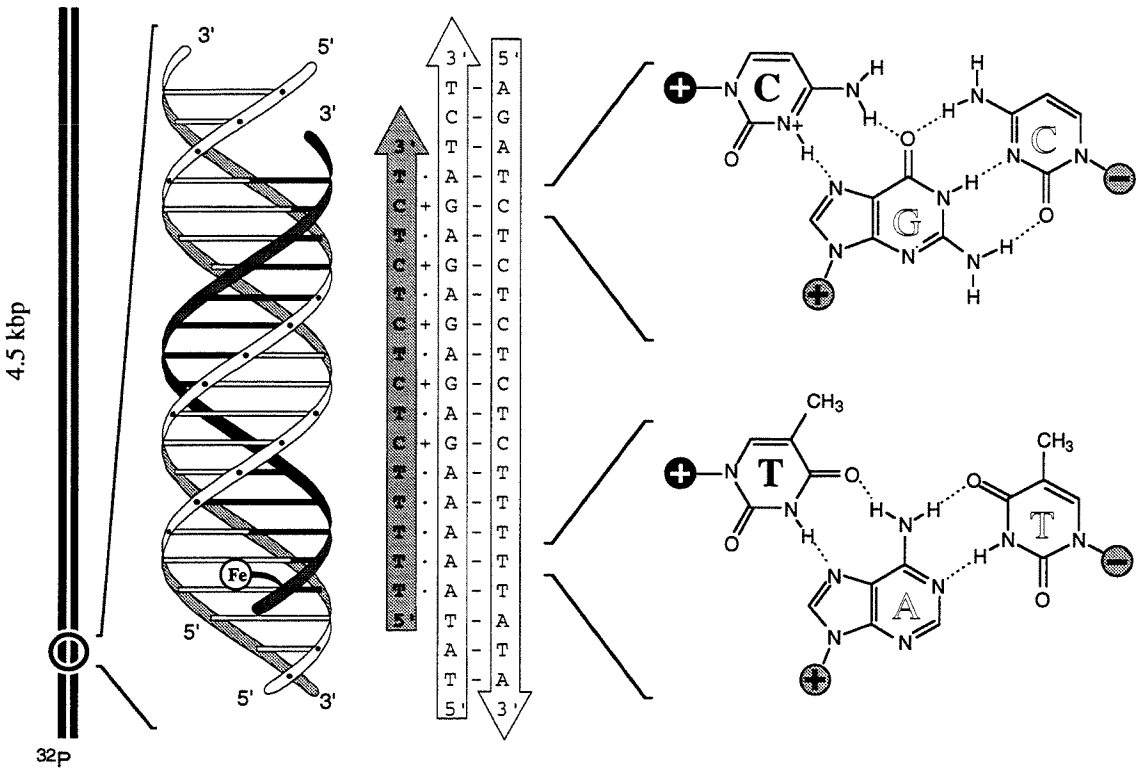


Figure 6. Ribbon model of local triple-helical complex formed by binding of a 15mer oligonucleotide in major groove of duplex DNA parallel to the purine strand.

for complex formation at ~ 1 μM oligonucleotide concentration at pH 7.0 and 22°C.

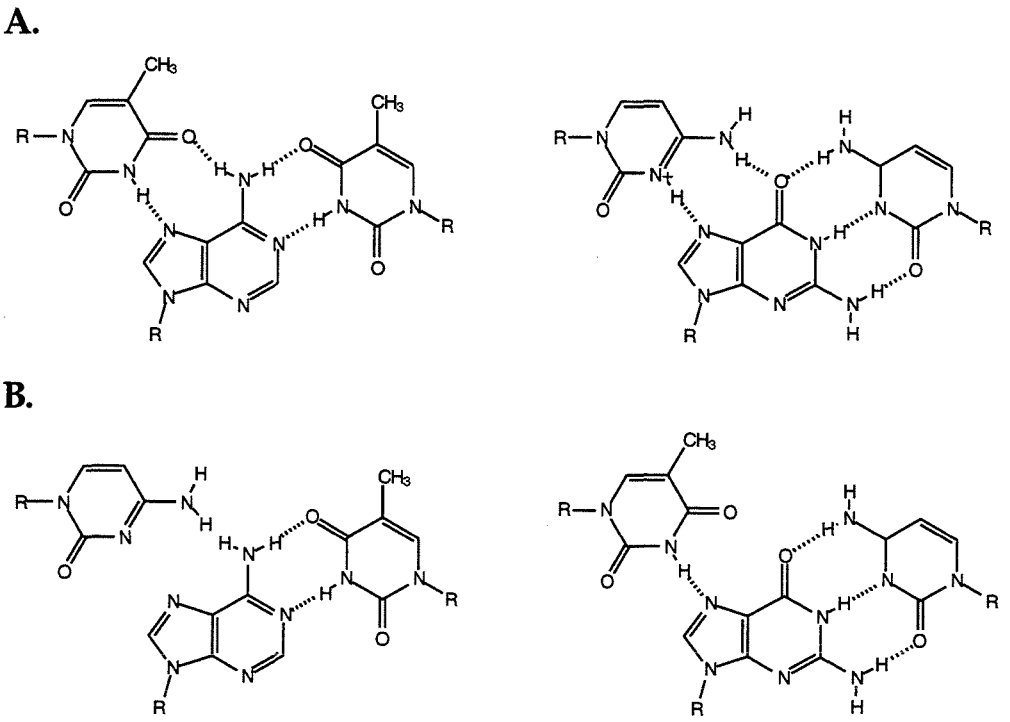


Figure 7. (A) The structures of T·AT and C+GC base triplets. (B) C·AT (left) and T·GC (right) base triplet mismatches.

Oligonucleotide-directed formation of local triple-helical structures provides a motif for the design of sequence-specific DNA-binding ligands. In principle, a triplex-forming oligonucleotide with any desired sequence preference can be designed because the principles of specificity are understood. The base triplet code requires a T for recognizing an A·T base pair, while a C is required for binding a G·C base pair. In addition, there are no complex tertiary folding requirements for the oligonucleotide. To a first approximation, the sequence of the

oligonucleotide determines its functional specificity. These characteristics are distinct from those of protein domains, for which the difficulty in predicting folding patterns precludes the *de novo* design of novel proteins with designed sequence specificities.

Triple Helix Mediated Function

The length of the DNA binding sites, and the remarkable apparent affinity and specificity of the local triple-helical structures, suggested that the formation of such complexes might ultimately allow oligonucleotides to compete with native DNA-binding proteins at single sites within genomic DNA. In fact, it has now been demonstrated that, when a 21mer is bound at a single site in plasmid DNA, the activities of restriction endonucleases and methylases at sequences occluded by the triplex could be completely inhibited (Figure 8).[66] The ability of oligonucleotides to inhibit protein binding[65,67-69] at DNA sequences overlapping the protein binding site has been employed to inhibit DNA transcription *in vitro*.[67,70,71]

The apparent sequence specificity of oligonucleotide-directed triple helix formation has been tested in the attempt to cleave total genomic DNA at a single location. The size of the genome studied was incrementally increased, so that bacteriophage,[72] yeast,[73] and mammalian DNA[74] have been cleaved at single designated sites. The latter was a particularly dramatic application of oligonucleotide inhibition of protein-DNA complex formation, involving the use of a 16mer oligonucleotide to mediate the cleavage of human chromosome 4. The absence of cleavage of the other chromosomes indicated that only the intended 18-bp duplex sequence was effectively occupied, providing clear

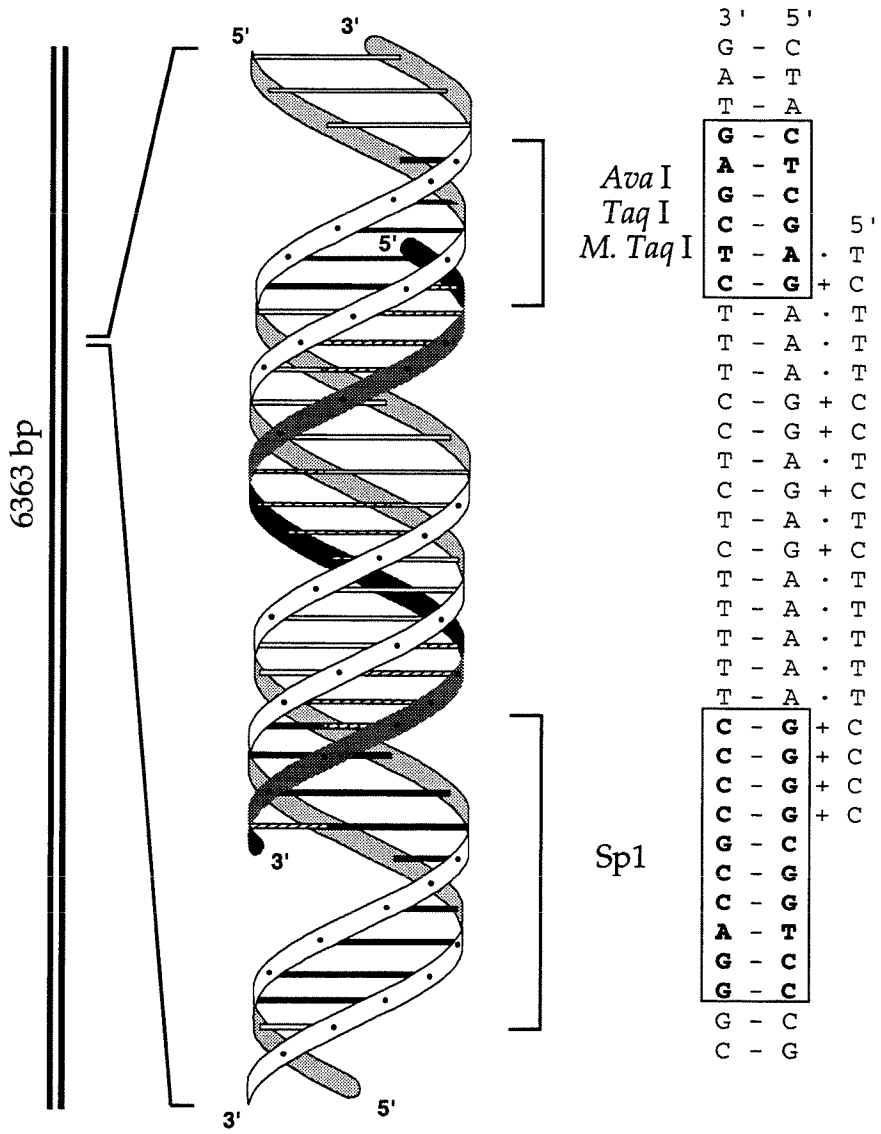


Figure 8. Schematic representation of a triple-helical complex capable of inhibiting restricition endonuclease activity. The structure forms at a single site within a 6.4-kbp plasmid. The active sites for *Ava* I and *Taq* I restriction endonucleases and *M. Taq* I methylase are indicated by brackets next to the ribbon model and by a box around the sequence. The base pairs of the protein binding sites are indicated by shaded rods on the ribbon model. Note that the third strand (dark ribbon) overlaps the protein binding sites.

evidence that triple-helical complexes can form at single, specific sequences within genomic DNA.

Other Research Directions

Oligonucleotide-directed triple helix formation is a versatile approach to sequence-specific DNA recognition. The ability to recognize a broad range of sequences by the formation of stable local triple-helical complexes which are highly sensitive to base triplet mismatches make this a powerful technique for binding single sites within large segments of double-helical DNA. However, to this point, only the recognition of two of the four base pairs has been addressed. The identification of other natural base triplets (Figure 9A),[75-79] the design of non-natural nucleosides for forming novel base triplets (Figure 9B),[48,80-82] and alternate strand triple helix formation[83] have each allowed progress toward the generalization of the base triplet code for the recognition of DNA sequences containing all four base pairs. However, because these strategies have been only partially successful, the synthesis of oligonucleotides capable of triplex formation at sites containing all four base pairs remains an area of intense effort.

A second area of investigation involves alternative triple-helical structures. Triple helices involving two pyrimidine strands and one purine strand constitute one of two general classes of triplex structures. The second family of triple-helical structures differs from that described above in the sequence composition of the third strand, the relative orientations and positions of the backbones of the three strands, and the base triplet interactions.[84-87] This triple-helix structure consists of two purine-rich strands and a single pyrimidine strand. One purine-rich strand binds in the major groove of DNA anti-parallel to the Watson-Crick purine strand.[86] Sequence specificity is derived from hydrogen bonds between

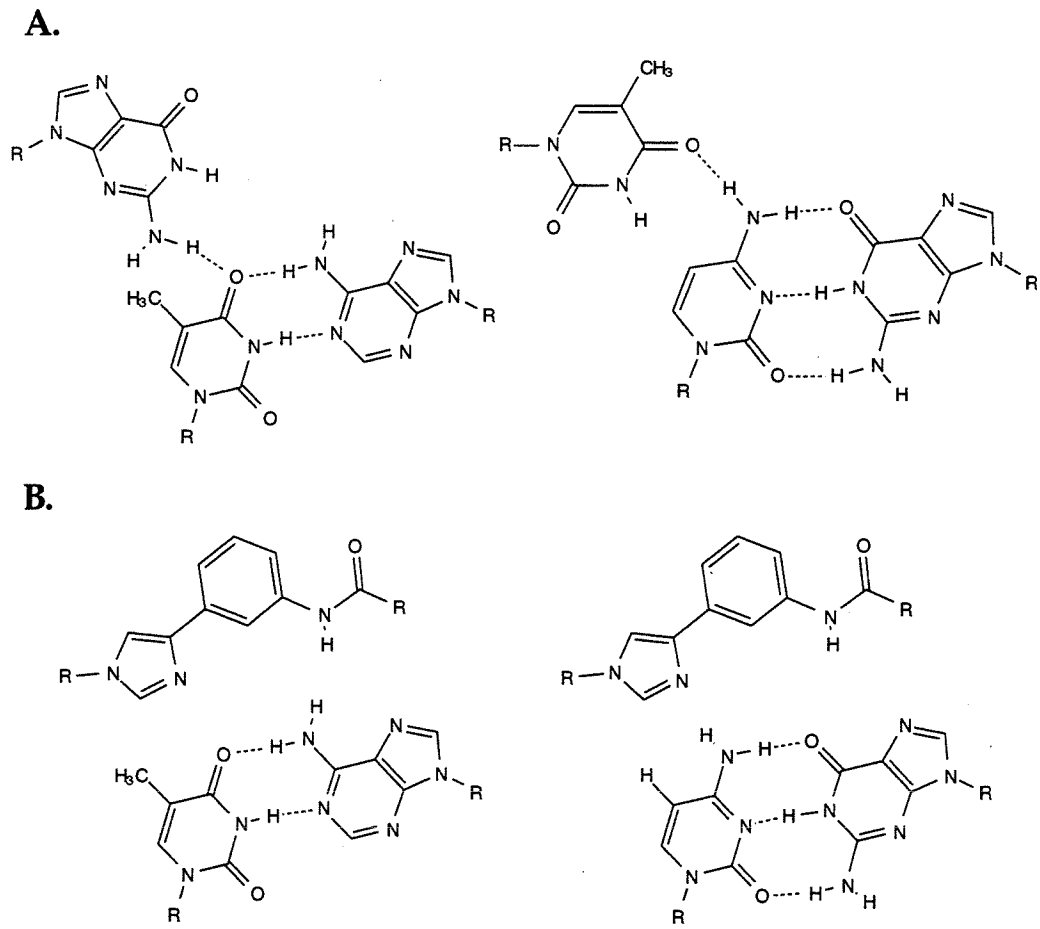


Figure 9. Two-dimensional schemes for interactions in the natural base triplets G·TA and T·CG (A), and the non-natural D_3·TA and D_3·CG base triplets (B).

the third strand and the purine-rich strand of the Watson-Crick duplex, so that G recognizes G·C base pairs (G·G·C base triplets; Figure 10A) and A or T recognizes A·T base pairs (A·A·T and T·A·T base triplets). The T·A·T base triplet that stabilizes (purine)$_2$·pyrimidine triplex structures is different from that based on Hoogsteen hydrogen bonding because the thymine heterocycle has been flipped over in the anti-parallel purine-rich third strand. As is the case for pyrimidine third strands, designed triplet interactions can be used to extend the recognition code for purine-rich third strands (Figure 10B).[88] Furthermore, the two structural classes can be combined in alternate-strand triplexes, in which a homopurine tract on one duplex strand is recognized using one structural motif, and an adjacent homopurine tract on the opposite strand is bound using the other structural motif.[89-91]

Description of the Thesis Work

The year before this work was initiated, Professor Dervan had characterized the chemical approach to sequence-specific DNA recognition in the following way.[92]

> "This entire exercise concerning the molecular recognition of DNA is directed at understanding the rules of non-covalent bonding. If one examines the productivity of synthetic and physical organic chemistry in defining a mechanistic framework for the rules of covalent bonding during the past thirty years, a prerequisite was a knowledge base of important reaction types (substitution reactions, Diels Alder, aldol, etc.). Importantly, the introduction of techniques borrowed from physical chemistry (electronic, infrared and NMR spectroscopy, VPC/HPLC, lasers, matrix isolation) to eluci-

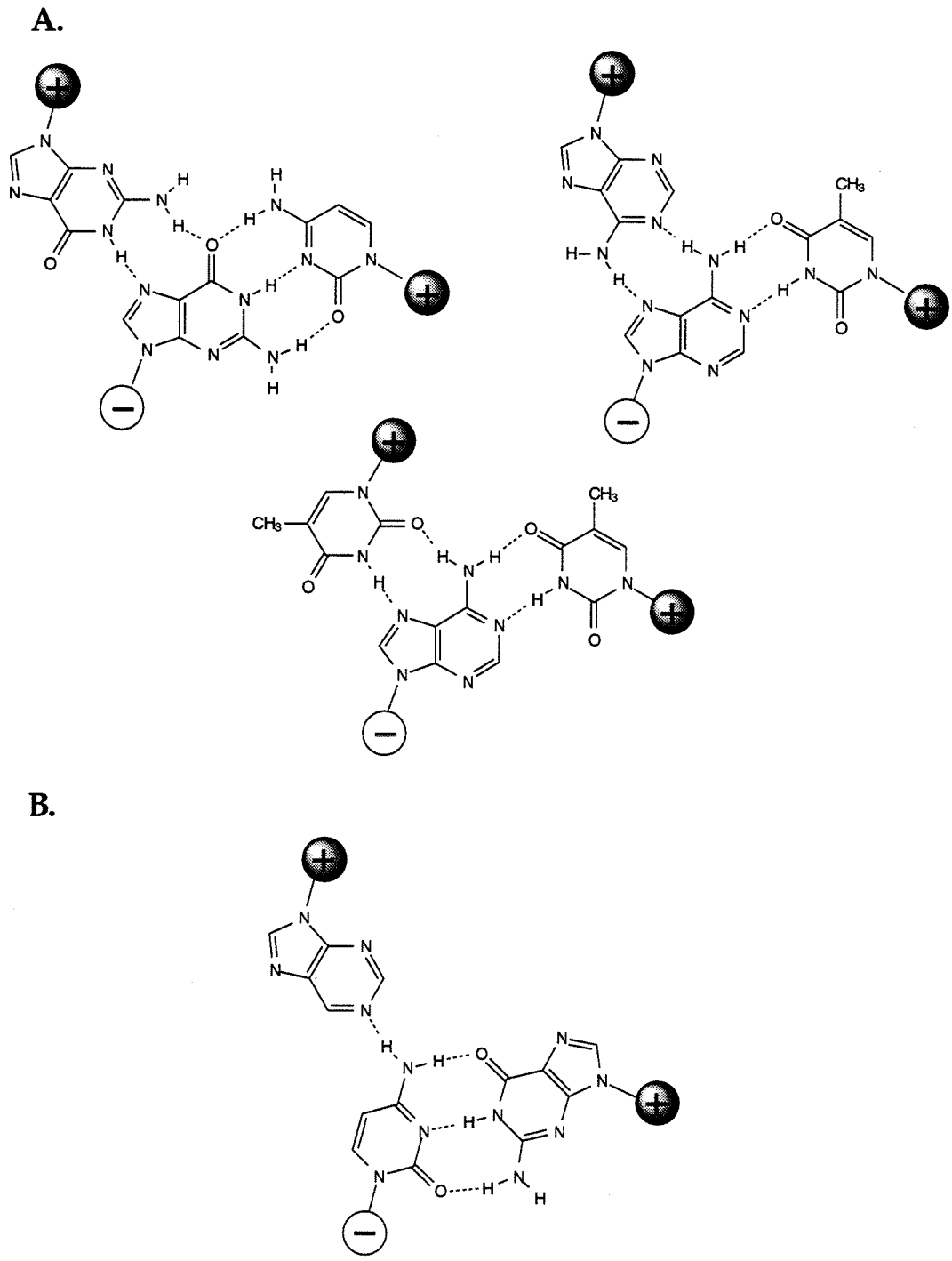


Figure 10. (A) Hydrogen bonding within base triplets formed by binding of purine-rich oligonucleotides anti-parallel to a purine-rich strand of the duplex. (B) The interaction of nebularine with a CG base pair.

> date the bond-making and bond-breaking processes responsible for these reaction types were pivotal for the productivity and advancement of this area. The work described here is at a stage of development not unlike that of covalency decades ago. We are documenting first what will work and what will not (i.e., the non-covalent *complex* as synthetic target) and developing (or inventing) the techniques necessary to analyze the data (in this case, exploiting sequencing gels borrowed from molecular biology). Our long range goal is to characterize the thermodynamics and kinetics for all these binding events and to reduce this exercise in the non-covalent arena to a predictive science."

Toward this long-range goal, we undertook the characterization of the stabilities of the local triple-helical structures formed by binding of pyrimidine oligonucleotides.

In their original report of oligonucleotide-directed triple helix formation, Moser and Dervan described the apparent dependence of the stability of the local triple-helical complex on the third-strand length and the solution conditions, including pH, salt concentrations, and temperature. Throughout the work described here, their report has guided our investigation of the effects of incremental changes in the triple-helical complex or in the solution conditions. The determination of thermodynamic properties represents the first step in enhancing our understanding of the local triple helix structure and our ability to rationally implement these complexes for novel function.

Chapter One introduces the method of quantitative affinity cleavage titration for measuring single-site binding constants on plasmid DNA fragments. The method combines the utility of established footprint titration techniques with

the power of affinity cleaving for the independent analysis of binding to individual sites within larger duplex DNA. Chapters Two through Five describe results using this method to measure equilibrium constants for oligonucleotide-directed triple helix formation and to characterize their linkage to properties of the third strand and to the solution conditions. The influence of oligonucleotide length and base triplet mismatches are discussed in Chapter Two. In Chapter Three, the effect of pH on the association constant is described. In Chapter Four, the linkage between triplex free energy and cation valence and concentration is explored. In Chapter Five, the influence of temperature on the equilibrium association constant is examined, and the results are analyzed using the van't Hoff equation to allow estimation of the enthalpic contribution to triplex stability. Finally, Chapter Six contains the results of a calorimetric measurement of the enthalpy change associated with triplex formation that was conducted in collaboration with Professor Kenneth J. Breslauer of Rutgers University. The enthalpy values measured by both van't Hoff analysis and calorimetry are compared with one another and with measured values reported in the literature.

References

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

(2) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(3) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861-6863.

(4) Vandyke, M. W.; Hertzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5470-5474.

(5) Schultz, P. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6834-6837.

(6) Schultz, P. G.; Dervan, P. B. *J. Am. Chem. Soc.* **1983**, *105*, 7748-7750.

(7) Schultz, P. G.; Dervan, P. B. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1133-1147.

(8) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457-465.

(9) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528-5529.

(10) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565-2569.

(11) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1986**, *108*, 5008-5009.

(12) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 6840-6842.

(13) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564-7566.

(14) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7586-7590.

(15) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892-9899.

(16) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572-2576.

(17) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochem.* **1993**, *32*, 11385-11389.

(18) Geierstanger, B. H.; Jacobsen, J. P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochem.* **1994**, *33*, 3055-3062.

(19) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 3663-3664.

(20) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129-1132.

(21) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572-7574.

(22) Graham, K. S.; Dervan, P. B. *J. Biol. Chem.* **1990**, *265*, 16534-16540.

(23) Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 4604-4606.

(24) Mack, D. P.; Sluka, J. P.; Shin, J. A.; Griffin, J. H.; Simon, M. I.; Dervan, P. B. *Biochem.* **1990**, *29*, 6561-6567.

(25) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847-850.

(26) Sluka, J. P.; Horvath, S. J.; Glasgow, A. C.; Simon, M. I.; Dervan, P. B. *Biochem.* **1990**, *29*, 6551-6561.

(27) Shin, J. A.; Ebright, R. H.; Dervan, P. B. *Nucleic Acids Res.* **1991**, *19*, 5233-5236.

(28) Mack, D. P.; Dervan, P. B. *Biochem.* **1992**, *31*, 9399-9405.

(29) Strobel, S. A. Ph.D. Thesis, California Institute of Technology, 1992.

(30) Dervan, P. B. *Biochem.* **1987**, *26*, 4171-4171.

(31) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

(32) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256-1260.

(33) Howard, F. B.; Frazier, J.; Lipsett, M. N.; Miles, H. T. *Biochem. Biophys. Res. Commun.* **1964**, *17*, 93-102.

(34) Miller, J. H.; Sobell, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1201-1205.

(35) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.

(36) Arnott, S.; Bond, P. J. *Nature New Biol.* **1973**, *244*, 99-101.

(37) Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucleic Acids Res.* **1976**, *3*, 2459-2470.

(38) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509-521.

(39) Gorenstein, D. G.; Luxon, B. A.; Goldfield, E. M.; Lai, K.; Vegeais, D. *Biochem.* **1982**, *21*, 580-9.

(40) de los Santos, C.; Rosen, M.; Patel, D. *Biochem.* **1989**, *28*, 7282-7289.

(41) Rajagapol, R.; Feigon, J. *Nature* **1989**, *339*, 637-640.

(42) Rajagapol, P.; Feigon, J. *Biochem.* **1989**, *28*, 7859-7870.

(43) Pilch, D. S.; Levenson, C.; Shafer, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1942-6.

(44) Mooren, M. M. W.; Pulleyblank, D. E.; Wijmenga, S. S.; Blommers, M. J. J.; Hilbers, C. W. *Nucleic Acids Res* **1990**, *18*, 6523-9.

(45) Umemoto, K.; Sarma, M. H.; Gupta, G.; Luo, J.; Sarma, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 4539-45.

(46) Live, D. H.; Radhakrishnan, I.; Misra, V.; Patel, D. J. *J. Am. Chem. Soc.* **1991**, *113*, 4687-8.

(47) Radhakrishnan, I.; Gao, X.; De los Santos, C.; Live, D.; Patel, D. J. *Biochem.* **1991**, *30*, 9022-30.

(48) Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908-7909.

(49) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucleic Acids Res.* **1979**, *6*, 3073-3091.

(50) Steely, H. T., Jr.; Gray, D. M.; Ratliff, R. L. *Nucleic Acids Res.* **1986**, *14*, 10071-90.

(51) Antao, V. P.; Gray, D. M.; Ratliff, R. L. *Nucleic Acids Res.* **1988**, *16*, 719-738.

(52) Pilch, D. S.; Brousseau, R.; Shafer, R. H. *Nucleic Acids Res.* **1990**, *18*, 5743-50.

(53) Manzini, G.; Xodo, X. E.; Gasparotto, D.; van der Marel, G. A.; van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833-843.

(54) Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436-40.

(55) Liquier, J.; Coffinier, P.; Firon, M.; Taillandier, E. *J. Biomol. Struct. Dyn.* **1991**, *9*, 437-45.

(56) Taillandier, E.; Firon, M.; Liquier, J. *Spec. Publ. R. Soc. Chem.* **1991**, *94*, 355-6.

(57) O'Connor, T.; Bina, M. *J. Biomol. Struct. Dyn.* **1984**, *2*, 615-625.

(58) Caruthers, M. H.; Barone, A. D.; Beaucage, S. L.; Dodds, D. R.; Fisher, E. F.; Mcbride, L. J.; Matteucci, M.; Stabinsky, Z.; Tang, J. Y. *Meth. Enzym.* **1987**, *154*, 287-313.

(59) Caruthers, M. H. *Acc. Chem. Res.* **1991**, *24*, 278-284.

(60) Sinha, N. D.; Biernat, J.; Köster, M. *Tetrahedron Lett.* **1983**, *24*, 5843-5846.

(61) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.

(62) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, *15*, 7749-7760.

(63) Praseuth, D.; Perrouault, L.; Trung, L. D.; Chassignol, M.; Nguyen, T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349-53.

(64) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D.; Cantor, C. R. *Nucleic Acids Res.* **1988**, *16*, 2165-2178.

(65) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucleic Acids Res.* **1990**, *18*, 157-61.

(66) Maher, L. J.; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

(67) Maher, L. J.; Dervan, P. B.; Wold, B. *Biochem.* **1992**, *31*, 70-81.

(68) Francois, J. C.; Saison-Behmoaras, T.; Nguyen, T. T.; Helene, C. *Biochem.* **1989**, *28*, 9617-19.

(69) Collier, D. A.; Nguyen, T. T.; Helene, C. *J. Am. Chem. Soc.* **1991**, *113*, 1457-8.

(70) Grigoriev, M.; Praseuth, D.; Robin, P.; Hemar, A.; Saison-Behmoaras, T.; Dautry-Varsat, A.; Nguyen, T. T.; Helene, C.; Harel-Bellan, A. *J. Biol. Chem.* **1992**, *267*, 3389-95.

(71) Duval-Valentin, G.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 504-8.

(72) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.

(73) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75.

(74) Strobel, S. A.; Doucettestamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-1642.

(75) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-971.

(76) Giovannangeli, C.; Rougee, M.; Garestier, T.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8631-5.

(77) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3840-4.

(78) Fossella, J. A.; Kim, Y. J.; Richards, E. G.; Fresco, J. R. *Nucleic Acids Res.* **1993**, *21*, 4511-15.

(79) Miller, P. S.; Cushman, C. D. *Biochem.* **1993**, *32*, 2999-3004.

(80) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976-7982.

(81) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochem.* **1992**, *31*, 2829-2834.

(82) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470-1478.

(83) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435-2437.

(84) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-9.

(85) Kohwi, Y.; Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781-5.

(86) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(87) Durland, R. H.; Kessler, D. J.; Gunnell, S.; Duvic, M.; Hogan, M. E.; Pettitt, B. M. *Biochem.* **1991**, *30*, 9246-55.

(88) Stilz, H. U.; Dervan, P. B. *Biochem.* **1993**, *32*, 2177-2185.

(89) Beal, P. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 4976-4982.

(90) Jayasena, S. D.; Johnston, B. H. *Biochem.* **1992**, *31*, 320-7.

(91) Jayasena, S. D.; Johnston, B. H. *Biochem.* **1993**, *32*, 2800-7.

(92) Dervan, P. B. In *Stereochemistry of Organic and Bioorganic Transformations*; W. W. Bartmann, Ed.; VCH: New York, 1987; pp 231-232.

CHAPTER ONE

Quantitative Affinity Cleavage Titration: A Technique for Measuring Equilibrium Constants for Single-Site DNA Binding

Introduction

A primary endeavor in our laboratory is the attempt to gain an understanding of the forces controlling the affinities and specificities of ligands for double-helical DNA. Our approach toward this goal has been to characterize the energetics of a variety of sequence-specific DNA recognition processes.[1-12] Although we have often focused on the importance of the experimental results rather than on the methodology used to gather them, the work described in this dissertation has relied almost exclusively on the development of the quantitative affinity cleavage titration technique. Throughout the thesis work, the details of this technique and our understanding of its limitations have been constantly refined. Although it precludes a chronologically correct description, I have chosen to describe the affinity cleavage titration methodology in this first chapter because of its central role in the chapters that follow. The intellectual orgins of the method and its theoretical basis will be discussed in detail in Part I. In Part II, a detailed protocol for performing a quantitative titration is presented. In Part III, potential sources of error and their propagated effects on the measured association constants will be examined. Where possible, experimental data will be in-

troduced to supplement the statistical treatment. Finally, in Part IV, our experience with the protocol and the error analysis will be used for a discussion of the limitations of the technique.

Part I. Theoretical Basis for Quantitative Affinity Cleavage Titrations

Equilibrium Constants for Single-Site DNA Binding

A full understanding of double-helical DNA recognition will require a complete thermodynamic profile of the relevant DNA-ligand complex. Such knowledge would make possible the elucidation of the individual contributions of different residues or functional groups to the overall stability of the complex. Furthermore, an understanding of the linkage between complex stability and solution conditions will allow the controlled implementation of DNA binding under solution conditions dictated by the *in vitro* or *in vivo* environment. Our studies have focused on the triple-helical complexes formed at single sites within larger duplex DNAs under conditions which approximate those found within cells. Such "local" complexes are models of those that might form, for example, on chromosomal DNA within cell nuclei, where they have the potential to modulate biological events at the level of the gene.[13,14] In addition to playing a potential role in human medicine, oligonucleotide-directed triple helix formation provides a method for the manipulation of genomic DNA *in vitro*,[15] and could also function as a diagnostic tool.[16] Because we are ultimately interested in controlling the relative amount of a local triple-helical complex formed at equilibrium, we chose to measure oligonucleotide binding free energies at single DNA sites.

A number of techniques have been developed for the estimation of enthalpies and entropies for various denaturation processes (see Chapter Six). For example, van't Hoff analysis of a melting curve allows one to determine the enthalpy and entropy of *dissociation*, from which the free energy can be calculated.

This method requires several important, and sometimes incorrect, assumptions to be made in order to allow thermodynamic parameters to be calculated. Moreover, the resulting parameters must be extrapolated from the melting temperature to 25°C with the assumption that the parameters are temperature-independent. This assumption is rarely rigorously correct, although the error is often small enough to allow comparison between different systems. A typical analysis results in a relative uncertainty in the reported enthalpy and entropy terms of 5 – 10% (*e.g.*, 75 ± 5 kcal/mol), without accounting for the uncertainty introduced by extrapolation of the parameters to 25°C. Gauss's Law describes the propagation of errors (eq 1), where σ represents an error or uncertainty, and x and y represent the independent and dependent variables, respectively:

$$\sigma_y^2 = \sum_i \sigma_{x_i}^2 \left(\frac{\partial y}{\partial x_i} \right)^2 \qquad (1)$$

From eq 1, the uncertainty in a free energy term, calculated from a van't Hoff enthalpy and entropy using $\Delta G = \Delta H - T\Delta S$, can be derived:

$$\sigma_{\Delta G} = \sqrt{\sigma_{\Delta H}^2 + (T\sigma_{\Delta S})^2} \qquad (2)$$

For a typical nucleic acid order-disorder transition for which the enthalpy change was estimated to be 75 ± 5.0 kcal/mol and the entropy change to be 220 ± 15 cal/mol·K, the free energy change would be 10 kcal/mol with a propagated uncertainty of 6.6 kcal/mol. Remarkably, the propagated error in the derived equilibrium *association* constant, $K_{assoc} = \exp(+\Delta G_{dis}/RT) = 2.1 \times 10^7\ M^{-1}$, would

$$\sigma_K = \frac{K}{RT} \cdot \sqrt{\frac{\sigma_{\Delta H}^2}{RT} + \frac{\sigma_{\Delta S}^2}{R}} \qquad (3)$$

be ~ 2.4 x 10^8 M^{-1} (eq 3). Clearly, a method for generating more reliable association constants is desirable.

Methods capable of measuring individual DNA site binding constants are rare. Most binding assays, such as equilibrium dialysis, filter binding, gel mobility shifting, and thermal disruption, measure macroscopic (average) properties, so that the determination of single site values requires the use of oligomeric duplexes containing only one site. However, the correspondence of oligonucleotide properties to those of larger DNA is not exact.[17-24] Individual DNA sequences within plasmid fragments (100 – 1000 bp in length) bound by proteins can be identified by "footprinting" with Deoxyribonuclease I (DNase I),[25] a natural enzyme that promotes hydrolysis of the phosphodiester bonds in double-stranded DNA. When a DNA ligand sterically blocks the site to which it is bound, that site's phosphodiester backbone is protected from cleavage by external footprinting agents, while DNA backbone positions removed from the site of complexation are cleaved normally. Because autoradiograms of sequencing gels used to separate radiolabeled DNA products show a dimunition in the signal intensities resulting from cleavage within the ligand binding site, a "footprint" is left by the ligand on the DNA surface (Figure 1.1). A number of footprinting agents, including DNase I, MPE·Fe, and dimethylsulfate (DMS),[26] have been used to study local triplexes formed by oligonucleotides at single DNA sites. The importance of adapting footprinting for characterizing the thermodynamics of ligand-DNA interactions is that it provides the ability to *independently* monitor binding to

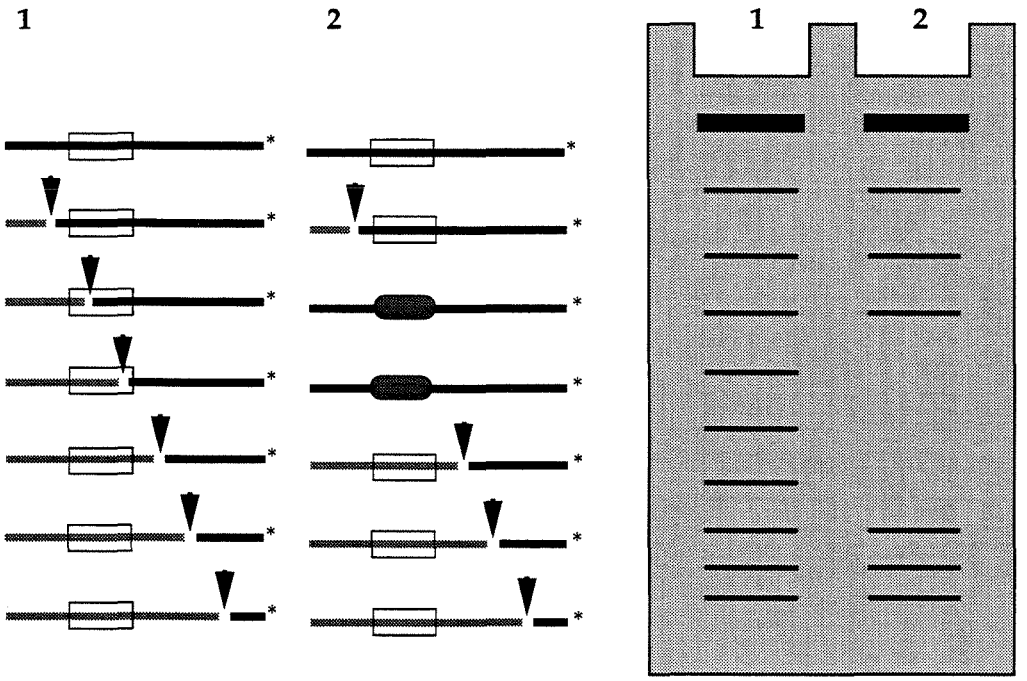


Figure 1.1. Schematic depiction of sampling a DNA-ligand complex by footprinting. (1) A labeled duplex DNA (*) is exposed to a footprinting reagent that nicks the DNA (arrowheads). (2) When the duplex is exposed in the presence of a ligand, complexation prevents cleavage within the binding site (rectangle), producing a footprint pattern upon autoradiography of an electrophoretic sample.

specific DNA sequences. Several methods for measuring single site binding constants using DNase I footprinting have been developed.[27-44]

DNase I Footprint Titrations

Hill has elegantly demonstrated the derivation of equilibrium binding expressions from the grand partition function of a system.[45] Based on such an analysis, the equilibrium association constant for a general ligand (**L**) binding to a specific DNA sequence (**D**) is defined by

$$\mathbf{L} + \mathbf{D} \rightleftharpoons \mathbf{DL} \tag{4a}$$

$$K_a = \frac{DL_{eq}}{D_{eq} \cdot L_{eq}} \tag{4b}$$

where DL_{eq}, D_{eq}, and L_{eq} represent the equilibrium concentrations of the specific DNA-ligand complex, the unbound DNA, and the unbound ligand, respectively. The association constant can be redefined using the fraction of DNA sites bound, $\theta_{eq} = DL_{eq}/(DL_{eq} + D_{eq}) = DL_{eq}/D_{tot}$, as follows:

$$K_a = \frac{\theta_{eq}}{1 - \theta_{eq}} \cdot \frac{1}{L_{eq}} \tag{5}$$

Eq 5 leads to two conclusions that are important for using footprinting for quantitative titrations. First, because K_a is independent of the DNA concentration, the amount of DNA does not need to be accurately measured (although, as is discussed below, an upper limit on the DNA concentration must be known). This allows the use of radiolabeled DNA of indeterminate specific activity. Second,

determination of the association constant only requires measurement of the fractional occupancy of DNA sites and the concentration of unbound ligand.

In a titration experiment, the ligand concentration is sytematically varied, and the amount of cleavage at a particular phosphodiester bond is a measure of the fraction of DNA molecules that are not bound by ligand at that site.[46] The amount of knicking can be described by a Poisson distribution when the cleavage events are independent, the length of the DNA molecules remains constant over the course of the experiment, and all labeled DNA molecules are equivalent targets. This set of conditions is referred to as the "single-hit kinetics regime," and can be practically achieved by limiting the extent of DNase I digestion so that, on average, each DNA molecule is cleaved once or less.[43] The fractional occupancy can then be determined from an autoradiogram by measuring the amount of cleavage at the binding site relative to the constant cleavage in DNA sequences where the ligand does not bind. A reduction in the concentration of radiolabeled DNA to ≤ 5% the total ligand concentration (L_{tot}) allows the approximation $L_{eq} \approx L_{tot}$, because even DNA site saturation will have no practical effect on the total ligand concentration. This can be achieved by maximizing the specific activity of the DNA.

In a footprint titration experiment, equilibrium mixtures of radiolabeled DNA and different concentrations of a ligand are exposed to the footprinting agent. The cleavage reactions are quenched, the DNA is denatured, and the cleavage products are separated by high resolution polyacrylamide gel electrophoresis under strand separating conditions (urea-PAGE) (Figure 1.1). The DNA fragments are visualized by autoradiography, a digital image is produced, and the images are analyzed to obtain fractional occupancies and individual site binding isotherms. The principle of the DNase I footprint titration method,

which has been used for the measurement of protein-DNA[42-44,47] and small molecule-DNA[27-41,48,49] stability constants, was extended to measuring association constants for small molecules binding to DNA using the footprinting agent methidiumpropyl-EDTA·Fe (MPE·Fe).[2,23] More recently, this principle has been used for the development of a technique wherein restriction endonuclease cleavage sites can be protected by oligonucleotide-directed triple helix formation at an overlapping sequence to yield a "low resolution footprint."[3] We wished to adapt these ideas to the analysis of affinity cleavage data.

Affinity Cleaving

Affinity cleaving, like footprinting, has been employed for identifying individual binding sites within radiolabeled plasmid DNA fragments.[1,50-54] The cleavage technique relies on the covalent attachment of a nonspecific DNA cleaving agent to a DNA-binding ligand of interest. The resulting affinity cleaving agent is pre-equilibrated in solution with a radiolabeled DNA fragment containing a putative binding sequence. The equilibrium mixture is then exposed to a reagent that activates the cleavage moiety, initiating the reaction. Following reaction quenching, the cleavage products and uncut DNA starting material can be separated by sequencing gel electrophoresis and the radiolabeled products visualized by autoradiography. A general scheme for an affinity cleaving experiment is shown in Figure 1.2.

Affinity cleaving has proven to be a powerful method for studying sequence-specific DNA recognition at independent sites.[1,50-54] While affinity cleavage of DNA could, in principle, be directed by any DNA-binding ligand attached to any appropriate DNA cleaving moiety, only oligonucleotide-EDTA·Fe affinity cleaving will be discussed in detail here. Such conjugates of EDTA·Fe, which

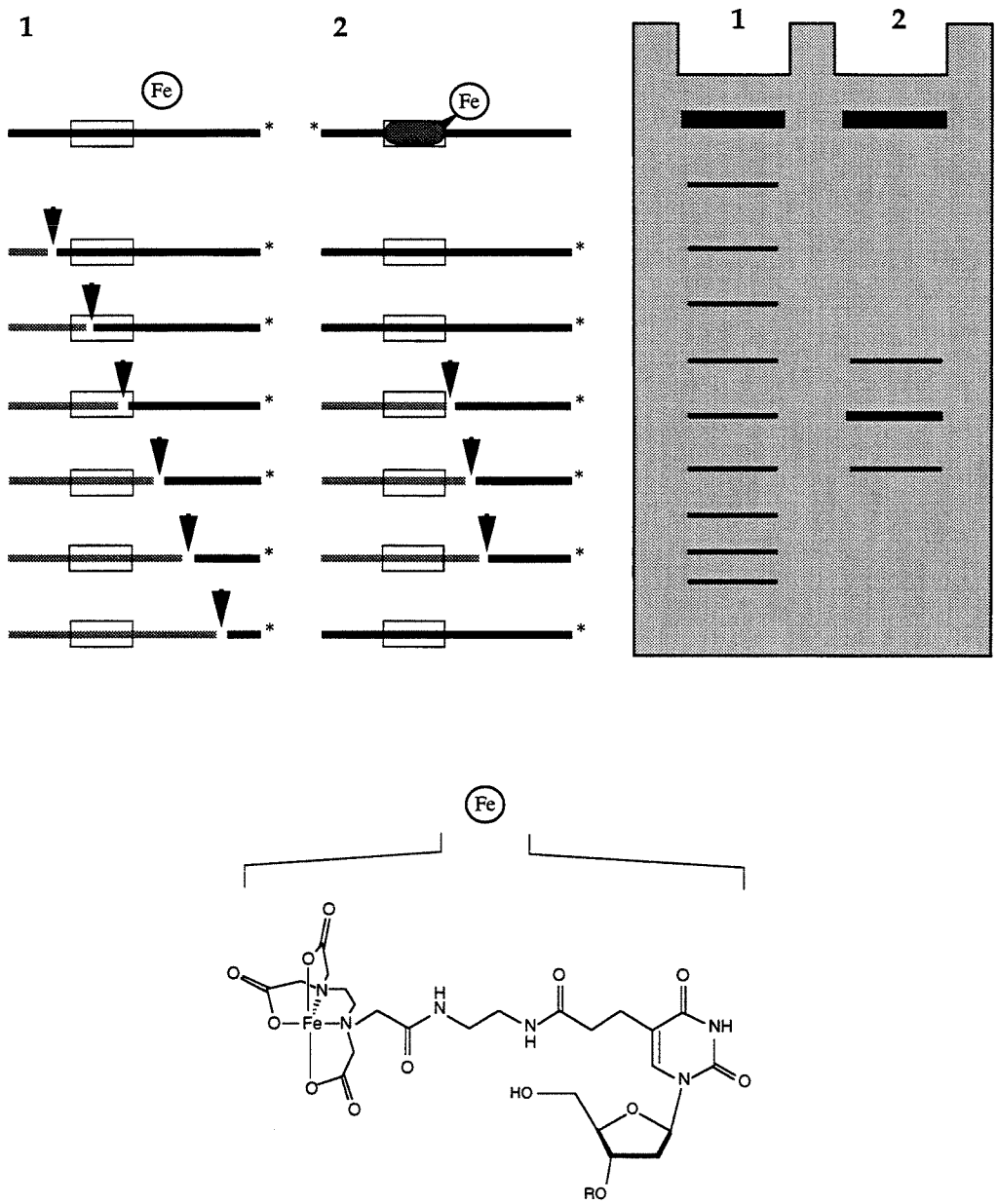


Figure 1.2. Schematic diagram of an affinity cleaving experiment. (1) A sequence indpendent DNA cleaving molecule (*e.g.*, EDTA·Fe) produces strand breaks all along a radiolabeld duplex. (2) When the DNA cleaving moiety is covalently attached to a DNA-binding ligand, the cleavage is localized near the biding site. The thymidine-EDTA (T*) nucleoside used for affinty cleaving studies involving oligonucleotides.

generate a diffusible oxidant species capable of sequence independent DNA cleavage near the bound ligand-DNA complex, have been employed to demonstrate groove locations and binding orientations, as well as to evaluate the relative binding strengths of a variety of DNA-binding molecules.[13,55-61] In addition to the identification of orientation and groove location, affinity cleaving experiments offer a number of potential advantages over the analogous footprinting experiments that could be useful in a quantitative titration.

(1) The use of affinity cleaving eliminates the possibility of direct interactions between components of the ligand-DNA complex and a footprinting agent that can reduce or eliminate a footprint. In principle, footprinting agents that bind DNA can perturb the energetics of the ligand-DNA interaction. For example, intercalating agents such as MPE·Fe bind preferentially at duplex-triplex junctions.[62-64] In addition, DNase I is inhibited by high concentrations of single-stranded DNA, including oligonucleotides.

(2) The use of a cleaving moiety that shows virtually sequence-indepndnet cleavage, such as EDTA·Fe, allows all potential binding sites to be measured with similar accuracy. In contrast, DNase I cleavage is somewhat sensitive to DNA sequence, so that some sites are inefficiently cleaved even in the absence of bound ligand.[46]

(3) The fact that affinity cleavage experiments involving tethered EDTA·Fe moieties have been performed between pH 5.5 and pH 10.0, at temperatures between 4°C and 45°C, and in the presence of a variety of cations and cation concentrations, suggests that affinity cleavage may be possible under conditions where enzymatic or chemical footprinting is not feasible.[65,66]

(4) Because the cleaving moiety is attached directly to the DNA-binding ligand, even proximal sites can be indpendently monitored. However, steric

overlap of the large DNase I protein[67] and the DNA-bound ligand produces a footprint region that is larger than the actual binding site, and prevents resolution of closely spaced sites. For oligonucleotide-directed formation of local triplexes, DNase I footprints extend 5 bp beyond both triplex termini.[26] In fact, affinity cleavage titrations have been used to measure the independent binding of oligonucleotides to abutting sites as short as 8-bp in length.[12]

(5) The EDTA·Fe-mediated cleavage chemistry is relatively independent of the nature of the nucleic acid target. Thus affinity cleavage can be used to investigate binding to single-stranded nucleic acids,[51,68] double-helical RNAs,[11]and RNA·DNA hybrid duplexes.[11]

(6) Finally, because the intensities of bands from an affinity cleavage experiment autoradiogram are positive signals (*i.e.*, increasing cleavage yields enhanced signals), they afford an increase in the signal-to-noise ratio in the primary data when compared to the observation of signal dimunition for a footprinting experiment.

A potential disadvantage of using affinity cleavage is that the EDTA·Fe moiety must be covalently attached to the ligand, and the association constants measured by affinity cleavage titrations could be somewhat altered from those of the parent ligand. It should be possible to calibrate the effect of the EDTA·Fe group on the binding constant under some conditions by comparison with a footprint titration (see Chapter Two).[6]

Analysis of Kinetic Factors in Affinity Cleavage

The theoretical development of the quantitative affinity cleavage titration method is based on the formulation, by Ackers and co-workers,[42,43] of the DNase I footprint titration technique. The experimental scheme for an affinity cleavage

titration is shown in Figure 1.3. Thymidine-EDTA[51] is incorporated into an oligonucleotide using standard solid phase synthetic methods.[69,70] Iron is loaded into the EDTA moiety, and the oligonucleotide-EDTA·Fe is pre-incubated with the target DNA. Equilibrium mixtures of different oligonucleotide-EDTA·Fe concentrations with radiolabeled DNA are exposed to dithiothreitol (DTT) in the presence of molecular oxygen, producing localized DNA cleavage. The cleavage reactions are quenched and the products separated electrophoretically. By measuring the cleavage intensity over a systematically varied concentration range, a titration isotherm can be constructed. We examine the kinetic foundations of such an experiment in the following paragraphs.

Scheme 1

$$\mathbf{O} + \mathbf{D} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \mathbf{T}$$

Triple Helix Association Kinetics. Consider the one-to-one association of an oligonucleotide, **O**, with a duplex DNA sequence, **D**, to form a local triple-helical complex, **T** (Scheme 1). Under equilibrium conditions (*e.g.*, following the pre-incubation period), the derivation of the concentration of triple-helical complex as a function of time is straightforward.[3] The rate equation describing the change in concentration of **T** with time, t, can be written as:

$$\begin{aligned} \frac{dT}{dt} &= k_{\text{on}}OD - k_{\text{off}}T \\ &= k_{\text{on}}O(D_{\text{tot}} - T) - k_{\text{off}}T \\ &= k_{\text{on}}OD_{\text{tot}} - (k_{\text{on}}O + k_{\text{off}})T \end{aligned} \tag{6}$$

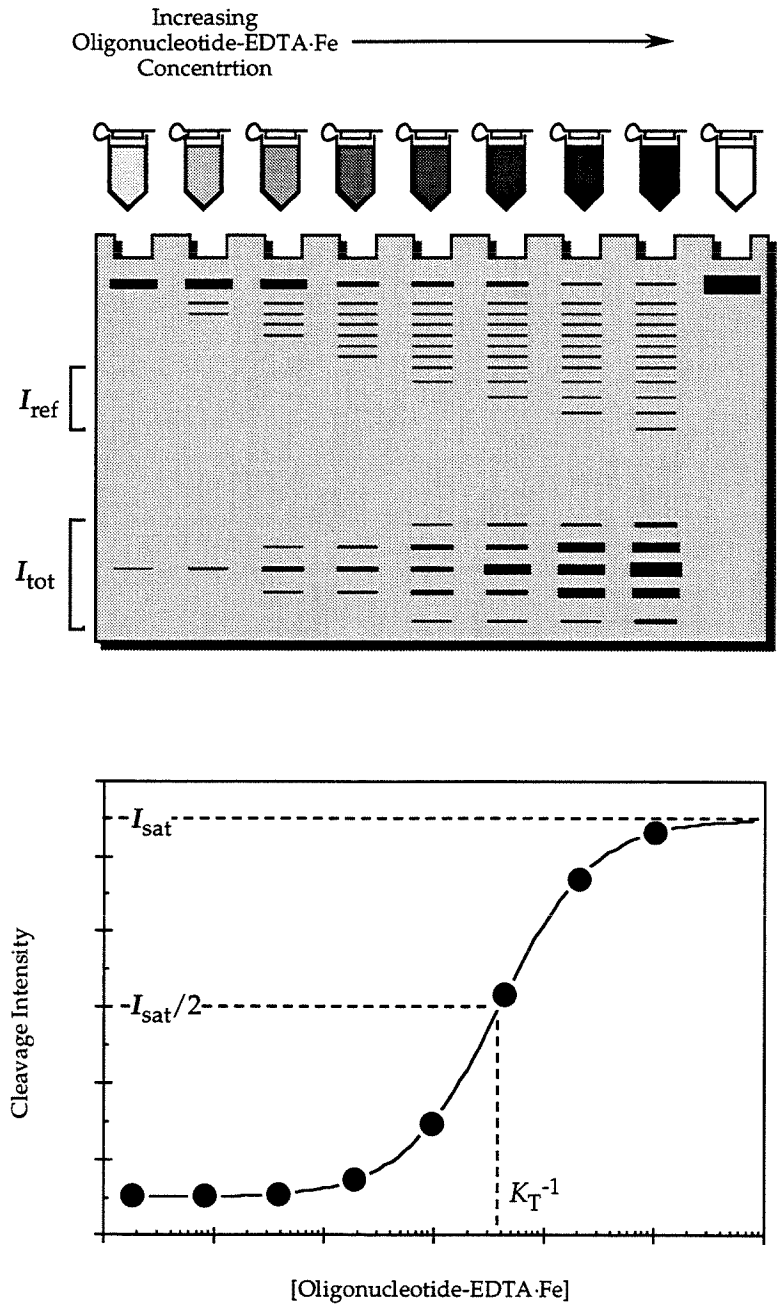


Figure 1.3. Schematic representation of the key steps in performing an affinity cleavage titration. The concentration of an oligonucleotide-EDTA·Fe is systematically vaired, and the site-specific cleavage is measured. Plotting the resulting data points allows fitting of a Langmuir titration isotherm.

where O, D, and T represent the concentrations of free oligonucleotide, uncomplexed duplex, and triplex, respectively. The subscript "tot" indicates a total concentration. Solution of the differential equation gives an expression for the concentration of complex after a given pre-incubation time, t_{inc}:

$$T = T_{eq}\left\{1 - \exp\left[-(k_{on}O + k_{off})t_{inc}\right]\right\} \tag{7}$$

Eq 7 simply relates T to the fractional formation of the equilbrium concentration of complex, T_{eq}. Upon division of both sides by the constant D_{tot}, this relation can be written for the fraction of duplex bound:

$$\theta_{inc} = \theta_{eq}\left\{1 - \exp\left[-(k_{on}O + k_{off})t_{inc}\right]\right\} \tag{8}$$

With the use of eq 8, the fraction of duplex sites bound by oligonucleotide can be calculated for any t_{inc}. By defining a parameter, ε, as the extent to which the system at t_{inc} reflects the equilibrium condition so that $(1\text{-}\varepsilon) \equiv (\theta_{eq} - \theta)$, the incubation time required for equilibration can also be calculated from eq 8:

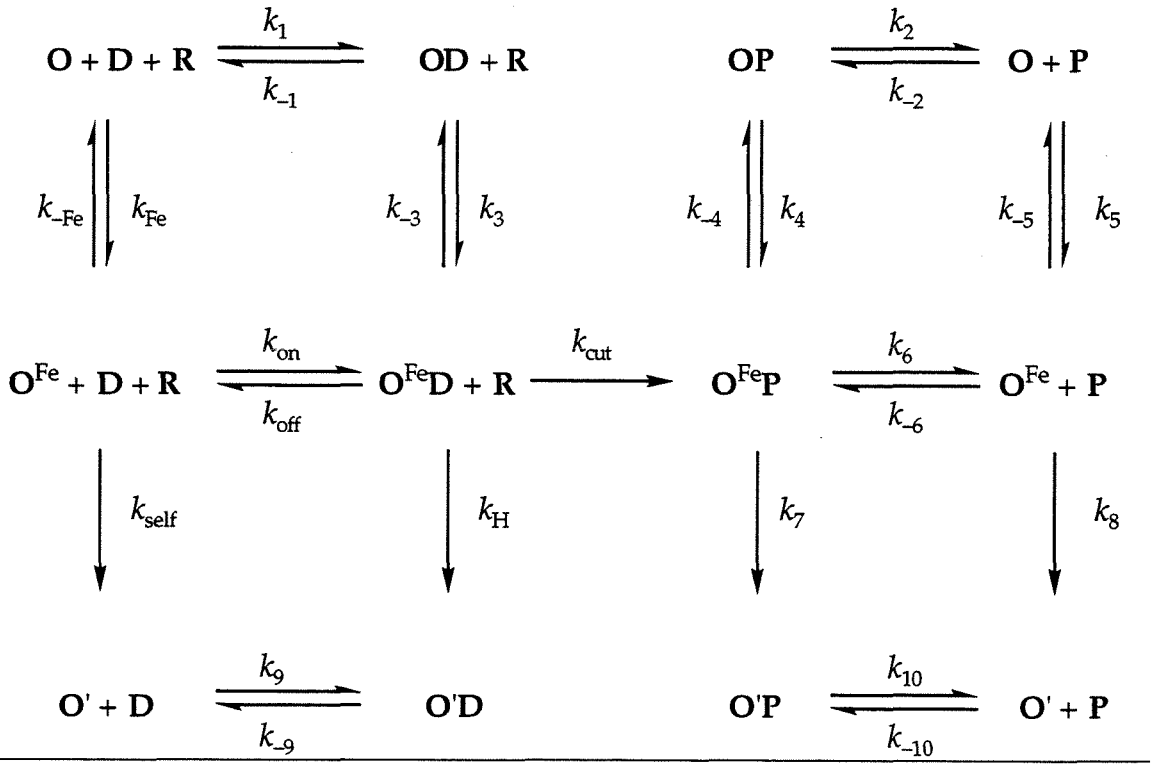
$$t_{inc} = -\left(1 - \theta_{eq}\right) \cdot \frac{K_T}{k_{on}} \cdot \ln(1 - \varepsilon) \tag{9}$$

where K_T is the equilbrium constant for triplex formation ($K_T \equiv k_{on}/k_{off}$).

Triple Helix Affinity Cleavage Kinetics. In considering an equilibrium for binding of an oligonucleotide-EDTA·Fe affinity cleaving agent ($\mathbf{O}^{Fe}$) to the target duplex, a new set of equilibria must be added to Scheme 1, namely those involving thymidine-EDTA and iron (Scheme 2). Moreover, when the cleavage reaction

is initiated, the complex can decay along the reaction pathway to products. These considerations contribute complexity to eq 6 by the addition of rate constants and another species whose concentration changes as a function of time. Scheme 2 represents the most reasonable "complete" set of transformations involved in the affinity cleavage mechanism.

Scheme 2



In Scheme 2, **O** is an oligonucleotide-EDTA, **O'** is an oligonucleotide which is no longer able to cleave the target site (as a result of degradation of EDTA or the linker between thymidine and EDTA, or the loss of some small oligonucleotide fragment containing thymidine-EDTA), **P** is a duplex molecule containing a single lesion on the labeled strand resulting from oxidative chemistry, **R** is a

coreactant (*e.g.*, DTT), and **OD**, **OP**, $\mathbf{O^{Fe}D}$, $\mathbf{O^{Fe}P}$, **O'D**, and **O'P** are the triple-helical complexes involving the corresponding oligonucleotide or fragment and a target or product duplex molecule. Even this scheme is an approximation: the equilibria and reactions are assumed to be conducted at fixed concentrations of DTT, spermine, salts, pH, temperature, etc. While these are reasonable approximations in general, the concentration of DTT, for example, changes during the course of the reaction and can become a limiting factor in the rate of the cleavage chemistry.

Scheme 3

$$\mathbf{O} + \mathbf{D} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathbf{T} \xrightarrow{k_d} \mathbf{P}$$

It would be impossibly difficult to derive a series of testable equations to describe the concentrations of all of the species depicted in Scheme 2. Instead, we have concentrated on solving the kinetic equations for a simplified series of equilibria and reactions (Scheme 3). In simplifying the model, several assumptions were made. First, based on the stability of EDTA·Fe complexes, we assume that the oligonucleotide-EDTA remains complexed with iron throughout the experiment ($k_{Fe} \gg k_{-Fe}$ and $k_3 \gg k_{-3}$), and therefore represent oligonucleotide-EDTA·Fe by **O** rather than $\mathbf{O^{Fe}}$. Second, the rate of autocleavage of the oligonucleotide (k_{self}) is assumed to be neglible. Third, based on the assumption that oligonucleotide "turnover" is not important over the short reaction time (relative to the lifetime of the triple-helical complex), cleavage of the oligonucleotide (Hoogsteen) strand within the triple-helical complex (k_H) is ignored. Fourth, the

reaction is performed in a large excess of DTT, so that the appearance of products is assumed to be a pseudo-unimolecular process from the molecular complex ($k_{cl} = k_{cut}[\mathbf{R}]$). Finally, we consider all product species to be identical. The last assumption is valid because, following reaction quenching, the products are fully denatured and then electrophoretically separated under strand-denaturing conditions. However, the prior assumptions are less robust and may limit the analysis under some conditions. Our justifications for making the first four assumptions and the scope of the limits they impose are discussed in more detail in Part IV of this chapter.

In analogy to the DNase I footprint titration formulation, we wish to derive an expression similar to eq 5. However, it is necessary to relate θ to an experimentally observable variable. For affinity cleaving, this can be accomplished by writing an equation relating the concentration of cleavage products to a function of the rate constants in Scheme 3 and the reaction time, t_{rxn}. From Scheme 3, the rate of product appearance is given by:

$$\frac{dP}{dt} = k_{cl}T \tag{10}$$

where k_{cl} is the pseudo-unimolecular rate constant for cleavage of the radiolabeled duplex strand within the triple-helical complex. The concentration of triple-helical complex will itself be a function of time, given by the appropriate rate equation:

$$\frac{dT}{dt} = k_{on}OD - (k_{off} + k_{cl})T \tag{11}$$

The concentration of duplex also changes as the full-length substrate is converted to products, but can be replaced by an equivalent expression containing terms of the two variables, T and t. Combining eq 11 with the mass balance, $D_{tot} = D + T + P$, yields the following:

$$\frac{dT}{dt} = k_{on}O(D_{tot} - T - P) - (k_{off} + k_{cl})T \tag{12}$$

Comparison of eq 10 with eq 12 reveals that P and T, both functions of the reaction time, cannot be isolated. Hence, it is impossible to find an explicit solution to either differential equation. We have considered three solutions to eq 12, each of which requires an approximation concerning the P term.

Approximation 1

$$P \approx 0$$

The first approximation requires that the cleavage reaction be run to a low conversion yield ($P \ll D_{tot}$). Then, from eq 12, we can write:

$$\begin{aligned}\frac{dT}{dt} &= k_{on}O(D_{tot} - T) - (k_{off} + k_{cl})T \\ &= k_{on}OD_{tot} - (k_{on}O + k_{off} + k_{cl})T\end{aligned} \tag{13}$$

Differential equation 13 can be solved using the initial condition that T is given by eq 8 following the pre-incubation ($t_{rxn} = 0$):

$$T = D_{\text{tot}} \frac{k_{\text{on}}O}{k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}}} \left\{1 - \exp\left[-(k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}})t_{\text{rxn}}\right]\right\} +$$
$$D_{\text{tot}} \frac{k_{\text{on}}O}{k_{\text{on}}O + k_{\text{off}}} \left\{1 - \exp\left[-(k_{\text{on}}O + k_{\text{off}})t_{\text{inc}}\right]\right\} \times \tag{14}$$
$$\exp\left[-(k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}})t_{\text{rxn}}\right]$$

Because we are interested in the fraction of duplex bound by oligonucleotide (eq 5), we define the following terms:

$$\theta_{\text{eq}} = \frac{k_{\text{on}}O}{k_{\text{on}}O + k_{\text{off}}} \tag{15}$$

$$\theta_{\text{inc}} = \frac{k_{\text{on}}O}{k_{\text{on}}O + k_{\text{off}}} \left\{1 - \exp\left[-(k_{\text{on}}O + k_{\text{off}})t_{\text{inc}}\right]\right\} \tag{16}$$

$$\theta_{\text{s}} = \frac{k_{\text{on}}O}{k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}}} \tag{17}$$

Eq 15 simply represents the fraction of duplex bound at equilibrium, while eq 16 involves the fraction of duplex bound after the pre-incubation period. Eq 17 defines the apparent fraction of duplex bound under steady-state conditions, and θ_{s} can be directly compared with a corresponding term in Michaelis-Menten kinetics.[71] These terms can be used to rewrite eq 14:

$$T = D_{\text{tot}}\theta_{\text{s}} \left\{1 - \exp\left[-(k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}})t_{\text{rxn}}\right]\right\} +$$
$$D_{\text{tot}}\theta_{\text{inc}} \exp\left[-(k_{\text{on}}O + k_{\text{off}} + k_{\text{cl}})t_{\text{rxn}}\right] \tag{18}$$

Finally, the concentration of **P** for Approximation 1 can be determined by solution of the rate law (eq 10) using the equality expressed by eq 18:

$$P = D_{tot}k_{cl}t_{rxn}\theta_s + \frac{D_{tot}k_{cl}}{k_{on}O + k_{off} + k_{cl}}(\theta_{inc} - \theta_s)\left\{1 - \exp\left[-(k_{on}O + k_{off} + k_{cl})t_{rxn}\right]\right\} \quad (19)$$

Approximation 2

$$P \approx T_{inc} - T$$

A second approximation can be made if the reaction is conducted such that the equilibrium between duplex and triplex is not altered during the reaction. Specifically, this will be true for short reaction times during which the triplex does not dissociate ($t_{rxn} \ll k_{off}^{-1}$) and the concentration of the triplex does not change enough to perturb the equilibrium position established during the pre-incubation. Substitution of the relation given by Approximation 2 into eq 12 yields:

$$\frac{dT}{dt} = k_{on}OD_{tot}(1 - \theta_{inc}) - (k_{off} + k_{cl})T \quad (20)$$

Solution of differential equation 20 gives the following:

$$T = D_{tot}\theta_{inc}\exp\left[-(k_{off} + k_{cl})t_{rxn}\right] + \frac{k_{on}}{k_{cl} + k_{off}}OD_{tot}(1 - \theta_{inc})\left\{1 - \exp\left[-(k_{off} + k_{cl})t_{rxn}\right]\right\} \quad (21)$$

The concentration of **P** can then be obtained by substitution of the expression for T (eq 21) into eq 10 followed by solution of the resulting differential equation:

$$P = \frac{k_{cl}}{k_{off} + k_{cl}} D_{tot}\left[\theta_{inc} - \frac{k_{on}}{k_{off} + k_{cl}} O(1 - \theta_{inc})\right] \times \left\{1 - \exp\left[-(k_{off} + k_{cl})t_{rxn}\right]\right\} + \frac{k_{on}}{k_{off} + k_{cl}} OD_{tot}(1 - \theta_{inc})k_{cl}t_{rxn} \quad (22)$$

Approximation 3

$$P \approx D_{tot} - \frac{T}{\theta_{inc}}$$

A third approximation can be employed to solve eq 12 when the cleavage reaction is slower than the equilibration and the products do not take part in the equilibration reaction. Under these conditions, θ is constant over the reaction time because the equilibrium shifts slightly toward **T** when products are formed. Substitution of this approximation into eq 12 yields:

$$\frac{dT}{dt} = \left[k_{on} O\left(\frac{\theta_{inc} - 1}{\theta_{inc}}\right) + k_{off} + k_{cl}\right] T \quad (23)$$

The solution of differential equation 19 gives the following:

$$T = D_{tot}\theta_{inc} \exp\left\{-\left[k_{on} O\left(\frac{\theta_{inc} - 1}{\theta_{inc}}\right) + k_{off} + k_{cl}\right] t_{rxn}\right\} \quad (24)$$

The concentration of **P** can then be obtained by substitution of the expression for *T* (eq 24) into eq 10, followed by solution of the resulting differential equation:

$$P = \frac{k_{cl} D_{tot} \theta_{inc}}{k_{on} O\left(\frac{\theta_{inc} - 1}{\theta_{inc}}\right) + k_{off} + k_{cl}} \times \left(1 - \exp\left\{-\left[k_{on} O\left(\frac{\theta_{inc} - 1}{\theta_{inc}}\right) + k_{off} + k_{cl}\right] t_{rxn}\right\}\right) \quad (25)$$

Model Testing. In order to test these three models (eqs. 19, 22, and 25), we first found reasonable estimates of the rate constants themselves. For the cleavage reaction at 22°C and pH 7.0, we know that the cleavage yields at saturating concentrations of oligonucleotide span the range of 5 – 20% for reaction times of 4 – 8 h.[6,10,11] From eq 10, the slope of a plot of relative cleavage yields (P/D_{tot}) at saturating concentrations of oligonucleotide versus time will be equal to k_{cl}. Such an analysis suggests that k_{cl} is in the range between 1×10^{-6} s^{-1} to 2×10^{-5} s^{-1}. Moser and Dervan have measured single- and double-strand cleavage efficiencies as a function of time. Their results are consistent with $k_{cl} \approx 5 \times 10^{-5}$ s^{-1}.[72] A reasonable value for the apparent pseudo-first-order rate constant (k_{cl}) is 1×10^{-5} s^{-1}.

Literature values of apparent pseudo-first order on-rate constants for DNA triplex formation vary from 10^2 $M^{-1}s^{-1}$ to 10^4 $M^{-1}s^{-1}$ with the actual value depending on the solution conditions. Helene and co-workers reported that a 22-nt oligopyrimidine binds to a 22-bp duplex with $k_{on} = 1.8 \times 10^2$ $M^{-1}s^{-1}$ at 15 °C, pH 6.8, 100 mM Na^+.[73] Crothers has reported that a 12-nt oligopyrimidine binds to a 12-bp hairpin duplex with $k_{on} = 2.2 \times 10^4$ $M^{-1}s^{-1}$ at pH 5.5, 100 mM Na^+.[74] Maher et al. reported that a 21-nt oligopyrimidine occupies a 21-bp target sequence

within a plasmid with k_{on} = 2.2 x 10^3 $M^{-1}s^{-1}$at 37 °C, pH 6.5, 70 mM Na^+, 20 mM Mg^{2+},and 0.4 mM spermine.[3] The latter result is likely the best starting point for estimation of k_{on} in our case because the target molecule and conditions are most similar. For our system, the oligonucleotide is shorter, the temperature is lower (larger k_{on}),[73] the pH is higher (smaller k_{on}),[3] [Na^+] is higher (smaller k_{on}),[3] [Mg^{2+}] is lower (smaller k_{on}),[3] and [spermine] is higher (larger k_{on}).[75] From this analysis, it is likely that k_{on} is within the range of 2 x 10^2 to 2 x 10^4 $M^{-1}s^{-1}$. The ratio k_{on}/K_T can be used to estimate the off-rate constants.

The test of the three models involved computer modeling of the behavior of the now well-characterized 15mer, 5'-T*TTTTCTCTCTCTCT-3',[4,6-9] using eqs 19, 22, and 25. Specifically, $\boldsymbol{P}$ was calculated using each equation for a range of $\boldsymbol{O}_{tot}$ (where $\boldsymbol{O}_{tot} \gg \boldsymbol{D}_{tot}$). All three models afforded appropriate titration curves for simulated reaction times from 2 – 8 h (Figure 1.4). However, eq 19 does not appear to be valid for reaction times longer than ~ 2 h because the approximation $\boldsymbol{P}$ = 0 is not valid for longer reaction times. The best fit of the known data for the 15mer (K_T = 3.7 x 10^6 M^{-1}, ~ 10% cleavage after 6 h, and ≥ 12 h required for equilibration)[6] resulted when k_{on} = 2.8 x 10^2 $M^{-1}s^{-1}$, k_{off} = 7.6 x 10^{-5} s^{-1}, and k_{cl} = 4.5 x 10^{-6} s^{-1}. These rate constants are all within the reasonable limits outlined above. The third approximation (eq 25) produced the most realistic results, likely because $k_{off} > k_{cl}$ (*vide infra*).

Further Simplifications. To analyze the differences between the approximations further, we considered the limiting behaviors of the models ($t_{rxn} \rightarrow 0$ and $t_{rxn} \rightarrow \infty$) and further simplifications. For Approximation 1, we note that for reaction times near zero, the second term approaches zero and the concentration of products is linearly proportional to the reaction time. At long reaction times, $\boldsymbol{P}$ becomes unbounded. Thus, while Approximation 1 holds for extremely short

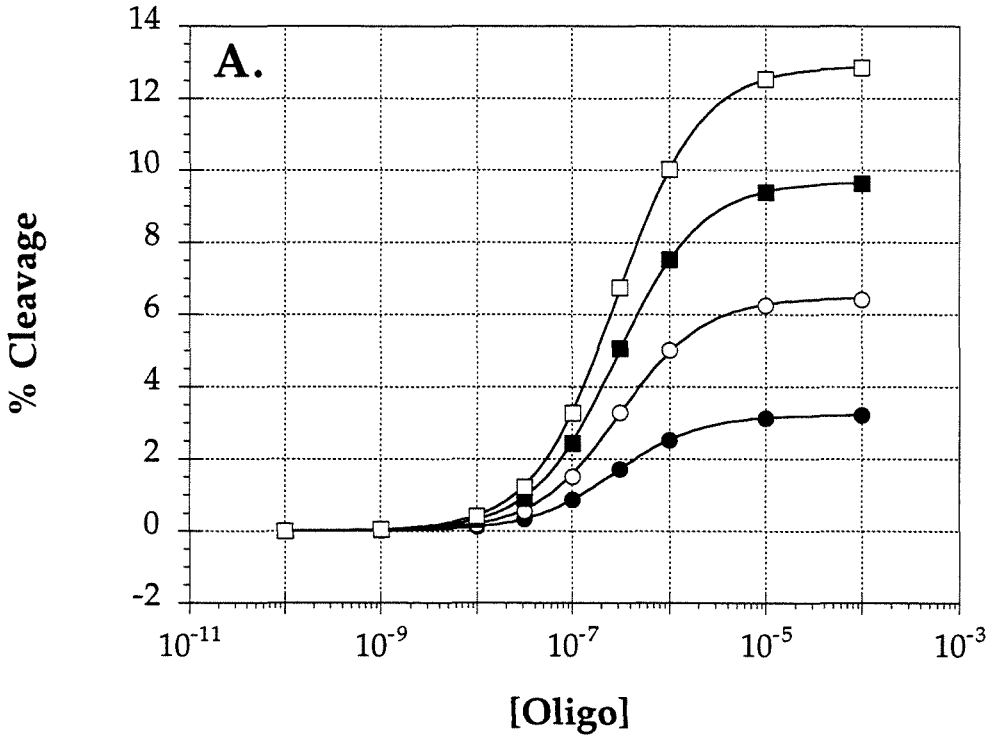
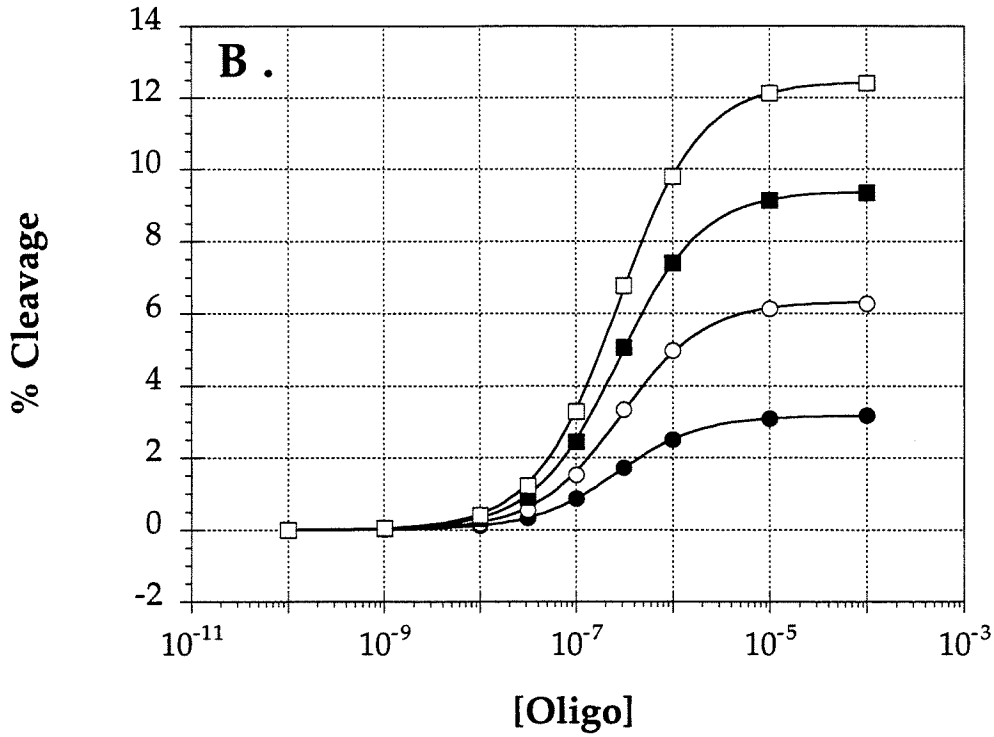
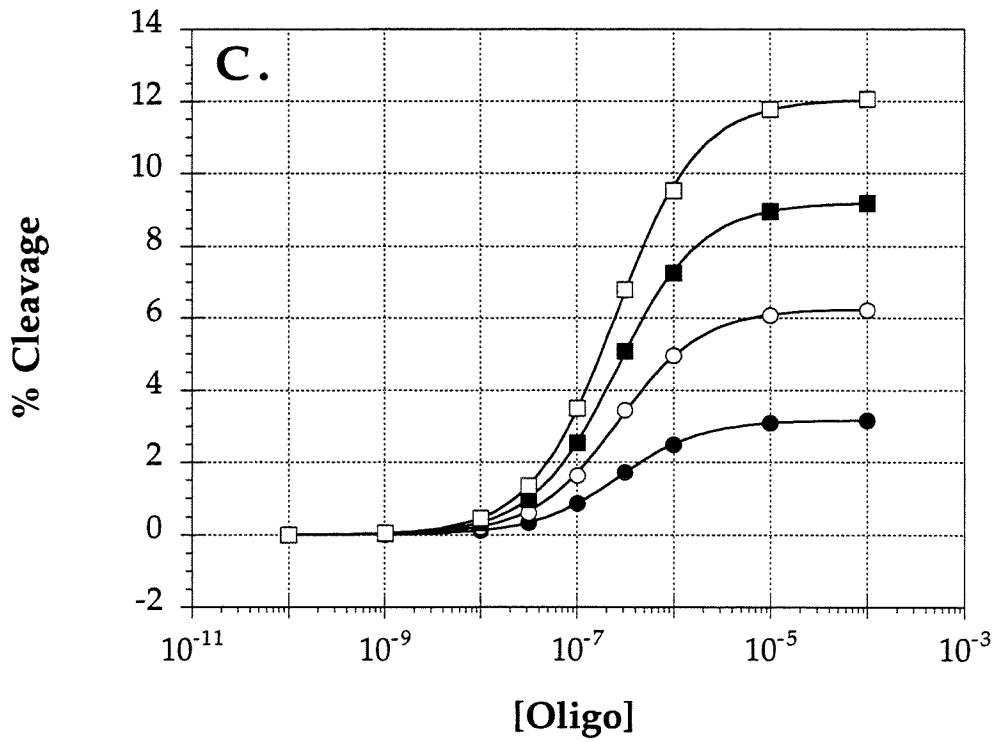


Figure 1.4. Results of testing three different approximate integrated rate equations (eqs 19, 22, and 25). Calcuated cleavage yields after 2 h (•), 4 h (○), 6 h (■), and 8 h (□) are plotted as a function of oligonucleotide-EDTA·Fe concentration. (A) Eq 19. (B) Eq 22. (C) Eq 25.


B .
% Cleavage
14
12
10
8
6
4
2
0
-2
10^{-11}
10^{-9}
10^{-7}
10^{-5}
10^{-3}
[Oligo]


C.
% Cleavage
14
12
10
8
6
4
2
0
-2
10^{-11}
10^{-9}
10^{-7}
10^{-5}
10^{-3}
[Oligo]

reaction times, the overall behavior described by the model does not correspond to the chemical reaction. For Approximation 2, when $t_{rxn} \ll k_{off}^{-1}$, $k_{cl} > k_{off}$, and $\theta_{inc} \approx \theta_{eq}$, eqs 21 and 22 are transformed into the following simplified equations, respectively:

$$T = D_{tot}\theta_{eq} \tag{26}$$

$$P = D_{tot}\theta_{eq}\left[1 - \exp(-k_{cl}t_{rxn})\right] \tag{27}$$

For longer reaction times, however, P becomes unbounded (eq 22). When the pre-incubation time is sufficiently long to allow equilibrium to be reached ($\theta_{inc} \approx \theta_{eq}$), Approximation 3 (eqs 24 and 25) yields the same simplified equations shown above (eqs 26 and 27, respectively). As the reaction time increases, the concentration of products calculated using Approximation 3 is bounded by the total concentration of duplex and the fraction of duplex bound at the start of the reaction. Although it is conceptually more complex than the second approximation, Approximation 3 represents a more reasonable model of the experimental conditions, requires fewer assumptions for simplification, and reproduces the real behavior of the system with the most fidelity. Therefore, we have employed Approximation 3 and the simplified equation 27 for the further development of the affinity cleavage titration method.

Site Occupation from Affinity Cleavage

Using Approximation 3 to solve eq 10 (eq 25), and assuming that the binding reaction has been fully equilibrated (eq 27), the amount of DNA cleavage products resulting from an oligonculeotide-EDTA·Fe bound to the radiolabeled

DNA fragment is dependent on the amount of triple helix formed during pre-incubation and the kinetics of the cleavage reaction:

$$P = T_{eq}\left[1 - \exp(-k_{cl}t_{rxn})\right] \tag{28}$$

The amount of radiolabeled product can be measured using autoradiography. The intensity of the signal at the target site (i_{site}) will be proportional to the amount of cleavage product present ($|\mathbf{P}|$) and the specific acitivity of the radiolabel (A):

$$I_{site} = k_{dens} A\,|\mathbf{P}| \tag{29}$$

where k_{dens}, the proportionality constant for the densitometric analysis, accounts for such variables as the autoradiographic exposure time and the translation of radioactive decays into measurable "counts." Eqs 28 and 29 can be combined to express I_{site} as a function of T:

$$I_{site} = k_{dens} A V_{rxn} T_{eq}\left[1 - \exp(-k_{cl}t_{rxn})\right] \tag{30}$$

where $|\mathbf{P}| = P \cdot V_{rxn}$ and V_{rxn} is the reaction volume. Using the definition of θ, eq 30 may be rewritten as

$$I_{site} = k_{dens} A V_{rxn} D_{tot}\left[1 - \exp(-k_{cl}t_{rxn})\right]\theta_{eq} \tag{31}$$

Thus, for a series of affinity cleaving reactions where the reaction time, reaction volume, and concentration of labeled duplex are constant, the fraction of duplex bound will be proportional to the signal intensity at the binding site

$$\theta = k_{ac} \cdot I_{site} \tag{32}$$

where k_{ac} is the empirical proportionality constant for affinity cleavage that accounts for D_{tot}, V_{rxn}, t_{rxn}, k_{dens}, and A of a specific titration experiment. In practice, this parameter is defined by measuring the signal intensity resulting from cleavage at saturation binding, $I_{sat} = k_{ac}^{-1}$, so that

$$\theta = I_{site} \cdot I_{sat}^{-1} \tag{33}$$

Combining eq 33 with eq 5, and $I_{site} = I_{sat}$ when $\theta = 1$, yields:

$$K_T = \frac{I_{site}}{I_{sat} - I_{site}} \cdot \frac{1}{O} \tag{34}$$

As is the case for footprint titrations (eq 5), eq 34 indicates that an accurate measurement of K_T requires only that cleavage signal intensities and concentrations of unbound oligonucleotide-EDTA·Fe be measured. The signal intensities can be measured from the intensity of luminescence from a storage phosphor autoradiogram.[6,76] A reduction in the total concentration of duplex DNA, such that $D_{tot} \ll O_{tot}$, allows the approximation $O \approx O_{tot}$ (*vide supra*).

The cleavage signal intensity that is actually observed at the binding site (I_{tot}) has a specific component (I_{site}) which is produced by oligonucleotide-

EDTA·Fe bound at the target site, and a nonspecific component (I_{uns}) which is produced by unbound oligonucleotide-EDTA·Fe (Scheme 4)

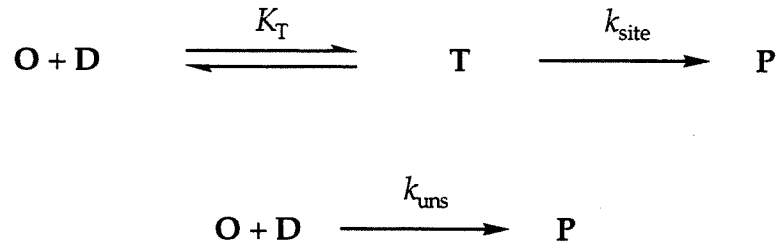
$$I_{tot} = I_{site} + I_{uns} \tag{35}$$

Although both components should vary with the concentration of the oligonucleotide-EDTA·Fe, eq 34 treats only the specific component. Nonspecific cleavage at the binding site is proportional to the cleavage at other sequences where no oligonucleotide-EDTA·Fe is bound. Measurement of the cleavage intensity at a reference site that is spatially removed from the binding site of interest, and displays no specific component of cleavage (I_{ref}), allows the specific cleavage to be determined by

$$I_{site} = I_{tot} - \lambda I_{ref} \tag{36}$$

where λ is a scaling parameter, empirically determined by the ratio I_{tot}/I_{ref} at $\theta = 0$, that accounts for any intrinsic differences between the target and reference sites (*e.g.*, the number of bands quantitated).

Scheme 4



Rearrangement of eq 34, followed by substitution with the relationship for the specific component of cleavage and the approximation $O = O_{tot}$ yields

$$I_{tot} - \lambda I_{ref} = I_{sat} \frac{K_T O_{tot}}{1 + K_T O_{tot}} \tag{37}$$

The equilibrium constant for triple helix formation, K_T, can be determined by performing a series of affinity cleavage reactions in which only the concentration of oligonucleotide-EDTA·Fe is systematically varied, measuring the amount of cleavage products generated at the target and reference sites over the concentration range, and fitting the data using a nonlinear least squares algorithm. Connors has demonstrated a general method for deriving binding expressions for equilibria of any molecularity.[71] In addition, a method for transforming binding models into testable binding isotherms such as eq 37, that can be evaluated experimentally, was presented. Because, in contrast to linear isotherms such as Lineweaver-Burke and Scatchard plots, the semi-logrithmic Bjerrum formation function maintains isolation of the variables, evenly distributes the data along a concentration range spanning several orders of magnitude, and allows the point of inflection (defined by $\theta = 0.5$ at $O_{tot} = K_T^{-1}$) to be accurately determined even

when the maximum y-value (I_{sat}) cannot,[71] we have used eq 37 along with semi-log plots of the (I_{site}, O_{tot}) data points to determine equilibrium association constants.

Part II. Protocol for Performing Quantitative Affinity Cleavage Titrations

Michael Brenowitz, formerly of Gary Ackers' laboratory, and his collaborators have reported a detailed protocol for DNase I footprint titrations.[46] This work serves as an excellent and highly recommended general reference for performing quantitative titration experiments. I have attempted to follow their exemplary text here.

Preparation of Oligonucleotide-EDTA and Radiolabeled Duplex DNA

The measurement of reliable association constants requires high quality quantitative titrations, which, in turn, require ligands and DNA of high purity. Thermodynamic information can be obtained only if the chemical activities of the reaction components are known.[46]

1. Oligonucleotide-EDTA Preparation.

Oligodeoxyribonucleotides can be synthesized using standard automated chemistry on controlled-pore glass (CPG) in an Applied Biosystems Model 380B DNA synthesizer.[69,70] We purchased suitably protected thymidine and 2'-deoxycytidine *O*-cyanoethyl-*N,N*-diisopropyl phosphoramidites from Cruachem. The 5'-*O*-DMT-thymidine–EDTA nucleoside phosphoramidite was prepared as described[51] and incorporated at the 5' terminus of oligonucleotides with the EDTA carboxylates protected as their ethyl esters. Following oligonucleotide synthesis, 5'-terminal DMT protecting groups are not removed and the oligonucleotide-EDTA conjugates are purified by the protocol described below.

1. Following a 1-μmol scale synthesis, dry the oligonucleotide on the solid support *in vacuo*, and transfer equal portions into two 1.5-mL screw-top Eppendorf tubes.
2. Suspend the oligonucleotide-CPG in each tube in 1.2 mL of 0.1 N NaOH, cap the tubes, and seal with parafilm. Carry out deprotection at 55°C for 36 h.
3. After the mixtures have cooled to room temperature, neutralize each solution to ~ pH 8 (determined by spotting pH paper) by adding a few μL of 1.0 M acetic acid.
4. Filter both mixtures using a 0.45-μ Centrex membrane unit in a clinical centrifuge. Freeze the filtrate and lyophilize.
5. Dissolve the lyophilized powder in 2.0 mL of 0.1 M triethylammonium acetate at pH 7.0 (FPLC Buffer A) and filter as above.
6. Purify the crude oligonucleotide product bearing the dimethoxytrityl protecting group on its 5'-terminal hydroxyl by reverse phase FPLC using a ProRPC HR10/10 (C_2/C_8) column (Pharmacia LKB) and a gradient of 0–40% acetonitrile in 0.1 M triethylammonium acetate, pH 7.0. Collect and pool fractions containing the DMT-oligonucleotide-EDTA in a 15-mL polystyrene Falcon tube. Freeze the solution and lyophilize.
7. To detritylate the oligonucleotide, dissolve the solid residue in 1.0 mL of 80% (V/V) acetic acid and vortex the solution thoroughly. Centrifuge the tube in a clinical centrifuge to collect the small volume in the bottom of the tube, and transferr to a 1.7-mL Eppendorf tube. After the reaction has proceeded for a *total* time of 20 min at room temperature, concentrate the solutions on a SpeedVac rotary evaporator.

8. Dissolve the fully deprotected oligonucleotide-EDTA in 2.0 mL of FPLC Buffer A, filter and chromatographically fractionate as above. Collect the fractions corresponding to the oligonucleotide product on a shallow gradient (~1% B/5 min). Pool the collected fractions in a 15-mL Falcon tube, freeze the solution, and lyophilize.
9. Desalt the FPLC-purified oligonucleotide-EDTA on a NAP-5 column (Pharmacia LKB) using MilliQ water as the eluent according to the manufacturer's protocol. Lyophilize the collected fraction, and lyophilize from water a second time.
10. Dissolve the powder in ~ 5 mL MilliQ water. From this solution, dilute 40 μL to a total volume of 600 μL with more water. Transfer the diluted solution into a semi-micro quartz cuvette with a 1-cm pathlength and record the UV absorbance spectrum from 200 to 350 nm. Determine the concentration of oligonucleotide using the absorbance at 260 nm and the molar extinction coefficient calculated from the nucleotide values (using the value for T in place of T* in the calculation).[77]
11. Transfer aliquots of the original solution corresponding to 5 – 10 nmol oligonucleotide into 1.5-mL tubes, lyophilize, and store dry at -20°C.

2. *DNA Preparation*.

The DNA must be radiolabeld to high specific activity in order to maintain the DNA concentration as low as possible. Normally, this can be achieved by inserting several radiolabeled nucleotides into the 3' recessed end created by restriction endonuclease cleavage.[23,43,46] However, because single-stranded oligonucleotides can be quantitatively phosphorylated using high specific activ-

ity [γ-^{32}P]ATP, the labeled oligonucleotide can be used as a primer for high-fidelity polymerase chain reaction (PCR) synthesis of labeled duplex DNA. We prepared the 5'-^{32}P-labeled duplex by PCR amplification of a 438-bp segment from the plasmid pDMAG10,[13] using a "Vector PCR" procedure and Vent DNA Polymerase.[78,79]

The error-rate of Vent polymerase has been measured to be 57 per 10^6 nucleotides.[80] Using 50 pmol of primer, 70 nmol of nucleotides will be added to the newly synthesized DNA. Based on the measured error rate, we calculate that less than 0.2% of the 15-bp triplex formation sites will contain a single mutation.

A typical yield was 2 μg of the desired fragment with a total Cerenkov radioactivity of 6 x 10^6 cpm. The primers used for amplification, 5'-GAGCCCTTTCGTCTTCAAG-3' (5'pDM-EcoRI) and 5'-TGCGGCGACGATAGTCATGC-3' (3'pDM-SalI), were synthesized and purified as described above, except that the initial deprotection step was performed in concentrated ammonium hydroxide rather than 0.1 N NaOH.

1. End-label 60 pmol of 5'pDM-EcoRI in two separte microfuge tubes. Each tube should contain 30 pmol primer, 32 pmol [γ-^{32}P]ATP (ICN Biomedicals end-labeling grade; 7000 Ci/mmol), and 30 units T4 Polynucleotide Kinase (New England Biolabs) in 50 μL of 1x Kinase Buffer (the 10x stock is supplied by New England Biolabs). Incubate the samples at 37°C for 30 min.
2. Combine the two solutions and add 5 μL of 0.5 M EDTA (pH 8.0; Gibco BRL). Deproteinize the solution by adding 100 μL TE-saturated phenol/chloroform (Gibco BRL), vortexing the mixture, touch-spinning

the tube, and transferring the aqueous phase to a new tube. Repeat the deproteinization three more times.

3. Remove excess ATP on a NICK column (Pharmacia LKB) using the manufacturer's protocol and water as the eluent.
4. Ethanol-precipitate the DNA by adding 50 µL of NaOAc/$MgCl_2$ (3 M and 0.1 M, repectively) at pH 5.2, and 1 mL of cold ethanol.
5. Pellet the precipitate by microfugation at 4°C for 30 min, and remove the supernatant.
6. Dissolve the pellet in a solution containing 50 pmol 3'pDMSalI primer, 50 nmol each of all four dNTPs (Pharmacia Ultrapure grade, Lithium salts), 10 µL 10x Vent Buffer (New England Biolabs), and enough H_2O to bring the final volume to 70 µL.
7. Transfer the solution to a fresh 0.6-mL microfuge tube and add an AmpliWax PCR Gem (Perkin Elmer-Cetus).
8. Heat the sample to 80°C for 5 min, and then cool to 20°C in a Thermal Cycler (Perkin Elmer-Cetus).
9. After the wax has hardened (~ 2 min at 20°C), add 30 µL of a solution containing 4 units Vent DNA Polymerase (New England Biolabs), 10 ng of the plasmid pDMAG10, 1 µL 100x acetylated-BSA (New England Biolabs), and H_2O to the reaction tube on top of the wax layer.
10. Perform PCR amplification over 30 cycles, with each cycle consisting of duplex denaturation at 94°C for 1 min, primer annealing at 54°C for 1 min, and primer extension at 72°C for 1 min. Allow an extension time of 10 min during the last cycle to ensure that all single stranded DNA had been copied or reannealed to form duplex.

11. Using a micropipettor, remove the reaction solution from under the wax pellet. Deproteinize the solution by extracting twice with equal volumes of phenol.
12. Desalt on a NICK column and ethanol-precipitate as above, with the exception that 3 M NaOAc (pH 5.2) is used in place of NaOAc/$MgCl_2$.
13. Digest the PCR products with 20 units *Eco NI* restriction endonuclease (New England Biolabs) in 40 µL at 37°C for 30 min.
14. Purify the labeled product on a 6% polyacrylamide (19:1 acrylamide:bis-acrylamide) gel.
15. Visualize the labeled DNA by autoradiography and excise the gel band corresponding to the labeled 339-bp fragment using a razor blade.
16. Transfer the gel slice to a piece of glacine weighing paper, crush the slice between your fingers, and transfer equal portions of the gelatinous material to two 1.5-mL screw-top tubes.
17. Add 1.0 mL Elution Buffer (25 mM Tris·HCl, 250 mM NaCl, 5 mM EDTA, 0.1% SDS, pH 8.0) to each tube and vortex the mixtures thoroughly. Allow the crushed gel slice to soak in Elution Buffer for at least 15 h at 37°C.
18. Filter the suspensions through Millipore UltraFree DuraPore filter units. The suspension can be quantitatively transferred by inverting a filter unit and placing it over the 1.5-mL tube. Tape the filter unit onto the tube, and filter the sample in a clinical centrifuge.
19. Transfer each filtrate sample to a new 2.0-mL tube and precipitate the DNA using 700 µL isopropanol with no added salt. Pellet the precipitate by ultracentrifugation for 30 min at 4°C. Remove and discard the supernatant.

20. Dissolve the pellet in 100 μL of 0.5 mM EDTA in 5 mM Tris, pH 8.0 (0.5x TE Buffer). Deproteinize the sample as in step 2.2.
21. Desalt the sample on a NICK column.
22. Ethanol-precipitate using 50 μL 2 M NaCl in 0.1 M Tris-chloride (pH 8.0) and 1 mL ethanol.
23. Dissolve pellet in 100 μL 0.5x TE.
24. Desalt on NICK column using 5 mM Tris, pH 8.0 as eluent.
25. Estimate the specific radioactivity of the DNA by measuring the Cerenkov radioactivity with a Beckman LS2801 liquid scintillation counter. If necessary, adjust the solution activity to ~ 30,000 cpm per μL. Store the solution at 4°C.

Affinity Cleavage Titration Assay

3. *General Stock Solutions.*

In order to maintain identical solution conditions between reactions within a single titration experiment, we employ stock solutions. Prior to the preparation of a series of equilibrium mixtures of ^{32}P-DNA and oligonucleotide-EDTA·Fe, several stock solutions containing common components, such as buffer, salts, and carrier DNA, are prepared. A pH near 7.0 can be maintained during equilibration and reaction with a Bis-tris [bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane] buffer system. Bis-tris (Fluka Biochemika MicroSelect), sodium chloride (Fluka BioChemika MicroSelect), and spermine tetrahydrochloride (Fluka BioChemika MicroSelect) were used as obtained from commercial suppliers. In order to obtain a pH 7.0-buffered solution of the desired salt composition, two different stock solutions containing 5x concentrations of the desired salts and 5x concentration of *either* Bis-tris free base or its conjugate

acid were prepared. The solution containing the conjugate acid of Bis-tris was prepared identically to that containing the free base except that a few drops of 1 M HCl were added to pH ≈ 4 (determined by spotting pH paper). The stock solutions were mixed in the appropriate proportions at room temperature to give pH 7.0. The pH of the equilibration solutions was measured as follows: a 1 mL solution containing all buffer and salt components (except radiolabeled DNA) at concentrations identical to those used in the reactions was prepared and equilibrated, and the pH of the buffer-salt solution was recorded using a digital pH/millivolt meter (Orion Research; model no. 611) and a ROSS semimicro combination pH electrode (Orion Research; model no. 81-15).

Sonicated, deproteinized calf thymus DNA (Pharmacia) was dissolved in unbuffered MilliQ water to a final concentration of 2.0 mM in base pairs (20x final concentration). All solutions were stored at 4°C prior to use.

4. Preparation of Oligonucleotide-EDTA·Fe Serial Dilution.

1. Dissolve 8 nmol dry oligonucleotide-EDTA in 50 μL MilliQ water and heat the solution to 55°C for 15 min to denature and disrupt aggregates.
2. After the solution cools to room temperature, add 50 μL of 320 μM $Fe(NH)_2(SO_4)_2$ to give solution **1**, containing 80 μM oligonucleotide-EDTA·Fe plus 80 μM free Fe(II).
3. Prepare a serial dilution of **1** with nearly constant spacing on log scale according to the following table:

Solution	[Oligo]	µL **1**	µL H_2O	Solution	[Oligo]	µL **5**	µL H_2O
2	40	10	10	**6**	4	10	10
3	20	10	30	**7**	2	10	30
4	10	10	70	**8**	1	10	70
5	8	10	90	**9**	0.8	10	90

Solution	[Oligo]	µL **9**	µL H_2O	Solution	[Oligo]	µL **13**	µL H_2O
10	0.4	10	10	**14**	0.04	10	10
11	0.2	10	30	**15**	0.02	10	30
12	0.1	10	70	**16**	0.01	10	70
13	0.08	10	90	**17**	0.008	10	90

5. Preparation of the Equilibrium Binding Mixtures.

The goal of the titration experiment is to prepare binding reactions over a range of oligonucleotide-EDTA·Fe concentrations that produce as little as 1% fractional saturation of the DNA binding site to as much as 99%. This requires a four order of magnitude range in the concentration. In a typical affinity cleaving titration using the serial dilution prepared above, 17 data reactions (containing oligonucleotide-EDTA·Fe) and 1 "no-oligonucleotide" control reaction should be prepared.

1. Prepare a stock solution containing labeled target DNA in buffer by first adding 152 µL of the 5x solution of buffer and salts (50 mM Bis-tris, 500 mM NaCl, and 5 mM spermine·4HCl) to 408.5 µL H_2O in a 1.5-mL tube. Next, add 38 µL of the calf thymus DNA solution ([CT-DNA] = 2.0 mM-bp). Vortex the solution thoroughly and touch-spin the solution down. Finally, add approximately 285,000 cpm of ^{32}P-DNA (~ 9.5

μL) to a final volume of 608 μL. Vortex the solution thoroughly and touch-spin the solution down. Allow this stock solution to equilibrate for 15 – 20 min at room temperature.

2. Aliquot 32 μL of the stock solution into each of 18 siliconized 0.6-mL microfuge tubes.
3. Add 4 μL of the appropriate oligonucleotide-EDTA·Fe solution to each reaction tube. Vortex each tube gently to mix the solutions and touch-spin centrifuge to collect all of the solutions to the bottoms of the microfuge tubes.
4. Allow the mixtures of oligonucleotide-EDTA·Fe and ^{32}P-DNA to equilibrate for at least 24 h at room temperature.

6. *Affinity Cleaving Reactions.*

1. In order to insure that iron is loaded into thymidine-EDTA, add 2 μL of 0.4 mM $Fe(NH_4)(SO_4)_2$ to each tube 30 min prior to starting the cleavage reaction. Vortex the tubes gently and touch-spin. Allow the equilibration to continue for 30 min at room temperature.
2. Initiate the cleavage reactions by adding 2 μL of a freshly prepared 10 mM aqueous DTT solution to each tube. Vortex gently, touch-spin, and allow the reactions to proceed for 4 h at room temperature. Final reaction conditions in 40 μL total volume are 10 mM Bis-tris buffer at pH 7.0, 100 mM NaCl, 1 mM spermine, 20 μM Fe(III), 0.1 mM-bp calf thymus DNA, 1 mM DTT, and approximately 15,000 cpm labeled duplex. The specific activity of the DNA may vary slightly from experiment to experiment, but should be the same for each reaction within a

given titration. The amount of ^{32}P-DNA used must result in a final target site concentration of less than 50 pM.

3. Quench the cleavage reactions by precipitating the DNA: add 4 μL of a solution containing 3 M NaOAc and 0.1 M $MgCl_2$ at pH 5.2, 2 μL of 2 mg/mL glycogen (Boehringer-Mannheim), and 115 μL of cold ethanol. Incubate the mixtures at -20°C or below for 0.5 – 4 h to precipitate the DNA. We have found that the presence of Mg^{2+} and glycogen, in conjunction with the cold incubation, are important for reproducible, quantitative precipitations.
4. Microfuge the samples for 30 min at 4°C to pellet the DNA.
5. Carefully remove all of the supernatant with a micropipettor. Hold the bottoms of the tubes to a Geiger counter to check that the DNA pellet remains.
6. Immediately add 20 μL of H_2O warmed to 55°C to each pellet. Vortex to dissolve the DNA and touch-spin to recollect the samples in the bottoms of the tubes. Samples may be used immediately or stored at -20°C for up to 48 h.
7. Concentrate the samples to dryness on a SpeedVac. We have found that dissolving the DNA in a relatively large volume of H_2O and drying the samples helps achieve quantitative resuspension in formamide loading buffer.
8. Add 5 μL formamide-TBE loading buffer (1x TBE in 80% aqueous formamide). Flick the tubes manually to distribute the buffer and dissolve the DNA, and then microfuge briefly. Repeat this process a second time. The samples can be electrophoresed immediately or stored for no more than 24 h.[46]

9. Determine the Cerenkov radioactivity for each sample by liquid scintillation counting, and dilute to 2500 cpm/μL with more formamide-TBE loading buffer. It is essential that the samples to be electrophoresed have nearly identical radioactivities.

7. *Denaturing Gel Electrophoresis.*

1. Prepare a 34 cm x 42 cm denaturing 8% polyacrylamide (29:1 acrylamide:bis-acrylamide) sequencing gel.
2. Pre-electrophorese the gel in 1x TBE until the surface of the front gel plate becomes warm. This usually takes around 30 min at 2000 V (~ 40 V/cm).
3. Denature the DNA in formamide-TBE buffer by heating the microfuge tubes in a heat block at 85°C for 10 min. Higher temperatures can result in sample degradation.[46] *Immediately* plunge the sample tubes into an ice bath.
4. Load 5 μL of each sample onto the gel.
5. Electrophorese the samples at 30-35 V/cm until the marker dyes have migrated the appropriate distance in order to visualize the DNA region of interest. Separate the plates and transfer the gel onto Whatman 3MM filter paper. Cover the gel surface with a single layer of plastic wrap and dry the gel using a vacuum slab dryer at 80°C for 40 – 60 min.
6. Determine residual radioactivity of each sample tube so that the radioactivity loaded onto each gel lane can be calculated.

Data Reduction and Analysis of Titration Binding Isotherms

8. Preparation of Storage Phosphor Autoradiograms.

Although photographic films can be used for autoradiography, the use of phosphor storage technology provides an increased sensitivity to ^{32}P decay and a linear response range covering *five orders of magnitude.* These improvements are crucial for affinity cleavage titration where the signal strengths at saturation and no binding vary more dramatically than in a footprinting experiment. Moreover, storage phosphor screens have an advantage in convenience because the exposure time does not have to be carefully controlled to produce a useful autoradiogram.

1. Completely erase a photostimulable storage phosphor imaging plate[76] (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) just prior to use.
2. Press a phosphor storage screen flat against a dried gel sample and expose in the dark at room temperature for 12 – 24 h. A brick can be used to maintain thorough contact between the gel and the screen.
3. Remove the screen from the gel in the dark and scan the gel using a Molecular Dynamics 400S PhosphorImager.

Densitometric Analysis of Storage Phosphor Autoradiograms.

1. Display the image of the gel on a computer graphics screen using 1x or 2x scaling.
2. Define a rectangular boundary that encloses the 5 – 10 most intense cleavage bands at the target site (the exact number of bands depends on the oligonucleotide, the position of thymidine-EDTA, the sequence

of the DNA duplex, and the labeled strand). This group of bands will be called the "site block." Draw the upper and lower sides of the rectangle at the minimum intensity levels between bands. Draw the left and right sides about 40% of the distance between the bands of the lane in question and the neighboring lanes.

3. Repeat operation 2 for each lane of the gel by copying the drawn rectangle onto each lane of the gel. Each block must encompass the same bands in each lane.
4. Define rectangular blocks in regions of the lanes where there is no specific component of cleavage. These reference blocks are generally chosen about 10 – 20 bands above the distal end (relative to thymidine-EDTA) of the target *binding* sequence. The area of a rectangle enclosing a reference block is generally 50 – 100% of the area of a site block rectangle. As above, each block must encompass the same bands in each lane.
5. Use the ImageQuant software (Molecular Dynamics version 3.0 running on an AST Premium 386/33 computer) to perform volume integrations of the blocks within the rectangles. This operation sums the intensity values at each pixel within the rectangular boundaries. The volumes can be transferred to a spreadsheet program for further reduction and analysis.

10. Construction of Equilibrium Titration Isotherms.

The following transformations are automated using a Microsoft Excel spreadsheet.

1. Copy the site and reference block volumes into the spreadsheet.
2. Enter the radioactivity of each sample before (C_{pre}) and after loading (C_{post}). Calculate the radioactivity in each lane by subtraction ($C_{load} = C_{post} - C_{pre}$) and compute a mean value ($\langle C_{load} \rangle$). Data from experiments for which the relative standard deviation ($S_C/\langle C_{load} \rangle$) is > 0.15 should be discarded.
3. Divide each site block volume and reference block volume by the product of $\langle C_{load} \rangle$ times the autoradiographic exposure time to obtain I_{tot} and I_{ref} values, respectively, for each oligonucleotide-EDTA·Fe concentration.
4. Calculate λ (eq 36) by averaging the ratios I_{tot}/I_{ref} ratios near $\theta = 0$. Generally, this is accomplished using the no oligo control data and the data for the lowest 1 – 3 oligonucleotide concentrations.
5. Calculate the site-specific cleavage intensity (I_{site}) for each oligonucleotide-EDTA·Fe concentration using eq 36.
6. Transfer the set of (O_{tot}, I_{site}) data points to KaleidaGraph (Synergy Software version 3.0.1 running on a Macintosh IIfx computer) and plot I_{site} versus O_{tot}, where the abscissa is scaled logarithmically.
7. Fit the following theoretical binding curve to the experimental data:

$$I_{fit} = I_{sat} \cdot \frac{K_T O_{tot}}{1 + K_T O_{tot}} \tag{38}$$

The difference between I_{fit} and I_{site} for all data points is minimized by adjusting the apparent maximum cleavage (I_{sat}) and K_T using the Levenberg-Marquardt algorithm implemented in KaleidaGraph's General Curve-Fit program. The titration isotherm fits should be performed

without weighting the data points. Include all data points in the fitting procedure unless one of the following conditions is met: (i) visual inspection of the computer image from a storage phosphor screen revealed a flaw at either the site or reference blocks; (ii) the I_{site} value for a single lane was greater than two standard errors away from both values from the neighboring lanes; or (iii) the I_{site} values at high oligonucleotide concentrations decreased because the DTT was depleted. Experiments for which fewer than 80% of the data lanes are usable should not be used further. Typically, three or four out of five gels will result in usable data, each with 14 or 15 acceptable data lanes out of 17 total. The goodness of fit of the binding curve to the data points is judged by the reduced chi-squared criterion, and fits were judged acceptable for $\chi_v^2 \leq 1.5$. (The standard error in each I_{site} value was calculated using eq 39. See step 13, below.) Nonlinear correlation coefficients for acceptable fits are ≥ 0.95.

11. Repeat Titrations and Data Reporting.

Experiments using a particular set of conditions were performed using different serial dilutions of oligonucleotide prepared from at least two different aliquots of the original solution (step 4.2), at least two different preparations of 5'-labeled duplex DNA, and at least two uniquely prepared buffer solutions. All reported K_T values are the means of three to five experimental observations plus or minus the standard error of the mean. All measured K_T values are included in determining mean values unless one of the following conditions was met: (i) the relative standard deviation in loading was > 15% (step 10.2); (ii) fewer than 80% of the data lanes were usable (step 10.7); (iii) the reduced χ^2 criterion for judging

the fit of the titration curve to the data points was > 1.5 (step 10.7); (iv) the measured K_T value failed the Q-test.[81]

12. Graphical Presentation.

For graphical representation only, I_{site} values are normalized to the range representing apparent fractional occupancy of the duplex binding site (θ_{app}) by dividing I_{site} by I_{sat}. Values of $I_{site} \cdot I_{sat}^{-1}$ from repeat experiments under the same conditions are averaged and plotted along with the isotherms constructed using mean K_T values and eq 37 with $I_{sat} \equiv 1$.

13. Minimal Error Analysis.

For affinity cleavage titrations, the error in the observed cleavage intensity signal is proportional to the uncertainty in the radioactivity of the DNA sample loaded onto the lane in question. Use of the Gauss's Law for propagation of errors (eq 1), allows the following equation describing the uncertainty in I_{site} to be derived:

$$\varepsilon_{I_{site}} = \rho\sqrt{I_{tot}^2 + 3(\lambda I_{ref})^2} \qquad (39)$$

where ρ is the fractional error in the loading ($\rho = S_C/\langle C_{load}\rangle$; step 10.2). The general form of this relationship was confirmed by comparison with the results of 28 independent titration experiments under identical conditions (see Figures 1.7A and 1.10 and Table 1.8 for details). Specifically, the standard deviation in each $I_{site} \cdot I_{sat}^{-1}$ value was plotted as function of the oligonucleotide-EDTA concentration (Figure 1.5). Comparison of the trend in the standard deviations with the binding isotherm for the oligonucleotide in Figure 1.5 demonstrates that the error

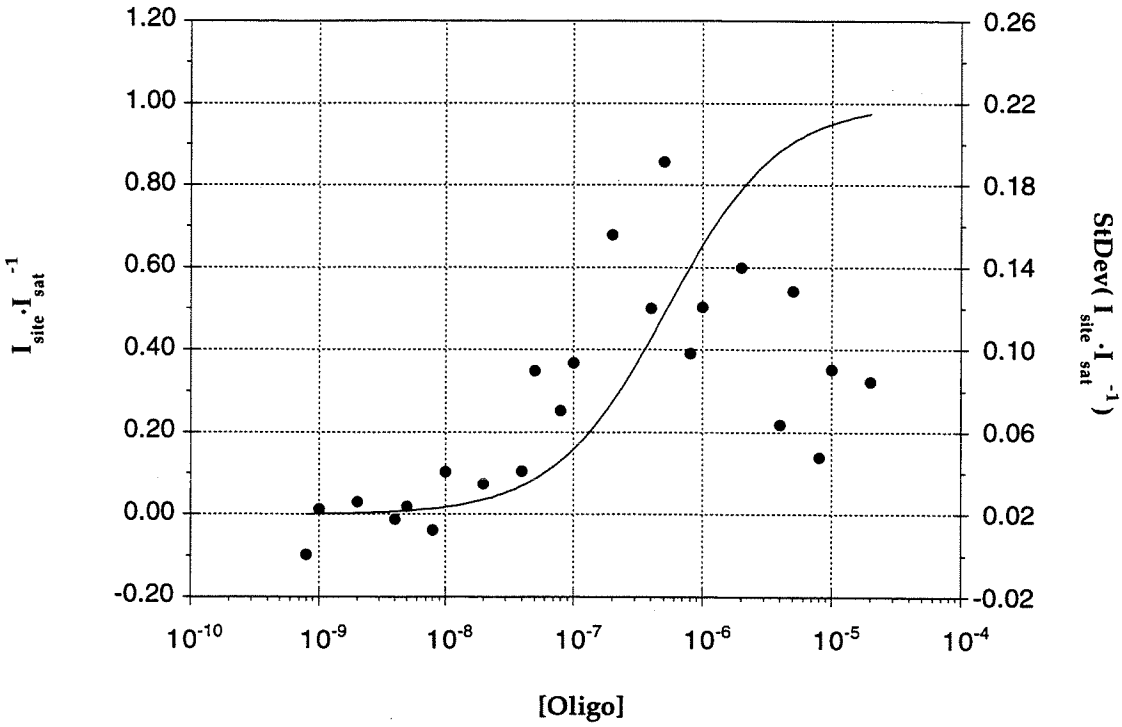


Figure 1.5. Plot of the standard deviation in each mean $I_{site} \cdot I_{sat}^{-1}$ value from 28 independent titrations as a function of oligonucleotide-EDTA concentration. The symbols (•) indicate the value of each standard deviation. The curve indicates the binding isotherm fit to the mean $I_{site} \cdot I_{sat}^{-1}$ data points (not shown).

in each data point increases with increasing concentration and then level out (or decrease slightly) at a relative error of ~ 10%.

There are two main sources of error in the oligonucleotide concentrations, one being the initial uncertainty from employing the additive nucleoside extinction coefficients for calculating the molar extinction coefficient, and the second from pipetting errors during the serial dilution. By using relatively large volumes (*i.e.*, ≥ 10 μL) and well-calibrated pipettors, the error resulting from the latter source can be minimized.

Fits of data points using statistical weighting and the described uncertainties (eq 39) results in standard errors in the fitting parameter K_T (σ_K) that are typically 20-30% of the magnitude of K_T.

Part III. Analysis of Uncertainty in Association Constants Measured by Quantitative Affinity Cleavage Titrations

A central issue in the development and further application of the affinity cleavage titration method is the need for an estimate of the uncertainty in observed equilibrium association constants. There are several approaches that have been traditionally employed in error analyses, but which are difficult to apply here. Common statistical estimates of variability, such as the standard deviation, can be used for rigorous statistical testing to determine whether two observed association constants are significantly different; however, such tests require the assumptions that the error distribution is normal, and that the sources of error are independent and random. Moreover, test results are reliable only when a relatively large sample size is used.[82] Because we have little knowledge concerning the validity of the required assumptions and beacause we have typically used three to five experimental observations to determine mean association constants, we have not employed formal statistical testing.

A second approach involves the estimation of the uncertainty introduced by each manipulation performed in a quantitative titration, and the subsequent propagation of the uncertainties to the observed equilibrium constant. Groups of variables could be isolated from each other by dividing reaction tubes into several equal volumes and then measuring association constants for each set. The standard deviation of the sample population would then reflect the random error introduced in the manipulations performed independently after division of reaction samples. Inspection of the protocol for performing a quantitiative titration suggests that such a task would be prohibitively time-consuming. While we have not systematically investigated this approach, we have divided titration ex-

periments into two new sets during the experiment in an effort to identify problematic steps. Association constants determined for such split experiments typically differ from one another by a factor of two or less (data not shown).

A third method of error analysis is often employed when the value of interest is obtained by curve-fitting. By estimating the uncertainty in each value of the independent and dependent variables (*e.g.*, O_{tot} and I_{site}, respectively), the uncertainty in the fitted parameter of interest can be calculated. Usually, this process involves making several measurements of the dependent variable, fitting the mean values, and using the standard deviation as an estimate of the uncertainty in that variable. This approach is difficult to apply in the case of affinity cleavage titration for a number of reasons. First, the calculation of reliable confidence limits from nonlinear functions is not straightforward.[83] In addition, we are not able to reasonably estimate the uncertainty in I_{site} *a priori*, and, because the number of experiments is minimal, the use of the standard deviation of I_{site} values will be of limited utility. Most importantly, however, the affinity cleaving technique does not lend itself to the analysis of mean I_{site} values. From eq 37 it is clear that the range of observed cleavage is not strictly bounded, but depends on I_{sat}. The cleavage observed at saturation binding depends, in turn, on the time of the reaction, the concentration of dithiothreitol, the DNA precipitation efficiency, the specific activity of the radiolabel, etc., so that small differences between experiments can afford relatively larger differences in I_{sat}. Hence, averaging data from different experiments will not generally result in reliable mean data values. It should be possible to overcome this limitation by measuring relative cleavage yields for tightly regulated reactions.

Although a formal error analysis based on one of the three traditional approaches cannot be conducted, we have performed experiments to identify po-

tential sources of error and to obtain a reasonable estimate of the uncertainty in apparent association constants. Eight members of the laboratory performed a series of identical quantitative affinity cleavage titrations, and the differences between association constants measured by each participant were analyzed in an effort to isolate potential sources of uncertainty. The analysis indicated that several experimental sources of error can contribute to a relatively large uncertainty in a measured association constant. However, by randomizing the contributing variables, the sampling bias of the observed mean association constant can be reduced to afford improved accuracy. The results of the "group experiment" also allowed rough estimates of the uncertainty to be made. These estimates were then refined by analyzing 28 independent quantitative titrations of the same oligonucleotide-EDTA under identical solution conditions.

Sources of Experimental Uncertainty

The Group Experiment. In order to identify and control sources of uncertainty in affinity cleavage titration experiments as well as to address general questions concerning the reproducibility, reliability, and inherent uncertainty in the technique, eight group members agreed to participate in a set of two control experiments. The researchers were Dr. Andrea B. Staubli, Dr. Kenneth D. Turnbull, Hogyu Han, George C. Best, E. Scott Priestley, Natalia Colocci, Michelle E. Parks, and Scott F. Singleton. For the first experiment (Experiment A), each researcher measured the association constant for binding of the 15mer, 5'-T*TTTTCTCTCTCTCT-3', in 100 mM NaCl and 1 mM spermine·4HCl, pH 7.0-buffered at room temperature by 50 mM Tris-acetate, using materials provided for him or her (Figure 1.7A). Each researcher was provided target [^{32}P]DNA, 5x buffer, 20x calf thymus carrier DNA, and a 10-nmol aliquot of the oligonu-

cleotide-EDTA. In the second experiment (Experiment B), each researcher measured the association constant for the 15mer, 5'-T*TTTTm^{5}CTm^{5}CTm^{5}CTm^{5}CT-m^{5}CT-3' (where m^{5}C is 5-methyl-2'-deoxycytidine), using his or her own reagents (Figure 1.7B).

Results and Discussion. Mean association constants (± 1 SEM) reported by each researcher from the two experiments, in addition to ratios of the association constants for the two oligonucleotides, are presented in Tables 1.2 (Experiment A), 1.3 (Experiment B), and 1.4 [K_T(B)/K_T(A)]. For both sets of experiments, there is some scatter around the mean values. This variability is contributed to by each of the following six sets of experimental parameters: (i) the aliquot of oligonucleotide used, (ii) the details of the experimental technique, (iii) the researcher, (iv) the preparation method of the oligonucleotide, (v) the preparation of the labeled target DNA, and (vi) the preparation of the buffer. In addition, there may be sources of error that are unidentified. The approximate contribution of each paramter to the observed range of K_T values can be determined by analysis of the experimental results.

In Experiment A, all of the measured K_T values were tightly clustered around the mean value (the spread is 2.6-fold over the whole range). This relatively small scatter, which can be attributed to differences in the researchers, their techniques, and the aliquot of oligonucleotide they received, indicates that the technique yields results that are reproducible from researcher to researcher when the reagents are equivalent. The fact that the group mean is 1.7-fold lower than the published value[6] suggests that reagent preparation may be important since the latter value was obtained from several different preparations of both [^{32}P]DNA and buffer solutions.

A. Common Materials

Oligonucleotide:

5'-*TTTTTCTCTCTCTCT-3'

Target Duplex:

5'-ATAT**AAAAAGAGAGAGAGA**TGGA-3'
3'-TATA**TTTTTCTCTCTCTCT**ACCT-5'

(Site within 230-bp EcoRI/PvuII restriction fragment of pGCBHM15)

Conditions:

50	mM Tris-acetate, pH 7.0
100	mM NaCl
1	mM spermine·4HCl
0.1	mM-bp calf thymus DNA

B. Individually Prepared Materials

Oligonucleotide:

5'-*TTTTT^{m}CT^{m}CT^{m}CT^{m}CT^{m}CT-3'

Target Duplex:

5'-ATAT**AAAAAGAGAGAGAGA**TGGA-3'
3'-TATA**TTTTTCTCTCTCTCT**ACCT-5'

(Site within 230-bp EcoRI/PvuII restriction fragment of pGCBHM15)

Conditions:

50	mM Tris-acetate, pH 7.0
100	mM NaCl
1	mM spermine·4HCl
0.1	mM-bp calf thymus DNA

Figure 1.7A. Experimental design for the two quantitative affinity cleavage titration group control experiments.

Table 1.1. Mean association constants (± 1 SEM) from Experiment A.

Researcher	Association Constant	Rank	Protocol
Chapter Two[a]	**3.7 (± 1.1) x 10^6**		
1	3.4 x 10^6	1	A
2	3.37 (± 0.46) x 10^6	2	A
3	2.8 x 10^6	3	D
$\langle K_T \rangle \pm \lambda_{95}$[b]	**2.2 (± 0.7) x 10^6**		
4	1.76 (± 0.20) x 10^6	4	A
5	1.74 (± 0.45) x 10^6	5	C
6	1.58 (± 0.30) x 10^6	6	A
7	1.57 (± 0.24) x 10^6	7	B
8	1.32 (± 0.34) x 10^6	8	B

[a] This value has been published, and its measurement is described in Chapter Two. [b] The global mean value ($\langle K_T \rangle$) and the 95% confidence limit (λ_{95}) for the eight tabulated association constants.

Table 1.2. Mean association constants (±1 SEM) from Experiment B.

Researcher	Association Constant	Rank	Protocol
1	3.0 x 10^7	1	A
4	2.25 (± 0.68) x 10^7	2	A
7	2.14 (± 0.50) x 10^7	3	B
$\langle K_T \rangle \pm \lambda_{95}$[a]	**1.3 (± 0.9) x 10^7**		
6	8.32 (± 1.10) x 10^6	4	A
Chapter Three[b]	**7.1 (± 2.8) x 10^6**		
2	5.94 (± 0.88) x 10^6	5	A
5	5.45 (± 1.66) x 10^6	6	C
8	5.20 (± 1.40) x 10^6	7	B
3	4.0 x 10^6	8	D

[a] The global mean value ($\langle K_T \rangle$) and the 95% confidence limit (λ_{95}) for the eight tabulated association constants. [b] This value has been published, and its measurement is described in Chapter Three.

Table 1.3. Ratios of observed association constants: $K_T(B)/K_T(A)$.

Researcher	Ratio	Protocol
7	13	B
4	13	A
1	8.8	A
Mean ± λ_{95}	**6.3 ± 4.7**	
6	5.3	A
8	3.9	B
5	3.1	C
Chapter Three	**1.9**	
2	1.8	A
3	1.4	D

[a] The global mean value and the 95% confidence limit (λ_{95}) for the eight tabulated association constants. [b] This value has been published, and its measurement is described in Chapter Three.

In Experiment B, the measured values were more highly scattered than those from Experiment A (7.5-fold spread over the entire range). If this scatter is attributed to all six sets of experimental variables and Experiment A suggests that variables (i) - (iii) account for a factor of ~2.6 in this spread (see Table 1.4), then differences in reagent preparation account for a factor of 2.9 (7.5/2.6). The fact that the group mean is 1.8-fold higher than the published value[7] may reflect that the latter value was obtained from a smaller range of materials preparations, so that the sampling bias was likely larger. In addition, because the observed association constants span nearly an order of magnitude in size, the group mean is weighted towards a larger magnitude by the higher values (*i.e.*, those near 2×10^7 M^{-1}).

In addition to effects from the researcher and the protocol, the range of values reported in Experiment A results from different oligonucleotide aliquots and differences in the initial resuspension of the dried samples. While the uncertainty resulting from these variables has not been measured, the theoretical uncertainty can be calculated using reasonable estimates of the possible errors in each of the contributing variables and eq 1 for propagating the errors. Following the protocol in Part II for determining the oligonucleotide-EDTA concentrations after synthesis and purification (step 1.10), distributing aliquots (step 1.11), and resuspending the dry oligonucleotide for serial dilution (step 4.1), we estimate that the original solution is prepared in 5.0 ± 0.1 mL, that 40 ± 2 μL of this solution are diluted with 560 ± 10 μL H_2O for spectroscopic analysis, that the absorbance is 0.5 ± 0.02 OD, that the extinction coefficient has a relative uncertainty of 10%, that aliquots of 40 ± 1 μL of the original solution are distributed for drying and storage, that the resuspension volume is 100 ± 2 μL, and that resuspension yields are 95 ± 5%. From these estimates and eq 1, the relative error in the

concentration of the first oligonucleotide-EDTA·Fe solution was determined to be ± 8%. Hence, this error contributes a factor of ~ 1.2 (1.08/0.92) to the observed range (Table 1.4).

Table 1.4. Partitioning of relative error values among the experimental sources.

Experiment	Source of Error	Relative Error	Uncertainty Multiplier	Overall Multiplier
A	i	8%	1.2	2.6
	ii, iii[b]	5%	1.1	
	unknown	33%	2.0	
B	iv	5%	1.1	2.9
	v, vi[c]	45%	2.6	

[a] The six sets of experimental parameters are (i) the aliquot of oligonucleotide used, (ii) the details of the experimental technique, (iii) the researcher, (iv) the oligonucleotide preparation method, (v) the preparation of the labeled target DNA, and (vi) the preparation of the buffer. [b] The experimental data do not allow distinction of these parameters. [c] The experiments do not allow distinction between buffer and target DNA preparations; however, the assignment of the other sources of variability yields a factor of 2.6 for these two parameters combined. This corresponds to an uncertainty of approximately ± 45%.

If the exact technical details of the experimental protocol are critical to the measurement, then researchers using the same protocol are expected to report similar K_T values. Discussion among the participants indicated that four differ-

ent protocols were used for performing titration experiments (A-D; Tables 1.2 and 1.3). Protocol A corresponds to that presented earlier in this chapter. Differences between the protocols include oligonucleotide-EDTA dilution schemes, reaction volumes, and precipitation techniques. The association constants measured using each protocol were averaged, and the average values ranked in order of increasing mean value (Figure 1.8). No strong correlation is observed in Figure 1.8, or in the distribution of individually reported values (Tables 1.2, 1.3, and 1.4). These facts suggest that differences in the protocols used are not a source of systematic error.

Beyond the protocol employed, the researcher may contribute subtle and specific factors leading to differences in observed association constants. These factors may also include physical parameters that coincide with the researcher, such as the pipettors used and the laboratory temperature. If the identity of the researcher is important to the measurement, then the relative rank of each researcher's values in the two experiments is expected to be similar. In order to evaluate the effect of researcher identity, the numerical rank for each researcher's Experiment A value (Table 1.1) was plotted versus the same researcher's Experiment B rank (Table 1.2). The result (Figure 1.9) suggests any possible correlation between researcher identity and observed K_T is small. A plot of the two measured K_T values reported by each researcher from the two experiments against each other (Figure 1.9) further corroborates the conclusion that any correlation is small at best. Furthermore, the ratios of the two values (Table 1.3) cover a wide (10-fold) range, which is also consistent with a small correlation between the researcher and the measured value. These facts indicate that systematic variability between researchers is minimal. From Table 1.4, the oligonucleotide aliquot, the protocol used, and the researcher combine to account for a factor of

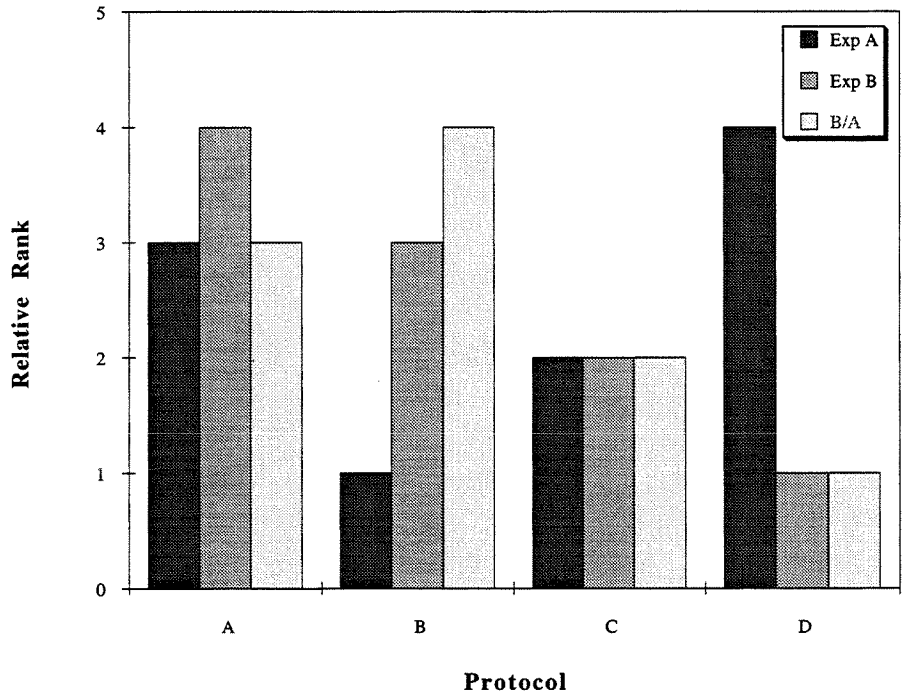


Figure 1.8. Column plot of the rankings of average association constants for each protocol. The average values were ranked from highest (4) to lowest (1) for Experiment A, Experiment B, and the ratio $K_T(B)$ to $K_T(A)$. Protocol A was described in Part II of this chapter. The other protocols contain small variations in many of the manipulations. Each researcher's protocol is listed with his or her measured values in Tables 1.2, 1.3, and 1.4.

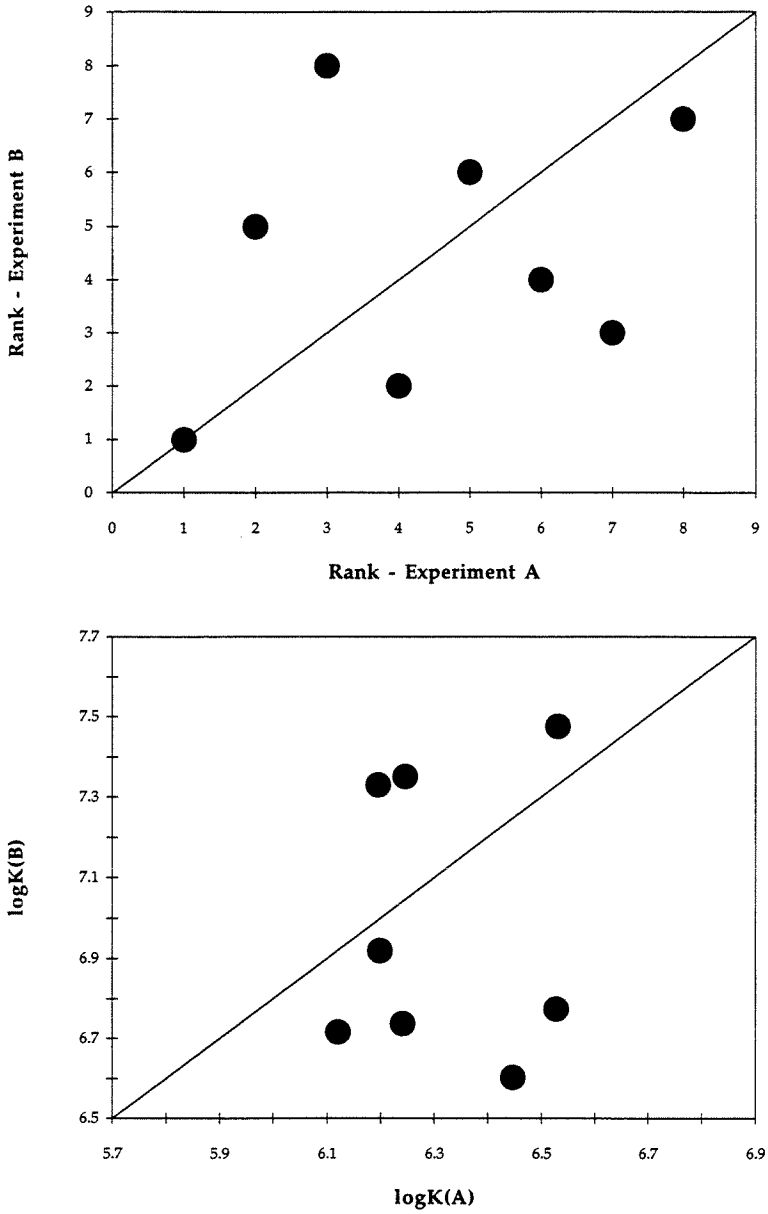


Figure 1.9. Scatter plots of the results from Experiment B versus those from Experiment A for each individual participant. *Upper.* The rank of each researcher's K_T(A) (see Table 1.1) is plotted against the rank of the same researcher's measured K_T(B) (Table 1.2). *Lower.* The log(K_T) values, rather than ranks, are plotted.

2.6 in the range of observed association constants. Since variable (i) alone contributes a factor of 1.2, the latter two parameters yield a maximum factor of 2.1. In reality, the researcher and the detailed protocol may be tightly linked, so that dissecting their individual influences is not reasonable. However, the absence of any observable correlations led us to place an estimate of 5% on the relative systematic error caused by researcher differences. Therefore, the remainder of the observed range must stem from unidentified sources (Table 1.4).

From the results of Experiment B, the effect of method of oligonucleotide-EDTA purification can be analyzed. The two purification methods involved chromatographic fractionation (FPLC) as described in Part II, or polyacrylamide gel electrophoresis (PAGE). The reported association constants were sorted by purification method and are listed in Table 1.5. The mean values for each method are very similar, suggesting that the method of purification is not an important determinant to the overall uncertainty. The small variation resulting from differences in oligonucleotide preparation is estimated to be ± 5% (a factor of ~1.2).

Multiplying together the factors contributed by each of the experimental sources of error discussed above suggests that a factor of ~ 2.6 (7.5/2.9) remains to be accounted for. Apparently, the source of this error is differences in the preparations of [^{32}P]DNA and the reaction solutions, including the stock solution of buffer and salts. The group experiments do not allow these variables to be analyzed separately; however, the size of the uncertainty factor contributed by the sum of the variables emphasizes their importance to the determination of reliable association constants.

Table 1.5. Influence of oligonucleotide purification method on observed association constant.

Method	Sample No.	Researchers	Mean K_T	Method Mean
PAGE	1	2	5.94×10^6	1.42×10^7
	2	3, 6	7.34×10^6	
	3	1	3.6×10^7	
FPLC	4	5, 7, 8	1.26×10^7	1.61×10^7
	5	4	2.68×10^7	

Esimate of the Uncertainty in Observed Association Constants. The experimental basis of the variability observed in the two experiments has been roughly partitioned among the six parameters described above. The relative uncertainties (expressed as percentages) and the corresponding multipliers contributed by the experimental sources of error are compiled in Table 1.4. From these data, the relative uncertainty in a mean association constant can be roughly estimated. Because the range of mean association constants spans a factor of 7.5, a measurement performed according to the protocol described here will typically be within a factor of 3 of the true apparent equilibrium constant [$(7.5)^{1/2} = 2.7$, which is rounded up to yield a conservative estimate]. In order for two measured association constants to be considered different, they should differ by at least a factor of 8 (7.5 rounded up). This factor is derived from the fact that, when two association constants are being compared, we will not know where they lie in the distribution of mean values. Therefore, when the two are different by less than a factor of 8, we cannot be certain that they are not two means from the same range.

Unidentified sources contribute heavily to the uncertainty. If these sources remain unaccounted for, the relative error contributed by these factors represents

the minimum level of uncertainty. Thus, no matter how carefully the experiments are conducted, a mean association constant will never be less than a factor of 1.5 from the true association constant [$(2.0)^{1/2} = 1.4$], and two association constants should never be considered different if the ratio $K_T(1)/K_T(2) \leq 2$. (The limitation for comparing two association constants could be eased by performing a sufficient number of repeat experiments to allow application of a statistical test of significance.)

Control of the Experimental Uncertainty. Controlling the contribution that each source makes to the ranges of association constants observed here will allow researchers to compensate for the variability among experiments to obtain the most reliable (accurate) association constants. In order to obtain mean association constants that are as free as possible from sampling bias, a few experimental measurements can be used to sample the variability inherent in each source of error. For example, the use of a unique sample of oligonucleotide-EDTA for each titration would prevent a large error in a particular sample from unduly influencing the mean association constant. Thus, by broad sampling of the sources of random error, the accuracy of the apparent association constants can be enhanced, albeit at the expense of statistical precision. The following guidelines are directed toward improving the accuracy of reported values:

1. Reported association constants should consist of mean values (± 1 SEM) of at least three independent experiments.

2. At least two different aliquots of oligonucleotide-EDTA should be used to prepare unique serial dilutions for experiments performed to obtain mean association constants.

3. At least two unique samples of target [^{32}P]DNA should be used to obtain each association constant. A unique sample is one that has been gel purified in a separate lane, extracted, and isolated independently from other samples.

4. Experiments used to obtain reported mean values should be performed using at least two uniquely prepared stock solutions of buffer, salts, and carrier DNA.

5. Each experiment performed to measure a single association constant should involve a *different combination* of the unique oligonucleotide-EDTA aliquot, [^{32}P]DNA, and buffer preparations.

If these requirements are fulfilled by conducting n independent titrations, then apparent random errors propagated to the observed association constant will be reduced by a factor of $n^{1/2}$. Random errors reduced in this manner, with $n = 3$, were calculated and are summarized in Table 1.6. From these error levels, we estimate that observed mean association constants will be within a factor of ~ 2.5 [$(4.6)^{1/2} = 2.1$] of the true association constant and that mean association constants which differ by a factor of less than ~ 5 should not be considered different.

Table 1.6. Estimated random errors influencing the association constant after three independent experiments have been performed to reduce sampling bias.[a]

Source of Error	Relative Error	Uncertainty Multiplier	Overall Error	Overall Multiplier
i	5%	1.1	58%	4.6
ii,iii	5%	1.1		
iv	5%	1.1		
v,vi	26%	1.7		
unknown	33%	2.0		

[a] The six sets of experimental parameters are (i) the aliquot of oligonucleotide used, (ii) the details of the experimental technique, (iii) the researcher, (iv) the preparation method of the oligonucleotide, (v) the preparation of the labeled target DNA, and (vi) the preparation of the buffer.

Uncertainty in Mean Association Constants

In order to analyze the uncertainty in a mean association constant measured using affinity cleavage titrations, we have analyzed the results of repeat titrations using the same oligonucleotide-EDTA under identical conditions. Specifically, we performed 28 independent measurements of K_T for the triple-helical complex depicted in Figure 1.7A under the conditions listed in the figure. These titration experiments were performed according to the protocol given in Part II of this chapter, and the guidelines delineated in the preceding section for conducting repeat titrations were adopted. The results of these experiments, namely the fit-

ted K_T parameters, are listed in Table 1.7 in order from the highest to the lowest value. Although from the results of the previous section we understand that only the first digit of each value is significant, we have listed three digits for individual values in the tables to preserve the data. The observed association constants for the series of individual titration experiments covers a range spanning a 7.5-fold change in magnitude, similar to that found for *mean* association constants from Experiment B described above.

For analysis of these data, standard statistical parameters were calculated (Table 1.8). Because of the relatively large sample size considered (n = 28), the mean association constant ($\langle K_T \rangle$) is our best estimate for the true association constant [$K_{true} = (2.4 \pm 0.5) \times 10^6\ M^{-1}$], and the standard deviation (S) represents our best estimate of the variability in affinity cleavage titration measurements (σ). Then, for mean values determined from smaller numbers of experiments, *e.g.*, n = 3, the relative confidence limits on the observed mean value ($\lambda/\langle K_T \rangle$) in each case can be calculated from

$$\lambda = t_{P,n-1} \cdot \frac{\sigma}{\sqrt{n}} \tag{40}$$

where t is the Student's t-value at probability level P for n–1 degrees of freedom.[82] We have performed these calculations for several confidence intervals with n = 3, 4, or 5 (Table 1.9). The uncertainty is dependent on both the number of observations and the desired confidence level. In the preceding section, we had roughly approximated that the uncertainty in the mean K_T

Table 1.7. Twenty-eight apparent equilibrium association constants measured for 5'T*TTTTCTCTCTCTCT-3' (22°C, 100 mM NaCl, 1 mM spermine·4HCl, and pH 7.0) using independent titrations.

Rank	K_T (M^{-1})	Rank	K_T (M^{-1})
1	5.94×10^6	15	2.04×10^6
2	4.81×10^6	16	1.70×10^6
3	4.11×10^6	17	1.69×10^6
4	3.84×10^6	18	1.64×10^6
5	3.49×10^6	19	1.61×10^6
6	3.48×10^6	20	1.58×10^6
7	3.25×10^6	21	1.57×10^6
8	3.02×10^6	22	1.48×10^6
9	2.54×10^6	23	1.42×10^6
10	2.52×10^6	24	1.41×10^6
11	2.31×10^6	25	1.08×10^6
12	2.25×10^6	26	1.01×10^6
13	2.16×10^6	27	9.54×10^5
14	2.10×10^6	28	7.90×10^5

Table 1.8. Results of analyses of 28 independent affinity cleavage titrations.

Parameter[a]	Value[b]	Relative Value[c]
$\langle K_T \rangle$	2.35×10^6	1.00
S	1.24×10^6	0.529
SEM	2.35×10^5	0.100
λ_{95}	4.81×10^5	0.205
K_{fit} [d]	1.76×10^6	0.749
σ_{fit}	3.60×10^5	0.153
$t_{0.05,21} \cdot \sigma_{fit}$ [e]	7.48×10^5	0.318
K_{fit}' [f]	2.52×10^6	1.07
σ_{fit}'	4.42×10^5	0.188
λ_{fit}'	9.20×10^5	0.391

[a] The symbols $\langle K_T \rangle$, S, SEM, λ_{95}, K_{fit}, and σ_{fit} indicate the mean of the association constants reported in Table 1.7, the standard deviation, the standard error of the mean, the 95% confidence limits, the fitted parameter corresponding to the association constant (Figure 1.10A), and the standard error in the parameter. [b] The units on the values are the same as for an association constant, namely $L \cdot mol^{-1}$. [c] The relative value is defined by the ratio of the value to the mean association constant. [d] This parameters and the two error estimates were obtained by the curve fit pictured in Figure 1.10A. [e] The product indicated does not correspond to a confidence interval. [f] This parameter and the two corresponding error estimates were obtained by fitting the double-reciprocal plot (Figure 1.10B).

derived from three experiments should be ~ 5-fold. This uncertainty coincides with a confidence level of 75% – 90% for n = 3 (the uncertainties on these intervals are 2.9-fold and 17-fold, respectively). The magnitudes of these confidence limits emphasizes the difficulty in achieving statistical precision in association constants measured by affinity cleavage titrations.

Table 1.9. Relative confidence limits ($\lambda/\langle K_T \rangle$) calculated for mean association constants derived from three to five titration experiments.

	Confidence Interval		
n	75%	90%	95%
3	0.49	0.89	1.3
4	0.38	0.62	0.84
5	0.31	0.50	0.61

The tabulated values are calculated using eq 40 as described in the text.

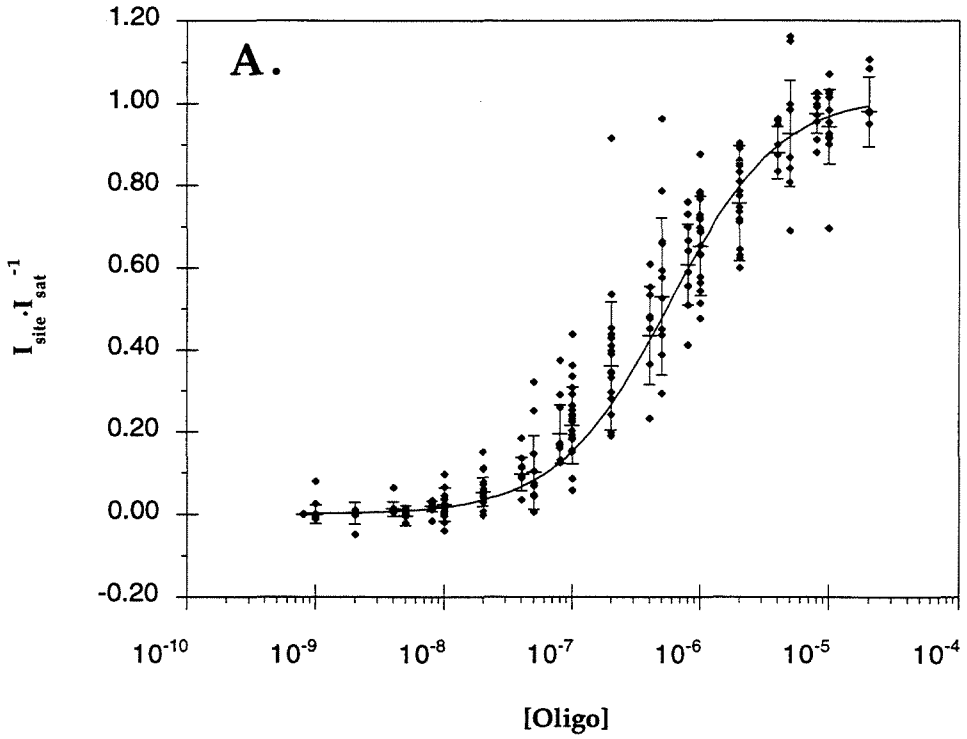
The data from the repeat experiments were used further in a different analytical approach. By dividing the observed I_{site} values by the I_{sat} parameter for each experiment, the I_{site} values are effectively normalized for differences between individual titrations. The $I_{site} \cdot I_{sat}^{-1}$ values for every experiment were averaged, and the resulting mean values and their standard deviations are reported for each oligonucleotide-EDTA concentration in Table 1.10. The (O_{tot}, $I_{site} \cdot I_{sat}^{-1}$) data points from each experiment are plotted in Figure 1.10A. The mean $I_{site} \cdot I_{sat}^{-1}$ values are idicated by crosses (+) and the error bars are derived from the standard deviation of each mean value. The curve indicates the titration binding isotherm that best fit the mean data points using eq 38 with statistical weighting based on the standard deviations. The values of the parameter K_{fit}

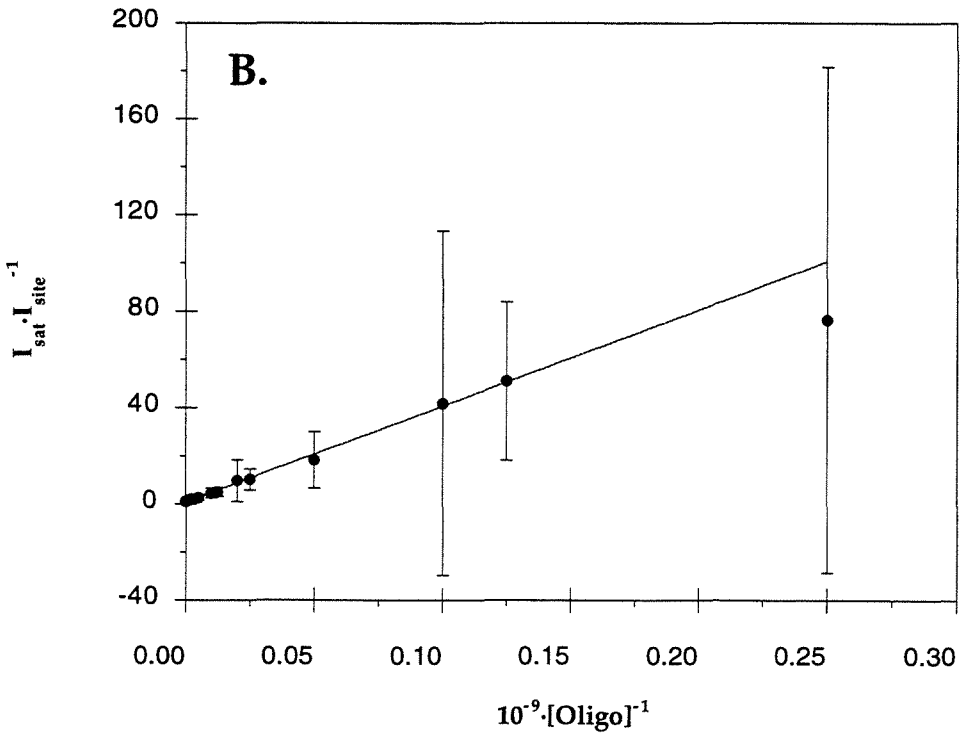
Table 1.10. Average quantitative affinity cleavage titration data compiled from 28 independent experiments.

Conc (μM)	Mean($I_{site} \cdot I_{sat}^{-1}$)	StDev($I_{site} \cdot I_{sat}^{-1}$)
8.0×10^{-10}	-0.001	0.001
1.0×10^{-9}	-0.001	0.023
2.0×10^{-9}	0.002	0.026
4.0×10^{-9}	0.013	0.018
5.0×10^{-9}	-0.003	0.024
8.0×10^{-9}	0.019	0.012
1.0×10^{-8}	0.024	0.041
2.0×10^{-8}	0.054	0.035
4.0×10^{-8}	0.097	0.041
5.0×10^{-8}	0.102	0.090
8.0×10^{-8}	0.195	0.071
1.0×10^{-7}	0.216	0.094
2.0×10^{-7}	0.361	0.156
4.0×10^{-7}	0.435	0.120
5.0×10^{-7}	0.530	0.191
8.0×10^{-7}	0.608	0.098
1.0×10^{-6}	0.653	0.121
2.0×10^{-6}	0.757	0.140
4.0×10^{-6}	0.880	0.064
5.0×10^{-6}	0.926	0.129
8.0×10^{-6}	0.975	0.048
1.0×10^{-5}	0.944	0.091
2.0×10^{-5}	0.980	0.085

See Figure 1.7A and Table 1.7 for experimental details.

Figure 1.10. (A) Semi-logarithmic plot of the $I_{site} \cdot I_{sat}^{-1}$ values from 28 independent affinity cleavage titrations versus the oligonucleotide-EDTA concentration (◆). The mean values at each concentration are also plotted (+) along with error bars representing the standard deviation in each mean value. The titration binding curve that best fits the mean data points is also plotted. (B) Double reciprocal plot of the mean data from Figure 1.10A. $I_{sat} \cdot I_{site}^{-1}$ is plotted as a function of the reciprocal of the oligonucleotide-EDTA concentration (•). Error bars indicate the standard deviations from above transformed onto the double reciprocal coordinate space (see text for details). The line indicates the best linear fit of the data points.


A.
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
$I_{site} \cdot I_{sat}^{-1}$
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
[Oligo]


B.
200
160
120
80
40
0
-40
$I_{sat} \cdot I_{site}^{-1}$
0.00
0.05
0.10
0.15
0.20
0.25
0.30
$10^{-9} \cdot [Oligo]^{-1}$

and the standard parametric error (σ_{fit}) are listed in Table 1.8. The fitted parameter and the mean association constant are in good agreement. Using the relationship

$$\lambda_{fit} = t_{.05,d-2} \cdot \sigma_{fit} \tag{41}$$

where d is the number of data points (23 in this case), we have estimated the 95% confidence limits of the fitted association constant; however, the limits are only approximate because the relationship between the confidence interval and the standard error from nonlinear equations is more complicated than indicated by eq 41, which is rigorously correct only for linear equations.[8,3] We determined more reasonable confidence limits on K_T by transforming the data into double-reciprocal form (O_{tot}^{-1}, $I_{sat} \cdot I_{site}^{-1}$) and curve-fitting using the following transformation of eq 38:[71]

$$\frac{1}{I'_{fit}} = \frac{1}{I'_{sat}K'_{fit}} \cdot \frac{1}{O_{tot}} + \frac{1}{I'_{sat}} \tag{42}$$

The transformed data points are plotted in Figure 1.10B, indicating the linear relationship between $I_{sat} \cdot I_{site}^{-1}$ and O_{tot}^{-1} suggested by eq 42. The transformed data were fitted to a linear equation with statistical weighting applied using transformed standard deviations [$S(I_{sat} \cdot I_{site}^{-1}) = S(I_{site} \cdot I_{sat}^{-1})/I_{site}^2$].[71] Note that because the data points were normalized using I_{sat} from each particular experiment, the y-intercept, $(1/I_{sat}') \approx 1$. The association constant K_{fit}' was obtained from the ratio of the y-intercept to the slope, and is in good agreement with both the mean association constant and the parametric value obtained by nonlinear curve-fitting. Both K_{fit}' and the calculated standard error are listed in Table 1.8.

Eq 41 can be directly applied using the standard error in the linear parameters to calculate λ_{fit}' (Table 1.8). Because the curve-fitting accounted for all of the experimental variability in the data points, the λ_{fit}' value is representative of the experimental error in the observed association constant. In this case, the error is ± 40% of the mean value, indicating that affinity cleavage titration data allow K_T to be determined to within a factor of 1.5 of the true association constant.

Summary

In this part of Chapter One, several estimates of the uncertainty in observed association constants have been presented. In general, following the protocol outlined in Part II and adopting the guidelines presented earlier in Part III, we observe that average association constants have a relatively low level of statistical precision as a result of the attempt to reduce experimental sampling bias and improve the accuracy. Based on the preceding discussion of the uncertainty in a mean association constant, we propose the following "rule of thumb" standards for the purpose of analyzing differences between individual average association constants. When two values differ by a factor of 3 or greater, the association constants are likely different at the 75% – 90% confidence level, depending on the number of indpendent titrations performed (see Table 1.10). If the values being compared are reported by the same researcher, then variables ii, iii, and iv no longer contribute to the uncertainty in the difference. Thus, from Table 1.6, the factor of 3 is reduced to a factor of 2.5 [3/(1.1×1.1)] in this case. When the two values differ by a factor of 5 or more, the association constants are likely different at the 90% – 95% confidence level. In the absence of a large number of repeat experiments and rigorous statistical significance testing, values that differ by a factor of 2 or less can never be taken as truly different with any confidence.

For much of the work that will be described in the following chapters, we were interested in changes in the association constant as a function of changes in solution parameters. In order to estimate the uncertainty in each K_T value used to evaluate the trends, we employed the following empirical relation:

$$\kappa = 2 \cdot \frac{t_{0.1,n-1}}{t_{0.1,\infty}} \cdot (1 + |S - \sigma|) \tag{43}$$

where κ is the uncertainty multiplier for K_T, t was defined above, n is the number of independent titrations performed, S is the standard deviation in $\langle K_T \rangle$, and σ is the standard deviation of the population. The parameter $t_{0.1,\infty}$ is equal to 1.645 and, based on the standard deviation derived from 28 independent measurements of the same K_T, σ is taken to be 0.5 (Table 1.8) The form chosen for eq 43 was based on the following reasoning. First, the κ function should reflect differences between S and σ. Means obtained from sample populations having a large standard deviation are not reliably determined. Because the independent titrations are conducted to minimize sampling bias, a standard deviation smaller than σ is taken to mean that the true range of experimental variables have not been sampled. Thus, the absolute value of (S–σ) is used. The ratio of t-values explicitly includes the influence of the number of different titrations performed, because, as n increases, $t_{0.1,n\text{-}1}$ decreases. The term, $t_{0.1,\infty}$ is the percentage point in the t-distribution corresponding to an infinite number of degrees of freedom. Student's t-values are used because they transform standard deviations into confidence limits (eq 40). The factor of 2 represents the lowest uncertainty level possible for an affinity cleavage titration measurement.

Values of the κ function for n = 1 to 3 are plotted versus the ratio of S to σ in Figure 1.11. When S = σ, the calculated uncertainty is 2.5- to 3.5-fold, depending

on n. When S = 1.5σ or S = 0.5σ, the calculated uncertainty is 4- to 5.5-fold. Because eq 43 reproduces the level of uncertainty obtained from more thorough analyses, we have used it to estimate uncertainties in K_T. The use of eq 43 standardizes the process of estimating the uncertainties in K_T values. Many trends involving concentration (activity) or temperature changes follow energetic differences between the systems under different conditions, so that $\ln K_T$ is often used. Since κ is a multiplicative uncertainty, the error limit on $\ln K_T$ can be calculated directly from lnκ.

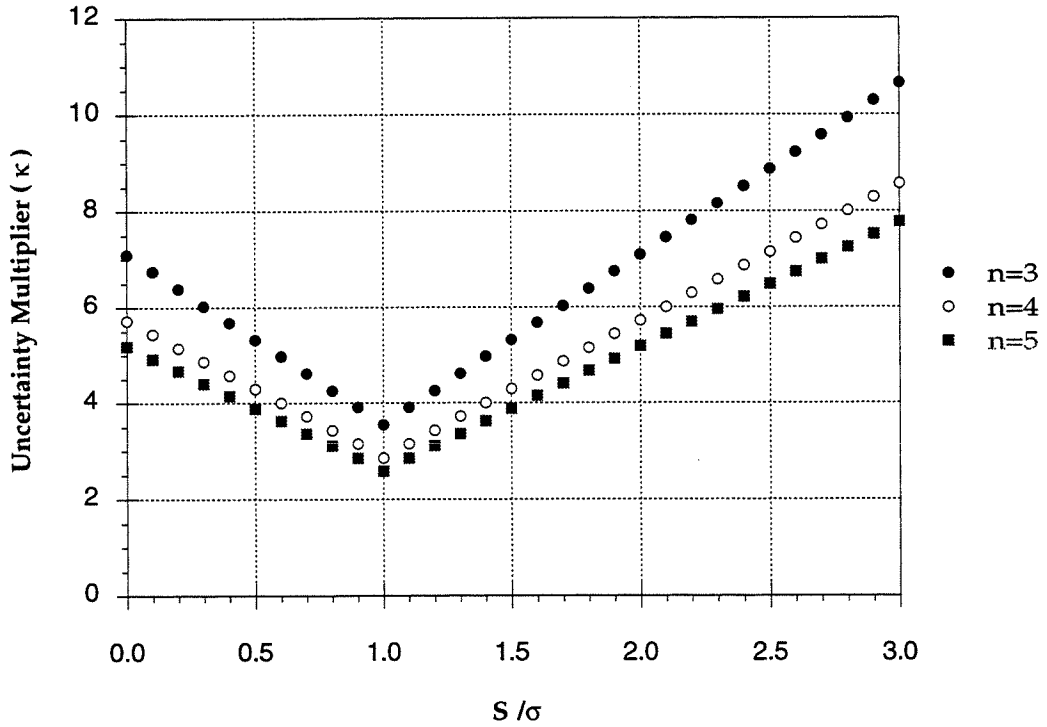


Figure 1.11. The estimated uncertainty multiplier for mean association constants is plotted as function of the ratio of S to σ. The values of κ are calculated using eq 43 for mean values obtained from three to five independent titrations (n = 3, 4, or 5).

Part IV. Technical Limitations

The Equilibrium Condition. In order to measure an equilibrium association constant, the binding reaction must have reached equilibrium. Because the affinity cleavage reactions can be initiated, independent of the binding reactions, by the addition of DTT, proper equilibration can be achieved by allowing the solutions to incubate for a sufficient length of time prior to the addition of DTT. A steep binding transition near $\theta = 0.5$ can indicate insufficient pre-incubation. Upon increasing the pre-incubation time, the transition steepness is reduced until a sufficient equilibration time is reached (Figure 1.12).

Thymidine-EDTA Loading with Iron. One of the most fundamental assumptions required for an affinity cleavage titration to produce true equilibrium binding isotherms is that each oligonucleotide-EDTA is capable of DNA cleavage. This requires that an iron atom be loaded into the EDTA moiety covalently attached to each oligonucleotide. Although the EDTA·Fe(III) stability constants are extremely high at pH 8.0 ($\sim 10^{25}$ M^{-1}), the stability rapidly decreases with decreasing pH. Moreover, one of the EDTA carboxylates has been covalently linked to the thymidine heterocycle (Figure 1.2), also reducing the apparent stability constant. In addition, iron must compete with other metal ions for thymidine-EDTA binding, and other chelators, such as spermine and Bis-tris, may also bind free iron.

Our initial experimental approach was to use as little excess iron as possible, so that iron was in excess of the oligonucleotide concentration by only 10 – 50 mol%. However, the measurement of high association constants ($> 10^7$ M^{-1}) indicated that the concentrations of iron at low oligonucleotide-EDTA concentrations limited the observed cleavage. For example, Figure 1.13 shows the

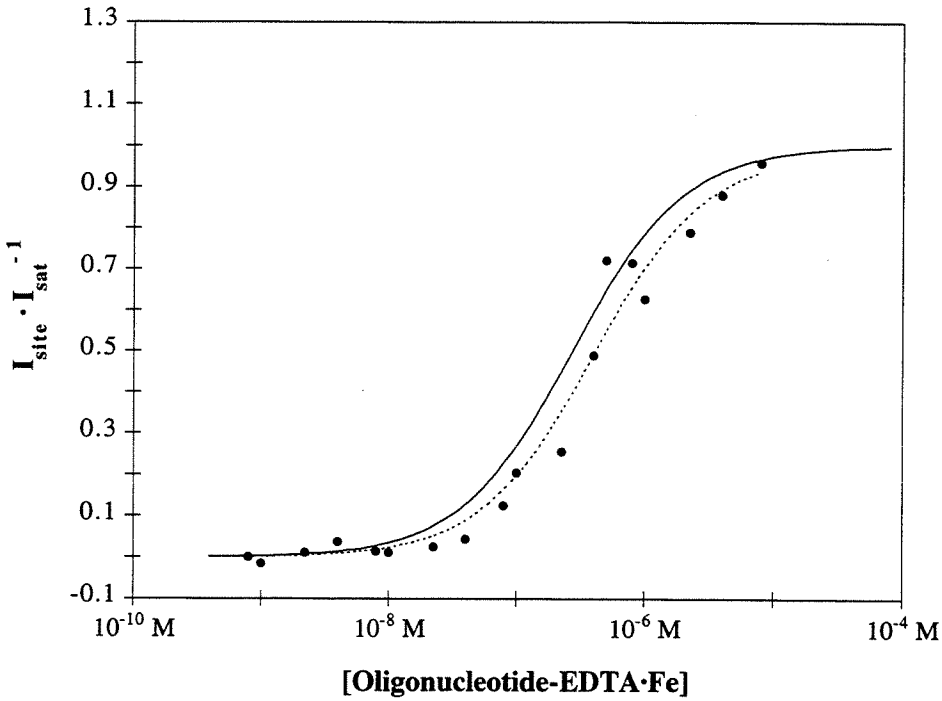


Figure 1.12. Influence of pre-incubation time on the shape of the experimental binding curve. The solid curve is the average binding curve for the 15mer 5'-T*TTTTCTCTCTCTCT-3' at pH 7.0 and 22°C after 24 h pre-incubation and 6.0 h cleavage reaction times (see Chapter Two). Data points (•) indicate results of affinity cleavage titration after 2.0 h pre-incubation and 6.0 h cleavage reaction times. The dotted curve is the best fit to the data points.

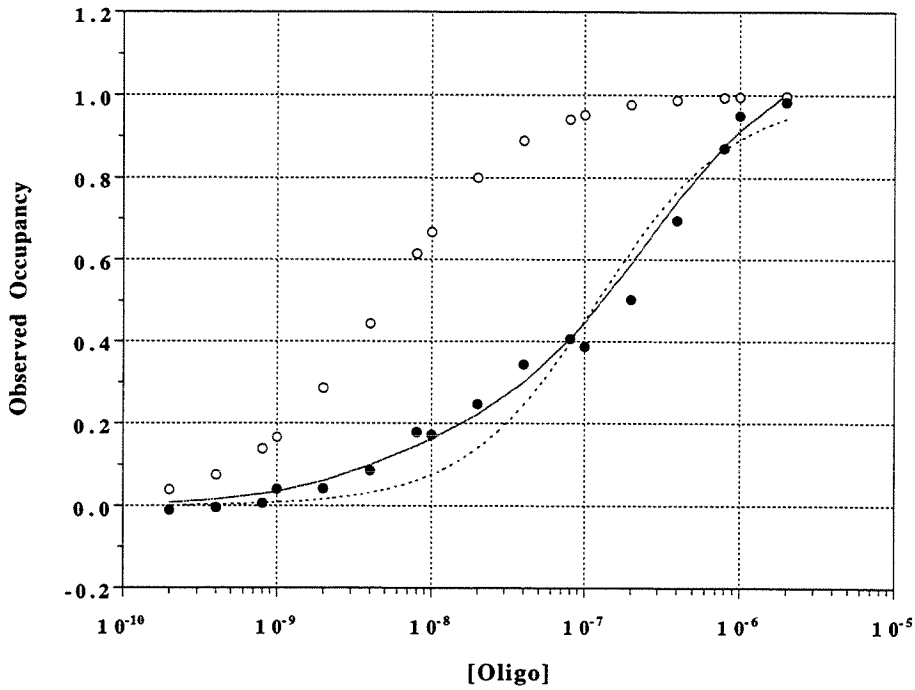


Figure 1.13. Difference between affinity cleavage (•) and footprint titration (○) data when the Fe concentration is limiting. The dotted curve is the best fit titration binding curve. The solid curve was obtained by adjusting the apparent occupancy ($I_{site} \cdot I_{sat}^{-1}$) for the fractional loading of Fe into the oligonucleotide-EDTA.

observed cleavage efficiency as a function of oligonucleotide-EDTA concentration and the "true" binding curve as measured by DNase I footprint titration. The association constant measured by footprint titration was ~ 2×10^8 M^{-1}, while the apparent association constant from affinity cleavage titration was only ~ 9×10^6 M^{-1}. In addition, the I_{site} data points do not fit the binding curve well, with the data points at low concentration lying above the curve, and those at high concentration falling below the curve.

In order to rationalize the observed difference between the affinity cleavage and footprint titration data, we considered a model which included simple one-to-one binding of iron to the oligonucleotide-EDTA independent of triple helix formation ($k_{Fe}/k_{-Fe} = k_3/k_{-3}$; Scheme 2), so that the apparent thymidine-EDTA·Fe stability constant was given by

$$K_{Fe} = \frac{O^{Fe}}{O \cdot [Fe]} \tag{44}$$

where brackets denote concentration and the symols O^{Fe} and O are defined as above. Using eq 44, the fraction of oligonucleotide-EDTA present with iron loaded can expressed by

$$\frac{O^{Fe}}{O_{tot}} = \frac{K_{Fe}O_{tot} + K_{Fe}[Fe]_{tot} + 1}{2K_{Fe}O_{tot}} + \frac{-\sqrt{\left(K_{Fe}O_{tot} + K_{Fe}[Fe]_{tot} + 1\right)^2 - 4K_{Fe}^2 O_{tot}[Fe]_{tot}}}{2K_{Fe}O_{tot}} \tag{45}$$

A theoretical curve calculated using the true apparent association constant (from footprint titration) adjusted for the fractional iron loading using eq 45 was fit to

the data points in Figure 1.13 with K_{Fe} as the adjustable parameter. The best-fit theoretical curve is displayed with the data points in Figure 1.12. Similar analysis of several other sets of data points yielded an average stability constant K_{Fe} = 2×10^6 M^{-1}. This stability constant limits the use of affinity cleavage titration to the measurement of association constants below 2×10^7 M^{-1} unless excess iron is employed to maximize the fractional iron loading at low oligonucleotide-EDTA concentrations.

It is experimentally trivial to add excess iron to affinity cleavage reactions to ensure iron loading. However, when excess iron is allowed to equilibrate in solution prior to the addition of DTT, a decrease in the cleavage signal is observed (Figure 1.14). In particular, we observe that when [Fe(II)] ≥ 2.0 μM or [Fe(III)] ≥ 40 μM is incubated in solution for 20 h or more, the relative cleavage yield is dramatically reduced. Unfortunately, long incubations times are required to ensure equilibration of the triplex formation reaction, and relatively high concentrations of iron are required to ensure complete loading into the oligonucleotide-EDTA. However, when 2 – 40 μM Fe(III) is allowed to remain in solution for up to 5.0 h, no diminution of the cleavage signal is observed. Based on these data, the protocol calls for loading the oligonucleotide-EDTA in the presence of a slight excess of Fe(II) during preparation of the serial dilution (protocol step 4.2, above), and the addition of Fe(III) to a final concentration of 20 μM just prior to (~ 30 min) the addition of DTT to initiate the cleavage reactions (protocol step 6.1).

No Self-Cleavage of the Oligonucleotide-EDTA. In simplifying Scheme 2, the approximations that $k_{self}[RO^{Fe}] \approx 0$ and $k_H[RT] \approx 0$ were made. The former assumption can be justified by the relatively low self-cleavage rates observed for EDTA·Fe complexes in solution. The rate of self-cleavage of the oligonucleotide-

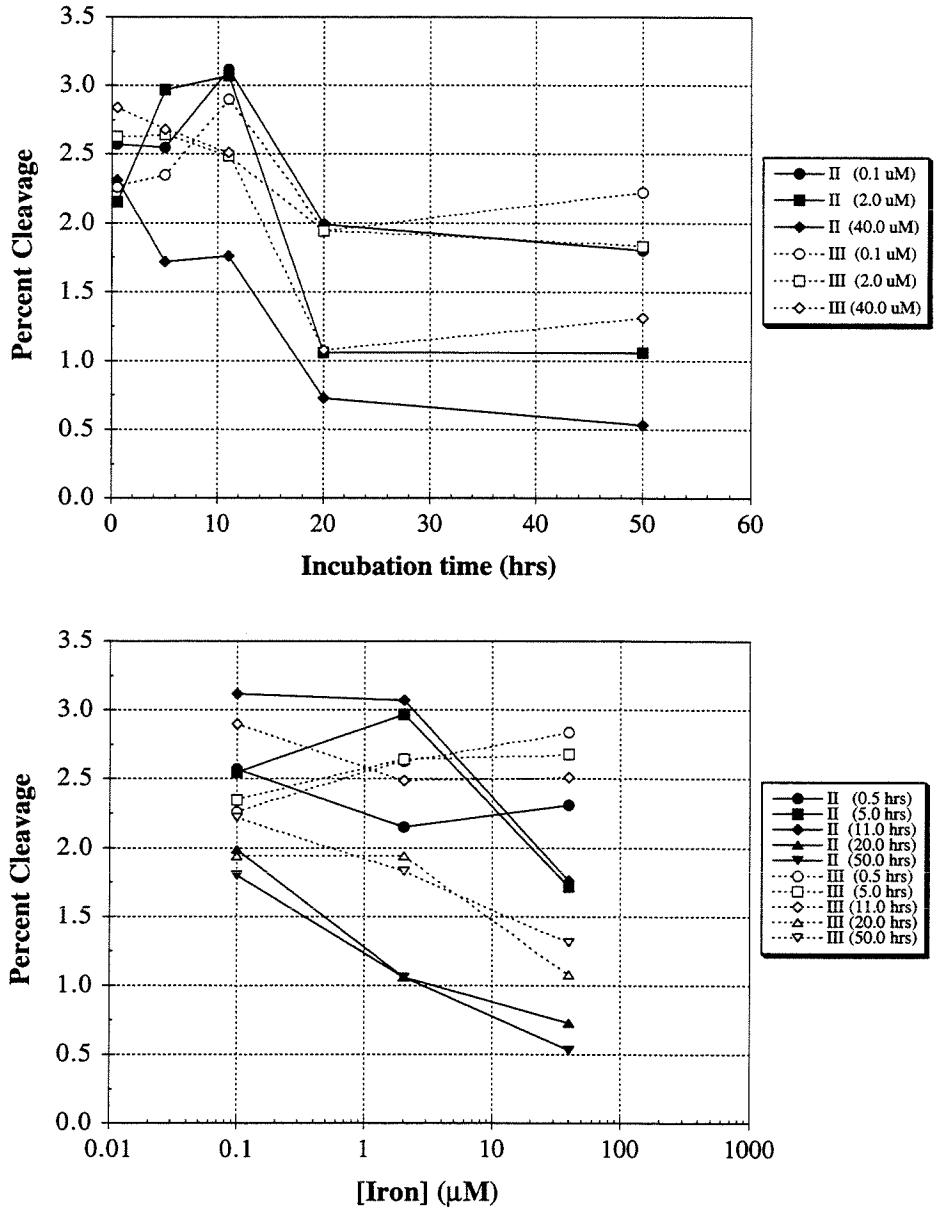


Figure 1.14. Effects of iron oxidation state, concentration, and incubation time in solution on the observed cleavage efficiency. *Upper,* Cleavage yield is plotted as a function of the length of time excess iron was allowed to equilibrate in solution with the binding mixtures. *Lower,* Cleavage yield is plotted as a function of the iron concentration in solution. In the legends, "II" indicates Fe^{2+} and "III" indicates Fe^{3+}.

EDTA·Fe bound in the triple-helical complex, however, has been shown to limit the observed cleavage efficiency of the double-helical DNA target.[84] In principle, third strand cleavage can reduce the apparent equilibrium fraction of double-helical DNA bound if the reaction is performed under conditions in which oligonucleotide-EDTA turnover occurs. This effect can be controlled by minimizing the reaction time so that triplex dissociation, and, therefore, turnover is reduced.

The Cleavage Reaction is Independent of the Dithiothreitol Concentration. The dependence of the cleavage reaction rates on the DTT concentration places another formal limit on the appearance of cleavage products over time. Because the DTT concentration is usually ≥ 1 mM while the highest oligonucleotide-EDTA concentration is ≤ 0.1 mM, we assume that the site-specific cleavage reaction is pseudo-first order. It is clear, however, that when the reaction time is long and the high concentrations of oligonucleotide-EDTA are in the range of 10 to 100 μM, this assumption can break down as the reaction continues and the DTT is depleted. In order to minimize this effect on the observed cleavage yield, the initial concentration of DTT should be kept as high as possible (1 to 4 mM) and the reaction times should be kept as short as possible.

Carrier DNA. In both footprinting and affinity cleaving reactions, it is typical to add non-specific DNA to concentrations near 0.1 mM in base pairs, in excess of the radiolabeled DNA whose concentration is typically picomolar. This "carrier DNA" minimizes any nonsequence-specific ligand-binding effects, serves as a "buffer" for binding cations, prevents overdigestion of the radiolabeled DNA, and provides a high enough DNA concentration to allow complete and reproducible precipitation following the reaction. The carrier DNA may also contain sequences capable of triple helix formation. When the concentration of carrier DNA is relatively high, binding of the oligonucleotide-EDTA to a carrier

site will reduce the equilibrium concentration of free oligonucleotide-EDTA below the total oligonucleotide-EDTA concentration. The fractional occupation of a carrier DNA site can be calculated:

$$\theta_C = \frac{K_C O_{tot} + K_C C_{tot} + 1}{2K_C C_{tot}} + \frac{-\sqrt{(K_C O_{tot} + K_C C_{tot} + 1)^2 - 4K_C^2 O_{tot} C_{tot}}}{2K_C C_{tot}} \quad (46)$$

If the concentration of carrier DNA (C_{tot}) is 100 μM-bp, the average binding constant to carrier DNA (K_C) is $\leq 10^5$ M^{-1}, and the binding site size is $\geq$ 12 bp in length, then $\leq$ 45% of the oligonucleotide-EDTA will be bound at any concentration, and the observed equilibrium constant will be within a factor of two from the true equilibrium constant. It is important to note that this effect will be influenced both by the concentration of carrier DNA and by the average binding constant to a site on carrier DNA. For the 15mer oligodeoxyribonucleotide investigated here, there is no effect on the apparent equilibrium constant when the concentration of calf thymus DNA is varied or when tRNA, a carrier for which a DNA oligomer has no measurable affinity,[11] is used (Figure 1.15).

Summary. Each potential limitation described in this part of Chapter One tends to decrease the observed binding constant, so that reported association constants likely represent lower bounds on the true equilibrium constants. Moreover, many of the limitations affect larger association constants relatively more than smaller ones. This differential effect results from the fact that lower oligonucleotide-EDTA concentrations, at which thymidine-EDTA·Fe and carrier

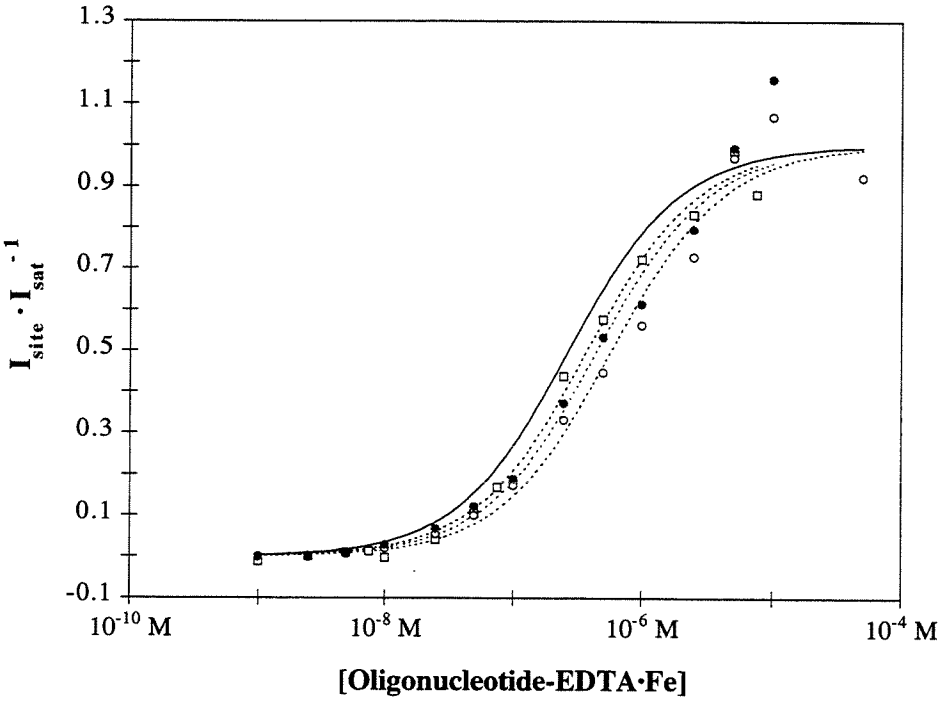


Figure 1.15. The influence of carrier nucleic acid on the observed titration binding curve. The concentration of calf thymus DNA was varied between 20 mM-bp (•) and 200 mM-bp (○). The open squares (□)indicate apparent fractional binding in the presence of tRNA-XX (Sigma). Dashed curves indicate the best fit isotherms to the three data sets. The solid line represents the average binding isotherm from Chapter Two (100 mM-bp).

DNA binding effects are more pronounced, are required for measuring larger association constants. Thus, relative ratios of observed equilibrium association constants are also likely lower bounds on the true ratios. In order to insure that relatively high association constants are measured with maximum accuracy, at least four different variables must be optimized. Specifically, the pre-incubation time should be increased, the affinity cleavage reaction time decreased, the concentration of DTT increased (to a maximum initial concentration of ~ 4 mM), and the concentration of carrier DNA decreased until both the shape and the translation of the binding curve along the abscissa no longer change.

References

(1) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(2) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochem.* **1993**, *32*, 11385-11389.

(3) Maher, L. J., III; Dervan, P. B.; Wold, B. J. *Biochem.* **1990**, *29*, 8820-6.

(4) Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436-40.

(5) Shin, J. A. Ph.D. Thesis, California Institute of Technology, 1992.

(6) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957-65.

(7) Singleton, S. F.; Dervan, P. B. *Biochem.* **1992**, *31*, 10995-1003.

(8) Singleton, S. F.; Dervan, P. B. *Biochem.* **1993**, *32*, 13171-9.

(9) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, submitted.

(10) Distefano, M. D.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1179-1183.

(11) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3806-10.

(12) Colocci, N.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 785-786.

(13) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-50.

(14) Hélène, C.; Toulmé, J. J. *Biochem. Biophys. Acta* **1990**, *1049*, 99-125.

(15) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-42.

(16) Ito, T.; Smith, C. L.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 495-8.

(17) Record, M. T., Jr.; Anderson, C. F.; Lohman, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103-178.

(18) Record, M. T., Jr.; Mazur, S. J.; Melançon, P.; Roe, J.-H.; Shaner, S. L.; Unger, L. *Annu. Rev. Biochem.* **1981**, *50*, 997-1024.

(19) Nadeau, J. G.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*,

(20) Olmsted, M. C.; Anderson, C. F.; Record, M. T., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7766-7770.

(21) Fenley, M. O.; Manning, G. S.; Olson, W. K. *Biopolymers* **1990**, *30*, 1191-1203.

(22) Olmsted, M. C.; Anderson, C. F.; Record, M. T., Jr. *Biopolymers* **1991**, *31*, 1593-1604.

(23) Wade, W. S. Ph.D. Thesis, California Institute of Technology, 1989.

(24) Mack, D. P. Ph.D. Thesis, California Institute of Technology, 1991.

(25) Schmitz, A.; Galas, D. J. *Nucleic Acids Res.* **1978**, *6*, 111-137.

(26) Povsic, T. J. Ph.D. Thesis, California Institute of Technology, 1992.

(27) Lane, M. J.; Dabrowiak, J. C.; Vournakis, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3260-3264.

(28) Lane, M. J.; Vestal, D. J.; Fish, E. L.; Vournakis, J. N.; Dabrowiak, J. C. *Biophys. J.* **1984**, *45*, 403-403.

(29) Rehfuss, R.; Ward, B.; Dabrowiak, J. C. *Biophys. J.* **1987**, *51*, 421-421.

(30) Ward, B.; Rehfuss, R.; Dabrowiak, J. C. *J. Biomol. Struct. Dyn.* **1987**, *4*, 685-695.

(31) Ward, B.; Dabrowiak, J. C. *Nucleic Acids Res.* **1988**, *16*, 8724-8724.

(32) Ward, B.; Rehfuss, R.; Goodisman, J.; Dabrowiak, J. C. *Nucleic Acids Res.* **1988**, *16*, 1359-1369.

(33) Ward, B.; Rehfuss, R.; Kissinger, K.; Goodisman, J.; Dabrowiak, J. C. *Biophys. J.* **1988**, *53*, 3-3.

(34) Dabrowiak, J. C.; Goodisman, J. In *Chemistry and Physics of DNA-Ligand Interactions*; N. R. Kallenbach, Ed.; Adenine Press: Guilderland, NY, 1989; pp 143-174.

(35) Dabrowiak, J. C.; Kissinger, K.; Goodisman, J. *Electrophoresis* **1989**, *10*, 404-412.

(36) Dabrowiak, J. C.; Ward, B.; Goodisman, J. *Biochem.* **1989**, *28*, 3314-3322.

(37) Lee, M.; Shea, R. G.; Hartley, J. A.; Lown, J. W.; Kissinger, K.; Dabrowiak, J. C.; Vesnaver, G.; Brelauer, K. J.; Pon, R. T. *J. Mol. Recog.* **1989**, *2*, 6-17.

(38) Rehfuss, R.; Goodisman, J.; Dabrowiak, J. C. *Biochem.* **1990**, *29*, 777-781.

(39) Goodisman, J.; Dabrowiak, J. C. *Biochem.* **1992**, *31*, 1058-1064.

(40) Goodisman, J.; Rehfuss, R.; Ward, B.; Dabrowiak, J. C. *Biochem.* **1992**, *31*, 1046-1058.

(41) Stankus, A.; Goodisman, J.; Dabrowiak, J. C. *Biochem.* **1992**, *31*, 9310-9318.

(42) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8462-8466.

(43) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Meth. Enzymol.* **1986**, *130*, 132-181.

(44) Senear, D. F.; Brenowitz, M.; Shea, M. A.; Ackers, G. K. *Biochem.* **1986**, *25*, 7344-7354.

(45) Hill, T. L. *Cooperativity Theory in Biochemistry: Steady State and Equilibrium Systems*; Springer: New York, 1985.

(46) Brenowitz, M.; Senear, D.; Jamison, E.; Dalma-Weiszhausz, D. In *Footprinting of Nucleic Acid-Protein Complexes*; A. Revzin, Ed.; Academic Press: San Diego, 1993.

(47) Oakley, M. G.; Mrksich, M.; Dervan, P. B. *Biochem.* **1992**, *31*, 10969-10975.

(48) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892-9899.

(49) Mrksich, M. Ph.D. Thesis, California Institute of Technology, 1994.

(50) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457-465.

(51) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(52) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528-5529.

(53) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565-2569.

(54) Dervan, P. B. In *Nucleic Acids and Molecular Biology*; F. Eckstein and D. M. J. Lilley, Ed.; Springer-Verlag: London, 1986; Vol. 2; pp 49-64.

(55) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-61.

(56) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-971.

(57) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(58) Distefano, M. D.; Shin, J. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 5901-5902.

(59) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochem.* **1992**, *31*, 2829-2834.

(60) Distefano, M. D.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 11006-11007.

(61) Beal, P. A.; Dervan, P. B. *Nucleic Acids Res.* **1992**, *20*, 2773-2776.

(62) Collier, D. A.; Mergny, J. L.; Thuong, N. T.; Hélène, C. *Nucleic Acids Res.* **1991**, *19*, 4219-24.

(63) Sun, J. S.; Lavery, R.; Chomilier, J.; Zakrzewska, K.; Montenay-Garestier, T.; Hélène, C. *J. Biomol. Struct. Dyn.* **1991**, *9*, 425-36.

(64) Chomilier, J.; Sun, J. S.; Collier, D. A.; Garestier, T.; Hélène, C.; Lavery, R. *Biophys. Chem.* **1992**, *45*, 143-52.

(65) Vandyke, M. W.; Dervan, P. B. *Cold Spring Harbor Symposia On Quantitative Biology* **1982**, *47*, 347-353.

(66) Hertzberg, R. P.; Dervan, P. B. *Biochem.* **1984**, *23*, 3934-3945.

(67) Oefner, C.; Suck, D. *J. Mol. Biol.* **1986**, *192*, 605-632.

(68) Griffin, L. C. Ph.D. Thesis, California Institute of Technology, 1990.

(69) Sinha, N. D.; Biernat, J.; Köster, M. *Tetrahedron Lett.* **1983**, *24*, 5843-5846.

(70) *Oligonucleotide Synthesis: A Practical Approach*; Gait, M. J., Ed.; IRL Press: Oxford, UK, 1984.

(71) Connors, K. A. *Binding Constants. The Measurement of Molecular Complex Stability.*; John Wiley & Sons: New York, 1987.

(72) Moser, H. E.; Dervan, P. B., unpublished result.

(73) Rougée, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Hélène, C. *Biochem.* **1992**, *31*, 9269-78.

(74) Roberts, R. W.; Crothers, D. M., personal communication.

(75) Hanvey, J. C.; Williams, E. M.; Besterman, J. M. *Antisense Res. Dev.* **1991**, *1*, 307-17.

(76) Johnson, R. F.; Pickett, S. C.; Barker, D. L. *Electrophoresis* **1990**, *11*, 355-360.

(77) Cantor, C. R.; Tinoco, I., Jr. *J. Mol. Biol.* **1965**, *13*, 65-71.

(78) Higuchi, R.; Krummel, B.; Saiki, R. K. *Nucleic Acids Res.* **1988**, *16*, 7351-7367.

(79) Runnebaum, I. B.; Syka, P.; Sukumar, S. *Biotechniques* **1991**, *11*, 446-452.

(80) Mattilla, P.; Ronka, J.; Tenkanen, T.; Pitkanen, K. *Nucleic Acids Res.* **1991**, *19*, 4967-4973.

(81) Shoemaker, D. P.; Garland, C. W.; Steinfeld, J. L.; Nibler, J. W. *Experiments in Physical Chemistry*; McGraw-Hill: New York, 1981.

(82) Bailey, N. T. J. *Statistical Methods in Biology;* 2nd ed.; Cambridge University Press: Cambridge, UK, 1981.

(83) Brown, D.; Rothery, P. *Models in Biology: Mathematics, Statistics, and Computing*; John Wiley & Sons: Chichester, UK, 1993.

(84) Strobel, S. A. Ph.D. Thesis, California Institute of Technology, 1992.

CHAPTER TWO

Thermodynamics of Oligonucleotide-Directed Triple Helix Formation: An Analysis Using Quantitative Affinity Cleavage Titration

The text of this chapter is taken from a published manuscript that was coauthored with Professor Peter B. Dervan.

(S. F. Singleton & P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 6957-6965.)

Introduction

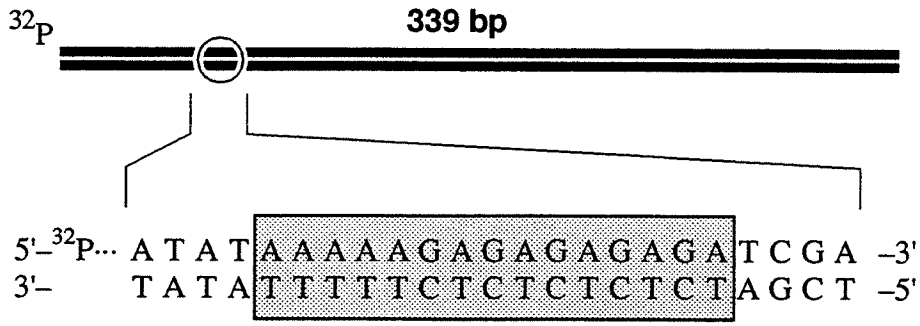
The stability of triple helices is dependent on the length, sequence composition, and functional modification of the oligonucleotide as well as solution conditions, including pH and cation concentrations. A full understanding of the factors contributing to the stability and specificity of the binding of oligonucleotides to double-helical DNA will require a complete characterization of the thermodynamics of complex formation. The measurement of binding constants at individual DNA sites represents the first step of this process. While some thermodynamic data has been reported for homopolymeric[1] and, more recently, oligonucleotide triplexes at acidic pH,[2,3] we wish to measure thermodynamic parameters for oligonucleotide-directed triple helix formation at single sites on relatively large DNA (> 200 bp) near pH 7.

The affinity cleaving technique, which has proven useful for studying DNA recognition, relies on the covalent attachment of a DNA cleaving agent to a DNA binding molecule and the subsequent generation of a nonspecific, diffusible oxidant near a bound ligand–DNA complex.[4] The affinity cleaving assay has been used previously to evaluate the relative binding strengths of oligonucleotide-EDTA·Fe conjugates.[5-7] These analyses relied on the assumption that the cleavage efficiency produced by an oligonucleotide-EDTA·Fe is proportional to the fraction of labelled duplex DNA bound by the oligonucleotide-EDTA·Fe. We report the extension of the affinity cleaving technique to the quantitative measurement of equilibrium association constants for the binding of oligonucleotide-EDTA conjugates to single sites on large DNA duplexes at neutral pH. By measuring the amounts of site-specific cleavage produced over four orders of magnitude of oligonucleotide-EDTA·Fe concentration, we can construct empirical titration binding curves and determine equilibrium association constants for oligonucleotide-directed triple helix formation. As an independent check of the method, we demonstrate that quantitative affinity cleavage titration affords association constants that are identical with those obtained from quantitative DNase I footprint titration. The affinity cleavage method has been used to measure the effects of covalent modification of an oligonucleotide with EDTA·Fe, oligonucleotide length, and internal base triplet mismatches on the stability of a local triple helical complex at a single site in large DNA.

Results and Discussion

Oligonucleotide-EDTA·Fe Association Constants by Quantitative Affinity Cleavage Titration. To assess the association constants for oligonucleotides binding to a 339-bp DNA duplex, ^{32}P-5'-end labelled DNA (< 100 pM) and various concen-

trations of an oligonucleotide-EDTA·Fe (1 nM – 10 μM) were mixed in Association buffer (100 mM Na^+, 1 mM spermine tetrahydrochloride, 50 mM Tris-acetate, pH 7.0) and incubated at 24 °C. The association reactions were allowed to equilibrate over 24 h. In fact, we observed no changes in the shapes of best-fit binding titration isotherms or measured values of K_T for experiments in which the equilibration time was greater than 18 h. The length of time required for the binding reaction to reach equilibrium is consistent with the saturation of a target duplex to 95% of its equilibrium occupancy within 3 min and 34 h in the presence of 10 μM and 10 nM, respectively, of a 21mer oligonucleotide at 37 °C and pH 6.8.[8] Only after the reactions had reached equilibrium was DTT added to initiate the EDTA·Fe-mediated cleavage chemistry. The duration of the cleavage reactions was constrained to allow for a maximum site-specific cleavage yield of about 5%. This constraint was enforced to maximize the statistical probability that specific cleavage of the target resulted from single-hit binding and cleaving, and to ensure that a majority of the nonspecific cleavage products had lengths greater than those products from specific cleavage.[9] The products were separated by polyacrylamide gel electrophoresis under strand denaturing conditions. The results of a typical experiment performed using the Fe complex of oligonucleotide-EDTA **1** (**1**·Fe, Figure 2.1) are shown in Figure 2.2. The amounts of radiolabelled DNA in the bands at the target cleavage site (Figures 2.2 and 2.3) and at a reference site (Figure 2.2) were measured from a photostimulable storage phosphor autoradiogram, and the ([**1**·Fe], I_{site}) data points were plotted semilogarithmically along with the Langmuir titration binding isotherm, which was obtained by performing a non-linear least squares fit (see Eq 38 in Chapter One), in Figure 2.4. The mean value of the equilibrium association constant extracted from three such experiments is equal to $3.7 \pm 1.1 \times 10^6$ M^{-1} (Table 2.1).



1 5'–*T T T T T C T C T C T C T C T

2 5'–*T T T T T C T C T C T C T

3 5'–*T T T T T C T C T C T

4 5'–*T T T T T C T C T C C C T C T

5 5'–*T T T T T C T C T T T C T C T

6 5'– T T T T T C T C T C T C T C T

Figure 2.1. The sequences of oligonucleotides **1 – 6** are shown, where T* indicates the position of thymidine-EDTA. The sequence of the 15-bp target site (shaded box) is shown at the top, along with the relative postion of the target site within the radiolabelled duplex. The nucleotides represented by outline type in oligonucleotides **4** and **5** indicate the positions where C·AT and T·GC mismatched triplets, repectively, are expected to be formed upon triple helix formation.

Figure 2.2. Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by **1**·Fe during a quantitative affinity cleavage titration experiment performed in Association buffer (100 mM Na^+, 1 mM spermine tetrahydrochloride, 50 mM Tris-acetate, pH 7.0) at 24 °C. The bar drawn on the left of the autoradiogram indicates the postion of the 15-bp duplex target site (white box), the bands used to measure I_{tot} (black box), and the bands used to measure I_{ref} (grey box). (Lane 1) Intact 5' labelled duplex obtained after incubation in Association buffer at 4 °C in the absence of oligonucleotide-EDTA·Fe and DTT. (Lane 2) Products of an adenine specific sequencing reaction. (Lane 3) Products of a guanine specific sequencing reaction. (Lanes 4–21) DNA affinity cleavage products produced by **1**·Fe at various concentrations: no oligonucleotide (lane 4); 10.0 μM (lane 5); 7.5 μM (lane 6); 5.0 μM (lane 7); 2.5 μM (lane 8); 1.0 μM (lane 9); 750 nM (lane 10); 500 nM (lane 11); 250 nM (lane 12); 100 nM (lane 13); 75 nM (lane 14); 50 nM (lane 15); 25 nM (lane 16); 10 nM (lane 17); 7.5 nM (lane 18); 5.0 nM (lane 19); 2.5 nM (lane 20); 1.0 nM (lane 21).

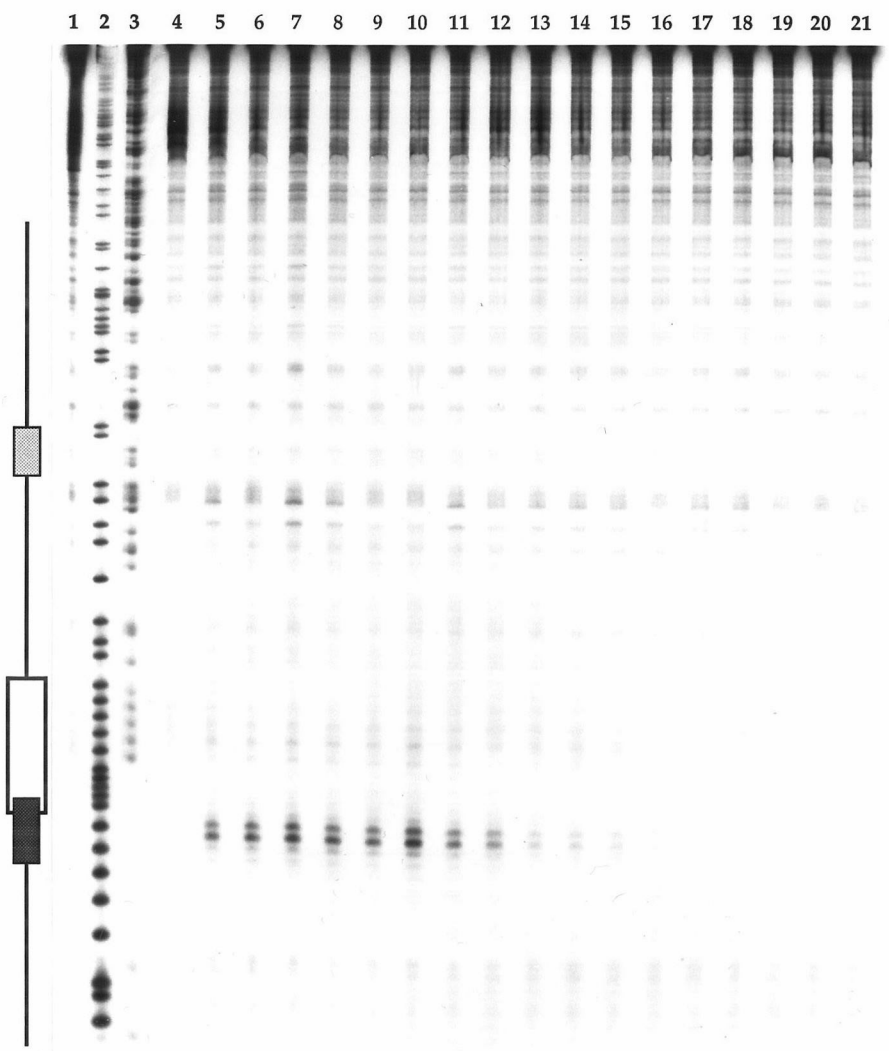
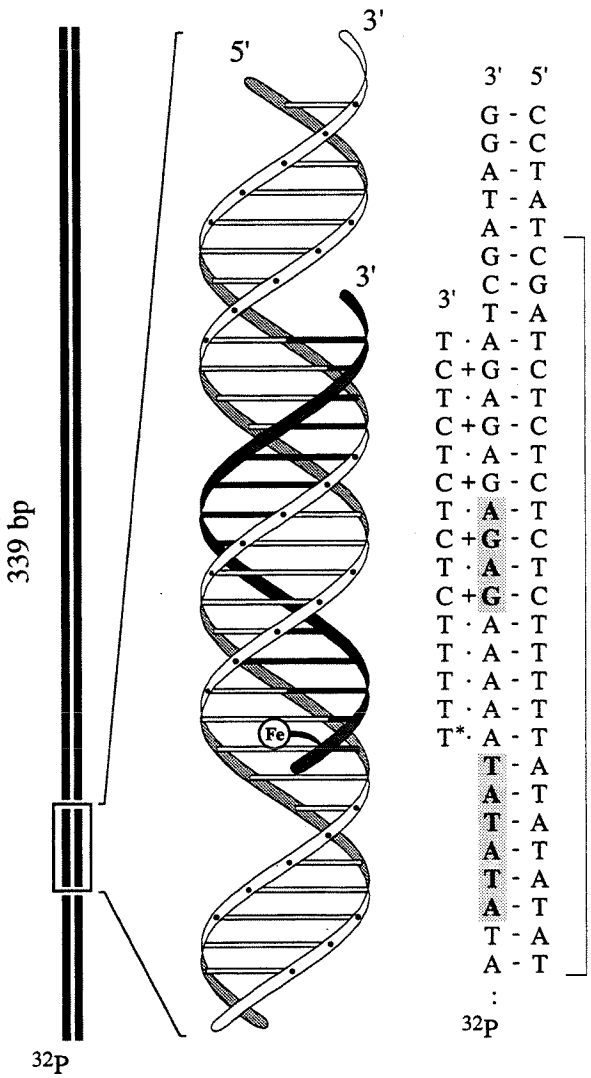
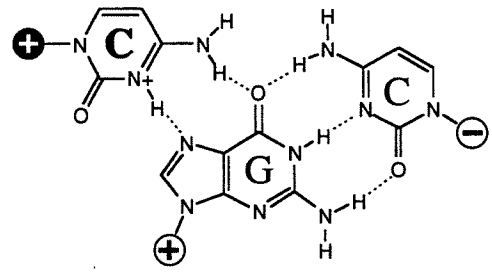

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

Figure 2.3. *Left,* Ribbon model of the local triple-helical structure formed by the binding of the 15mer **1**·M^{n+} (M^{n+} = Fe^{2+} for affinity cleaving and In^{3+} for footprinting experiments, respectively) to the 15-bp target sequence within the 339-bp end-labeled duplex. The Watson-Crick duplex strands are depicted as white ribbons while the oligonucleotide-EDTA is depicted as a dark ribbon. The sequences modelled by the ribbons are shown in the center where gray boxes have been drawn around the ten nucleotide positions cleaved most efficiently by the oligonucleotide-EDTA·Fe conjugate and used to obtain I_{site} values. The nucleotide positions represented by outline type indicate the positions where mismatched base triplets will form when oligonucleotides **4** and **5** bind to the target duplex. The bracket to the right of the sequences indicates the nucleotide positions of the duplex protected from DNase I digestion when **1**·In or **6** is bound. *Right,* Two-dimensional models depicting the C+GC and T·AT base triplets formed by Hoogsteen-type hydrogen bonding of *N*3-protonated C to a Watson-Crick GC base pair (top) and by Hoogsteen hydrogen bonding of T to a Watson-Crick AT base pair (bottom). The bases of the third strand are labeled with bold type and bases of the Watson-Crick duplex are labeled with normal type. The circles attached to *N*1 of the pyrimidines and *N*9 of the purines indicate the position of attachment to the sugar-phosphate backbone, while the plus and minus signs designate the relative 5'-to-3' polarity of the strands.


339 bp
5'
3'
3'
Fe
32P
3' 5'
G - C
G - C
A - T
T - A
A - T
G - C
C - G
3' T - A
T · A - T
C + G - C
T · A - T
C + G - C
T · A - T
C + G - C
T · A - T
C + G - C
T · A - T
T · A - T
T · A - T
T · A - T
T*· A - T
T - A
A - T
T - A
A - T
T - A
A - T
T - A
A - T
:
32P


C
G
C

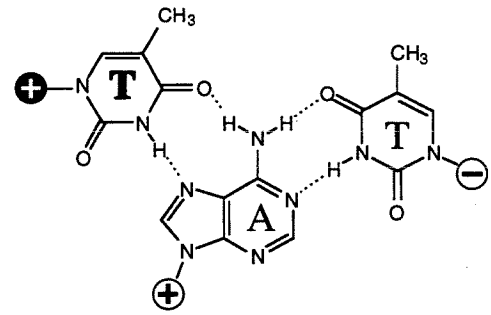

CH3
T
CH3
T
A

Table 2.1. Comparison of Footprint and Affinity Cleavage Titration Measurements of K_T at 24 °C and pH 7.0.[a]

oligonucleotide	method[b]	K_T (M^{-1})
1·Fe	AC	3.7 (± 1.1) x 10^6
1·In	FP	2.8 (± 0.5) x 10^6
6	FP	3.7 (± 0.4) x 10^6

[a] Affinity cleavage (AC) and footprint (FP) titration experiments were performed at 24 °C in Association buffer (100 mM Na^+, 1 mM spermine·4HCl, 50 mM Tris-acetate, pH 7.0) and Footprint buffer (10 mM Na^+, 20 mM Mg^{2+}, 20 mM Ca^{2+}, 1 mM spermine·4HCl, 50 mM Tris-acetate, pH 7.0), respectively.[10] [b] The K_T value reported in row one was measured by quantitative affinity cleavage titration (AC), and the K_T values reported in rows two and three were measured by quantitative DNase I footprint titration (F).

In order to validate the affinity cleavage titration method, we compared the association constant measured by affinity cleavage with that obtained from quantitative DNase I footprint titration, a technique which has been previously used to examine the energetics of protein–DNA interactions and has given results that are consistent with other physical methods.[11] We performed the footprint titration essentially according to the published protocol. Briefly, oligonucleotide-EDTA **1** was allowed to chelate In^{3+}, which is redox inactive, and then various concentrations of the complex were mixed with end-labelled duplex in Footprint buffer (10 mM Na^+, 20 mM Mg^{2+}, 20 mM Ca^{2+}, 1 mM spermine tetrahydorchloride, 50 mM Tris-acetate, pH 7.0)[10] and incubated at 24 °C. After

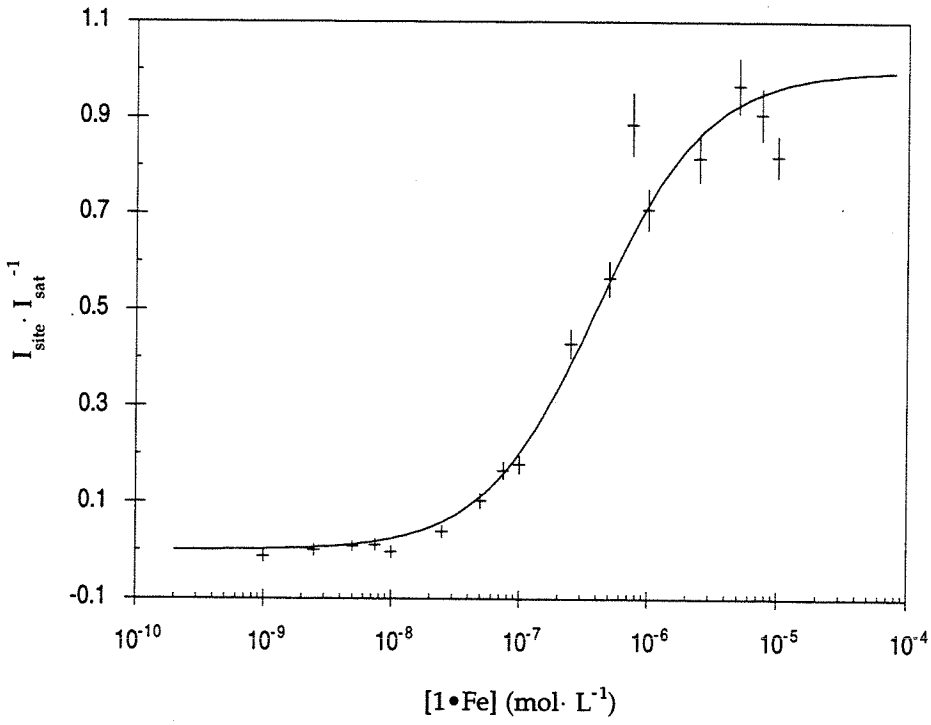


Figure 2.4. Data from the quantitative affinity cleavage titration experiment whose auto-radiogram is shown in Figure 2.2. The I_{site} data were obtained using photostimulable storage phosphor autoradiography and eq 36 (Chapter One), and have been normalized to a maximum value of 1 using the fit value of I_{sat}. The data points are represented by crosses whose horizontal and vertical bars indicate the ranges covered by the estimated uncertainties in $[\mathbf{1}\cdot Fe]_{tot}$ and I_{site} (eq 39 from Chapter One), respectively. The solid curve is the best fit Langmuir binding titration isotherm obtained from a non-linear least squares algorithm using eq 38 (Chapter One) and a value of 1 for I_{sat}.

24 h, DNase I was added to each solution and the digestion products were separated by polyacrylamide gel electrophoresis under denaturing conditions. The results of such a footprint titration are shown in Figure 2.5, and the region of the duplex footprinted is indicated on the model in Figure 2.3. The amounts of DNase I digestion at both the 15-bp target site and at a reference site (Figure 2.5) were measured by integrating the bands in a storage phosphor autoradiogram and the ([**1**·In], θ_{app}) data points were fit to the published equation for a titration binding isotherm (see ref. 13a and the text under the "Footprint Titration Fitting Procedure" heading in the Experimental Section). The data points and the best fit curve for the titration shown in Figure 2.5 are plotted in Figure 2.6A. The mean value of the binding constant for association of **1**·In with the duplex target, as measured by footprint titration, is $2.8 \pm 0.5 \times 10^6$ M^{-1} (Table 2.1). This is within the experimental uncertainty of the value $3.7 \pm 1.1 \times 10^6$ M^{-1} obtained from affinity cleavage titration.

In a similar manner, we measured the association constant for oligonucleotide **6** (Figure 2.1), which lacks the metal-chelator functionality (*i.e.*, the natural thymidine nucleoside has been substituted for T* at the 5′-end of the oligonucleotide), using quantitative DNase I footprint titration (Figure 2.6B). We obtained a mean value of $K_T = 3.7 \pm 0.4 \times 10^6$ M^{-1} (Table 2.1) for the equilibrium binding of **6** to the duplex target site.

The results of the three types of experiments described in this section are summarized in Table 2.1. The fact that the values of K_T measured by affinity cleavage titration and the established footprint titration method agree within the experimental uncertainty indicates that quantitative affinity cleavage will be useful for the reliable measurement of equilibrium binding constants for oligonucleotide-directed triple helix formation. Furthermore, a comparison of the mea

Figure 2.5. Autoradiogram of an 8% denaturing polyacrylamide gel used to separate the products generated by DNase I digestion during a quantitative footprint titration experiment using **1**·In and performed in Footprint buffer (10 mM Na^+, 20 mM Mg^{2+}, 20 mM Ca^{2+}, 1 mM spermine tetrahydrochloride, 50 mM Tris-acetate, pH 7.0) at 24 °C. The bar drawn on the left of the autoradiogram indicates the bands used to measure I_{tot} within the 15-bp duplex target site (white box) and the bands used to measure I_{ref} (grey box). (Lane 1) Intact 5' labelled duplex obtained after incubation in Footprint buffer at 4 °C in the absence of oligonucleotide-EDTA·In and DNase I. (Lane 2) Products of an adenine specific sequencing reaction. (Lane 3) Products of a guanine specific sequencing reaction. (Lanes 4–20) DNase I digestion products produced in the presence of various concentrations of **1**·In: no oligonucleotide (lane 4); 8.0 μM (lane 5); 4.0 μM (lane 6); 2.0 μM (lane 7); 1.0 μM (lane 8); 800 nM (lane 9); 400 nM (lane 10); 200 nM (lane 11); 100 nM (lane 12); 80 nM (lane 13); 40 nM (lane 14); 20 nM (lane 15); 10 nM (lane 16); 8.0 nM (lane 17); 4.0 nM (lane 18); 2.0 nM (lane 19); 1.0 nM (lane 20).

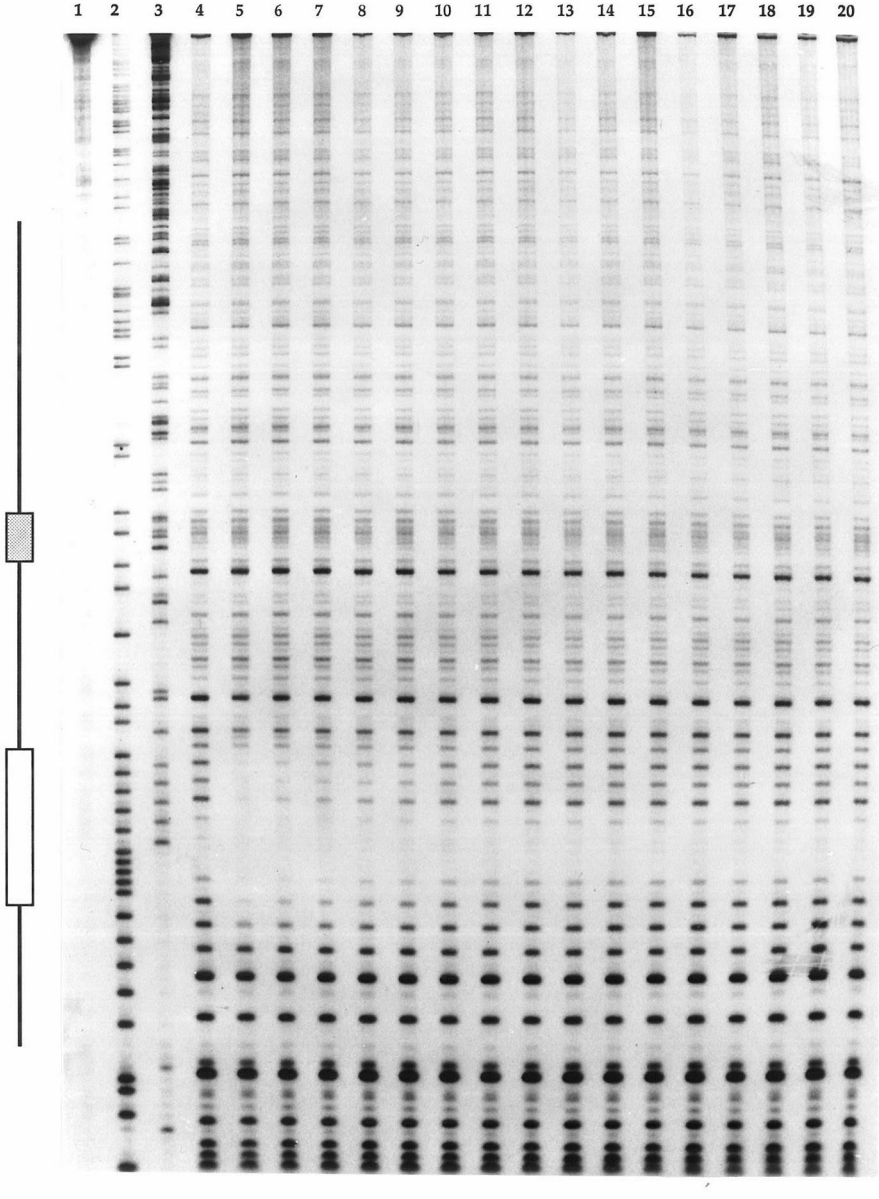
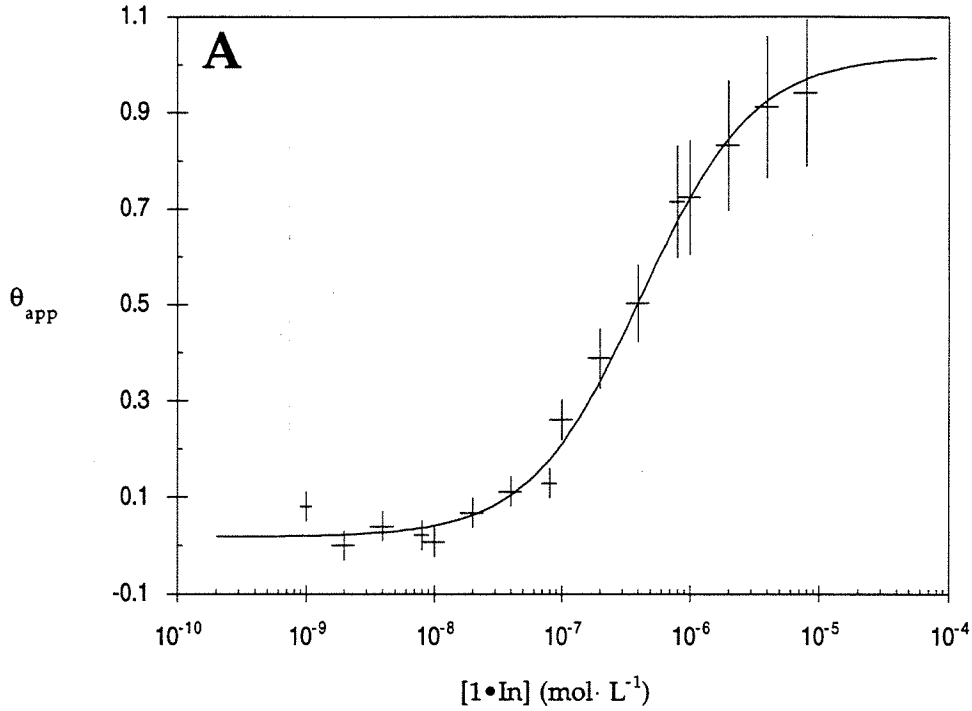
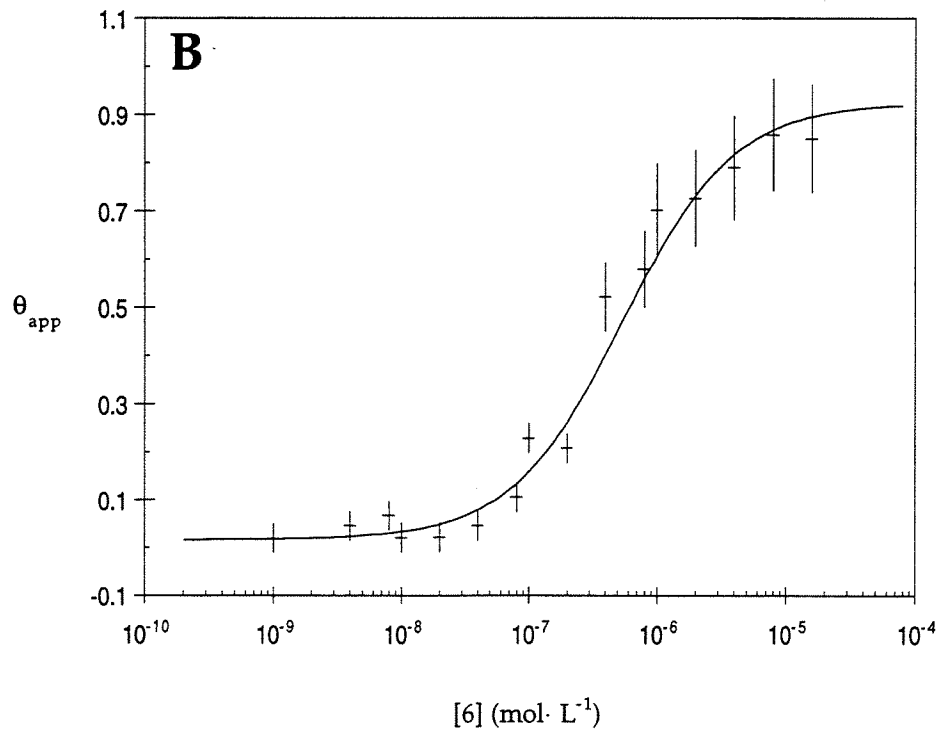

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

Figure 2.6. (A) Data from the quantitative DNase I footprint titration experiment whose autoradiogram is shown in Figure 2.5. The θ_{app} data were obtained using photostimulable storage phosphor autoradiography and eq 1. The data points are represented by crosses whose horizontal and vertical bars indicate the ranges covered by the estimated uncertainties in $[\mathbf{1}\cdot In]_{tot}$ and θ_{app} (eq 3), respectively. The solid curve is the best fit Langmuir binding titration isotherm obtained from a non-linear least squares algorithm using eq 2. (B) Data from a quantitative DNase I footprint titration experiment using **6**. The experiment was performed exactly as the one whose results are presented in Figures 2.5 and 2.6A. The data are plotted as described for Figure 2.6A.


A
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
θapp
10-10
10-9
10-8
10-7
10-6
10-5
10-4
[1•In] (mol· L-1)

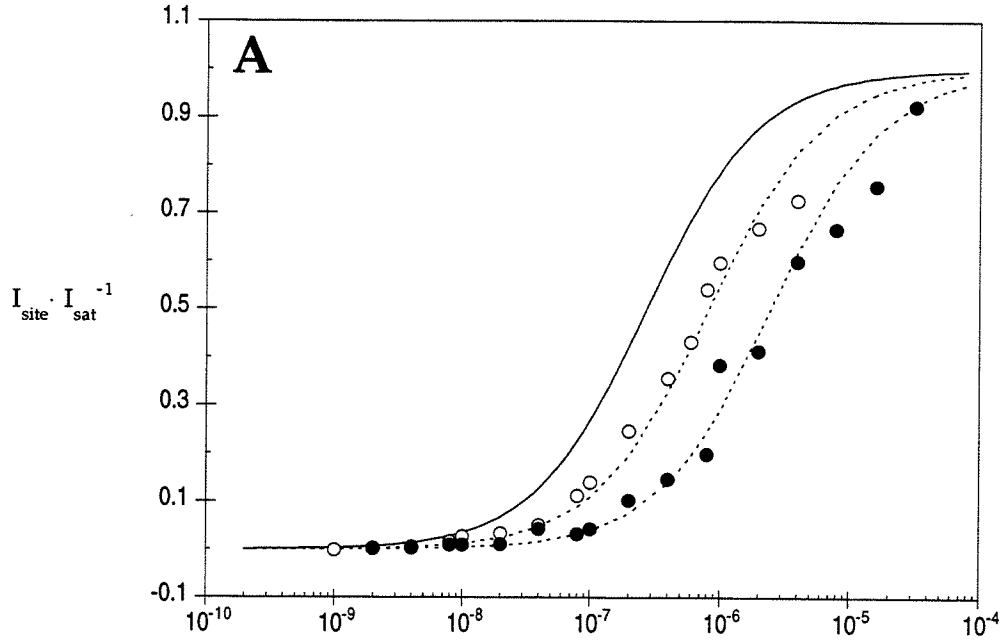

B
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
θapp
10-10
10-9
10-8
10-7
10-6
10-5
10-4
[6] (mol· L-1)

sured K_T values from rows one and two of Table 2.1 with that from row three suggests that the covalent attachment of the EDTA moiety to a thymine heterocycle at the 5′-end of an oligonucleotide has little effect (*i.e.*, one that is not measurable within experimental limitations) on the association of the oligonucleotide with its duplex target at 24 °C and pH 7.0.

The Influence of Oligonucleotide Length. Comparison of the K_T values for the equilibrium binding of **1**·Fe, **2**·Fe, and **3**·Fe (Figure 2.1) to the duplex target in Association buffer at 24 °C reveals quantitative information concerning the effect of oligonucleotide length on binding affinity. The binding isotherms measured in Association buffer are plotted in Figure 2.7A, and the K_T values are reported in the first three rows of Table 2.2. The association constants for binding of the 13-mer **2**·Fe and the 11-mer **3**·Fe are clearly reduced from that of **1**·Fe, a 15-mer. The quantitative difference in the binding constants for **1**·Fe, **2**·Fe, and **3**·Fe corroborate the relative ranking of the affinities of these three modified oligonucleotides that was reported earlier and was based on the amount of specific cleavage generated at a single concentration of each oligonculeotide.[5a]

An analysis of the energetics of binding for **1**·Fe, **2**·Fe, and **3**·Fe (Table 2.2) reveals that removal of one thymidine residue and one cytidine residue from **1**·Fe reduces K_T by a factor of two ($\Delta G_2 - \Delta G_1 = 0.5$ kcal·mol^{-1}), and removal of two thymidine and two cytidine residues from **1**·Fe reduces K_T by a factor of four ($\Delta G_3 - \Delta G_1 = 1.1$ kcal·mol^{-1}). Surprisingly, the changes in the free energy of binding with decreasing strand length are six-fold smaller than the changes in the free energy of duplex formation predicted from the sequences of **1**, **2**, and **3**.[12] The relatively modest dependence of triple helix stability on third-strand length likely arises from several sources. One probable factor is a reduced role in triple helix stabilization played by the cytidine residues in the third strand at pH 7.0

Figure 2.7. (A) Data for quantitative affinity cleavage experiments involving oligonucleotides **2**·Fe (○) and **3**·Fe (•) in Association buffer at 24 °C. The data points represent the average site specific cleavage signal intensities from three experiments. The dotted sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T for **2**·Fe and **3**·Fe (Table 2.2) and eq 38 (Chapter One). The data points were normalized using I_{sat} from each experiment and the binding curves were subsequently normalized using I_{sat} = 1 for eq 38 (Chpater One). The solid curve shows the normalized titration binding isotherm obtained using the mean value of K_T for **1**·Fe. (B) Data for quantitative affinity cleavage experiments involving oligonucleotides **4**·Fe (◇) and **5**·Fe (♦) in Association buffer at 24 °C. The data points represent the average site specific cleavage signal intensities from three experiments. The dotted sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T for **4**·Fe and **5**·Fe (Table 2.2) and eq 38 (Chpater One). The data points and fit binding curves were normalized as above. The solid curve shows the normalized titration binding isotherm obtained using the mean value of K_T for **1**·Fe.


A
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
$I_{site} \cdot I_{sat}^{-1}$
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
[Oligo•Fe] (mol· L^{-1})

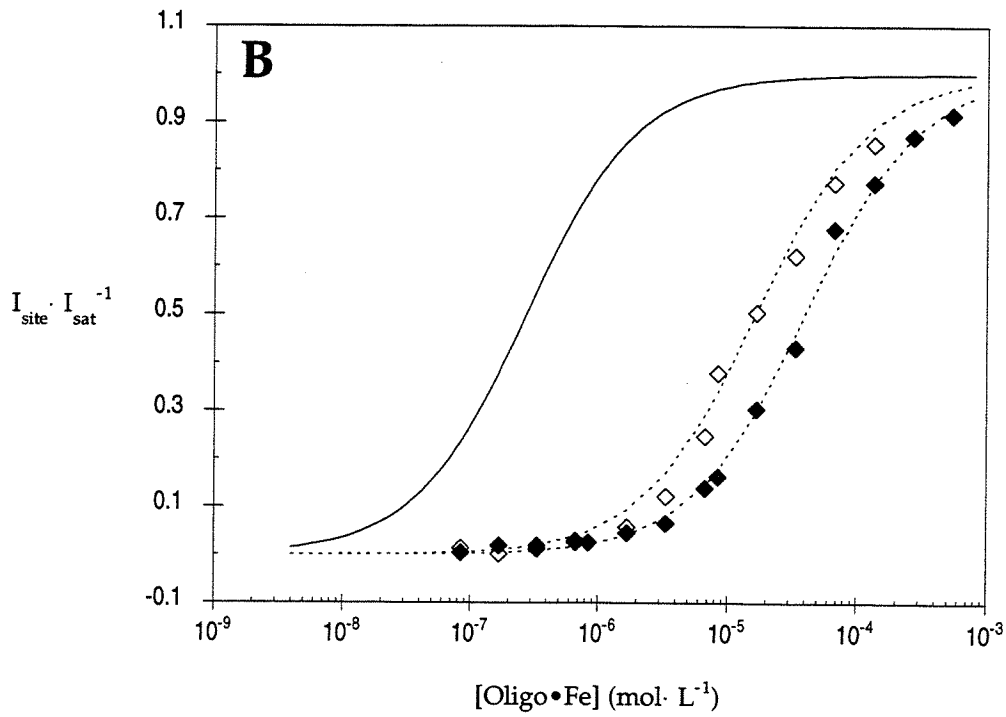

B
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
$I_{site} \cdot I_{sat}^{-1}$
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
[Oligo•Fe] (mol· L^{-1})

Table 2.2. Length and Sequence Mismatch Effects on K_T at pH 7.0 and 24 °C.[a]

Oligo-nucleotide	Length (nts)	No. T's (nts)	No. C's (nts)	Base Mismatch	K_T (M^{-1})	ΔG_T (kcal·mol^{-1})
1	15	10	5	—	3.7 (±1.1) x 10^6	−9.0 (±0.2)
2	13	9	4	—	1.8 (±0.4) x 10^6	−8.5 (±0.1)
3	11	8	3	—	6.2 (±3.8) x 10^5	−7.9 (±0.6)
4	15	9	6	C·AT	6.1 (±2.3) x 10^4	−6.5 (±0.2)
5	15	11	4	T·GC	2.5 (±0.1) x 10^4	−6.0 (±0.1)

[a] Values reported in the table are mean values measured from affinity cleavage titration experiments performed in Association buffer (100 mM Na^+, 1 mM spermine tetrahydrochloride, 50 mM Tris-acetate, pH 7.0).

where these bases may not be protonated. In fact, preliminary evidence indicates that removing two thymidine residues from the 5'-terminus of **1** reduces K_T by a factor of nearly ten, suggesting that T·AT triplets confer higher energetic stabilization to the triplex than do C+GC triplets at pH 7.0.[13] Clearly, the actual energetic contributions to triple helix stability by single T·AT and C+GC base triplets will be distinct and will depend on the identity of nearest neighbor triplets[5i] as well as the pH of the buffer.

Influence of Base Sequence Mismatches. A comparison of the equilibrium binding constants for **4**·Fe and **5**·Fe, which contain single, internal C·AT and T·GC base triplet mismatches,[5a,e] respectively (Figures 2.1 and 2.3), with that for **1**·Fe provides information on the effect of sequence mismatches between an oligonucleotide and its target duplex. The data obtained from reactions in Association buffer at 24 °C and the best fit titration binding isotherms are plotted in Figure 2.7B. The measured K_T values for **1**·Fe, **4**·Fe, and **5**·Fe are reported in rows one, four, and five of Table 2.2, respectively. The results show that a single, internal base triplet mismatch can reduce the equilibrium binding constant for an oligonucleotide by two orders of magnitude, a result that is consistent with the relative amounts of cleavage generated by these three oligonucleotides at a single concentration.[5a]

An analysis of the differences in binding free energy between **1**·Fe and **4**·Fe and between **1**·Fe and **5**·Fe (Table 2.2) indicates that a single internal C·AT mismatch results in a 2.5 kcal·mol^{-1} loss of stability for the triple helical complex, while a single internal T·GC mismatch results in a 3.0 kcal·mol^{-1} stability loss. Although we expect that the energetic penalties for a mismatch will be affected by nearest neighbor base triplets, the energetic changes observed here are consistent with those recently derived from equilibrium competition experiments and

shape analysis of UV melting curves for oligonucleotide triplexes at pH 5.0[2g] and with the energetic destabilization range estimated by two-dimensional gel electrophoresis studies of H-DNA triplexes.[14] Moreover, thermal denaturation studies at pH 5.6[15] and NMR spectroscopic investigations at pH 5.2[16] are consistent with a large degree of destabilization resulting from mismatched C·AT and T·GC base triplets. Perhaps of most interest is the observation that our measured free energies for triplex destabilization are near those reported for the destabilization of DNA[17] and RNA[18] duplexes with mismatched base pairs. The high specificity for a designated duplex target sequence afforded by oligonucleotide-directed triple helix formation demonstrates clearly the potential functional application of oligonucleotides or their analogues to problems of DNA recognition in biology and human medicine.[19]

Conclusion

A method for the determination of equilibrium constants for the formation of local triple-helical complexes at single sites within large DNA using affinity cleavage titration experiments has been presented. Equilibrium constants for oligonucleotide-EDTA binding measured using this method are identical within experimental uncertainty with those measured by DNase I footprint titration and indicate that covalent attachment of EDTA·Fe to the 5′-terminal base of an oligonucleotide does not have a measurable effect on the binding constant. The order of thermodynamic stabilities obtained by comparing the magnitudes of the oligonucleotide binding constants is consistent with earlier qualitative rankings obtained from comparison of the amounts of site-specific cleavage produced at a single oligonucleotide-EDTA·Fe concentration.[5a] An analysis of the thermodynamics of oligonucleotide binding at pH 7.0 revealed only a modest energetic

dependence on oligonucleotide length, but a substantial penalty (2.5–3.0 kcal·mol^{-1}) for a single internal base triplet mismatch. Moreover, prior energetic analyses of triple helices using short oligonucleotide duplex targets and acidic solution conditions[2,12] can now be compared with the affinity cleavage thermodynamic data for single site triple helix formation on large DNA at pH 7.0.

Because the synthesis of affinity cleaving molecules has proven to be general[4] and because affinity cleavage experiments can be performed under a wide range of conditions,[4,5] the use of quantitative affinity cleavage titration should allow further thermodynamic analyses of DNA binding by oligonucleotides or other ligands under solution conditions where experiments involving oligonucleotide duplexes or competitive footprinting agents may not be successful.

Experimental Section

Materials. Oligonucleotides **1-6**, labeled duplex DNA, and buffered solutions were prepared as described in Chapter One.

Quantitative Affinity Cleavage Titrations. Affinity cleavage titrations were prepared essentially as described in Chapter One.

Quantitative DNase I Footprint Titrations. DNase I footprinting experiments were run essentially as described,[11] in a manner analogous to the quantitative affinity cleaving experiment. Only the differences are described below. Oligonucleotide-EDTAs were reconstituted with $InCl_3$ rather than $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$. The stock solution of radiolabelled DNA in Footprint buffer was prepared from 0.5 M Tris-acetate buffer at pH 7.0, 2.5 M NaCl , 1.5 M $MgCl_2$, 1.5 M $CaCl_2$, calf thymus DNA at a concentration of 2.0 mM in bp, 10 mM spermine·4HCl at pH 7.0, approximately 300,000 cpm 5'-end labelled target DNA, and enough H_2O to bring the total volume to 1.36 mL. Following equili-

bration of the oligonucleotide with the target duplex, DNase I and the non-specific oligonucleotide $(dA)_{10}$ (both from Pharmacia LKB), which was used to maintain uniform DNase I reactivity, were added and the digestion allowed to proceed at 24 °C for 10 min. Final reaction conditions in 100 μL Footprint buffer were 50 mM Tris-acetate at pH 7.0, 20 mM NaCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 1 mM spermine, 0.12 mM-bps calf thymus DNA, 1 μM $(dA)_{10}$, approximately 17,500 cpm labelled duplex, and 0.5 units DNase I. Reaction work-up and electrophoresis were performed as above.

Affinity Cleavage Titration Fitting Procedure. Titration binding isotherms were constructed and equilibrium association constants were determined as described in Chapter One.

Footprint Titration Fitting Procedure. The analysis of DNase I footprint titrations was performed according to the previously described method.[11] Briefly, θ_{app} was determined using the following equation:

$$\theta_{app} = 1 - \frac{I_{tot}/I_{ref}}{I^{o}_{tot}/I^{o}_{ref}} \qquad (1)$$

where I_{tot} and I_{ref} are defined as above, and $I_{tot}°$ and $I_{ref}°$ indicate the amount of digestion at the target site and at the reference site in a DNase I control to which no oligonucleotide was added, respectively. The ($[\mathbf{O}]_{tot}$, θ_{app}) data points were fit by minimizing the difference between θ_{app} and θ_{fit}, using eq 2, where θ_{min} and θ_{max} represent the experimentally observed values of the degree of saturation when $\theta = 0$ and $\theta = 1$, respectively:

$$\theta_{fit} = \theta_{min} + (\theta_{max} - \theta_{min}) \cdot \frac{K_T[\mathbf{O}]_{tot}}{1 + K_T[\mathbf{O}]_{tot}} \qquad (2)$$

By making use of the formula for propagation of errors (see Chapter One), we can dervive the following equation expressing the uncertainty in θ_{app} obtained from footprint titration:

$$\varepsilon_{\theta_{app}} = 2\rho\theta_{app} \quad (3)$$

where ρ is the standard deviation in the average number of counts loaded per lane. The error for each data point was employed for analysis of the goodness of fit using the reduced χ^2 criterion (see Chapter One).

References and Notes

(1) (a) Ross, P. D.; Scruggs, R. L. *Biopolymers* **1965**, *3*, 491-496. (b) Krakauer, H.; Sturtevant, J. M. *Biopolymers* **1968**, *6*, 491-512. (c) Cassani, C.; Bolum, F. J. *Biochemistry* **1969**, *8*, 3928-3936.

(2) (a) Plum, E. G.; Park, Y.-W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436-9440. (b) Pilch, D. S.; Levenson, C.; Shafer, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 1942–1946. (c) Pilch, D. S.; Brousseau, R.; Shafer, R. H. *Nucleic Acids Res.* **1990**, *18*, 5743–5750. (d) Manzini, G.; Xodo, X. E.; Gasparotto, D.; van der Marel, G. A.; van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833-843. (e) Mooren, M. M. W.; Pulleyblank, D. E.; Wijmenga, S. S.; Blommers, M. J. J.; Hilbers, C. W. *Nucleic Acids Res.* **1990**, *18*, 6523-6529. (f) Xodo, L. E.; Manzini, G.; Quadrifoglio, F. *Nucleic Acids Res.* **1990**, *18*, 3557-3564. (g) Roberts, R. W.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9397–9401.

(3) Pilch, D. S.; Levenson, C.; Shafer, R. H. *Biochemsitry* **1991**, *30*, 6083–6087.

(4) (a) Dervan, P. B. *Science* **1986**, *232*, 464-471. (b) Dervan, P. B. In *Nucleic Acids and Molecular Biology, Vol. 2* ; F. Eckstein and D. M. J. Lilley, Eds.; Springer-Verlag: London, 1986; pp 49–64.

(5) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650. (b) Strobel, S. A.; Moser, H.E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, 110, 7927–7929. (c) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061. (d) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 7286–7287. (e) Griffin, L. C.; Dervan, P. B.*Science* **1989**, *245*, 967-971. (f) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75. (g) Distefano, M. D.; Shin, J. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 5901–5902. (h) Koh, J. S.; Dervan, P. B. *J. Am.*

Chem. Soc., **1992**, *114*, 1470–1478. (i) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochemistry* **1992**, *31*, 2829–2834.

(6) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(7) (a) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435-2437. (b) Beal, P. A.; Dervan, P. B. *J. Am. Chem. Soc.*, in press.

(8) Maher III, L. J.; Dervan, P. B.; Wold, B. *Biochemistry* **1990**, *29*, 8820-8826.

(9) Additional quantitative affinity cleavage titration experiments have revealed no changes in the shapes of binding isotherms or measured values of K_T when the cleavage reactions are allowed to proceed to yields of 20 – 25%. This independence likely arises from a combination of the relatively long triplex lifetimes (see ref. 2b) and the high specificity of the oligonucleotide-EDTAs for the target sequence.

(10) Quantitative affinity cleavage titrations performed in Footprint buffer (10 mM Na^+, 20 mM Mg^{2+}, 20 mM Ca^{2+}, 1 mM spermine·4HCl, 50 mM Tris·OAc, pH 7.0) yielded apparent association constants identical to those obtained from experiments using Association buffer (100 mM Na^+, 1 mM spermine·4HCl, 50 mM Tris·OAC, pH 7.0).

(11) (a) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132-181. (b) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8462-8466. (c) Senear, D. F.; Brenowitz, M.; Shea, M. A.; Ackers, G. K. *Biochemistry* **1986**, *25*, 7344-7354.

(12) Breslauer, K. H.; Frank, R.; Blöcker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746–3750.

(13) Singleton, S. F.; Dervan P. B., unpublished observations.

(14) Belotserkovskii, B. P.; Veselkov, A. G.; Filipov, S. A.; Dobrynin, V. N.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1990**, *18*, 6621-6624.

(15) Mergny, J.-L.; Sun, J.-S.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791-9798.

(16) Macaya, R. F.; Gilbert, D. E.; Malek, S.; Sinsheimer, J. S.; Feigon, J. *Science* **1991**, *254*, 270–274.

(17) (a) Nelson, J. W.; Martin, F. H.; Tinoco, I., Jr. *Biopolymers* **1981**, *20*, 2509-2531. (b) Thibanyenda, N.; De Bruin, S.; Haasnoot, C. A. G.; van der Marel, G. A.; van Boom, J. H.; Hilbers, C. W. *Eur. J. Biochem.* **1984**, *139*, 19–27. (c) Aboul-Ela, F.; Koh, D.; Tinoco, I., Jr. *Nucleic Acids Res.* **1985**, *13*, 4811-4824.

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

(19) (a) Dervan, P. B. In *Human Genome Initiative and DNA Recombination*; R. H. Sarma and M. H. Sarma, Eds.; Adenine Press: Guilderland, NY, 1989; pp 37-49. (b) Helene, C.; Thuong, N. T.; Saison-Behmoaras, T.; Francois, J.-C. *Trends Biochem. Sci.* **1989**, *7*, 310-315. (c) Dervan, P. B. In *Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression*; J. S. Cohen, Ed.; Macmillan Press, Ltd.: London, 1989; pp 197–210. (d) Sigman, D. S.; Chen, C. B. *Annu. Rev. Biochem.* **1990**, *59*, 207-236. (e) Nielsen, P. E. *Bioconjugate Chem.* **1991**, *2*, 1-12.

CHAPTER THREE

Influence of pH on the Equilibrium Association Constants for Oligonucleotide-Directed Triple Helix Formation at Single DNA Sites

The text of this chapter is taken from a published manuscript that was coauthored with Professor Peter B. Dervan.

(S. F. Singleton & P. B. Dervan, *Biochemistry* **1992**, *31*, 10995-11003.)

Introduction

The stabilities of local pyrimidine·purine·pyrimidine triple-helical complexes at single sites on relatively large DNA (> 200 bp) are dependent on length, sequence composition, and base sequence mismatches (Moser & Dervan, 1987; Mergny et al., 1991; Roberts & Crothers, 1991; Singleton & Dervan, 1992; Kiessling et al., 1992). Moreover, the solution conditions, including pH, temperature, and the identity and concentration of cations, contribute to the absolute stabilities of these complexes (Moser & Dervan, 1987; Maher et al., 1990). Triple-helical nucleic acids containing cytosines in the third strand are stable in acidic to neutral solutions, but they dissociate as the pH increases (Lee et al., 1979; Moser & Dervan, 1987; Lyamichev et al., 1988; Povsic & Dervan, 1989; Maher et al., 1990; Plum et al., 1990). Apparently, an important factor in the stability of triple-helical

complexes is the protonation at cytosine *N*3 in the third strand which enables the formation of two hydrogen bonds and provides for partial charge neutralization.

The potential functional application of oligonucleotide-directed triple helix formation to the specific repression of eukaryotic gene expression *in vivo*, where the pH is strictly regulated, has stimulated recent efforts to stabilize the protonated C+GC triplets by functional modification of the oligonucleotide (Povsic & Dervan, 1989) and to design nucleosides capable of GC base pair recognition without protonation (Ono et al., 1991; Koh & Dervan, 1992, Krawczyk et al., 1992). For example, replacing the cytosines in the third strand with 5-methylcytosines (m^5C's) increased both the apparent stabilities and upper pH-limits of complexes resulting from oligonucleotide-directed triple helix formation (Povsic & Dervan, 1989). This substitution, in which a single –H to $-CH_3$ change has been affected on the opposite side of the heterocycle from the hydrogen-bonding sites (Figure 3.1), maintains the same array of potential hydrogen bond donors and acceptors found in cytosine, and results in minimal perturbation of the electronic structure (Sober, 1970). Despite the apparent similarity of C and m^5C, the incorporation of m^5C has proven to be important in the use of oligonucleotides as inhibitors of site-specific DNA binding by proteins near pH 7 (Maher et al., 1989, 1990; Strobel & Dervan, 1991; Strobel et al., 1991; Collier et al., 1991; Grigoriev et al., 1992; Duval-Valentin et al., 1992).

It has been demonstrated that the substitution of m^5C for C increases the thermal stabilities of both polynucleotide (Lee et al., 1984) and oligonucleotide triplexes (Plum et al., 1990). More recently, the heats of denaturation of triple-helical complexes formed by an oligonucleotide hairpin duplex and oligonucleotide third strands containing C or m^5C have been reported (Xodo et al., 1991). We are interested in the measurement of equilibrium association constants for

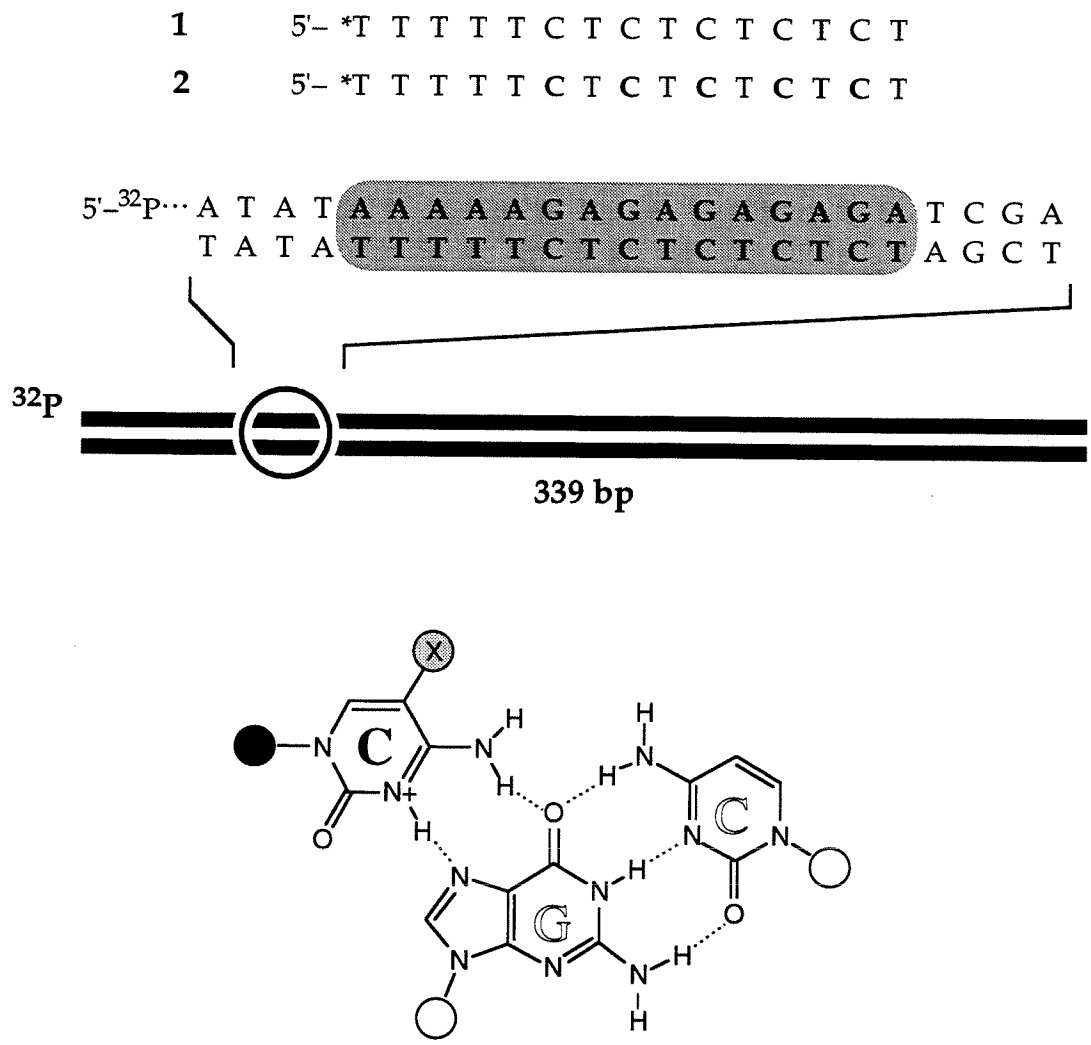


Figure 3.1. The sequences of oligonucleotides **1** and **2** are shown, where T* indicates the position of thymidine-EDTA·Fe and bold type indicates 5-methylcytidine nucleosides (X = CH_3, below). The sequence of the 15-bp target site (shaded box) is shown below, along with the relative position of the target site within the radiolabeled duplex.

oligonucleotide-directed triple helix formation at individual DNA sites as a first step in understanding the factors contributing to the stability and specificity of triple-helical complexes. We report the influence of pH on the energetics of oligonucleotide-directed triple helix formation using quantitative affinity cleavage titration (Singleton & Dervan, 1992). By measuring the amounts of site-specific cleavage produced by two oligonucleotide-EDTA·Fe conjugates, 5′-TTTTTCTCTCTCTCT-3′ (**1**, Figure 3.2) and 5′-TTTTTm^{5}CTm^{5}CTm^{5}CTm^{5}CT-m^{5}CT-3′ (**2**), over a range of concentrations spanning four orders of magnitude, we have constructed empirical titration binding isotherms and determined equilibrium constants for site-specific DNA binding at 22 °C (100 mM Na^+, 1 mM spermine) over the pH range of 5.8 – 7.6. These affinity cleavage titration experiments have been used to measure the effects of pH and C → m^{5}C substitution on the stability of a local triple-helical complex at a single site on large DNA.

Materials and Methods

Oligonucleotides **1** and **2**, labeled duplex DNA, and buffered solutions were prepared as described in Chapter One. Affinity cleavage titrations were prepared essentially as described in Chapter One. Titration binding isotherms were constructed and equilibrium association constants were determined as described in Chapter One.

Results

Association Constants as a Function of pH. Previous experiments demonstrated that the equilibrium association constant for the binding of an oligonu-

cleotide-EDTA·Fe to an individual DNA site can be measured using quantitative affinity cleavage titration (Singleton & Dervan, 1992). To assess the influence of pH on the energetics of oligonucleotide-directed triple helix formation, the equilibrium association constants for oligonucleotide-EDTA conjugates binding to a 339-bp DNA duplex containing a 15-bp target sequence (Figure 3.1) were measured in buffers of different pH. At a given pH, ^{32}P-5′-end labeled DNA (< 5 pM) and various concentrations of oligonucleotide-EDTA·Fe (100 pM – 10 μM) were mixed in Association buffer (100 mM Na^+, 1 mM spermine, and 50 mM MES or MOPS) at 22 °C. Only after the association reactions had been allowed to reach equilibrium over 24 h was DTT (1 mM final concentration) added to initiate the EDTA·Fe-mediated cleavage chemistry. The reactions were allowed to proceed for 8 h at 22 °C, allowing a maximum site-specific cleavage yield of about 15%. The products were separated by PAGE under strand-denaturing conditions. A storage phosphor autoradiogram from a typical experiment using of oligonucleotide-EDTA·Fe **1** (Figure 3.1) at pH 7.6 is shown in Figure 3.2.

The amounts of radiolabeled DNA in the bands at the target cleavage site (Figures 2.1 and 3.2) and at a reference site (Figure 3.2) were measured from storage phosphor autoradiograms and I_{site} for each $[\mathbf{O}]_{tot}$ was calculated using eq 36 (Chapter One). The ($[\mathbf{O}]_{tot}$, I_{site}) data points were fitted using a nonlinear least-squares method and eq 38 (Chpater One), with K_T and I_{sat} as adjustable parameters. The data points obtained for **1** at pH 7.6, 6.6, 6.2, and 5.8 were averaged from several experiments and are plotted along with average best-fit titration binding isotherms in Figures 3.3A and 3.3C–E. The average data points and best fit curve for **1** at pH 7.0 were published previously (Singleton & Dervan, 1992) and have been reproduced in Figure 3.3B. The corresponding averaged data and best-fit isotherms for **2** are presented in Figures 3.3F-J. The mean K_T values

Figure 3.2 Gray-scale representation of a storage phosphor autoradiogram of an 8% denaturing polyacrylamide gel used to separate the cleavage products generated by **1** during a quantitative affinity cleavage titration experiment performed in Association buffer (100 mM Na^+, 1 mM spermine, 50 mM MOPS) at 22 °C and pH 7.6. There are 224 levels of gray representing a 50-fold change in the signal from the lowest (120 arbitrary units, white) to highest (6,000 arbitrary units, black) intensities. The bracket on the right side of the autoradiogram encloses the nucleotide positions of the 15-bp duplex target site. The bar drawn on the right of the autoradiogram indicates the bands used to measure I_{tot} (black boxes), and the bands used to measure I_{ref} (gray box). (Lanes 1–16) DNA affinity cleavage products produced by **1** at various concentrations: no oligonucleotide (lane 1); 10 nM (lane 2); 20 nM (lane 3); 40 nM (lane 4); 80 nM (lane 5); 100 nM (lane 6); 200 nM (lane 7); 400 nM (lane 8); 800 nM (lane 9); 1.0 μM (lane 10); 2.0 μM (lane 11); 4.0 μM (lane 12); 8.0 μM (lane 13); 10 μM (lane 14); 20 μM (lane 15); 40 μM (lane 16). (Lane 17) Products of an adenine-specific sequencing reaction. (Lane 18) Intact 5'-labeled duplex.

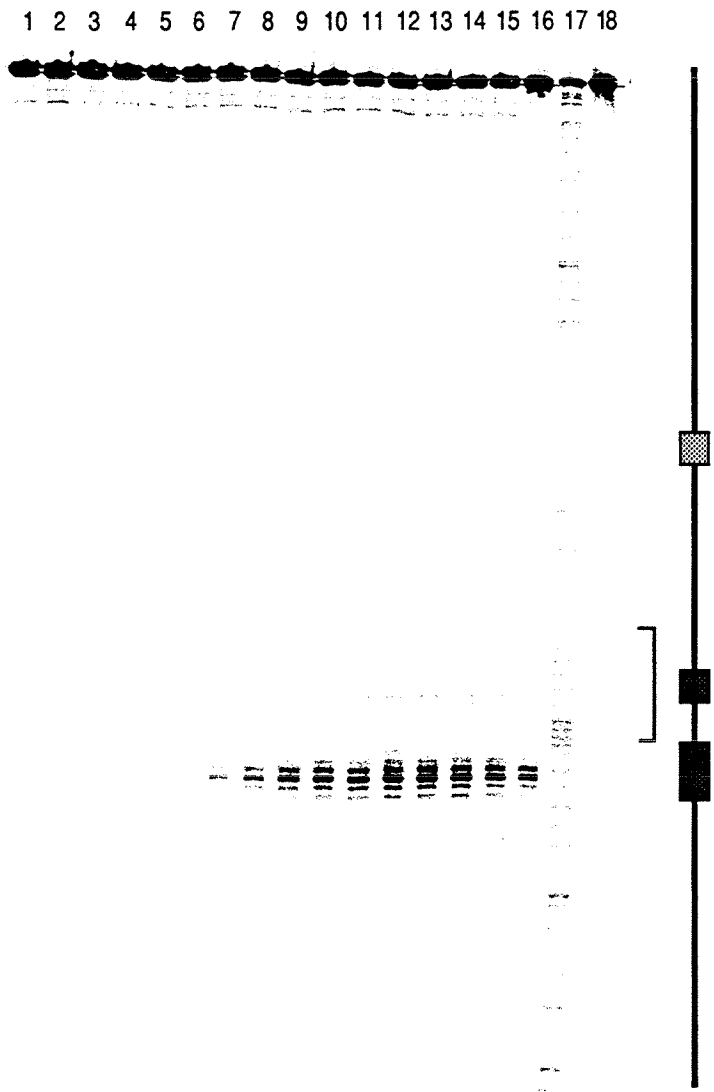
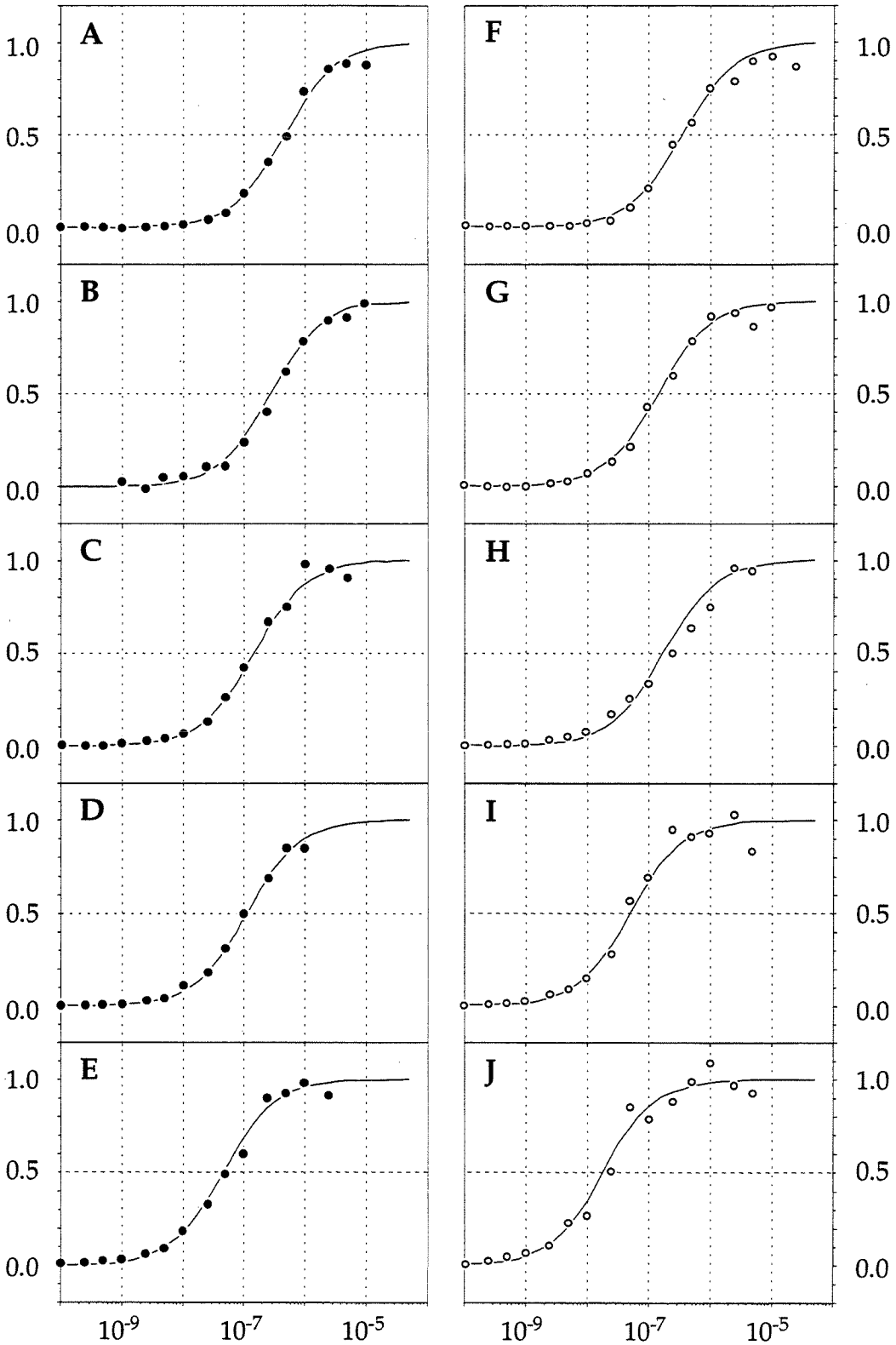

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Figure 3.3. Data for quantitative affinity cleavage experiments involving oligonucleotides **1** (•) or **2** (○) in Association buffer (22 °C) at pH 7.6 (A, F), 7.0 (B, G), 6.6 (C, H), 6.2 (D, I), and 5.8 (E, J). The data points represent the average site-specific cleavage signal intensities from three to five experiments. The sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T for **1** and **2** (Table 3.1) and eq 38 (Chapter One). The data points were normalized using I_{sat} from each experiment and the binding curves were subsequently normalized using $I_{sat} = 1$ for eq 38 (Chapter One).


A
B
C
D
E
F
G
H
I
J
1.0
0.5
0.0
10^{-9}
10^{-7}
10^{-5}

obtained from the analyses of these data are contained in Table 3.1. The isotherms (Figure 3.3) and the K_T values (Table 3.1) measured at different pH's demonstrate that the association of both **1** and **2** are driven towards triple helix formation at lower pH. The trend of increasing triple helix stability with increasingly acidic pH, which is consistent with the presumed requirement for cytosine *N*3 protonation, is more clearly indicated by the plots of K_T *versus* pH displayed in Figure 3.4.

Table 3.1. Influence of pH on K_T and ΔG_T at 22 °C.[a]

	1 (C, T)		**2** (m^5C, T)	
	K_T	ΔG_T	K_T	ΔG_T
pH	(M^{-1})	(kcal·mol^{-1})	(M^{-1})	(kcal·mol^{-1})
7.6	2.2 (±0.4) x 10^6	–8.5 (±0.2)	2.8 (±1.0) x 10^6	–8.7 (±0.4)
7.0	3.7 (±1.1) x 10^6 [b]	–8.9 (±0.3)[b]	7.1 (±2.8) x 10^6	–9.2 (±0.4)
6.6	7.0 (±0.9) x 10^6	–9.2 (±0.1)	5.6 (±1.3) x 10^6	–9.1 (±0.2)
6.2	8.8 (±2.6) x 10^6	–9.4 (±0.3)	2.0 (±0.4) x 10^7	–9.8 (±0.2)
5.8	2.2 (±0.5) x 10^7	–9.9 (±0.2)	5.5 (±1.7) x 10^7	–10 (±0.3)

[a] Values reported in the table are mean values measured from quantitative affinity cleavage experiments performed in Association buffer (100 mM Na^+, 1 mM spermine, 50 mM MOPS or MES buffer) at the indicated pH. [b] These values were taken from previously published experiments performed under the same conditions (Singleton & Dervan, 1992).

An expression for the pH-dependence of triple helix formation follows from a general theoretical description of a pH-dependent equilibrium (Edsall & Gutfreund, 1983; Lyamichev et al., 1988). The association constant may be factored

into two terms, one of which arises from the stability constant for the unprotonated triple-helical complex and the other from stabilization of the complex resulting from equilibrium protonation of the complex:

$$K_T = K_T^o \cdot \left(1 + 10^{pK_{aT} - pH}\right)^p \qquad (1)$$

where $K_T°$ is the association constant for the unprotonated oligonucleotide, p is the number of potential protonation sites (*i.e.*, the number of cytosine residues in the oligonucleotide), and K_{aT} is the mean acid dissociation constant of those sites (*i.e.*, C+GC triplets). The (pH, K_T) data points in Figure 3.4 were fitted to eq 1 with $p = 5$, and $K_T°$ and pK_{aT} as adjustable parameters. The best-fit curves to the data for both **1** and **2** are plotted in Figure 3.4.

Table 3.2. Parameters Describing the pH-Dependence of K_T at 22 °C.[a]

oligo	site[b]	$K_T°$ (M^{-1})	pK_{aT}	K_{aO}/K_{aT} [c]	$\Delta\Delta G_a$ [d] (kcal·mol^{-1})
1	C·GC	3.8×10^6	5.5	13	–1.5
2	m^{5}C·GC	5.4×10^6	5.7	20	–1.8

[a] The parameters were obtained by fitting the (K_T, pH) data points shown in Figure 5 to eq 1. [b] The site of protonation formed upon binding of the oligonucleotide listed in the first column to the double-helical target site. [c] The ratio of acid dissociation constants is obtained using $pK_{aO} = 4.4$ as described in the text (eq 4). [d] The difference in free energies for protonating a free nucleoside versus protonating a base triplet. The value was calculated from $\Delta\Delta G_a = -RT\ln(K_{aO}/K_{aT})$.

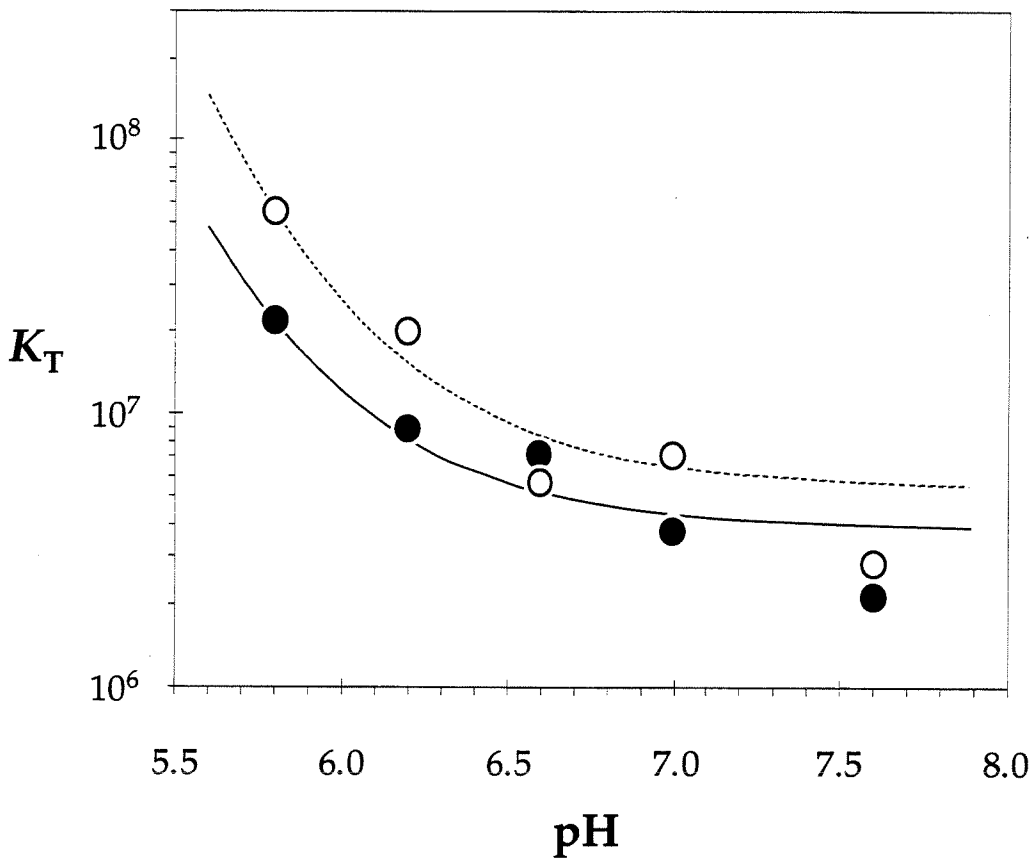


Figure 3.4. Plots of the mean values (logarithmic scale) of K_T for **1** (•) and **2** (◦) as a function of pH, and the nonlinear least-squares fitted curves (solid and dotted lines, respectively) generated using eq 1 with K_T° and pK_{aT} as adjustable parameters. The parameters obtained from the fitting procedure are reported in Table 3.2.

The parameters given by the fitting, summarized in Table 3.2, yield an estimated value of 5.5 for the pK_{aT} of the triplex helix formed upon binding of **1** to its double-helical target, and a value of 5.7 for the pK_{aT} of the triple helix with **2** bound. Interestingly, a pK_{aT} in this range suggests that the triple helix is not fully protonated near pH 7. A model of the triple helix in which the degree of protonation changes with pH is consistent with the large decrease in the lifetime of a triple-helical complex observed when the solution pH was raised from 6.8 to 7.2 (Maher et al., 1990).

Discussion

The Influence of pH on the Association Constant. The energetics of site-specific oligonucleotide-directed triple helix formation have been analyzed by quantitative affinity cleavage titration as a function of pH. Over the pH range of 5.8 – 7.6, K_T for **1** decreases by a factor of 10, while K_T for **2** decreases 20-fold over the same range. Least-squares fitting analysis of the K_T *versus* pH data afforded estimates of $pK_{aT} = 5.5$ and $K_T^\circ = 3.8 \times 10^6\ M^{-1}$ for the triple helix formed by association of **1** to the double-helical target site. The pK_{aT} value is near the semi-protonation points for oligonucleotide triplexes measured to be 5.6, 5.8, and 6.2 by ethidium bromide fluorescence (Callahan et al., 1991), UV absorption spectroscopy (Xodo et al., 1991), and CD spectroscopy (Callahan et al., 1991), respectively. A similar analysis revealed that $pK_{aT} = 5.7$ and $K_T^\circ = 5.4 \times 10^6\ M^{-1}$ for the triple-helical structure formed upon binding of **2**. This value is considerably lower than the value of 6.8 measured using a UV spectroscopic technique for an oligonucleotide triplex (Xodo et al., 1991). The physical model characterized by eq 1 describes the effect of pH on a global equilibrium process rather than an intrinsic pK_a at a single atom as could be measured using NMR spectroscopy. The

fact that other pH-dependent processes, including spermine binding and conformational changes in both the duplex and triplex, may be occurring under the conditions described herein adds complexity to the analysis. The polyelectrolyte behavior of the 339-bp double-helical DNA molecule containing the target site may result in a buffering of the pH of the local solution environment near the duplex (Slonitskii & Kuptsov, 1988). Such behavior would likely result in a decreased dependence of the triplex stability on bulk pH near pH 7 and an increased value for the apparent $K_T°$. Despite the complexity inherent in this analysis, the apparent pK_{aT} values measured using quantitative affinity cleavage titration allow an accurate characterization of the pH-dependence of oligonucleotide-directed triple helix formation on large DNA.

A quantitative estimate of the stabilization of the triple helix resulting from protonation was derived from the pH-dependence data and the following thermodynamic cycle:

$$\begin{array}{ccc} \mathbf{OH}_m + \mathbf{D} + \mathrm{H}^+ & \underset{\Delta G_1}{\overset{K_{Tm}}{\rightleftharpoons}} & \mathbf{T}_m + \mathrm{H}^+ \\ K_{aO} \Big\| \Delta G_3 & & K_{aT} \Big\| \Delta G_4 \\ \mathbf{OH}_{m+1} + \mathbf{D} & \underset{\Delta G_2}{\overset{K_{Tm+1}}{\rightleftharpoons}} & \mathbf{T}_{m+1} \end{array} \qquad (2)$$

where $\mathbf{OH}_m$ and $\mathbf{OH}_{m+1}$ represent oligonucleotides with m and m+1 protonated cytosines, respectively, $\mathbf{T}_m$ and $\mathbf{T}_{m+1}$ represent triplexes with m and m+1 protonated C+GC base triads, respectively, K_{aO} and K_{aT} are the acid dissociation con-

stants for the oligonucleotide and triplex, respectively, and K_{Tm} and K_{Tm+1} are the association constants for triple helix formation involving oligonucleotides with m and m+1 protonated cytosines, respectively ($0 \leq m \leq p$). Because the equilibria in eq 2 represent a cycle, the following relationships between the free energies and the equilibrium constants may be written:

$$\Delta G_4 - \Delta G_3 = \Delta G_2 - \Delta G_1 \tag{3}$$

$$\frac{K_{aO}}{K_{aT}} = \frac{K_{Tm+1}}{K_{Tm}} \tag{4}$$

Equations 10 and 11 equate the change in the pK_a of cytosine *N*3 upon binding of the oligonucleotide to its double-helical target site with the stabilization of the triple helix resulting from the addition of one proton to an acceptor site in the triplex to form a C+GC base triplet. Xodo et al. (1991) have determined the pK_a's for both of two single-stranded pyrimidine oligonucleotides containing C or m^5C residues to be 4.4 ± 0.2. Because the pK_as of cytosine residues in a single-stranded oligonucleotide are not expected to be strongly perturbed by the adjacent nucleotides, the ratio of equilibrium constants (eq 4) for **1** and **2** were estimated using pK_{aO} = 4.4 and included in Table 3.2.

The K_{Tm+1}/K_{Tm} ratios suggest that the protonated state of cytosine is stabilized by 1.5 kcal·mol^{-1} when the proton is placed in the acceptor site between cytosine *N*3 and guanine *N*7 in the triple helix (Figure 3.1), while *N*3-protonated 5-methylcytosine is stabilized by 1.8 kcal·mol^{-1} in the triplex. At least two sources give rise to this stabilization. Clearly, interactions such as hydrogen bonding between the cytosine heterocycle and proximal bases will stabilize the protonated state of cytosine. Charge-charge interactions may also play a key role in this

phenomenon. Investigations by Callahan et al. (1991) on the pH-dependent formation of oligonucleotide triplexes suggest that the stabilization energy resulting from Coulombic attraction between the positively charged *N*3 of a protonated cytosine and an anionic phosphate group of the duplex may be up to 0.5 kcal·mol^{-1}. The dependence of the stability of the triple helix on the solution pH is remarkable. The addition of one proton to a C·GC acceptor site in a local triple-helical structure stabilizes the structure by 1.5 kcal·mol^{-1}, while the addition of one C+GC and one T·AT triplet to the structure at pH 7.0 confers a stabilization of only 0.5 kcal·mol^{-1} (Singleton & Dervan, 1992).

The Influence of Substitution at C5 of Cytosine. An analysis of the stabilization conferred by C → m^{5}C substitution on oligonucleotide-directed triple helix formation is achieved by comparing the K_T values for **1** and **2**. The plot of K_T *versus* pH (Figure 3.4) demonstrates that, for the binding of **1** over the range of pH 5.8 to pH 7.6, the use of oligonucleotide **2** extends the pH range for forming a triple helix of equivalent stability by 0.5 – 1.0 units. This finding is consistent with earlier studies comparing relative cleavage efficiencies produced at single concentrations of oligonucleotide-EDTA·Fe (Povsic & Dervan, 1989). The methyl substituent at *C*5 of cytosine increases the apparent pK_{aT} by 0.2 units and increases $K_T°$ by a factor of 1.5. Although the uncertainty in the fitted pK_{aT} values is likely similar in magnitude to the difference between them (ΔpK_{aT} = 0.2), the increase in the pK_{aT} afforded by the C→ m^{5}C substitution is modest compared to that observed (ΔpK_{aT} = 1.0) for the triplexes formed upon binding of 5′-CTTCCTC-CTCT-3′ or 5′-m^{5}CTTm^{5}Cm^{5}CTm^{5}Cm^{5}CTm^{5}CT-3′ to an oligonucleotide duplex target (Xodo et al., 1991). Among the possible explanations for the apparent difference in ΔpK_{aT} are (i) the methods used to measure pK_{aT}; (ii) the sequence-dependence of pK_{aT}; and (iii) a combination of both. Recently, Hampel et al. (1991)

demonstrated that the method used to measure the apparent pK_aT of the polymeric triplex $[d(TC)]_n \cdot [d(GA)]_n \cdot [d(CT)]_n$ strongly affected the measured value. In fact, the pK_aT determined by adding base to the triplex formed at acidic pH was ~ 1.5 units higher than that measured under equilibrium conditions. Presumably, once formed at acidic pH, the triplex remains as a "metastable" complex for a length of time dependent on the dissociation rate of the complex. While it has been shown that 5-methylcytosine is required for a 21-mer oligonucleotide to protect an overlapping restriction endonuclease site from cleavage by the enzyme in a kinetic assay (Maher et al., 1990), data is not yet available to compare the dissociation rates of triplexes containing cytosines or 5-methylcytosines in the third strand. An alternative explanation for the observed differences in ΔpK_aT is the possibility of a sequence-dependent pK_aT. The earlier work (Xodo et al., 1991) involved triple-helical complexes wherein four of the six X^5C+GC triplets were flanked on one side by X^5C+GC triplets and a fifth one is at the 5'-terminus, while the complexes described here contain only internal X^5C+GC triplets flanked by T·AT triplets. Because the stabilities of triplexes containing C+GC triplets have been shown to be dependent on the base triplets flanking the C+GC triplets (Kiessling et al., 1992), we anticipate that the pK_aT values for these complexes may be sequence-dependent and that the effect of the functional group at C5 of the cytosine heterocycle may also be influenced by the nearest neighbor triplets.

A comparison of the ΔG_T values for formation of the triple-helical structures involving **1** and **2** (Table 3.1) reveals that substitution of a methyl group at C5 of cytosine stabilizes the triple-helical complex by 0.1 – 0.4 kcal·mol^{-1} over the pH range of 5.8 – 7.6. Replacement of the hydrogen atom normally attached to C5 of cytosine with an electron-donating methyl substituent is expected to slightly in-

crease the pK_a of the base in both the free and complexed oligonucleotides, making protonation of the m^{5}C residue easier while reducing its hydrogen bond donating ability from the protonated *N*3; however, the differences in the pK_a's of cytosine and 5-methylcytosine in nucleosides and oligonucleotides are measured to be ≤ 0.1 units (Sober, 1970; Xodo et al., 1991). Based on differential scanning calorimetric measurements which demonstrated that a C → m^{5}C substitution in the third strand results in a decreased denaturation enthalpy, Xodo et al. (1991) reasoned that methylation induces a release of water molecules from the helix, thereby contributing a positive entropy change. While differences in solvation in the complexes formed upon binding of **1** and **2** may play a role in the differential stabilities of the two triple-helical structures, it is interesting that the free 5-methyl-cytosine base is more soluble in water than cytosine (Sowers et al., 1987b).

In addition to changes in hydrogen bonding and solvation, another likely source of the stabilizing influence of a *C*5-methyl group may arise from increased base stacking interactions. The importance of base stacking in nucleic acid helical complexes was demonstrated by early investigations concerning the role of organic solvents in denaturing double-helical DNA (Hanlon, 1966). These studies implicated dispersion forces, a key component of stacking, as a stabilizing factor in the double helix. Stacking energies, which arise from dispersion and induced dipole interactions (DeVoe & Tinoco, 1962; Ornstein et al., 1978), have been shown to correlate reasonably well with the polarizabilities of the interacting heterocycles (Broom et al., 1967; Ts'o, 1968; Nakano & Igarichi, 1970; Lawaczek & Wagner, 1974; Sowers et al., 1987b; Darzynkiewicz & Lönnberg, 1989). Moreover, methylation of the stacked bases increases the interaction energies (Broom et al., 1967; T'so, 1968; Nakano & Igarichi, 1970; Sowers et al., 1987b). In fact, sub-

stitution of a methyl group at C5 of cytosine increases its molecular polarizability (Miller & Savchik, 1979; Sowers et al., 1987b) and increases the free energy of stacking of cytosine with adenine in water by 0.2 kcal·mol^{-1} (Sowers et al., 1987b). Models of stacked base triplets constructed using A'-form (Figure 3.5A) or B-form (Figure 3.5B) helical turn angles suggest that the C5-substituent is positioned to play a role in stacking with an adjacent pyrimidine in the third strand of a triple helix. The observation that local triple-helical structures containing m^{5}C in the third strand are stabilized relative to those containing C in the third strand by a similar amount over a range of pH, including those where the triplexes may not be fully protonated, is consistent with the interpretation that the methyl group may stabilize the complex without *directly* influencing the protonation event (*e.g.*, through increased stacking energy).

In addition to an equilibrium thermodynamic stabilization, the C5-methyl substitution may have a dynamic effect on the triple helical structure. It has been recognized that protonation of pyrimidines diminishes pyrimidine-pyrimidine stacking interactions in the crystalline state (Bugg et al., 1971; Saenger, 1984; Padmaja et al., 1991). This proton-induced destacking may account for the observation that third-strand cytosine *N*3 protons exchange more rapidly with solvent and that their NMR signals disappear more rapidly with increased temperature than thymine *N*3 protons (Rajagopal & Feigon, 1989). Furthermore, an unprotonated C·GC base triplet may have a shorter intrinsic lifetime than a T·AT base triplet, since the former has only a single hydrogen bond. Increased stacking through the interaction of the C5-methyl group with the polarizable π-system of the neighboring heterocycle (Figure 3.5) may hold an unprotonated m^{5}C base in better position for maintaining its lone hydrogen bond and may reduce any tendency for a protonated m^{5}C heterocycle to unstack. A similar dynamic phe-

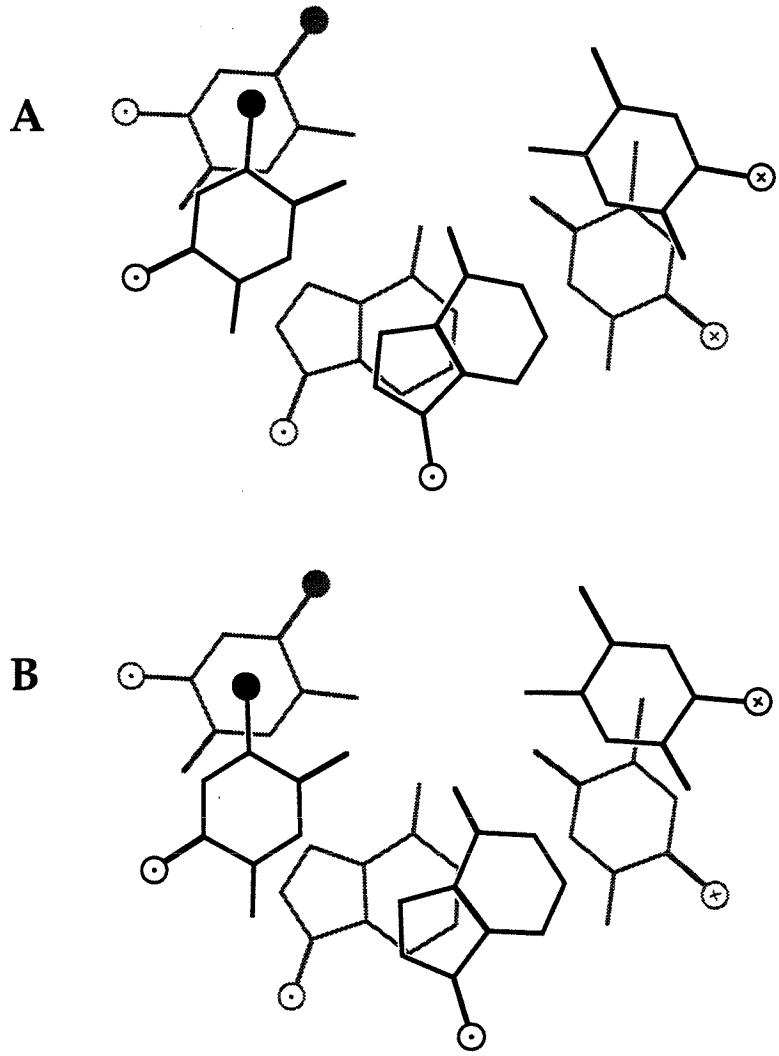


Figure 3.5. (A) Base stacking configurations drawn using the idealized helical parameters generated from diffraction data for fibers of the polymeric triplex $(T{\cdot}AT)_n$(Arnott et al., 1976). The C1' atoms of the deoxyribose sugars are represented by the open circles. The 5'→3' polarity of the strands is indicated by the symbols ⊗ (into the plane of the page) and ⊗ (out of the plane of the page). The filled circles indicate the positions of the methyl groups at C5 of thymines in the third strand. (B) Base stacking model generated as above except that the helical turn angle per residue for B-form nucleic acids was used (Saenger, 1984).

nomenon has been observed in a comparison of the NMR behavior of double-helical DNAs containing thymine or 5-fluorouracil bases (Sowers et al., 1987a). A complete elucidation of the equilibrium and dynamic influences of the C → m^5C substitution on a triple-helical structure must await more direct physical characterization.

Conclusions

Quantitative affinity cleavage titration has been used to measure the equilibrium constants for the formation of a local triple-helical complex a single site within large DNA over the pH range of 5.8 – 7.6 at 22 °C (100 mM Na^+ and 1 mM spermine). The triple helix formed upon binding of **1** to the double-helical target site is stabilized by 1.4 kcal·mol^{-1} as the pH is lowered from 7.6 to 5.8. Replacement of the five cytosine residues in the third strand with 5-methylcytosine residues stabilizes the triple-helical DNA complex by 0.1 – 0.4 kcal·mol^{-1} over this pH range, resulting in a modest increase in the pK_{aT} of the triple helix from 5.5 to 5.7. The importance of third-strand protonation and the potential contributions of base stacking forces to the energetics of pyrimidine oligonucleotide-directed triple helix formation offer intriguing possibilities for the direction of synthetic efforts toward novel nucleosides as well as for experiments to elucidate further the structure and stability of triple-helical complexes.

References

Arnott, S., Bond, P. J., Selsing, E., & Smith, P. J. C. (1976) *Nucleic Acids Res. 3*, 2459-2470.

Broom, A. D., Schweizer, M. P., & Ts'o, P. O. P. (1967) *J. Am. Chem. Soc. 89*, 3612-3622.

Bugg, C. E., Thomas, J. M., Sundaralingam, M., & Rao, S. T. (1971) *Biopolymers 10*, 175-219.

Callahan, D. E., Trapane, T. L., Miller, P. S., Ts'o, P. O. P., & Kan, L.-S. (1991) *Biochemistry 30*, 1650-1655.

Collier, D. A., Thuong, N. T., & Helene, C. (1991) *J. Am. Chem. Soc. 113*, 1457-1458.

Darzynkiewicz, E., & Lönnberg, H. (1989) *Biophys. Chem. 33*, 289-293.

DeVoe, H., & Tinoco, I. (1962) *J. Mol. Biol. 4*, 500-517.

Duval-Valentin, G., Thuong, N. T., & Helene, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 504-508.

Edsall, J. T., & Gutfreund, H. (1983) *Biothermodynamics: The Study of Biochemical Processes at Equilibrium*, J. Wiley & Sons, New York.

Grigoriev, M., Praseuth, D., Robin, P., Hemar, A., Saison-Behmoaras, T., Dautry-Varsat, A., Thuong, N. T., Helene, C., & Harel-Bellan, A. (1992) *J. Biol. Chem. 267*, 3389-3395.

Hampel, K. J., Crosson, P. & Lee, J. S. (1991) *Biochemistry 30*, 4455-4459.

Hanlon, S. (1966) *Biochem. Biophys. Res. Commun. 23*, 861-867.

Kiessling, L. L., Griffin, L. C., & Dervan, P. B. (1992) *Biochemsitry 31*, 2829-2834.

Koh, J. S., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470-1478.

Krawczyk, S. H., Milligan, J. F., Wadwani, S., Moulds, C., Froehler, B. C., & Matteucci, M. D. (1992) *Proc. Natl. Acad. Sci. U. S. A. 89*, 3761-3764.

Lawaczeck, R., & Wagner, K. G. (1974) *Biopolymers 13*, 2003-2014.

Lee, J. S., Johnson, D. A., & Morgan, A. R. (1979) *Nucleic Acids Res. 6*,

Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603-6614.

Lyamichev, V. I., Mirkin, S. M., Frank-Kamenetskii, M. D., & Cantor, C. R. (1988) *Nucleic Acids Res. 16*, 2165-2178.

Maher, L. J., III, Wold, B., & Dervan, P. B. (1989) *Science 245*, 725-730.

Maher, L. J., III, Dervan, P. B., & Wold, B. (1990) *Biochemistry 29*, 8820-8826.

Mergny, J.-L., Sun, J.-S., Rougee, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J., & Helene, C. (1991) *Biochemistry 30*, 9791-9798.

Miller, K. J., & Savchik, J. A. (1979) *J. Am. Chem. Soc. 101*, 7206-7213.

Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645-650.

Nakano, N. I., & Igarashi, S. J. (1970) *Biochemistry 9*, 577-583.

Ono, A., Ts'o, P. O. P., Lou-sing, K. (1991) *J. Am. Chem. Soc. 113*, 4032-4033.

Ornstein, R. L., Rein, R., Breen, D. L., & MacElroy, R. D. (1978) *Biopolymers 17*, 2341-2360.

Padmaja, N., Ramakumar, S., & Viswamitra, M. A. (1991) *Acta Cryst. 47*, 1445-1448.

Plum, G. E., Park, Y.-W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436-9440.

Povsic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059-3061.

Rajagopal, P., & Feigon, J. (1989) *Nature 339*, 637-640.

Roberts, R. W., & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9397-9401.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Singleton, S. F., & Dervan, P. B. (1992) *J. Am. Chem. Soc.*, in press.

Slonitskii, S. V., & Kuptsov, V. Y. (1988) *Molecular Biology 22*, 593-608.

Sober, H. A. (1970) *Handbook of Biochemistry. Selected Data for Molecular Biology*, CRC Press, Cleveland, OH.

Sowers, L. C., Eritja, R., Kaplan, B. E., Goodman, M. F., & Fazakerley, G. V. (1987) *J. Biol. Chem. 262*, 15436-15442.

Sowers, L. C., Shaw, B. R., & Sedwick, W. D. (1987) *Biochem. Biophys. Res. Commun. 148*, 790-794.

Strobel, S. A., & Dervan, P. B. (1991) *Nature 350*, 172-174.

Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E., & Dervan, P. B. (1991) *Science 254*, 1639-1642.

Ts'o, P. O. P. (1968) in *Molecular Associations in Biology* (B. Pullman, Ed.) pp 39-75, Academic Press, New York.

Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625-5631.

CHAPTER FOUR

Equilibrium Association Constants for Oligonucleotide-Directed Triple Helix Formation at Single DNA Sites: Linkage to Cation Valence and Concentration

The text of this chapter is taken from a published manuscript that was coauthored with Professor Peter B. Dervan.

(S. F. Singleton & P. B. Dervan, *Biochemistry* **1993**, *32*, 13171-13179.)

Introduction

In addition to their dependence on length (Moser & Dervan, 1987; Singleton & Dervan, 1992a), sequence composition (Kiessling et al., 1992), base triplet mismatches (Singleton & Dervan, 1992a), and functional groups on the heterocycle (Povsic & Dervan, 1989; Plum et al., 1990; Singleton & Dervan, 1992b; Froehler et al., 1992), the stabilities of local triple-helical complexes are sensitive to solution conditions, including temperature, pH, and the identities and concentrations of counterions (Moser & Dervan, 1987; Maher et al., 1990; Pilch et al., 1990; Plum et al., 1990; Hanvey et al., 1991; Rougée et al., 1992; Singleton & Dervan, 1992b). The experimental determination of oligonucleotide association constants as a function of these solution conditions is necessary to characterize the noncovalent forces which contribute to the affinity and specificity of binding. Moreover, any application of oligonucleotide-directed triple helix formation to control specific

gene expression *in vivo*, where the solution composition is tightly regulated, will require an understanding of the functional linkage between binding free energy and solution conditions.

Previous reports of the effects of cation concentrations on the stability of polymeric (Krakauer & Sturtevant, 1968; Krakauer, 1974; Lee et al., 1984; Latimer et al., 1989) and oligomeric (Pilch et al., 1990; Plum et al., 1990; Shea et al., 1990; Durand et al., 1992; Rougée et al., 1992) triple-helical DNAs, as well as on the kinetics of triple helix formation (Maher et al., 1990; Hampel et al., 1991; Rougée et al., 1992), stimulated us to examine the influence of cations and cation concentrations on the equilibrium constants for oligonucleotide-directed triple helix formation. Because potassium, magnesium, and spermine are thought to be the principle mono-, di-, and multivalent cations in eukaryotic cells, respectively, we have chosen to vary each of their concentrations separately in solutions containing all three cations. Using quantitative affinity cleavage titration (Singleton & Dervan, 1992a), we have measured association constants for the binding of the oligonucleotide 5′-d(T*TTTTCTCTCTCTCT)-3′ to a single 15-bp site within a 339-bp plasmid fragment (see Figure 2.1) in solutions containing various concentrations of KCl, $MgCl_2$, and $SpmCl_4$ at pH 7.0 and 22°C. Specifically, we have varied the concentration of potassium ion from 5 mM to 140 mM, that of magnesium ion from 0.1 mM to 10 mM, and that of spermine from 0.4 mM to 4 mM, because the cellular concentrations of these cations has been estimated to be near 140 mM potassium (Darnell et al., 1986), 1 mM magnesium (Darnell et al., 1986), and 1 mM spermine (Tabor & Tabor, 1976; Sarhan & Seiler, 1989).

Experimental Procedures

Stock Solutions of Buffer, Salts, and Carrier DNA. Spermine was purchased from Sigma as its tetrahydrochloride salt. A 5x stock solution of Tris-acetate (100 mM), NaCl (50 mM), KCl (25 mM), $MgCl_2$ (0.50 mM), and $SpmCl_4$ (2.0 mM) at pH 7.0 in Millipore water was prepared. Separate unbuffered aqueous stock solutions containing KCl (675 mM), $MgCl_2$ (49.5 mM), or $SpmCl_4$ (18.0 mM) were prepared. Sonicated, deproteinized calf thymus DNA (Pharmacia) was dissolved in unbuffered water to a final concentration of 2.0 mM in base pairs. All solutions were stored at 4°C prior to use.

DNA Preparation and Affinity Cleavage Titrations. The oligonucleotide-EDTA and labeled duplex DNA were prepared as described in Chapter One. Affinity cleavage titrations were performed essentially as described in Chapter One. Titration binding isotherms were constructed and equilibrium association constants were determined as described in Chapter One.

Results

It is well established that the association of single-stranded nucleic acids in solutions of monovalent cations is strongly enhanced by increasing cation concentrations (Manning, 1978; Record et al., 1978; Record et al., 1981). There is similar evidence demonstrating the stabilizing influence of monovalent cations on polymeric (Krakauer & Sturtevant, 1968; Lee et al., 1984; Latimer et al., 1989) and oligomeric (Plum et al., 1990; Shea et al., 1990; Durand et al., 1992; Rougée et al., 1992) triple-helical complexes. To assess the influence of monovalent counterion concentration on the energetics of oligonucleotide-directed triple helix formation in mixed salt solutions, the equilibrium association constant for

an oligonucleotide-EDTA conjugate binding to a 339-bp DNA duplex containing a 15-bp target sequence (see Figure 2.1) was measured in buffers containing different concentrations of KCl. At a given KCl concentration, ^{32}P-5′-end labeled DNA (< 5 pM) and various concentrations of oligonucleotide-EDTA·Fe (1 nM – 16 μM) were mixed in buffer (5.0 – 140 mM KCl, 1.0 mM $MgCl_2$, 1.0 mM $SpmCl_4$, 10 mM NaCl, and 20 mM Tris-acetate at pH 7.0) at 22°C. Only after the association reactions had been allowed to reach equilibrium over 24 h was DTT (1 mM final concentration) added to initiate the EDTA·Fe-mediated cleavage chemistry. The reactions were allowed to proceed for 6 h at 22°C, allowing a maximum site-specific cleavage yield of about 15%, and the products were separated by PAGE under strand-denaturing conditions. The amounts of radiolabeled DNA in the bands at the target cleavage site (see Figure 2.1) and at a reference site were measured from storage phosphor autoradiograms and I_{site} for each $[\mathbf{O}]_{tot}$ was calculated using eq 36 (Chapter One). The ($[\mathbf{O}]_{tot}$, I_{site}) data points were fitted using a nonlinear least-squares method and eq 2, with K_T and I_{sat} as adjustable parameters.

The data points obtained for the 15mer oligonucleotide at KCl concentrations of 5, 25, 90, and 140 mM (in the presence of 20 mM Tris-acetate, 10 mM NaCl, 1.0 mM $MgCl_2$, and 1.0 mM $SpmCl_4$) were averaged from four experiments and are plotted along with average best-fit titration binding isotherms in Figure 4.1A. The mean K_T values obtained from the analyses of these data are contained in Table 4.1. The value of K_T measured in the presence of 90 mM KCl (100 mM total monovalent cation concentration), 1 mM $MgCl_2$, and 1 mM $SpmCl_4$ is similar to that reported previously for the same triple helix at the same pH and temperature in the presence of 100 mM NaCl and 1 mM $SpmCl_4$ (Singleton & Dervan, 1992a). Remarkably, this value is 20-fold higher

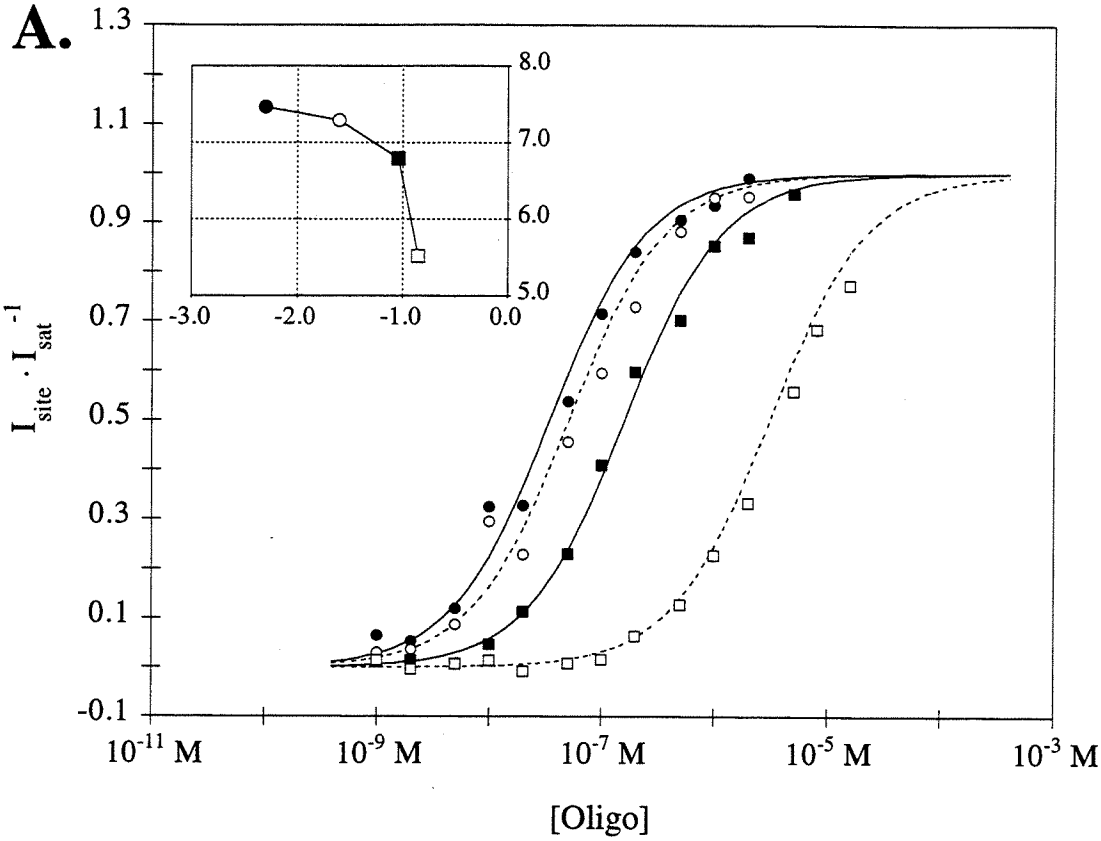


Figure 4.1. Data for quantitative affinity cleavage titrations performed in solutions containing 20 mM Tris-acetate (pH 7.0), 10 mM NaCl, 5 – 140 mM KCl, 0.1 – 10 mM $MgCl_2$, and 0.4 – 4 mM $SpmCl_4$ at 22 °C. The data points represent the average site-specific cleavage signal intensities from four experiments. The sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T (Table 4.1) and eq 38 (Chapter One). The data points were normalized using I_{sat} from each experiment and the binding curves were subsequently normalized using I_{sat} = 1 for eq 38 (Chpater One). (A) The experimental binding isotherms were measured in solution containing 1.0 mM $MgCl_2$, 1.0 mM $SpmCl_4$ and either 5.0 mM (filled circles), 25 mM (open circles), 90 mM (filled squares), or 140 mM (open squares) KCl. (A, inset) The apparent equilibrium association constant is plotted as a function of KCl concentration on a log-log scale.

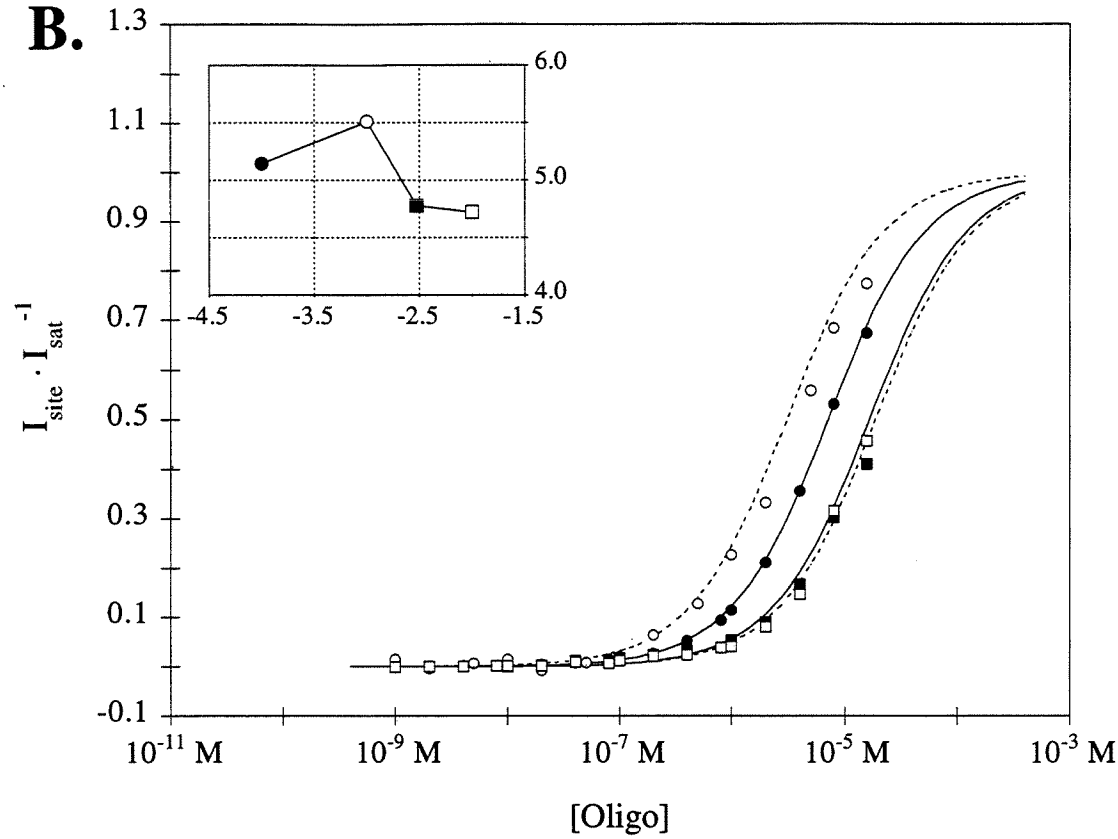


Figure 4.1 (cont'd) (B) The experimental binding isotherms were measured in solution containing 140 mM KCl, 1.0 mM $SpmCl_4$ and either 0.1 mM (filled circles), 1.0 mM (open circles), 3.0 mM (filled squares), or 10 mM (open squares) $MgCl_2$. (B, inset) The apparent equilibrium association constant is plotted as a function of $MgCl_2$ concentration on a log-log scale.

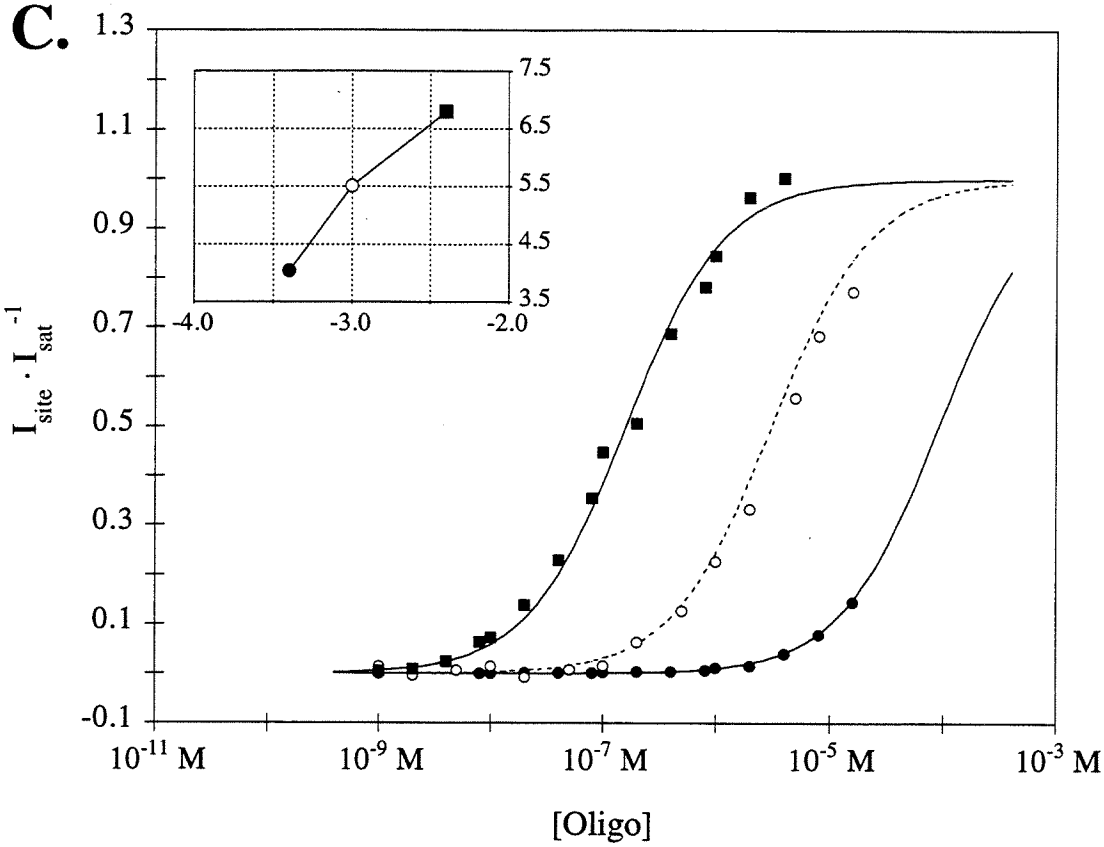


Figure 4.1 (cont'd) (C) The experimental binding isotherms were measured in solution containing 140 mM KCl, 1.0 mM $MgCl_2$, and either 0.40 mM (filled circles), 1.0 mM (open circles), or 4.0 mM (filled squaes) $SpmCl_4$. (C, inset) The apparent equilibrium association constant is plotted as a function of $SpmCl_4$ concentration on a log-log scale.

Table 4.1. Counterion Concentration Dependence of the Equilibrium Association Constant for Oligonucleotide-Directed Triple Helix Formation at pH 7.0 and 22 °C.[a]

Variation of [KCl][b]		Variation of [$MgCl_2$][c]		Variation of [$SpmCl_4$][d]	
[KCl]	K_T	[$MgCl_2$]	K_T	[$SpmCl_4$]	K_T
5.0	3.0 (± 0.8) x 10^7	0.10	1.4 (± 0.3) x 10^5	0.40	1.1[e] x 10^4
25	1.9 (± 0.8) x 10^7	1.0	3.3 (± 1.4) x 10^5	1.0	3.3 (± 1.4) x 10^5
90	6.2 (± 1.8) x 10^6	3.0	6.0 (± 2.1) x 10^4	4.0	6.2 (± 1.0) x 10^6
140	3.3 (± 1.4) x 10^5	10	5.3 (± 0.3) x 10^4		

[a] The K_T values in the table are mean values (± SEM) of four independent measurements. The K_T values and salt concentrations are reported in M^{-1} and $mmol \cdot L^{-1}$ units, respectively. [b] The concentration of KCl was varied in aqueous buffer containing 20 mM Tris-acetate, 10 mM NaCl, 1.0 mM $MgCl_2$, and 1.0 mM $SpmCl_4$. [c] The concentration of $MgCl_2$ was varied in aqueous buffer containing 20 mM Tris-acetate, 10 mM NaCl, 140mM KCl, and 1.0 mM $SpmCl_4$. [d] The concentration of $SpmCl_4$ was varied in aqueous buffer containing 20 mM Tris-acetate, 10 mM NaCl, 140 mM KCl, and 1.0 mM $MgCl_2$. [e] Of the four titration experiments performed to measure this association constant, three resulted in apparent association constants which were too small to be accurately measured ($\leq 10^4 M^{-1}$) and one resulted in an apparent association constant of 4.0 x 10^4 M^{-1}.

than that measured in the presence of 140 mM KCl. Overall, the measured K_T values demonstrate that increasing the concentration of KCl 28-fold, from 5.0 mM to 140 mM, causes a 100-fold *decrease* in the association constant. This trend of decreasing triple helix stability with increasing concentration of KCl is more clearly indicated by the plots of log(K_T) *versus* log[KCl] displayed in the inset of Figure 4.1A. Previous investigations have shown that θ_{app}, extracted from the extent of restriction endonuclease cleavage protection, decreases with increasing concentration of NaCl in the presence of $MgCl_2$ and $SpmCl_4$ at 37°C (Maher et al., 1990; Hanvey et al., 1991).

To clarify the origin of this result, the concentrations of $MgCl_2$ and $SpmCl_4$ were independently varied between 0.10 mM and 10 mM and between 0.40 mM and 4.0 mM, respectively, in solutions containing 10 mM NaCl, 140 mM KCl (least stabilizing concentration), and either 1.0 mM $SpmCl_4$ (for the magnesium experiments) or 1.0 mM $MgCl_2$ (for the spermine experiments) at pH 7.0 and 22°C. The results of the experiments in which the magnesium ion concentration was varied are shown in Figure 4.1B. The observed equilibrium association constant (Table 4.1) decreased whether the concentration of $MgCl_2$ was raised or lowered from 1.0 mM. Overall, as the concentration of $MgCl_2$ is increased 100-fold, from 0.10 mM to 10 mM, the association constant decreases 3-fold (Figure 4.1B, inset). In contrast, the result of increasing the spermine concentration 10-fold, from 0.40 mM to 4.0 mM (Figure 4.1C), is a greater than 500-fold *increase* in the observed association constant. Thus, in a solution containing potassium, magnesium, and spermine cations at or above millimolar concentrations, potassium ions are strongly inhibitory to triple helix formation, magnesium ions are slightly inhibitory, and the tetravalent spermine ion strongly enhances the stability of local triple-helical structures.

Discussion

The influence of the cationic environment on triple helix stability in solutions containing several cations of varying valence, conditions expected to be found within the cellular matrix (Darnell et al., 1986), is of considerable interest. In this series of experiments, we have used quantitative affinity cleavage titration to measure the effects of various cations on the stability of a local triple-helical complex in mixed valence salt solutions of varying composition at pH 7.0 and 22°C. Binding isotherms were measured in solutions containing sodium, potassium, magnesium, and spermine ions as the concentrations of K^+, Mg^{2+}, and Spm^{4+} were individually varied at constant concentrations of the other three cations. Analysis of the trends in the apparent equilibrium association constant allows us to elucidate the influence of each cation on triplex formation in the context of basic polyelectrolyte theory. The results indicate that the effect of a given counterion's concentration on the stability of a local triple-helical complex in solutions containing more than one type of counterion depends on the valence of that counterion. Specifically, increasing the concentration of the monovalent potassium ion from 5.0 mM to 140 mM in a solution containing 10 mM NaCl, 1.0 mM $MgCl_2$, and 1.0 mM $SpmCl_4$ at pH 7.0 and 22°C, results in a 100-fold decrease in the apparent equilibrium association constant for a 15mer binding to a single 15-bp homopurine·homopyrmidine site in a 339-bp plasmid fragment. Increasing the concentration of divalent magnesium ions from 0.1 mM to 10 mM in a solution containing 10 mM NaCl, 140 mM KCl, and 1.0 mM $SpmCl_4$ results in a three-fold decrease in the association constant for the same system. The equilibrium constant for this association reaction is increased nearly 500-fold when the concentration of the spermine tetracation is raised from 0.40 mM to 4.0 mM in the presence of 10 mM NaCl, 140 mM KCl, and 1.0 mM $MgCl_2$.

It is important to note that related trends have been observed in the thermal denaturation behavior of duplex DNA in solutions containing both Na^+ and either Mg^{2+} (Dove & Davidson, 1962; Manning, 1972; De Marky & Manning, 1975; Record, 1975) or a polyamine (Thomas & Bloomfield, 1984). In these solutions, the melting temperature of the DNA decreases with increasing [Na^+] until the melting temperature reaches a minimum at a critical sodium ion concentration (near 10 mM and 100 mM in the presence of Mg^{2+} and Spm^{4+}, respectively). For increasing [Na^+] above the critical concentration, the DNA melting behavior is similar to that of the DNA in solution free of oligovalent cations. The results for both oligonucleotide-directed triple helix formation and duplex DNA melting are consistent with competition among the different cations for nucleic acid phosphate binding sites and valence-specific abilities of the cations to stabilize the triple helix.

The Counterion Condensation Model. The molecular Counterion Condensation model developed by Manning (1978) and elaborated by Record and coworkers (Record et al., 1978), although an approximation to the real behavior of linear polyelectrolytes in solution, is a theoretical framework that provides reasonable descriptions of the interactions of small cations with nucleic acids and their thermodynamic consequences for nucleic acid conformational transitions (Record et al., 1981; Anderson & Record, 1982; Lohman, 1985). In the presence of a single type of counterion, the association of a single-stranded nucleic acid molecule with a duplex to form a triple-helical complex results in an overall increase in the linear charge density of the nucleic acid species in solution. Hence, triple helix formation is accompanied by the condensation of free cations (decreased entropy) and a decrease in the repulsive electrostatic free energy. At monovalent cation concentrations below 1 M, the first term is dominant and, al-

though it destabilizes the triple helix, decreases in magnitude with increasing cation concentration. Therefore, the triplex is stabilized relative to its unbound components when the bulk concentration of the counterion is increased. In accord with these expectations, the stabilities of polymeric (Krakauer & Sturtevant, 1968; Krakauer, 1974; Lee et al., 1984; Latimer et al., 1989) and oligomeric (Pilch et al., 1990; Plum et al., 1990; Shea et al., 1990; Durand et al., 1992; Rougée et al., 1992) triplexes in buffers containing single alkali or alkaline earth metal ions are found to increase with increasing concentrations of the metal ions.

In mixed valence salt solutions, the situation is more complicated than that found in solutions containing only nucleic acid and a single metal cation. The equilibrium between oligonucleotide, duplex, and triplex in the presence of three counterions of different valence can be written:

$$\mathbf{O} + \mathbf{D} + i\mathrm{M}^{+} + j\mathrm{Mg}^{2+} + k\mathrm{Spm}^{4+} \overset{K_T^{\circ}}{\rightleftharpoons} \mathbf{T} \qquad (1)$$

where M^+ represents both potassium and sodium cations (Manning, 1984), i, j, and k represent the numbers of M^+, Mg^{2+}, and Spm^{4+} ions thermodynamically bound per phosphate during the association reaction, respectively, and n is the length of the oligonucleotide. The thermodynamic equilibrium constant for this reaction can be written:

$$K_{\mathrm{T}}^{o} = \frac{a_{\mathbf{T}}}{a_{\mathbf{D}} a_{\mathbf{O}}} \cdot a_{\mathrm{M}^+}^{-ni} \cdot a_{\mathrm{Mg}^{2+}}^{-nj} \cdot a_{\mathrm{Spm}^{4+}}^{-nk} \qquad (2)$$

were a values are activities. The apparent equilibrium constant for eq 1 is simply:

$$K_T = \frac{[T]}{[D][O]} \tag{3}$$

It follows from eqs 7 and 8 that

$$\ln K_T = \ln K_T^o + n\left\{i \ln\left[M^+\right] + j \ln\left[Mg^{2+}\right] + k \ln\left[Spm^{4+}\right]\right\} - \ln \frac{\gamma_T}{\gamma_D \gamma_O} \tag{4}$$

were values of γ are activity coefficients. Record (1975) has shown that, when the magnitude of the activity coefficients originate from the purely electrostatic interactions between the nucleic acid species and its counterions, eq 4 can be rewritten

$$\ln K_T = \ln K_T^o + n\left\{i \ln\left[M^+\right] + j \ln\left[Mg^{2+}\right] + k \ln\left[Spm^{4+}\right] - \eta' \ln \kappa\right\} \tag{5}$$

where κ is the Debye-Hückel screening parameter (proportional to the square root of the ionic strength) and the factor η' is defined by

$$\eta' \equiv \left[\frac{1}{3}\left(\frac{\xi_{net}^2}{\xi}\right)_O + \frac{2}{3}\left(\frac{\xi_{net}^2}{\xi}\right)_D\right] - \left(\frac{\xi_{net}^2}{\xi}\right)_T \tag{6}$$

where ξ is the dimensionless structural charge density parameter for the nucleic acid molecule, and ξ_{net} is the effective value of ξ after counterion condensation has been considered. For idealized single-, double-, and triple-stranded DNA, the linear charge spacings are 4.3 Å, 1.7 Å, and 1.1 Å, respectively. Hence, the values of ξ were taken to be 1.7, 4.2, and 6.6 for each of the conformations, respectively. In the mixtures of mono-, di-, and tetravalent cations, the value of

ξ_{net} should be in the range $1 \geq \xi_{net} \geq 0.25$, and depends on the exact counterion composition (Manning, 1972; De Marky & Manning, 1975). The factor η' indicates differential screening of the phosphate charges between the free and bound states of the nucleic acid system, and larger positive values of η' represent better relative screening of the charge repulsions in the triple helix.

The apparent equilibrium constant K_T will be a function of all three cation concentrations:

$$\frac{\partial \ln K_T}{\partial \ln[M^+]} = i + \frac{\partial i}{\partial \ln[M^+]} \ln[M^+] + \frac{\partial j}{\partial \ln[M^+]} \ln[Mg^{2+}] + \frac{\partial k}{\partial \ln[M^+]} \ln[Spm^{4+}] - \frac{\partial \ln \kappa}{\partial \ln[M^+]} \eta' - \frac{\partial \eta'}{\partial \ln[M^+]} \ln \kappa \tag{7a}$$

$$\frac{\partial \ln K_T}{\partial \ln[Mg^{2+}]} = j + \frac{\partial i}{\partial \ln[Mg^{2+}]} \ln[M^+] + \frac{\partial j}{\partial \ln[Mg^{2+}]} \ln[Mg^{2+}] + \frac{\partial k}{\partial \ln[Mg^{2+}]} \ln[Spm^{4+}] - \frac{\partial \ln \kappa}{\partial \ln[Mg^{2+}]} \eta' - \frac{\partial \eta'}{\partial \ln[Mg^{2+}]} \ln \kappa \tag{7b}$$

$$\frac{\partial \ln K_T}{\partial \ln[Spm^{4+}]} = k + \frac{\partial i}{\partial \ln[Spm^{4+}]} \ln[M^+] + \frac{\partial j}{\partial \ln[Spm^{4+}]} \ln[Mg^{2+}] + \frac{\partial k}{\partial \ln[Spm^{4+}]} \ln[Spm^{4+}] - \frac{\partial \ln \kappa}{\partial \ln[Spm^{4+}]} \eta' - \frac{\partial \eta'}{\partial \ln[Spm^{4+}]} \ln \kappa \tag{7c}$$

The effect of a particular cation's concentration on K_T originates from a combination of the negative entropy of condensation of cations during triple helix forma-

tion and the electrostatic free energy change which results from phosphate charge neutralization during the transition (*vide infra*). The former effect is represented in the first four terms in eq 7, while the latter effect is represented by the last two terms.

In order to examine the observed counterion effects semi-quantitatively using eq 12, we have estimated binding densities for each of the cations on idealized triple-helical, double-helical, and single-stranded DNA under the experimental salt conditions. By analyzing the changes in the estimated cation binding densities at different cation concentrations, it is possible to compare the signs and relative magnitudes of the terms in eq 7, as well as the overall signs of the derivatives of $\ln K_T$. The values of the association constants for Mg^{2+} and Spm^{4+} binding to each nucleic acid species were calculated from the following equation:

$$\log K_{Mg} = 2\psi(1-4\theta_{Spm})\log[M^+] + \log K_{Mg}^\circ \qquad (8a)$$

$$\log K_{Spm} = 4\psi(1-2\theta_{Mg})\log[M^+] + \log K_{Spm}^\circ \qquad (8b)$$

where ψ is the thermodynamic monovalent cation binding parameter for a particular nucleic acid conformation, θ-values are oligovalent cation binding densities, and values of K° represent oligovalent cation binding constants at 1 M monovalent cation concentration (Record et al., 1978). From eq 8 it is clear that cation binding is competitive, *i.e.*, binding of one oligovalent cation reduces the apparent binding constants of other cations. Initial estimates of K_{Mg} and K_{Spm} were made using eq 8 with $\theta_{Mg} = \theta_{Spm} = 0$ and $\log K_{Mg} = \log K_{Spm} = 0$. Based on the computed values of the association constants and assumed binding site sizes of $n_{Mg} = 2$ and $n_{Spm} = 4$, binding densities for Mg^{2+} and Spm^{4+} were calculated

using the model of McGhee and von Hippel (1974) for ligands binding to overlapping sites on a uniform linear lattice of phosphates. This model has been shown to fit experimental data for the binding of Mg^{2+} and Spm^{4+} to duplex DNA adequately (Braunlin et al., 1982). The values of K_{Mg} and K_{Spm} were then adjusted to account for the binding densities of Spm^{4+} and Mg^{2+}, respectively. This process was repeated until values of K_{Mg} and K_{Spm} from successive iterations differed by less than 1%. Equilibrium constants for Spm^{4+} binding to duplex DNA calculated in this manner were in good agreement with those predicted using an empirically derived relation (Braunlin et al., 1982). The spermine-single strand binding constants were about 10-fold lower than those for duplex binding, in agreement with experiment (Morgan et al., 1986), while spermine-triplex binding was about 1.5-fold stronger than spermine-duplex binding, in accord with the difference in the charge densities of the DNA complexes (Murray & Morgan, 1973; Record et al., 1978). As a result of competition with Spm^{4+}, the estimated binding constants for Mg^{2+} interacting with all three DNA species were about 10-fold lower than predicted (Braunlin et al., 1982; Morgan et al., 1986). Following calculation of the binding densities of Spm^{4+} and Mg^{2+}, the thermodynamic binding density of K^+ was corrected for the number of potassium ions displaced by Mg^{2+} and Spm^{4+} (Record et al., 1978).

Influence of Potassium Ion Concentration on the Association Constant. It is clear that neither Na^+ nor K^+ are intrinsically destabilizing to double- or triple-helical nucleic acids because, in the presence of these ions alone and in the presence of high concentrations of these ions with oligovalent cations, the complex stabilities are enhanced by increasing the monovalent cation concentration. Rather, the observed trend of decreasing triplex stability with increasing $[K^+]$ up to 140 mM in the presence of millimolar magnesium and spermine results from the influence of

these oligovalent cations on the triplex formation equilibrium. The ability of a counterion to stabilize a nucleic acid complex of higher charge density increases with the charge on the counterion (Manning, 1972; Morgan et al., 1986). Thus, Spm^{4+} is more stabilizing to the triple helix than is Mg^{2+}, which, in turn, is more stabilizing than K^+. The effect of spermine is great enough that its presence at micromolar concentration effects the complete dismutation of polymeric DNA (Hampel et al., 1991) and RNA (Glaser & Gabbay, 1968) duplexes to the corresponding triplexes plus single strands at neutral pH.

At low ionic strengths, spermine binds with high affinity ($K_{Spm} > 10^5\ M^{-1}$) to all three DNA species and dominates the counterion condensation. At $[Spm^{4+}] = 1.0$ mM and $[M^+] = 15$ mM, all three DNA species are nearly saturated by Spm^{4+} and the triple helix is particularly stable ($K_T \approx 10^8\ M^{-1}$). Because there is little change in Spm^{4+} condensation upon triplex formation, there is little change in the overall degree of counter-ion condensation or in the electrostatic free energy, and K_T is near a maximum. As the concentration of K^+ increases, the equilibrium constant for Spm^{4+} binding to each of the DNA species decreases; however, because the spermine-triplex binding constant is always higher than the other two, there is a net condensation of Spm^{4+} and the entropy of condensation becomes more negative (fourth term of eq 7a). Moreover, the net difference in counterion binding to the triplex versus the duplex and oligonucleotide results in a decrease in the electrostatic free enenergy of association ($\eta' > 0$ and increases with increasing $[K^+]$). This effect (terms five and six of eq 7a) stabilizes the triplex but decreases in magnitude with increasing ionic strength. The result of increasing $[K^+]$ in the presence of magnesium and spermine is therefore a reduction in the apparent equilibrium constant K_T.

Above 0.5 M potassium ion concentration, Spm^{4+} binds poorly to all three DNA species, K^+ significantly contributes to the net counterion condensation, and K_T should increase with further increases in $[K^+]$. However, K^+ is sufficiently weaker than Spm^{4+} for triple helix stabilization that no site-specific affinity cleavage is detected at a potassium ion concentration of 500 mM (data not shown). In fact, it has been demonstrated that the oligonucleotide used here binds to 15 bp within a 21-bp target duplex with an equilibrium constant of only 10 M^{-1} at 25°C and pH 6.5 in the presence of 200 mM NaCl without spermine (Plum et al., 1990).

Influence of Spermine Ion Concentration on the Association Constant. At all concentrations of Spm^{4+} investigated, the tetracation is preferentially condensed upon triple helix formation, resulting in higher values of K_T at higher $[Spm^{4+}]$. The extent of net Spm^{4+} condensation decreases from 0.40 to 4.0 mM because the ion begins to saturate all forms of the DNA, and the entropic penalty (term four of eq 7c) is reduced. Concurrently, the differential screening parameter (term six of eq 7c) becomes smaller. These two effects add together to give the large positive slope for the data plotted in Figure 4.2. The condensation of Spm^{4+} is predicted to reach a minimum ($K \approx 0$) at a spermine concentration above 10 mM; however, precipitation of the DNA at elevated spermine concentrations prevents measurement of the association constant.

Influence of Magnesium Ion Concentration on the Association Constant. The stabilization of the triple helix by Mg^{2+} is intermediate between the stabilizations afforded by Spm^{4+} and K^+, and the influence of $[Mg^{2+}]$ is likewise intermediate between those of $[Spm^{4+}]$ and $[K^+]$. Over the range of magnesium ion concentrations studied, five of the six terms in eq 7b have the same sign and magnitude as the corresponding terms in eq 7a. The exception is the screening parameter η'

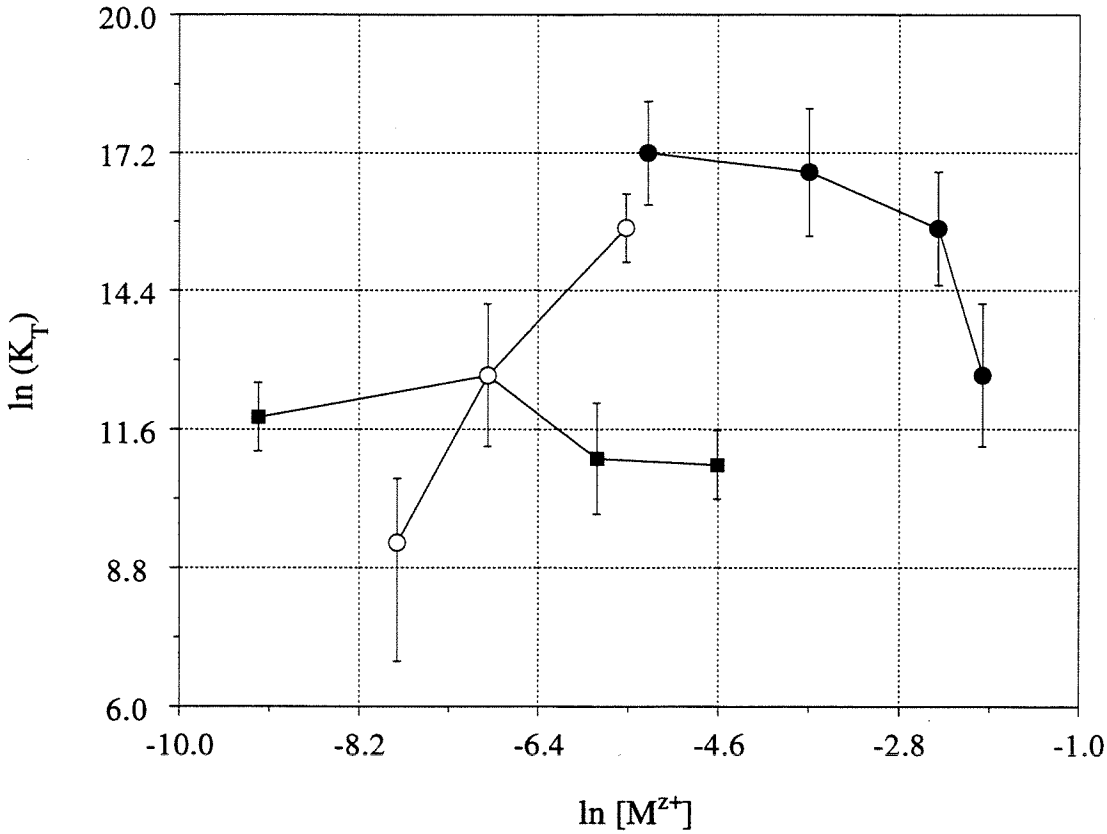


Figure 4.2. The natural logarithm of the mean association constants are plotted as a function of the natural logarithm of the concentration of NaCl plus KCl (filled circles), $MgCl_2$ (filled squares), or $SpmCl_4$ (open circles) for comparison with eq 7 (see text). The error bars represent estimated confidence limits.

(the last term), whose magnitude changes less dramatically with changes in $[Mg^{2+}]$ than with changes in $[K^+]$. For $[Mg^{2+}] \subseteq [Spm^{4+}]$, the condensation is dominated by Spm^{4+} and the effect of $[Mg^{2+}]$ is neglible ($\partial\eta / \partial \ln[Mg^{2+}] \approx 0$). As $[Mg^{2+}]$ increases, it competes effectively with Spm^{4+} and the stability of the triplex is diminished; however, because Mg^{2+} is more stabilizing than K^+, this decrease in triple helix stability is less dramatic than that caused by increasing $[K^+]$. At concentrations of Mg^{2+} above 10 mM, the triplex is expected to be stabilized by increasing the concentration of this cation because the apparent equilibrium fraction of a target duplex bound by a 21mer oligonucleotide was observed to increase from 20 mM to 160 mM $[Mg^{2+}]$ in the presence of 600 mM $[Na^+]$ and 0.4 mM $[Spm^{4+}]$ (Maher et al., 1990).

Implications for Oligonucleotide-Directed Triple Helix Formation. The observed influence of changes in the cationic environment on K_T demonstrate the importance of millimolar spermine concentrations or greater than ten millimolar magnesium concentrations for achieving equilibrium constants on the order of 10^6 M^{-1} or higher for oligonucleotide-directed triple helix formation near neutral pH in the presence of monovalent cations at a concentration of 150 mM. This dramatic linkage between the oligonucleotide association constant and the concentrations and valences of the counterions in solution is a consequence of the increase in charge density that accompanies oligonucleotide binding. The observations that the dependence of polymer triplex thermal denaturation temperature on $[Na^+]$ decreases as the fraction of C+GC triplets increases (Lee et al., 1984; Latimer et al., 1989) and that protonation of C GC sites plays a crucial role in triplex stability (Singleton & Dervan, 1992b; Völker et al., 1993) demonstrate the importance of reducing the increase in charge density in order to stabilize a triplex. It is important that protonation of C GC sites in the triplex

accompanies binding of an unprotonated single strand near neutral pH because the protonation specifically stabilizes the triplex relative to the unbound components.

References

Anderson, C. F., & Record, M. T., Jr. (1982) *Annu. Rev. Biochem. 33*, 191-222.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Braunlin, W. H., Strick, T. J., & Record, M. T., Jr. (1982) *Biopolymers 21*, 1301-1314.

Darnell, J., Lodish, H., & Baltimore, D. (1986) *Molecular Biology of the Cell*, p 618, Scientific American Books, New York.

De Marky, N., & Manning, G. S. (1975) *Biopolymers 14*, 1407-1422.

Dove, W. F., & Davidson, N. (1962) *J. Mol. Biol. 5*, 467-468.

Durand, M., Peloille, S., Thuong, N. T., & Maurizot, J. C. (1992) *Biochemistry 31*, 9197-9204.

Froehler, B. C., Wadwani, S., Terihorst, T. J., & Gerrard, S. R. (1992) *Tet. Lett. 33*, 5307-5310.

Glaser, R., & Gabbay, E. J. (1964) *Biopolymers 6*, 243-254.

Hampel, K. J., Crosson, P. & Lee, J. S. (1991) *Biochemistry 30*, 4455-4459.

Hanvey, J. C., Williams, E. M., & Besterman, J. M. (1991) *Antisense Res. Dev. 1*, 307-317.

Kiessling, L. L., Griffin, L. C., & Dervan, P. B. (1992) *Biochemsitry 31*, 2829-2834.

Krakauer, H. (1974) *Biochemistry 13*, 2579-2589.

Krakauer, H., & Sturtevant, J. M. (1968) *Biopolymers 6*, 491-512.

Latimer, L. J. P., Hampel, K., & Lee, J. S. (1989) *Nucleic Acids Res. 17*, 1549-1561.

Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603-6614.

Lohman, T. M. (1985) *CRC Crit. Rev. Biochem. 19*, 191-245.

Maher, L. J., III, Dervan, P. B., & Wold, B. (1990) *Biochemistry 29*, 8820-8826.

Manning, G. S. (1972) *Biopolymers 11*, 951-955.

Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179-246.

Manning, G. S. (1984) *J. Phys. Chem. 88*, 6654-6661.

McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469-489.

Morgan, J. E., Blankenship, J. W., & Matthews, H. R. (1986) *Arch. Biochem. Biophys. 246*, 225-232.

Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645-650.

Murray, N. L., & Morgan, A. R. (1973) *Can. J. Biochem. 51*, 436-449.

Pilch, D. S., Brousseau, R., & Shafer, R. H. (1990) *Nucleic Acids Res. 18*, 5743-5750.

Plum, G. E., Park, Y.-W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436-9440.

Povsic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059-3061.

Record, M. T., Jr. (1975) *Biopolymers 14*, 2137-2158.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103-178.

Record, M. T., Jr., Mazur, S. J., Melançon, P., Roe, J.-H., Shaner, S. L., & Unger, L. (1981) *Ann. Rev. Biochem. 50*, 997-1024.

Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannageli, C., Garestier, T., & Hélène, C. (1992) *Biochemistry 31*, 9269-9278.

Sarhan, S., & Seiler, N. (1989) *Biol. Chem. Hoppe-Seyler 370*, 1279-1284.

Shea, R. G., Ng, P., & Bischotberger, N. (1990) *Nucleic Acids Res. 18*, 4859-4866.

Singleton, S. F., & Dervan, P. B. (1992a) *J. Am. Chem. Soc. 114*, 6957-6965.

Singleton, S. F., & Dervan, P. B. (1992b) *Biochemistry 31*, 10995-11003.

Tabor, C. W., & Tabor, H. (1976) *Ann. Rev. Biochem. 45*, 285-306.

Thomas, T. J., & Bloomfield, V. A. (1984) *Biopolymers 23*, 1295-1306.

Völker, J., Botes, D. P., Lindsey, G. G., & Klump, H. H. (1993) *J. Mol. Biol. 230*, 1278-1290.

CHAPTER FIVE

Temperature Dependence of the Energetics of Oligonucleotide-Directed Triple Helix Formation at a Single DNA Site

The text of this chapter is taken from a manuscript that was coauthored with Professor Peter B. Dervan and submitted for publication in The Journal of the American Chemical Society.

Introduction

A continuing goal of our research efforts is the characterization of the kinetic[1] and thermodynamic[2-5] parameters that characterize the formation and stabilities of local triple-helical complexes at single sites on relatively large double-helical DNA (> 200 bp) near physiological solution conditions. The experimental determination of oligonucleotide association constants as a function of solution conditions is necessary to characterize the non-covalent forces which contribute to the affinity and specificity of DNA recognition. Moreover, such information will be important for the rational application of oligonucleotides to such tasks as the manipulation of large DNAs[6-8] or RNAs[9,10] *in vitro* and the artificial modulation of biological events.[11-16]

Triple-helical nucleic acids are enthalpically stabilized and the dissociation of a single-stranded nucleic acid from polymeric and oligomeric triple helices is

therefore strongly promoted by increasing temperature.[17] Early differential scanning calorimetric (DSC) measurements at pH 6.5 showed a triple-helical complex containing both T·AT and C+GC base triplets to be enthalpically stabilized by an average of 2.0 kcal per mole of base triplets (kcal·mol-bt^{-1}).[2] In contrast, the corresponding Watson-Crick duplex was stabilized by 6.3 kcal per mole of base pairs. Subsequent measurements on triplexes of different sequence compositions using a number of techniques, including the analysis of the concentration dependence of melting temperatures and fitting of optical melting curves, under a variety of conditions indicated the triplex to be enthalpically stabilized by a broader range of values, 2.5 – 8.1 kcal·mol-bt^{-1} (*vide infra*), suggesting that the original DSC result could be unreasonably low. Alternatively, these values are average enthalpies for T·AT, protonated C+GC, and unprotonated C·GC base triplets, and different sequence compositions would yield different average values. Van't Hoff analysis of the temperature-dependence of equilibrium association constants measured using quantitative affinity cleavage titrations[3] would provide an independent estimate of the enthalpic stabilization of a local triple-helical complex.

To assess the influence of temperature on the association free energy for an oligonucleotide binding a single site in large duplex DNA near intracellular solution conditions, we determined the temperature-dependence of the equilibrium association constant for binding of the oligonucleotide-EDTA conjugate 5′-d(T*TTTTCTCTCTCTCT)-3′ to a single 15-bp site within a 339-bp plasmid fragment (see Figure 2.1) at pH 7.0. Specifically, we varied the temperature from 8°C to 37°C in three different Bis-tris-buffered solutions characterized by differences in their concentrations of NaCl, KCl, $MgCl_2$, and spermine tetrahydrochloride ($SpmCl_4$), and measured the equilibrium association constant under each set of

conditions. Van't Hoff analysis of the observed temperature-dependence in each case allowed the extraction of enthalpy and entropy changes during oligonucleotide-directed triple helix formation. The observed thermodynamic parameters for single-site triple helix formation are compared to prior energetic analyses of oligomeric triplexes under simple solution conditions.

Results

Affinity Cleavage Titrations. At a given temperature, ^{32}P-5′-end labeled DNA (< 50 pM) and various concentrations of oligonucleotide-EDTA·Fe (1 nM – 20 μM) were mixed in Bis-tris-buffered salt solutions at pH 7.0. The time required to reach equilibrium depended on the temperature at which the association reactions occurred. As expected based on the heat of helix initiation,[18] lower temperatures yielded faster apparent association rate constants (k_{on}); however, this effect was mitigated by increased equilibrium association constants (*vide infra*) which required binding at lower oligonucleotide concentrations. Hence, the apparent triplex formation rates, which are controlled by $k_{on}[\mathbf{O}]_{tot}$, were only modestly influenced by temperature. After minimal equilibration times at each temperature were established,[3] equilibration times of 24 h at 8°C and 22°C, and for 48 h at 29°C and 37°C, were chosen for experimental convenience.

After the association reactions reached equilibrium, DTT was added (1 mM final concentration) to initiate the thymidine-EDTA·Fe-mediated cleavage chemistry. The reactions were allowed to proceed to a maximum site-specific cleavage yield of 10–15%. The time required for the cleavage reactions was also temperature-dependent: cleavage times were 8 h, 6 h, 4 h, or 3 h at 8°C, 22°C, 29°C, and 37°C, respectively. Cleavage was stopped by ethanol precipitation and subsequent denaturation of the DNA strands, and the products were separated from

uncleaved strands using urea-PAGE. The signal intensities of the bands at the target cleavage site (see Figure 2.1) and at a reference site were measured from storage phosphor autoradiograms and I_{site} for each $[\mathbf{O}]_{tot}$ was calculated using eq 36 (Chapter One). The ($[\mathbf{O}]_{tot}$, I_{site}) data points were fitted using a nonlinear least-squares method and eq 7, with K_T and I_{sat} as adjustable parameters (see Experimental Section).

Equilibrium Association Constants as a Function of Temperature. The sets of ($[\mathbf{O}]_{tot}$, I_{site}) data obtained for the 15mer oligonucleotide at 8°C, 22°C, 29°C, and 37°C in BTNS buffer at pH 7.0 (10 mM Bis-tris, 100 mM NaCl, and 1.0 mM $SpmCl_4$) were averaged from three to five independent titrations and are plotted in Figure 5.1A. The mean K_T values obtained from analyses of these data (Table 5.1) were used to construct binding isotherms which are plotted along with the averaged data points in Figure 5.1A. The value of K_T measured at 22°C is identical within experimental uncertainty to that reported previously for the same triple helix at the same pH, temperature, and salt concentrations in the presence of 50 mM Tris-acetate.[3] The room temperature value is 7.6-fold lower than that measured at 8°C, and 24-fold higher than that at 37°C. Overall, the measured K_T values demonstrate that raising the temperature from 8°C to 37°C causes a 180-fold decrease in the association constant. This trend of decreasing triple helix stability with increasing temperature is consistent with the negative enthalpies of triple-helix formation reported under a number of solution conditions.

In order to explore this result further, the temperature dependence of the equilibrium association constant was measured under two new solution conditions that differed from each other and from BTNS buffer in their salt compositions. The first buffer, BTP1 (10 mM NaCl, 140 mM KCl, 1 mM $MgCl_2$, and 1 mM $SpmCl_4$ in 10 mM Bis-tris at pH 7.0), was selected to provide a mini-

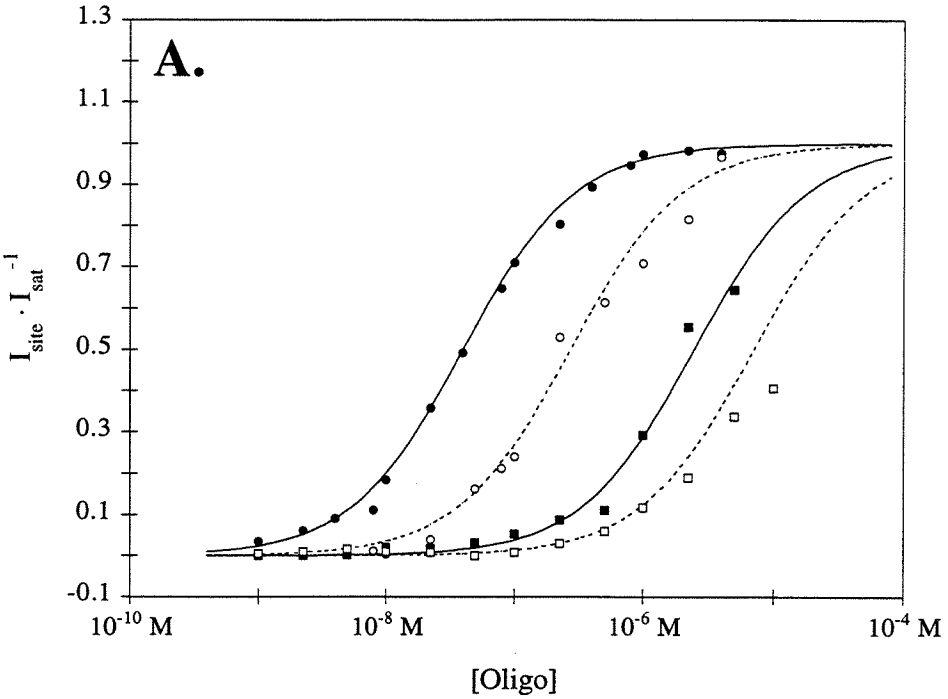
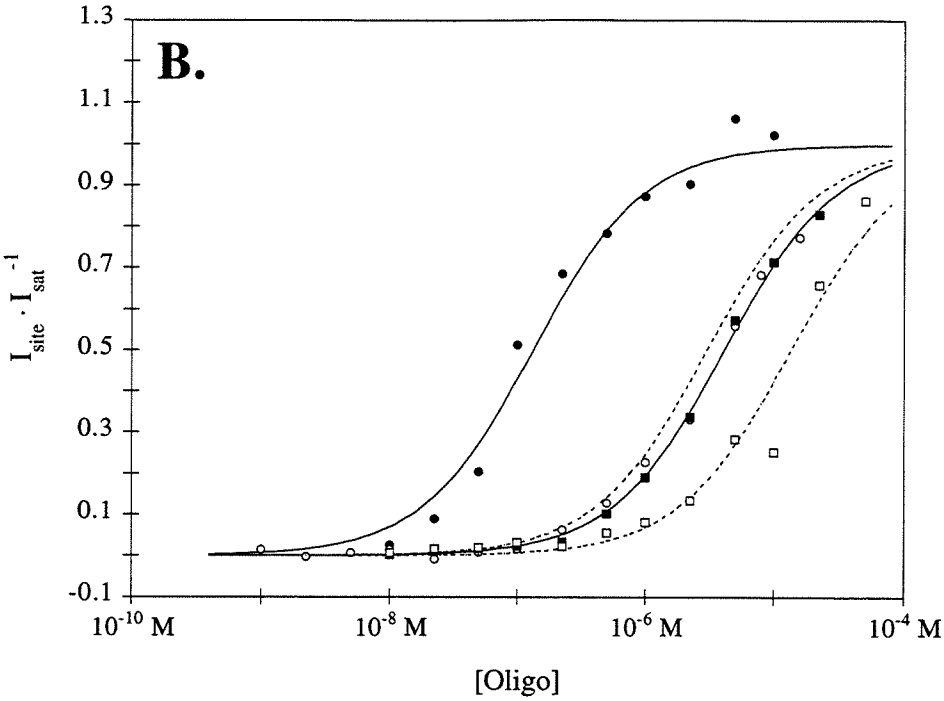


Figure 5.1. Data for quantitative affinity cleavage titrations performed in 10 mM Bis-tris (pH 7.0) solutions at either 8°C (●), 22°C (○), 29°C (■), or 37°C (□). The data points represent the average site-specific cleavage signal intensities from three to five experiments. The sigmoidal curves show the titration binding isotherms plotted using the mean values of K_T (Table 1) and eq 38 (Chapter One). The data points were normalized using I_{sat} from each experiment and the isotherms were subsequently normalized using I_{sat} = 1 for eq 38 (Chpater One). (A) Binding titrations performed in BTNS buffer at pH 7.0 (10 mM Bis-tris, 100 mM NaCl, and 1 mM $SpmCl_4$). (B) Binding titrations performed in BTP1 buffer at pH 7.0 (10 mM Bis-tris, 10 mM NaCl, 140 mM KCl, 1 mM $MgCl_2$, and 1 mM $SpmCl_4$). (C) Binding titrations performed in BTP4 buffer at pH 7.0 (10 mM Bis-tris, 10 mM NaCl, 140 mM KCl, 1 mM $MgCl_2$, and 4 mM $SpmCl_4$).


B.
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
$I_{site} \cdot I_{sat}^{-1}$
10^{-10} M
10^{-8} M
10^{-6} M
10^{-4} M
[Oligo]

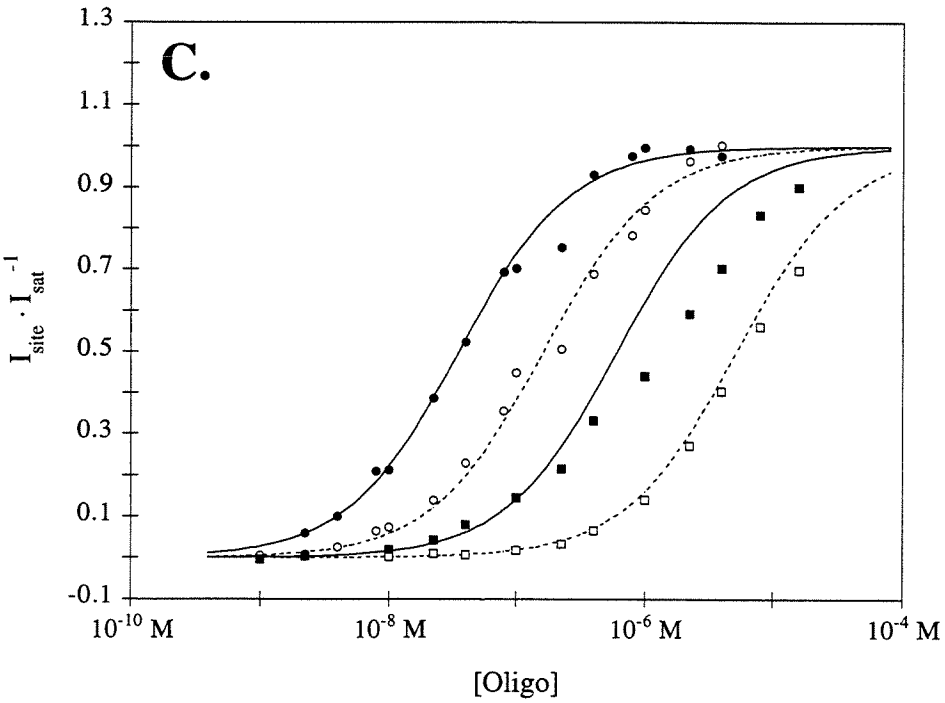

C.
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
$I_{site} \cdot I_{sat}^{-1}$
10^{-10} M
10^{-8} M
10^{-6} M
10^{-4} M
[Oligo]

Table 5.1. Temperature Dependence of the Equilibrium Association Constant for Oligonucleotide-Directed Triple Helix Formation at pH 7.0.[a]

	100 mM NaCl 1 mM $SpmCl_4$ [b]		10 mM NaCl 140 mM KCl 1 mM $MgCl_2$ 1 mM $SpmCl_4$ [c]		10 mM NaCl 140 mM KCl 1 mM $MgCl_2$ 4 mM $SpmCl_4$ [d]	
Temp	K_T	ΔG_T	K_T	ΔG_T	K_T	ΔG_T
8	2.5 (± 0.7) x 10^7	− 9.5 (± 0.2)	7.6 (± 0.6) x 10^6	− 8.8 (± 0.1)	3.7 (± 0.7) x 10^7	− 9.7 (± 0.1)
22	3.3 (± 0.4) x 10^6	− 8.8 (± 0.1)	3.3 (± 1.4) x 10^5	− 7.6 (± 0.2)	6.3 (± 1.0) x 10^6	− 9.2 (± 0.1)
29	4.1 (± 1.4) x 10^5	− 7.8 (± 0.2)	2.4 (± 0.4) x 10^5	− 7.4 (± 0.1)	1.7 (± 1.2) x 10^6	− 8.6 (± 0.5)
37	1.4 (± 0.8) x 10^5	− 7.3 (± 0.4)	7.4 (± 1.3) x 10^4	− 6.9 (± 0.1)	1.9 (± 0.7) x 10^5	− 7.5 (± 0.2)

[a] The K_T values in the table are mean values (± SEM) of three to five independent measurements in 10 mM Bis-tris·HCl buffer (pH 7.0) at the indicated temperatures (in °C) and salt concentrations. Standard errors in the free energies are propagated from the SEM of each K_T value. The K_T and derived ΔG_T values are reported in units of M^{-1} and kcal·mol^{-1}, respectively. [b] BTNS buffer. [c] BTP1 buffer. [d] BTP4 buffer.

mum of likely intracellular cation concentrations and was found to reduce the stability of the local triple-helical complex 10-fold relative to that in BTNS buffer because of its higher concentration of monovalent cations.[5] The second buffer, BTP4 (10 mM NaCl, 140 mM KCl, and 4 mM $SpmCl_4$ in 10 mM Bis-tris at pH 7.0), contains the same concentrations of mono- and divalent cations but a higher concentration of the spermine tetracation, at which the equilibrium association constant is increased to a value similar to that found in BTNS buffer.[5] The results of the titrations in which the temperature was varied in BTP1 buffer are shown in Figure 5.1B. The observed equilibrium association constant (Table 5.1) increased when the temperature was cooled below 22°C and decreased as the temperature was raised to 37°C. Overall, as the temperature is increased from 8°C to 37°C, the association constant decreases 100-fold. Similarly, the result of increasing the temperature from 8°C to 37°C in BTP4 buffer is a 190-fold decrease in the observed association constant (Figure 5.1C and Table 5.1).

Thermodynamic Parameters from van't Hoff Analyses. The observed trends of increasing triple helix stability with decreasing temperature in BTNS, BTP1, and BTP4 are clearly indicated by the van't Hoff plots ($\ln K_T$ *versus* T^{-1}) displayed in Figures 3A, B, and C, respectively. Linear least squares analyses of the data allowed the enthalpies to be extracted from the slopes of the lines based on a form of the van't Hoff expression

$$\ln K_{eq} = (-\Delta H°/R)\cdot T^{-1} + (\Delta S°/R) \qquad (1)$$

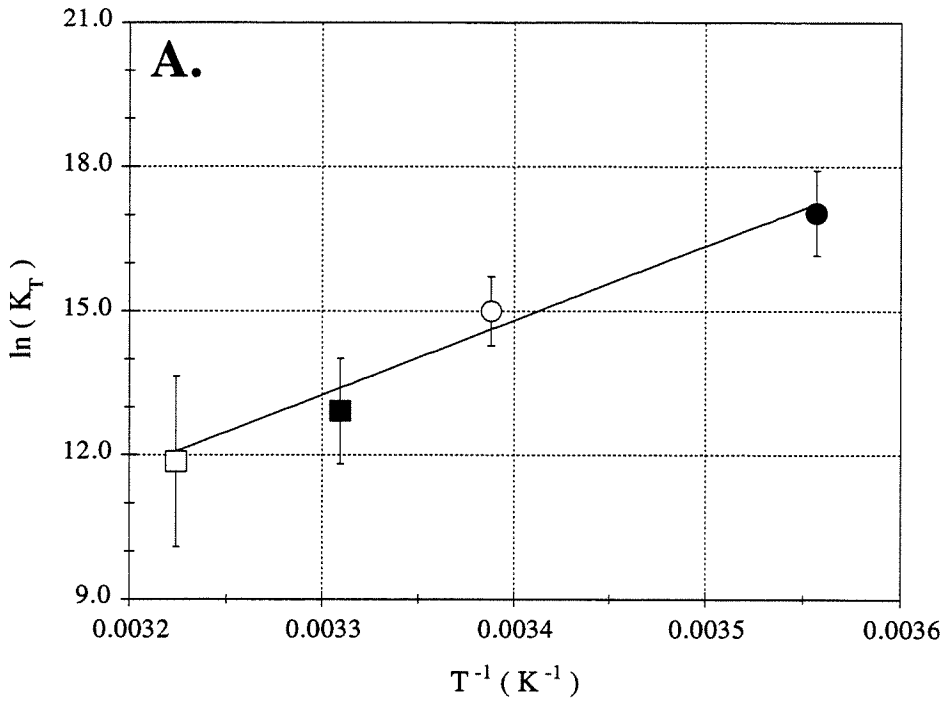
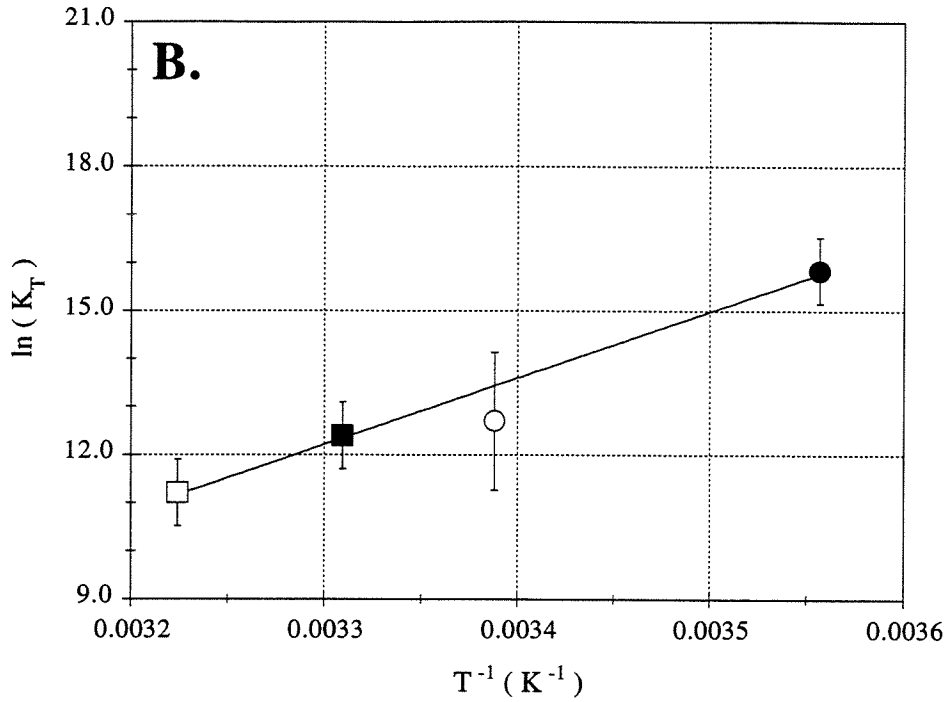
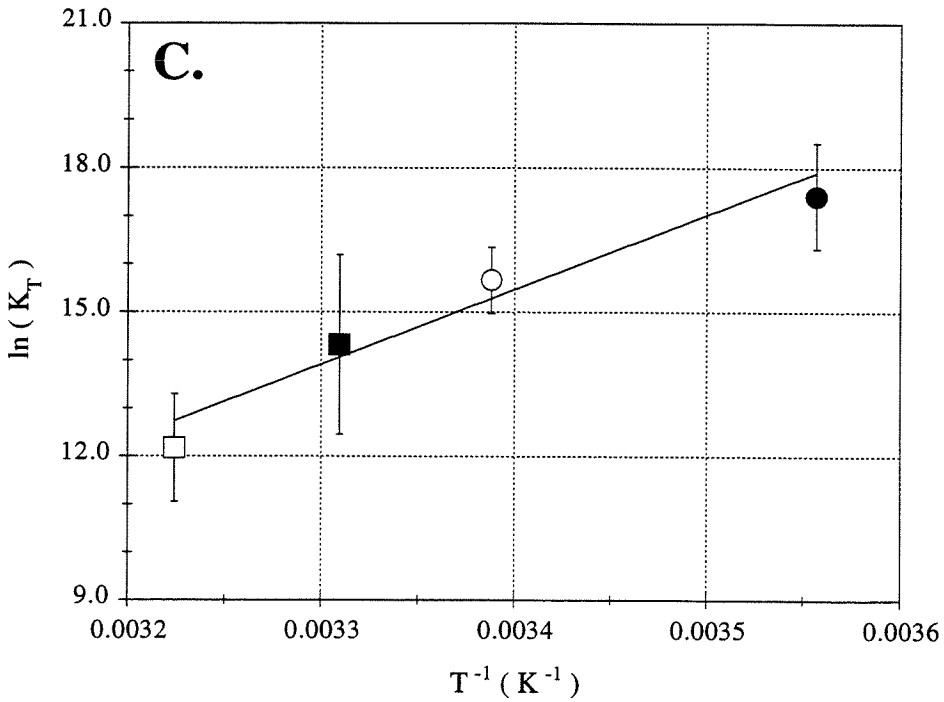


Figure 5.2. The natural logarithms of the mean association constants (Table 5.1) are plotted as a function of the reciprocal of the absolute temperature. The symbols represent the same temperatures as in Figure 5.1. Error bars represent estimated confidence limits. (A) Apparent equilibrium association constants measured in BTNS buffer. (B) Apparent equilibrium association constants measured in BTP1 buffer. (C) Apparent equilibrium association constants measured in BTP4 buffer.


B.
21.0
18.0
15.0
12.0
9.0
ln (K_T)
0.0032
0.0033
0.0034
0.0035
0.0036
T^{-1} (K^{-1})


C.
21.0
18.0
15.0
12.0
9.0
ln (K_T)
0.0032
0.0033
0.0034
0.0035
0.0036
T^{-1} (K^{-1})

(eq 1) and the entropies to be calculated from the intercept. Estimated uncertainties in each K_T were used to weight the linear fits and the uncertainties were propagated through the fitted parameters to the thermodynamic parameters. The linear fits are of high quality, having correlation coefficient values > 0.96 and χ^2 values < 0.8 in every case. The thermodynamic parameters derived in this manner are compiled in Table 5.2.

Table 5.2. Apparent Thermodynamic Parameters for Oligonucleotide-Directed Triple Helix Formation in 10 mM Bis-tris·HCl at pH 7.0.[a]

	100 mM NaCl 1 mM $SpmCl_4$ [e]	10 mM NaCl 140 mM KCl 1 mM $MgCl_2$ 1 mM $SpmCl_4$ [f]	10 mM NaCl 140 mM KCl 1 mM $MgCl_2$ 4 mM $SpmCl_4$ [g]
ΔG_T [b]	−8.8 (±0.4)	−7.6 (±0.9)	−9.2 (±0.2)
$\Delta H_T°$ [c]	−31 (±9.1)	−28 (±5.6)	−31 (±9.2)
$\Delta S_T°$ [d]	−75 (±31)	−67 (±19)	−74 (±31)

[a] The ΔG_T and $\Delta H_T°$ values are reported in units of kcal·mol^{-1}, while those of $\Delta S_T°$ are given in cal·mol^{-1}·K^{-1}. Estimated uncertainties in the thermodynamic parameters are given in parentheses. [b] The ΔG_T values are the values derived at 22°C and the indicated salt concentrations from Table 5.1. [c] The $\Delta H_T°$ values are derived from the slopes of the van't Hoff plots (Figure 5.2). [d] The $\Delta S_T°$ values are calculated from the intercepts of the van't Hoff plots. [e] BTNS buffer. [f] BTP1 buffer. [g] BTP4 buffer.

For each set of salt conditions at pH 7.0, the enthalpic contribution to the free energy of triple helix formation is relatively large and negative, in agreement

with previous measurements using oligonucleotide triplexes. In contrast, the entropic contribution from the $-T\cdot\Delta S$ term is large and positive, also as found in previous investigations of oligonucleotide triplexes. Under the solution conditions employed here, the entropy and enthalpy changes are indistinguishable within experimental uncertainty. Although affinity cleavage titrations provide accurate free energies, the small uncertainties in each K_T are magnified to produce larger relative uncertainties in $\Delta H_T°$ and $\Delta S_T°$. On average, binding of the 15mer oligonucleotide to the duplex target site under the range of conditions studied here is accompanied by an enthalpy change of –2.0 kcal per mole of base triplets and an entropy change of –4.8 cal·K^{-1} per mole of base triplets. It is important to point out that these *average* values contain contributions from both T·AT and C+GC base triplets, whose individual stabilities may be distinct.

Discussion

Thermodynamics of Oligonucleotide-Directed Triple Helix Formation. A minimal equilibrium expression for the formation of a triple helix ($\mathbf{S_1\cdot S_2\cdot S_3}$) from single-stranded ($\mathbf{S_1}$) and double-stranded ($\mathbf{S_2\cdot S_3}$) nucleic acid components in a solution containing buffer (**B**) and mono-, di-, and multivalent cations ($\mathbf{M^+}$, $\mathbf{N^{2+}}$, and $\mathbf{Q^{z+}}$, respectively) can be written as

$$\mathbf{S_1} + \mathbf{S_2\cdot S_3} + i\mathbf{M^+} + j\mathbf{N^{2+}} + k\mathbf{Q^{z+}} + l\mathbf{BH^+} \rightleftharpoons \mathbf{S_1H}_l\mathbf{\cdot S_2\cdot S_3} + l\mathbf{B} \qquad (2)$$

where i, j, and k represent the numbers of $\mathbf{M^+}$, $\mathbf{N^{2+}}$, and $\mathbf{Q^{z+}}$ ions, respectively, thermodynamically bound per phosphate during the association reaction,[5] and l is the number of protons transferred from the buffer in the formation of any pro-

tonated C+GC base triplets.[4,19] Evidence obtained by a number of laboratories indicates that the factors influencing the stability of triplexes include the following: (i) the nature of $\mathbf{S}_1$, including its length,[3,20] sequence composition,[19-21] and functional groups on both the heterocycles[2,22,23] and backbone;[9,10,24,25] (ii) the 2'-substituents on the ribosyl moieties of $\mathbf{S}_2$ and $\mathbf{S}_3$;[9,10,25] (iii) the identities and concentrations of cations;[1,2,5,18,2026,27] (iv) the solution pH;[1,2,4,19,20,22,27,28] and (v) the identity and ionization heat of **B**.[19] To date, most thermodynamic data has been reported for oligomeric triplexes in solutions containing single cationic species at low pH. We have reported the extension of the affinity cleavage titration method to the measurement of enthalpic and entropic contributions to the stabilities of single-site local triple-helical complexes formed within large DNA at neutral pH in the presence of physiologically important salt concentrations.

In each set of salt conditions at pH 7.0, the enthalpic contribution to the free energy of triple helix formation is relatively large and negative (Table 5.2), in accord with expectations based on the formation of hydrogen bonds and π-π interactions between the bases. In contrast, the entropic contribution to the free energy is large and positive as would be anticipated for the loss of translational and rotational degrees of freedom for the oligonucleotide[28,29] and for the condensation of cations around the triple helix.[5,30] These results and expectations also agree with previous investigations of oligonucleotide triplexes. Changes in salt composition among the three buffers were expected to be reflected in the entropic contribution from counterion condensation.[5,30] Moreover, by increasing the concentrations of mono- and divalent cations (from BTNS buffer to BTP1 buffer) or by increasing the concentration of Spm^{4+} (from BTP1 to BTP4), we anticipated small changes in ΔH_T° based on differential changes in the charge screening parameter.[5,26,31] In the three solutions studied here, these small differ-

ences in $\Delta H_T°$ and $\Delta S_T°$ are indistinguishable within the experimental uncertainties.

Comparison with Enthalpy Measured by DSC. In collaboration with Breslauer and co-workers, we characterized the thermodynamic stability of a triplex formed by the same oligonucleotide sequence as that studied here using differential scanning calorimetry and optical spectroscopy.[2] In the previous work, the triple helix formed by binding of the oligonucleotide to the 15-bp sequence within a 21-bp duplex at pH 6.5 (phosphate buffer) in the presence of 200 mM Na^+ was enthalpically stabilized by 30.4 (± 2) kcal·mol^{-1}. The apparent enthalpy measured in each salt solution used in the present study is identical within experimental uncertainty to that measured calorimetrically. While there are undoubtedly differences in the true enthalpies based on differences in the length and sequence of the duplex neighboring the target site, the salt compositions,[26] the pH,[19,28] and the buffering species (see eq 2),[19] the magnitudes of the differences are likely reduced by fortuitous cancellation of the contributing factors.

Comparison with Other Enthalpy Measurements. In order to compare the $\Delta H_T°$ measured here with values reported in the literature for oligonucleotide triple helices, we considered that there are at least three different kinds of oligodeoxyribonucleotide triplexes based on their sequence composition and the solution pH. First, triplexes containing only T·AT base triplets are distinct from those containing both C+GC and T·AT triplets for the following reasons: (i) enthalpies for T·AT-only triplexes result from only a single type of base triplet or nearest neighbor triplet pair (T·AT/T·AT) rather than an average over several types; (ii) as previously suggested, oligo(dA)·oligo(dT) duplexes may have unusual conformations and/or hydration states;[19] (iii) the protonation of C+GC

triplets contributes to the observed enthalpy of triple helix formation, and ionized cytosine heterocycles likely alter the enthalpies associated with hydrogen-bonding and base stacking as well as reduce the charge density of the triple-helical complex. Based on reason (iii), triplexes containing both C+GC and T·AT triplets were further divided into two categories depending on the pH used in the measurement. Over the pH range 4.5 – 6.0, the C+GC triplets are mostly protonated, while over the range 6.5 – 7.0, the protonation state likely varies among different sequence and solution compositions. The ranges of reported *dissociation* enthalpies per base triplet are compiled in Table 5.3.[2,18,19,26,28,32-38]

The magnitude of the average *association* enthalpy change per base triplet measured here (2.0 kcal per mole of base triplet) is on the low end of the range of values reported for triplexes containing C+GC and T·AT triplets from pH 6.5 to pH 7.0 (2.0 – 5.9 kcal·mol-bt^{-1}). One potential contributing source of the relative placement is the requirement for the formation of junctions between double- and triple-helical regions. The triple-helical complex reported here, and previously,[2] occurs within a larger double-helical DNA, and the conformational differences between the two domains results in the formation of two duplex-triplex junctions. Although there is an enthalpic penalty associated with junction formation, it is unlikely that the magnitude of such a penalty could account for $\Delta\Delta H°$ = 3.9 kcal·mol-bt^{-1}.[12,39-41] In addition to enthalpic differences resulting from differences in the identities of cation species, cation concentrations, pH, and buffering species (eq 2), the effect of sequence composition on triplex thermodynamics has not been systematically investigated. For example, the arrangement of cytosine residues in the sequence studied here, in which each cytosine is separated from the next by a single thymidine residue (see Figure 5.1.1), may render the complex

particularly sensitive to the protonation state and stability of the individual C+GC triplets.

Table 5.3. Literature Values of the Apparent Triple Helix Dissociation Enthalpy Change.[a]

		Calorimetric		Non-calorimetric	
Nts [b]	pH [c]	$\Delta H°_{app}$ [d]	Refs.	$\Delta H°_{app}$ [d]	Refs.
T	5.5 – 7.1	2.5	*e*	1.9 – 4.5	*f*
C,T	4.5 – 6.0	4.7 – 6.2	*g*	3.4 – 8.1	*h*
	6.5 – 7.0	2.0 – 4.2	*i*	4.8 – 5.9	*j*

[a] The ranges of enthalpies previously measured using calorimetric or non-calorimetric methods. [b] The nucleotide compositions of the third strands in the triplexes studied. "T" indicates the triplex contained T·AT base triplets only; "C,T" indicates the triplex contained both C+GC and T·AT base triplets. [c] The pH range indicated contains each solution pH used for measurements of the tabulated enthalpy changes. [d] The apparent dissociation enthalpy changes are reported in kcal per mole of base triplets. [e] References 119 and 133. [f] References 119, 134, and 135. [g] References 119, 128, and 134. [h] References 132, 126, 128, 137, and 138. [i] References 102, 119, and 128. [j] References 118 and 128.

Each of the variables discussed above (under eq 2) contributes to the enthalpic stabilization of triple-helical complexes and to differences between individual complexes, making comparisons among experiments difficult. Further, discrepancies between enthalpies measured using different techniques have also

been reported. For two triplexes — one, the intermolecular triplex formed between the 15mer and 21-bp duplex discussed above,[2] and the other, a 10mer intramolecular triplex formed by folding of a 38mer[28] — the enthalpy determined by direct integration of calorimetric excess heat capacity curves (ΔH_{cal}) was significantly less than that determined by van't Hoff analysis of the shapes of calorimetric or optical melting curves (ΔH_{vH}) for the same complex. It is important to note that this behavior is seen at pH 6.7, but not at pH 4.5, for the intramolecular triplex. The observation that $\Delta H_{vH} > \Delta H_{cal}$ is usually interpreted as evidence for aggregation of the complex undergoing the transition; however, in the cases of the intermolecular triplex, there is evidence inconsistent with aggregate formation,[42] and there is no corroborating experimental evidence for aggregation in either case.[2,28] Alternatively, the assumption that the conformational transition is a true two-state process, which is required for the extraction of thermodynamic parameters from the analysis of melting curve shapes, may not hold. Although the resolution of the discrepancy between these two methods must await the application of more precise biophysical techniques, the results of the van't Hoff analyses described herein provide independent corroboration of the enthalpy measured using calorimetry for one system and may be important for understanding the origins of the discrepancy.

Conclusions

In summary, we have extended the affinity cleavage titration method to the measurement of van't Hoff enthalpy and entropy parameters for oligonucleotide-directed triple helix formation. The significance of this work is that it allows the comparison between thermodynamic parameters for local triple-helical complexes formed at single sites within larger duplex DNA with those for oligomeric

triplexes. Moreover, the analysis presented here provides an independent corroboration of a prior measurement made using DSC. We observe that, on average, the local triple-helical complex is enthalpically stabilized by ca. 2 kcal per mole of base triplets at pH 7.0 in solutions containing several cations, including spermine. However, the average value reflects contributions from T·AT, protonated C+GC, and unprotonated C·GC base triplets, each of whose enthalpic stablizations are likely distinct. Future biophysical measurements will reveal "true" enthalpies for each base triplet, which will undoubtedly depend on sequence context effects.

Experimental Section

Stock Solutions of Buffer, Salts, and Carrier DNA. Bis-tris [bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane; Fluka BioChemika MicroSelect], sodium chloride (Fluka BioChemika MicroSelect), potassium chloride (Mallinckrodt), magnesium chloride (Mallinckrodt), and spermine tetrahydrochloride (Fluka BioChemika MicroSelect) were used as obtained from commercial suppliers. In order to obtain three pH 7.0-buffered solutions containing various salt compositions at each of four different temperatures, 12 different stock solutions were required. First, six stock solutions containing 5x concentrations of the desired salts and 5x concentration of *either* Bis-tris free base or its conjugate acid were prepared. These six stock solutions (and the concentration of solutes) were BTNSa (50 mM Bis-tris·H^+, 500 mM NaCl, 5 mM $SpmCl_4$), BTNSb (50 mM Bis-tris, 500 mM NaCl, 5 mM $SpmCl_4$), BTP1a (50 mM Bis-tris·H^+, 50 mM NaCl, 700 mM KCl, 5 mM $MgCl_2$, 5 mM $SpmCl_4$), BTP1b (50 mM Bis-tris, 50 mM NaCl, 700 mM KCl, 5 mM $MgCl_2$, 5 mM $SpmCl_4$), BTP4a (50 mM Bis-tris·H^+, 50 mM NaCl, 700 mM KCl, 5 mM $MgCl_2$, 20 mM $SpmCl_4$), and

BTP4b (50 mM Bis-tris, 50 mM NaCl, 700 mM KCl, 5 mM $MgCl_2$, 20 mM $SpmCl_4$). The solutions containing the conjugate acid of Bis-tris were prepared identically to those containing the free base except that a few drops of 1 M HCl were added to a pH ≈ 4. Based on the known temperature-dependence of the pK_a of Bis-tris, the pairs of stock solutions were mixed at room temperature to give twelve solutions whose pH's were a few tenths of a unit above the room temperature value corresponding to pH 7.0 at the desired temperature. Each of the solutions were then equilibrated in a water bath at the appropriate temperature and the pH adjusted to 7.0 at that temperature using 1 M HCl. The pH of the solutions was measured using a digital pH/millivolt meter (Orion Research, model no. 611) and a ROSS semimicro combination pH electrode (Orion Research, model no. 81-15). The final 12 buffers ready for use at the appropriate temperature were BTNS-08, -22, -29, and -37 (50 mM Bis-tris·HCl, 500 mM NaCl, and 5 mM $SpmCl_4$), BTP1-08, -22, -29, and -37 (50 mM Bis-tris·HCl, 50 mM NaCl, 700 mM KCl, and 5 mM $SpmCl_4$), and BTP4-08, -22, -29, and -37 (50 mM Bis-tris·HCl, 50 mM NaCl, 700 mM KCl, and 20 mM $SpmCl_4$).

Temperature Control. Triple-helix formation and affinity cleavage reactions performed at low temperature were carried out in 0.6-mL tubes suspended in a circulating cooling bath (Lauda) containing ≈ 20% aqueous ethylene glycol thermostated at 8.0 (± 0.2)°C. Reaction tubes for high-temperature titrations were placed in aluminum heating blocks (VWR) at 29.0 (±3.0) or 37.0 (±2.0)°C. Room temperature experiments were performed on the laboratory bench top at a temperature of 22.0 (±2.0)°C.

DNA Preparation and Affinity Cleavage Titrations. The oligonucleotide-EDTA and labeled duplex DNA were prepared as described in Chapter One. Affinity cleavage titrations were performed essentially as described in Chapter One.

Titration binding isotherms were constructed and equilibrium association constants were determined as described in Chapter One.

At low concentrations of oligonucleotide, iron loading in the thymidine-EDTA can limit the observed cleavage intensity and, thus, the apparent fraction of duplex sites occupied. Therefore, mean K_T values $\geq 3 \times 10^6$ M^{-1} were confirmed by performing titrations in which 20 μM Fe^{2+} was included in the reaction solutions to insure maximal oligonucleotide-EDTA loading.[42] Moreover, quantitative DNase I footprint titrations, which do not rely on the EDTA·Fe cleavage chemistry, were employed to independently confirm these values.

All titrations included carrier DNA (from calf thymus) at a final concentration of 100μM in base pairs. The inclusion of a relatively high concentration of potential binding sites can, in general, reduce the true concentration of free ligand significantly below the total concentration of ligand (see Chapter One Part IV), and is thus not recommended for all ligand-binding studies of this type. However, affinity cleavage titrations in the absence of carrier, in the presence of various concentrations of calf thymus DNA (10 – 100 μM-bp), or in the presence of tRNA all give mean equilibrium association constants that are identical within experimental uncertainty.[42] Likely, the presence of duplex carrier DNA does not affect the observed equilibrium association constant because the oligonucleotide binding site is relatively long and triple helix formation occurs with high sequence selectivity. We chose to include carrier DNA in order to maximize the quality and reproducibility of cleavage results from individual experiments.

Quantitative DNase I Footprint Titrations. Using the solution conditions described above, footprint titration experiments and data reduction and analysis were performed according to published protocols.[3]

References

(1) Maher, L. J., III; Dervan, P. B.; Wold, B. J. *Biochem.* **1990**, *29*, 8820-6.

(2) Plum, G. E.; Park, Y. W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 9436-40.

(3) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957-65.

(4) Singleton, S. F.; Dervan, P. B. *Biochem.* **1992**, *31*, 10995-1003.

(5) Singleton, S. F.; Dervan, P. B. *Biochem.* **1993**, *32*, 13171-9.

(6) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172-4.

(7) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-42.

(8) Strobel, S. A.; Dervan, P. B. *Methods Enzymol.* **1992**, *216*, 309-21.

(9) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463-6.

(10) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3806-10.

(11) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-30.

(12) Collier, D. A.; Nguyen, T. T.; Hélène, C. *J. Am. Chem. Soc.* **1991**, *113*, 1457-8.

(13) Maher, L. J., III; Dervan, P. B.; Wold, B. *Biochem.* **1992**, *31*, 70-81.

(14) Grigoriev, M.; Praseuth, D.; Robin, P.; Hemar, A.; Saison-Behmoaras, T.; Dautry-Varsat, A.; Nguyen, T. T.; Hélène, C.; Harel-Bellan, A. *J. Biol. Chem.* **1992**, *267*, 3389-95.

(15) Duval-Valentin, G.; Thuong, N. T.; Hélène, C. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 504-8.

(16) Hacia, J. A.; Wold, B.; Dervan, P. B. *Biochem.* **1994**, in press.

(17) Cheng, Y.-K.; Pettit, B. M. *Prog. Biophys. Molec. Biol.* **1992**, *58*, 225-257.

(18) Rougée, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochem.* **1992**, *31*, 9269-78.

(19) Wilson, W. D.; Hopkins, H. P.; Mizan, S.; Hamilton, D. D.; Zon, G. *J. Am. Chem. Soc.* **1994**, *116*, 3607-3608.

(20) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-50.

(21) Kiessling, L. L.; Griffin, L. C.; Dervan, P. B. *Biochem.* **1992**, *31*, 2829-34.

(22) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-61.

(23) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; Van der Marel, G. A.; Van Boom, J. H. *Nucleic Acids Res* **1991**, *19*, 5625-31.

(24) Escude, C.; Sun, J. S.; Rougée, M.; Garestier, T.; Hélène, C. *C. R. Acad. Sci., Ser. Iii* **1992**, *315*, 521-5.

(25) Escude, C.; Francois, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Hélène, C. *Nucleic Acids Res.* **1993**, *21*, 5547-5553.

(26) Pilch, D. S.; Brousseau, R.; Shafer, R. H. *Nucleic Acids Res.* **1990**, *18*, 5743-50.

(27) Hanvey, J. C.; Williams, E. M.; Besterman, J. M. *Antisense Res. Dev.* **1991**, *1*, 307-17.

(28) Völker, J.; Botes, D. P.; Lindsey, G. G.; Klump, H. H. *J. Mol. Biol.* **1993**, *230*, 1278-90.

(29) DeVoe, H.; Tinoco, I., Jr. *J. Mol. Biol.* **1962**, *4*, 500-517.

(30) Manning, G. S. *Q. Rev. Biophys.* **1978**, *11*, 179-246.

(31) DeMarky, N.; Manning, G. S. *Biopolymers* **1975**, *14*, 1407-1422.

(32) Roberts, R. W.; Crothers, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 9397-401.

(33) Hopkins, H. P.; Hamilton, D. D.; Wilson, W. D.; Zon, G. *J. Phys. Chem.* **1993**, *97*, 6555-6563.

(34) Sugimoto, N.; Shintani, Y.; Tanaka, A.; Sasaki, M. *Bull. Chem. Soc. Jpn.* **1992**, *65*, 535-40.

(35) Durand, M.; Peloille, S.; Thuong, N. T.; Maurizot, J. C. *Biochem.* **1992**, *31*, 9197-9204.

(36) Xodo, L. E.; Manzini, G.; Quadrifoglio, F. *Nucleic Acids Res.* **1990**, *18*, 3557-64.

(37) Manzini, G.; Xodo, L. E.; Gasparotto, D.; Quadrifoglio, F.; Van der Marel, G. A.; Van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833-43.

(38) Mooren, M. M. W.; Pulleybank, D. E.; Wijmenga, S.; Blommers, M. J. J.; Hilbers, C. W. *Nucleic Acids Res.* **1990**, *18*, 6523-6529.

(39) Sun, J. S.; Lavery, R.; Chomilier, J.; Zakrzewska, K.; Montenay-Garestier, T.; Hélène, C. *J. Biomol. Struct. Dyn.* **1991**, *9*, 425-36.

(40) Chomilier, J.; Sun, J. S.; Collier, D. A.; Garestier, T.; Hélène, C.; Lavery, R. *Biophys. Chem.* **1992**, *45*, 143-52.

(41) Yoon, K.; Hobbs, C. A.; Walter, A. E.; Turner, D. H. *Nucleic Acids Research* **1993**, *21*, 601-606.

(42) Singleton, S. F. Ph.D. Thesis, California Institute of Technology, 1994.

CHAPTER SIX

Thermodynamic Characterization of the Stability and the Melting Behavior of a DNA Triplex: A Spectroscopic and Calorimetric Study

The text of this chapter (with the exception of the General Background and the Addendum) is taken from a published manuscript that was co-authored with Professor Peter B. Dervan and our collaborators at Rutgers University, Professor Kenneth J. Breslauer, Dr. Eric Plum , and Young-Whan Park.
(G. E. Plum et al., *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436-9440.)

General Background

The *rational* design and use of an oligonucleotide (or analog) for the modulation of a biologically important event at any arbitrary DNA sequence requires an understanding of how the stability of the triple-helical complex depends on its component double- and single-stranded nucleic acids and the solution conditions. The characterization of the linkage between association free energies and solution conditions represents the first steps toward the goal of evaluating the energetics accompanying the formation of local triple-helical complexes at single sites within larger duplex DNAs (1-4). A knowledge of how the fundamental properties of a triplex lead to its relative stability or instability will be of even greater practical utility for the *de novo* design of oligonucleotides. For example, it

is important to link the presence of substituents on third strand nucleotides with either enthalpic changes (including those resulting from hydrogen bonding, base stacking, or van der Waals interactions) or entropic changes (such as those originating from hydrophobic effects or counterion condensation). Furthermore, the enthalpy changes associated with nucleic acid conformational transitions are generally less sensitive to changes in temperature and cation concentrations than the free energies, and can be more easily compared between different experiments (5). We have collaborated with Professor Kenneth Breslauer of Rutgers University to measure the enthalpic and entropic contributions to triple helix stability using both spectroscopic methods and differential scanning calorimetry (DSC) (6).

Methods for Obtaining Thermodynamic Data. Various spectroscopic observables have been employed for the indirect determination of transition enthalpies for nucleic acid conformational changes (7,8). In general, these methods rely on the measurement of the dependence of the equilibrium constant on the temperature and the use of either form of the van't Hoff equation shown below.

$$\Delta H_{\mathrm{vH}} = \mathrm{R}T^2 \cdot \frac{\partial \ln K}{\partial T} \tag{1}$$

$$\Delta H_{\mathrm{vH}} = -\mathrm{R} \cdot \frac{\partial \ln K}{\partial (1/T)} \tag{2}$$

Consider the following two-state equilibrium involving an oligonucleotide, **O**, a duplex, **D**, and the corresponding triple-helical complex, **T**:

$$\mathbf{O} + \mathbf{D} \underset{}{\overset{K_{\mathrm{T}}}{\rightleftharpoons}} \mathbf{T} \tag{3}$$

At the start of a spectrophotometric measurement, **O** and **D** are usually present at the same concentration. If C_{tot} represents the total concentration of molecules, then $O_0 = D_0 = C_{tot}/2$, where italicized symbols represent concentrations and the subscript "0" indicates initial conditions. The equilibrium constant (K_T) and the fraction of molecules bound as the triplex (α) can be written as

$$K_T = \frac{T_{eq}}{O_{eq} \cdot D_{eq}} \tag{4}$$

$$\alpha = \frac{T_{eq}}{C_{tot}/2} \tag{5}$$

where the subscript "eq" designates an equilibrium concentration. Eqs 4 and 5 can be combined to express K_T as a function of α

$$K_T = \frac{\alpha}{(1-\alpha)^2} \cdot \frac{1}{C_{tot}/2} \tag{6}$$

Upon taking the natural logarithm of both sides and differentiating with respect to temperature, eq 6 is transformed into

$$\frac{\partial \ln K_T}{\partial T} = \left(\frac{1}{\alpha} + \frac{2}{1-\alpha}\right) \cdot \frac{\partial \alpha}{\partial T} \tag{7}$$

Substitution of the right-hand-side of eq 7 into eq 1, followed by use of the definition $\alpha \equiv 0.5$ at the melting temperature yields the following expression for the van't Hoff transition enthalpy:

$$\Delta H_{vH}(T_m) = 6RT_m^2 \cdot \left(\frac{\partial \alpha}{\partial T}\right)_{T=T_m} \quad (8)$$

Eq 8 provides for the calculation of ΔH_{vH} at the melting temperature by determining the slope of a plot of α versus T at T_m. The temperature-dependence of α can be obtained from a melting curve by measuring the fractional progression of a spectroscopic property from one state (*e.g.*, **T**) to the other (*e.g.*, **O** + **D**) as the temperature is increased (Figure 6.1). Breslauer has presented general equations that allow the extraction of van't Hoff enthalpies for transitions of any molecularity from a variety of spectroscopic and non-spectroscopic melting curves (7,8).

The direct measurement of the heat change accompanying a transition using a calorimeter is an alternative approach to the shape analysis of optical melting curves. Several detailed reviews have presented the technical features and uses of differential scanning calorimeters (9-11). DSC can be used to follow thermally-induced order-disorder transitions in nucleic acids; however, with DSC, the excess heat capacity (ΔC_p) rather than an optical property of the solution is monitored, and the calorimeter allows one to continuously measure the heat capacity of a system relative to an external reference as a function of temperature. Figure 6.2 shows a schematic of a typical DSC cell compartment, in which are suspended two adiabatically shielded cells connected via a thermopile. In a typical experiment, the sample is placed in one cell, a solution containing only buffer and salts is placed in the reference cell, and the temperature is scanned from about 4°C to about 100°C. When a thermally induced endothermic transition occurs in the sample cell, its temperature will lag behind that of the reference cell because some of the heat energy will be absorbed in the transition rather than by

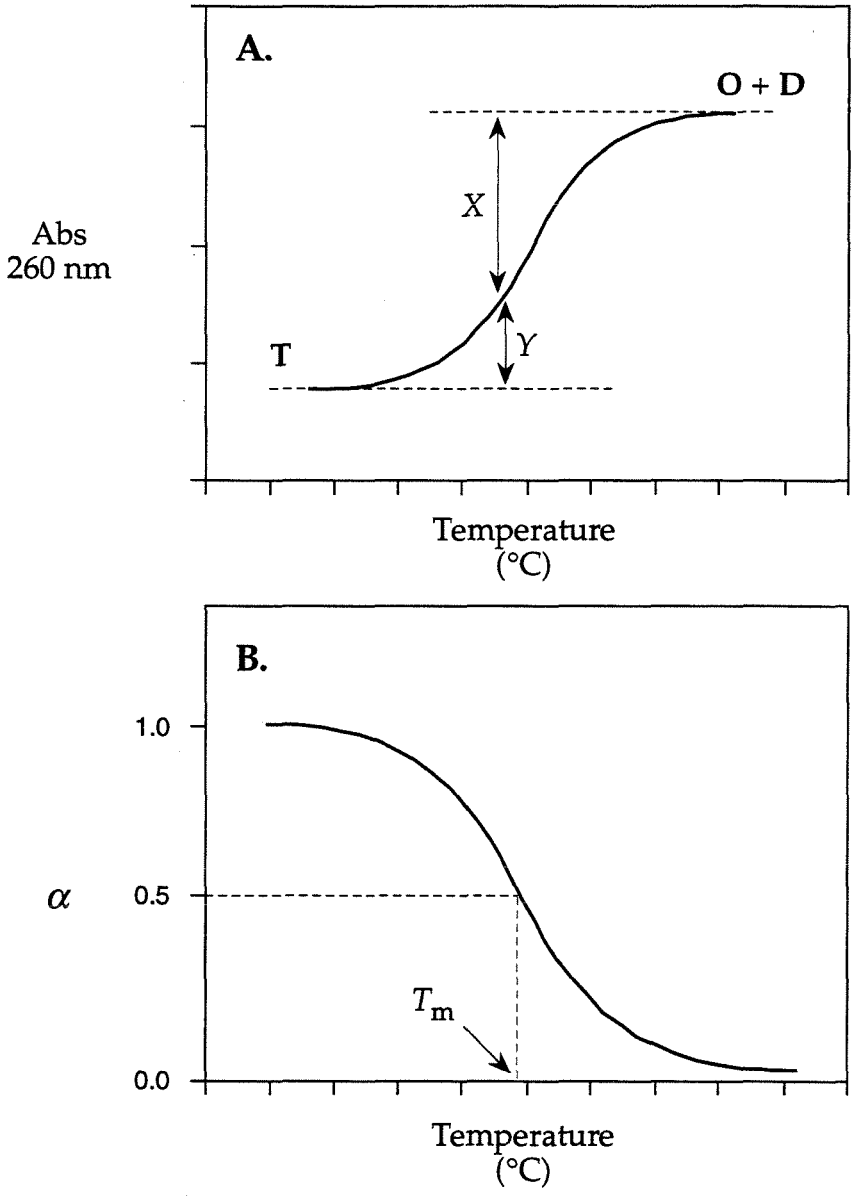


Figure 6.1. (A) A typical absorbance versus temperature melting curve. X represents the distance between the curve and the upper baseline (corresponding to absorption by the products), and Y represents the distance between the curve and the lower baseline (corresponding to absorption by the initial complex). The fraction of strands in the complex, α, is given by the ratio $X/(X + Y)$. (B) A plot of α versus temperature. The temperature corresponding to $\alpha \equiv 0.5$ is defined as the melting temperature of the transition (T_m).

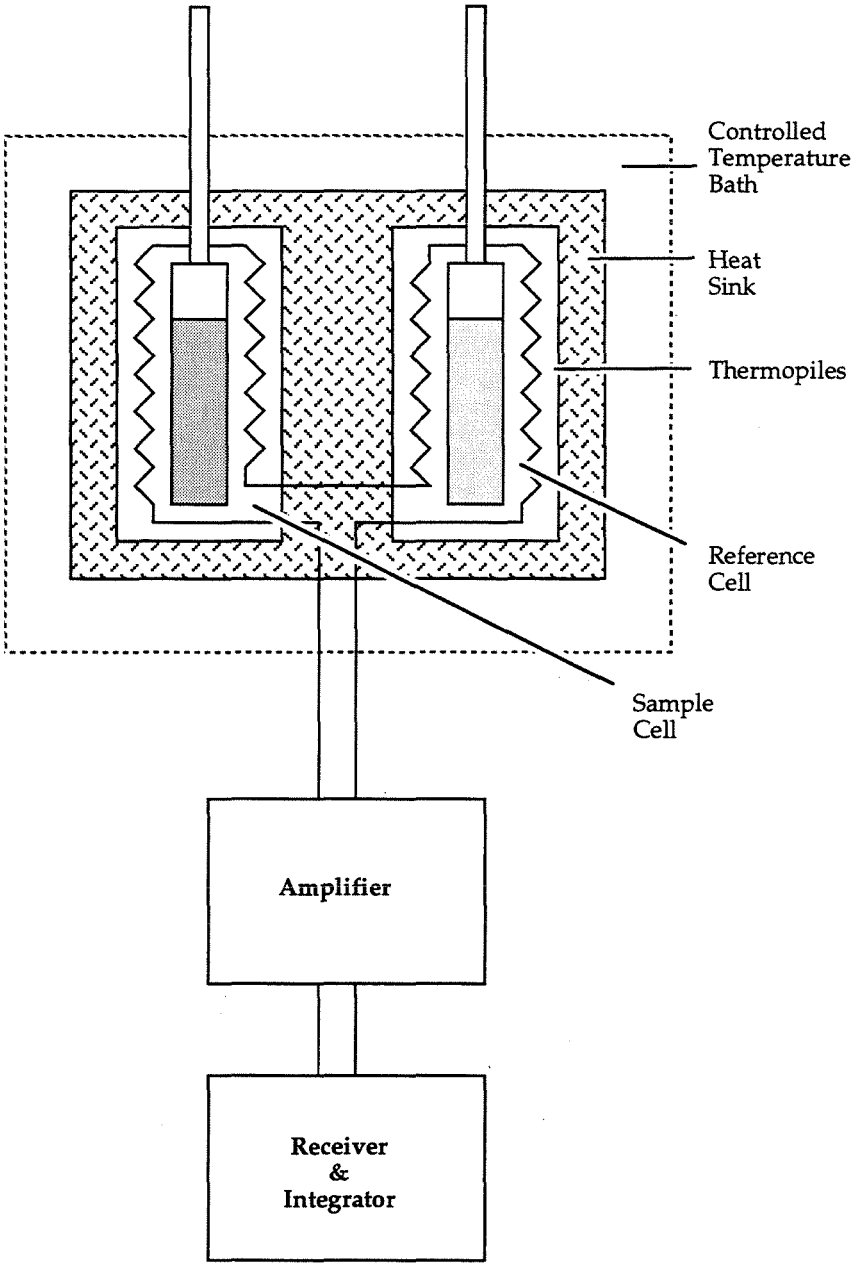


Figure 6.2. A schematic diagram showing the standard components of a typical differential scanning calorimeter.

the solution. The temperature difference is detected by the thermopile, extra heat is added to the sample cell to maintain the same temperature in the two cells, and the amount of extra heat is recorded. Because $\Delta H^\circ = \int \Delta C_p \cdot dT$, the area under the calorimetric transition curve is proportional to the energy of the transition. Calibration of the calorimetric area using the area produced by an electrical heating pulse of known energy, followed by division of the calibrated area by the concentration of complex present affords the transition enthalpy. Furthermore, a ΔC_p versus T profile can be derived from the experimental calorimetric curve, and, because $\Delta S^\circ = \int (\Delta C_p / T) \cdot dT$, the area under the new curve can be used to calculate the transition entropy. Thus, whereas ΔH° and ΔS° can be estimated indirectly using a van't Hoff equation and an assumption of a two-state process, thermodynamic parameters can be directly measured by integration of experimental calorimetric profiles without regard to transition models or assumptions.

Selection of the Experimental System. In order to define a thermodynamic baseline for triple helix formation, we chose to investigate the complex formed by binding of the 15mer 5'-TTTTTCTCTCTCTCT-3' (y15; see Figure 6.3B, below) to 15 bp within a 21-bp duplex (u21·y21). The sequence of the third strand and its target duplex were chosen because they had been characterized using affinity cleavage experiments over a range of sequence conditions (12,13). The 15-bp target duplex sequence was appended with three base pairs on each side in order to minimize end-effects resulting from fraying. Moreover, this complex is a minimal model of one formed by binding of an oligonucleotide to a duplex sequence within a larger double-helical DNA.

One of the most important concerns in selecting a system for investigation was the choice of solution conditions. The requirements included the following:

(i) relatively high complex stability; (ii) two well-resolved transitions for the triplex and duplex melting steps; (iii) solution conditions approximating intracellular conditions; and (iv) simple conditions. The first two considerations are important for analysis of the melting transitions. If the complex is not fully formed at the initial temperature or if the triplex and duplex melting transitions overlap, extraction of melting temperatures and thermodynamic parameters will be unnecessarily complicated. Because our interest lies ultimately in characterizing the stability of triplexes under intracellular conditions, the buffered salt solution should reflect those conditions as nearly as possible. Finally, in order to prevent complications in the calorimetric analysis, the solution conditions were kept as simple as possible.

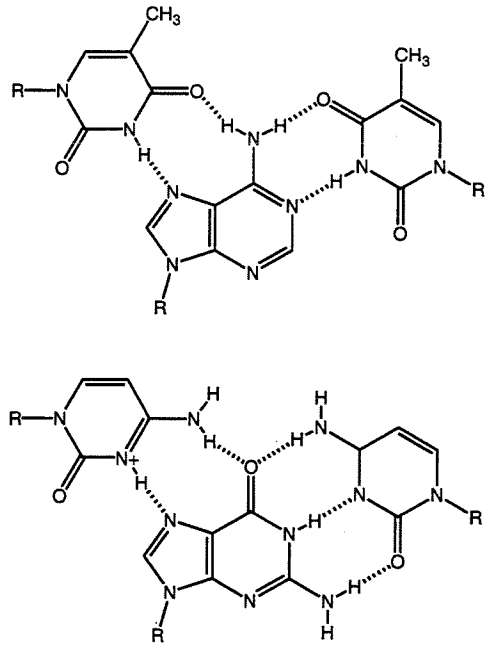
Sodium phosphate was chosen as the buffer because it has a low heat of ionization and a pK_a of 6.2 (14). The use of phosphate buffer precluded the use of magnesium ions because Mg^{2+} and PO_4^{3-} coprecipitate at high temperature. Polyamines such as spermine were excluded because their pK_a's (and, hence, their charges) generally change with temperature and they often possess high heats of ionization, characteristics which would complicate the melting profiles. The melting temperature (T_m) of the triplex was measured over the pH range 6.1 to 6.8 and over the NaCl concentration range 0.02 to 1.0 M. The requirements listed above were best fulfilled at pH 6.5 with $[Na^+]$ = 0.2 M, where the T_m of the triplex was 30°C and that of the duplex was 65°C. Although cytoplasm contains many different ions, and the identity and concentration of each can affect triplex stability (3), a salt concentration of 0.2 M affords an ionic strength that is similar to that of the intracellular matrix.

Introduction

Over three decades have passed since the first description of polynucleotide triple helices (15). In the ensuing years a small number of investigators interested in the fundamental properties of nucleic acids have studied the structure (16) and physical properties (17-19) of triple-helical nucleic acids. Most of the work in this area has focused on triple helices composed of one polypurine strand and two polypyrimidine strands; however, triplexes of (polypurine)$_2$·polypyrimidine also are known (20-23).

The widely accepted structural model for polypurine·(polypyrimidine)$_2$ triple helices is based on x-ray fiber diffraction studies on poly(A)·poly(U)$_2$ (16,24) and poly(dA)·poly(dT)$_2$ (16,25). In this structure, an A-form polypurine·polypyrimidine duplex joined by Watson-Crick base pairs binds the second polypyrimidine strand in its major groove. This second polypyrimidine strand interacts with the Watson-Crick duplex by means of Hoogsteen base pairing to form the base triplets T·AT and C+GC (Figure 6.3A). The formation of a triple helix from cytosine containing third strands has been shown to be pH dependent, suggesting that the C+GC triplet requires protonation of the cytosine (12,13,26). Oligonucleotide-directed site-specific cleavage of double-helical DNA establishes that the third strand lies in the major groove parallel to the purine strand of the Watson-Crick duplex (12). Recent NMR studies on oligomeric DNA triplexes are consistent with some of the structural details ascertained by less direct techniques and have begun to provide a more detailed picture of the

A.



B.

```
y15      5'-TTTTTCTCTCTCTCT-3'
u21  5'-GCTAAAAAGAGAGAGAGATCG-3'
y21  3'-CGATTTTTCTCTCTCTCTAGC-5'
```

Figure 6.3. (A) Nucleotide base triplets: T·A-T and C+G-C. (B) Sequences of the three oligonucleotides along with their designations. The sequences are displayed to emphasize the complementarity and polarity of the strands in the triplex.

DNA triple helix (27-30). Additional studies are required before a complete picture of triplex structures and their sequence dependence emerges.

Recent interest in triple-helical DNA has been stimulated by the discovery of novel triplex-containing structures, such as H-form DNA, which have been proposed to explain the enhanced sensitivity of mirror repeat polypurine-polypyrimidine sequences to chemical modification (31-33). Moreover, oligonucleotide-directed triple helix formation has the potential to be a general solution for DNA recognition, which has implications for physical mapping of chromosomes and site-specific inhibition of transcription *in vivo* (12,13,34-39).

An essential requirement for predicting the relative affinities of target duplex domains towards third strand hybridization is direct thermodynamic data on the temperature- and sequence-dependent stabilities of DNA triplex structures as a function of solution conditions. Although a few such studies for an RNA triplex exists (40-42), the relevant thermodynamic data for DNA triplexes are nonexistant. To determine the thermodynamics of a DNA triplex, three oligodeoxyribonucleotides were synthesized, a pyrimidine 15mer sequence designated y15, and two complementary 21mer sequences, u21 and y21, where the "y" prefix indicates a pyrimidine-rich strand and the "u" prefix designates a purine-rich strand (Figure 6.3B). We selected these sequences, in part, because the strand orientation within the y15·u21·y21 triplex is well characterized (12,13). From mixing curves we demonstrate that, under the appropriate conditions, the three strands combine to form a triplex containing one strand of each oligodeoxyribonucleotide. Temperature-dependent uv absorbance spectroscopy and differential scanning calorimetry (DSC) were used to determine the stability and the melting behavior of the oligomeric DNA triplex as a function of solution

conditions, thereby providing a complete thermodynamic characterization of a DNA triple-helical structure.

Materials and Methods

Oligodeoxyribonucleotide Synthesis and Purification. Oligodeoxyribonucleotides were synthesized by the DNA Synthesis Network Lab (Dept. Molecular Genetics & Microbiology, UMDNJ, Piscataway, NJ) on an Applied Biosystem 380B synthesizer (Applied Biosystem, Foster City, CA) using standard solid-phase cyanoethylphosphoamidite methods (43). Purification of all the unmodified sequences was accomplished by HPLC using established protocols (44), while purification of the one modified sequence was accomplished using both HPLC and gel electrophoresis. HPLC analysis of the enzyme degradation products of the oligomers revealed that each sequence exhibited the expected ratios of nucleosides.

Extinction coffecients (ε) for each oligomer were determined directly by phosphate analysis (45) yielding the following values at 260 nm and 25°C: 1.166×10^5 $M^{-1}cm^{-1}$ for y15; 1.846×10^5 $M^{-1}cm^{-1}$ for u21; and 1.965×10^5 $M^{-1}cm^{-1}$ for y21. All concentrations, unless otherwise noted, are designated on a per strand basis and were determined spectrophotometrically using the experimentally measured ε values.

Buffers. Unless otherwise indicated, all measurements were conducted in a 10 mM phosphate buffer containing 0.1 mM EDTA, 200 mM NaCl, at pH 6.5. The final pH of the buffer was adjusted to the desired value using 0.1 M HCl or 0.1 M NaOH (volumetric standards, Aldrich Chemical Co., WI).

UV Mixing Curves. Stock solutions of the u21 and the y21 single strands were prepared at equal concentrations (2.5×10^{-6} M). To construct a u21·y21 du-

plex mixing curve, 100 µl aliquots of either the y21 or u21 solution were added to a 1 cm path-length cuvette containing 600 µl of the stock solution of the complementary 21mer strand. After each addition, the cuvette was inverted repeatedly to insure complete mixing followed by equilibration at 15°C for 20 min. After equilibration, the absorbance at 260 nm was measured. Construction of a triplex mixing curve was accomplished similarly. In this case, the two solutions used were the previously constituted u21·y21 duplex (2.5×10^{-6} M *duplex*) and a solution of the y15 single strand (2.5×10^{-6} M *strand*).

UV Spectroscopy. Absorbance versus temperature profiles were obtained at 260 nm using a computer-interfaced, Perkin-Elmer 575 spectrophotometer equipped with a thermoelectrically controlled cell holder. Melting profiles were obtained by increasing the temperature from 5° to 100°C at a rate of 0.5°C/min and recording the absorbance and the temperature every 30 seconds. Melting temperatures, T_m, for each transition were obtained from the optical melting curves using previously described protocols (7,8,46,47). Optical cells with path-lengths of 0.035, 0.1, 0.5 and 1.0 cm were used to make measurements over a range of oligomer concentrations (1.6×10^{-6} to 5.0×10^{-5} M). The oligomers were dissolved in 10 mM phosphate buffers containing 0.1 mM EDTA and various sodium ion concentrations ranging from 0.02 to 1.0 M. Melting curves also were measured over a range of pH between 6 and 7.

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded on a computer-interfaced AVIV model 60DS spectropolarimeter (AVIV Associates, Lakewood, NJ) using the same samples employed in the uv melting studies. The temperature of the cell holder was thermoelectrically controlled. Spectra were obtained by scanning from 320 to 220 nm at 0.5 nm intervals. Each spectrum reported is an average of at least three scans.

Differential Scanning Calorimetry. Apparent excess heat capacity (ΔC_p) versus temperature (T) profiles for the order-disorder transitions of the duplex and the triplex were obtained using a Microcal MC-2 differential scanning calorimeter (Microcal Inc., Amherst, MA). In a typical experiment, the excess heat capacity was recorded continuously while the temperature was increased from 7° to 100°C at a rate of 60°C/hr. The triplex concentration in the DSC experiments was 5.8×10^{-5} M. The transition enthalpy ($\Delta H°$) was calculated from the area under the calorimetric ΔC_p versus T curve ($\Delta H° = \int \Delta C_p \, dT$). The transition entropy ($\Delta S°$) was calculated by integrating the area under the corresponding $\Delta C_p/T$ versus T curve ($\Delta S° = \int (\Delta C_p/T) \, dT$)).

Results and Discussion

Stoichiometry of the Complexes. The stoichiometry of each complex was established by the method of continuous fractions (48,49). Figure 6.4A shows the mixing curve for the two 21mer sequences, u21 and y21. An inflection point is observed at a mole fraction of 0.5, thereby establishing the formation of the expected 1:1 complex for the duplex. An extinction coefficient of 3.15×10^5 $M^{-1}cm^{-1}$ can be estimated for the duplex from the absorbance at the inflection point.

The stoichiometry for the complex formed by addition of the 15mer sequence, y15, to the duplex formed by the two 21mer sequences, u21 and y21, can be obtained in a similar fashion. As shown in Figure 6.4B, an inflection point at a mole fraction of 0.5 is observed. Therefore, the complex of the two 21mer strands and the 15mer strand results in the formation of a triplex. From the absorbance at the inflection point, the extinction coefficient for the triplex can be estimated to be 4.14×10^5 $M^{-1}cm^{-1}$.

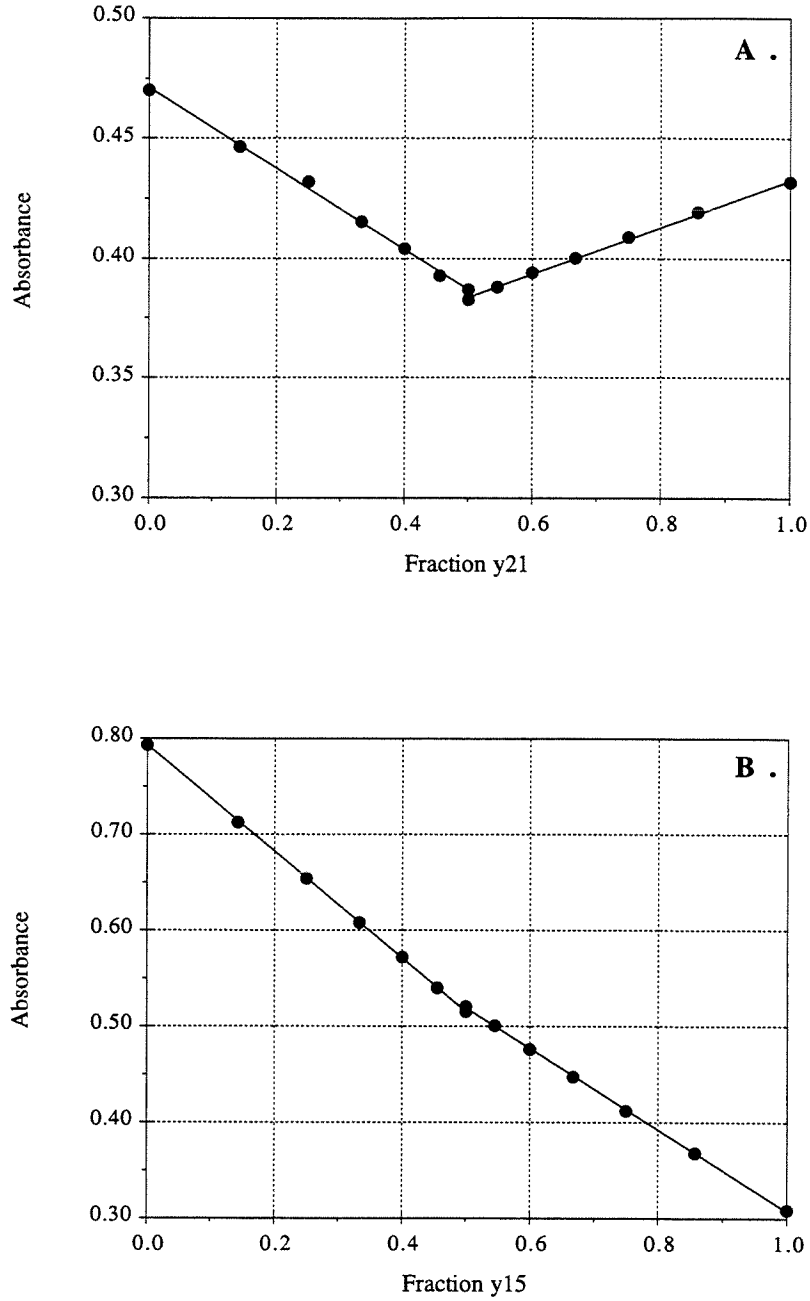


Figure 6.4. UV mixing curves for formation of complexes. (A) Mixing of the u21 single strand with the y21 single strand to form the u21·y21 duplex. Values on the abscissa are mole fraction y21. (B) Mixing of the u21·y21 duplex with the y15 single strand to form the y15·u21·y21 triplex. Values on the abscissa are mole fraction y15.

These mixing experiments establish that the three sequences chosen for study (u21, y21, and y15) can be added together stoichiometrically to form either a Watson-Crick duplex (u21·y21) or a Hoogsteen-Watson-Crick triplex (y15·u21·y21). This knowledge provides us with a framework for conducting the physical studies described below.

The Triplex Melts in Two Steps: Loss of the y15 Third Strand Followed by Disruption of the Released 21mer Duplex. We monitored the thermally-induced disruption of the triplex structure by measuring the temperature-dependent absorbance change at 260 nm as shown in Figure 6.5A. Inspection of this curve reveals that the triplex melts in two sequential, well-resolved steps. We assign the high temperature transition to the melting of the 21mer duplex since, as described below, this transition coincides exactly with the transition we measure under comparable conditions for the melting of the free 21mer duplex in the absence of the y15 single strand (see Figure 6.5B). We assign the low temperature transition to the thermally-induced release of the y15 strand from the triplex to form the 21mer duplex and the free y15 single strand. These assignments are supported by the observations described below.

We have measured the uv melting profile for the thermally-induced disruption of the isolated 21mer duplex. This melting curve is shown in Figure 6.5B. A comparison between Figure 6.5A and 6.5B reveals that under comparable conditions the melting transition of the free duplex coincides exactly with the high temperature melting transition of the triplex. This observation strongly supports our conclusion that the second transition in Figure 6.5A corresponds to duplex disruption.

Our studies of the pH dependence of the two melting transitions also assist us in defining the events associated with each transition. Figure 6.6 shows the

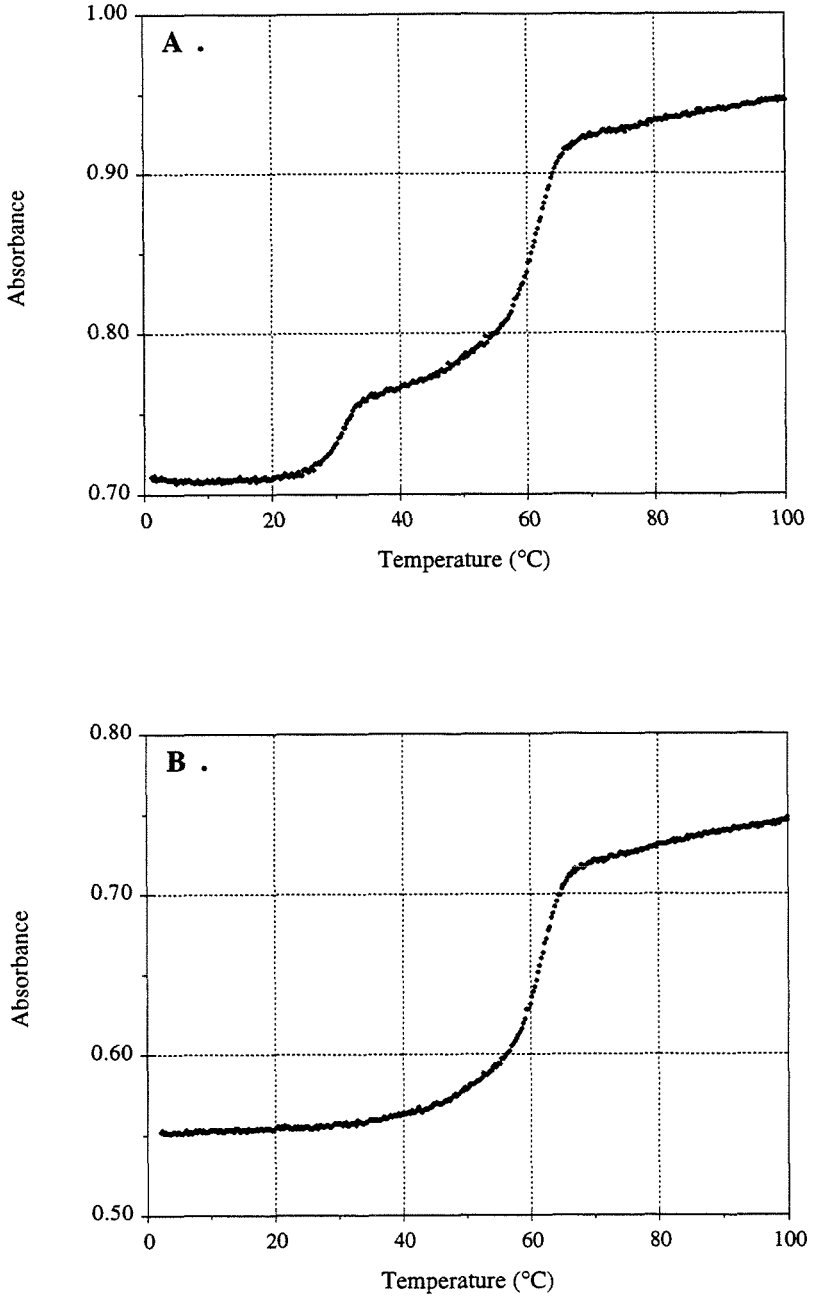


Figure 6.5. UV absorbance versus temperature profiles of complexes: (A) the y15•u21•y21 triplex. (B) the u21•y21 duplex.

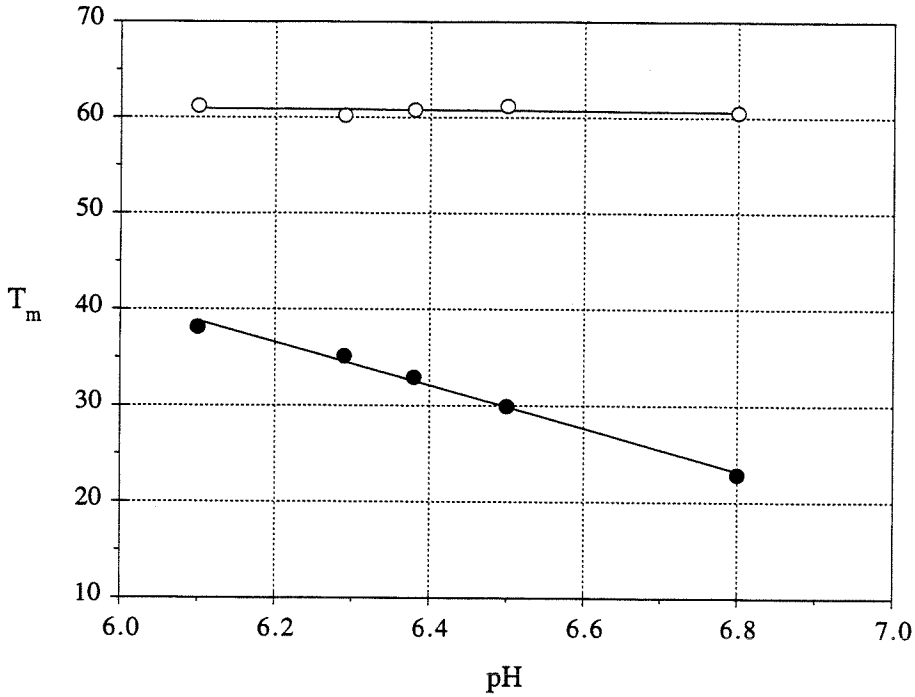


Figure 6.6. Effects of pH on thermal melting temperatures, T_m: (A) triplex disruption (y15·u21·y21 → u21·y21 + y15); (B) duplex disruption (u21·y21 + y15 → u21 + y21 + y15).

pH dependence we measure for the melting temperatures, T_m, of both transitions. Inspection of the upper line in this figure reveals that the high temperature transition is independent of pH between 6 and 7. This pH-independent melting behavior is consistent with the known pKa's of the bases in a Watson-Crick DNA duplex and further supports our assignment of the second transition to the melting of the u21·y21 duplex. By contrast, the lower line in Figure 6.6 reveals that the T_m of the low temperature transition exhibits a clear pH dependence. This observation is consistent with at least partial protonation of the cytosines on the third strand, thereby supporting our assignment of the first transition to triplex expulsion of the y15 single strand. Thus, the pH-dependence of T_m supports our assignments of the molecular events associated with each transition.

Dependence of Triplex Formation and Melting Temperature on Salt Concentration. To define optimal hybridization conditions for triplex formation, it is necessary to evaluate the influence of salt concentration on triplex thermal stability. We find that a Na^+ concentration of 200 mM is sufficient to induce complete y15·u21·y21 triplex formation, even in the absence of Mg^{2+}. Below 200 mM Na^+, triplex formation is incomplete as judged by a reduction in hyperchromicity of the first melting transition. To be specific, at 24 mM Na^+ no triplex forms. At 111 mM Na^+, approximately 2/3 of the triplex forms. At Na^+ concentrations above 200 mM, the apparent hyperchromicity of the first transition is constant, thereby suggesting complete formation of the triplex. By contrast, over this same range of salt concentrations (0.024 to 1 M), the hyperchromicity of the second transition is independent of Na^+ concentration, thereby suggesting that the 21mer duplex is fully formed over this entire range of solution conditions. Thus, for this system, changes in the Na^+ concentration between 24 and 200 mM can be used to modulate the degree of triplex formation without disrupting the target host duplex.

For Na^+ concentrations where the triplex is fully formed (0.2 to 1 M), we find the T_m of the first triplex transition (y15·u21·y21 → u21·y21 + y15) to be linearly dependent on log[Na^+], with $\partial T_m/\partial \log[Na^+] = 12.1$. This derivative provides us with a quantitative measure of how much one can thermally stabilize the triplex by increasing the Na^+ concentration. Higher ionic strengths and lower pH's can be used to enhance triplex thermal stability, but at the risk of diminished sequence specificity.

The Triplex CD Spectrum is Reproduced by the Sum of the Component Duplex and Single Strand Spectra. CD spectra between 220 and 320 nm were measured for the duplex, the triplex, and their component single strands. The spectrum for the 21mer duplex (u21·y21) is composed of two bands of comparable intensity and is consistent with a B-like double helix (46,50). At 20°C, a negative band with maximum intensity at 247 nm and a positive band with maximum intensity at 280 nm are observed. When the temperature is increased to 45°C, the positive band shifts to 276 nm and grows in intensity while the negative band declines slightly in intensity and shifts to about 246 nm. At 80°C, the positive band shifts further to 273 nm and decreases in intensity while the negative band intensity decreases precipitously. This temperature dependent CD behavior is consistent with the expected CD melting properties of a B-form DNA duplex (50).

The circular dichroism spectrum observed for the triplex can be duplicated by the sum of the CD spectrum of the y15 single strand and the CD spectrum of the isolated duplex. At 80°C, the CD spectrum is reproduced by the sum of the individual strand spectra at 80°C. At 20°C, which is below the first transition, and at 45°C, which is between the two transitions, the sum of the duplex CD spectrum and the y15 CD spectrum satisfactorily reproduce the CD spectrum we measure for the triplex. This additivity is both interesting and surprising since

recent NMR studies on oligonucleotides (27,29) suggest that some or all of the strands adopt an A-DNA-like conformation upon triplex formation. Perhaps fortuitous compensation masked CD changes for this system. Alternatively, the final conformation of a triplex may prove to be sequence dependent. Unfortunately, from the available x-ray fiber diffraction and NMR data, it is not yet possible to evaluate the dependence of triplex structure on sequence. However, our measurements demonstrate that triplex formation is not necessarily accompanied by a change in the near-uv circular dichroism relative to its component duplex and single strand.

The Triplex is Thermodynamically Much Less Stable than its Host Duplex. The utility of DNA triplex formation as a method to modulate a range of biological activities will, in large part, depend on the ability to assess if a given oligonucleotide will form selectively a stable triplex structure under specific hybridization conditions. Such an assessment requires thermodynamic data on triplex structures as a function of base sequence, chainlength, salt conditions, pH, and temperature. Differential scanning calorimetry is used to obtain directly some of the requisite thermodynamic data for the DNA triplex studied in this work.

The calorimetric heat capacity versus temperature profiles for the triplex and the isolated duplex which serves as the host for the third strand of the triplex are shown in Figures 6.7A and 6.7B. Inspection of these figures reveals that the triplex melts via two well-resolved transitions (Figure 6.7A), while the isolated duplex melts via a single transition (Figure 6.7B). Both of these melting behaviors are fully consistent with that observed in the optical melting profiles (see Figures 6.5A and 6.5B). The T_m's of each DSC transition agree exactly with those derived from the optical measurements under comparable conditions. Repeated heating and cooling of the DSC samples produced superimposable thermograms,

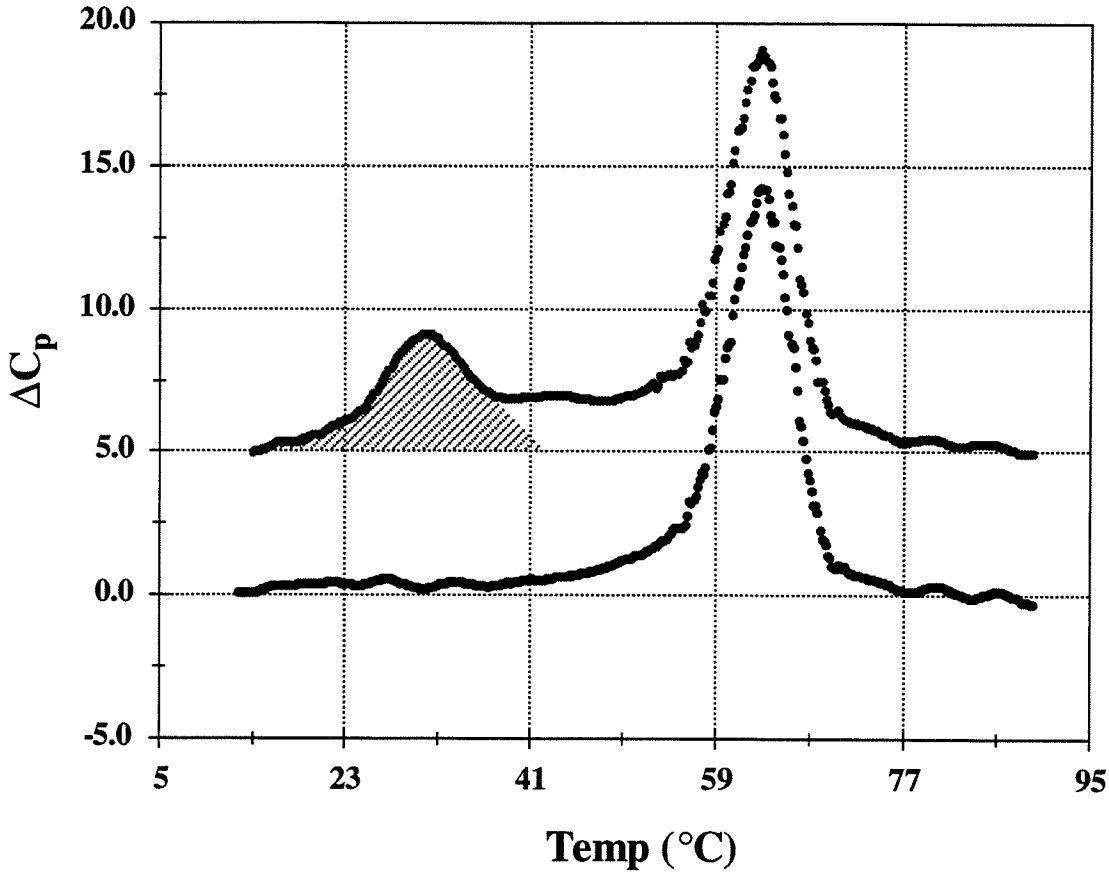


Figure 6.7. Calorimetric apparent excess heat capacity, ΔC_p, versus temperature profiles: (A) the y15·u21·y21 triplex. (B) the isolated u21·y21 duplex. For the sake of clarity, curve A is displaced by 5 kcal/mole·K.

thereby demonstrating the reversibility of the equilibria associated with triplex formation. We also found that changing the heating rate does not alter the thermodynamic parameters, thereby demonstrating that the equilibria under study are not kinetically limited.

Calorimetric transition enthalpies were calculated by integrating the area under each heat capacity vs. temperature curve. These data are summarized in Table 6.1. Inspection of these enthalpy data reveals a number of interesting features. First, the enthalpy change associated with the second transition of the triplex (128 kcal/mol) is in excellent agreement with the enthalpy change for disruption of the isolated duplex (133 kal/mol), which, in turn, is in good agreement with the predicted enthalpy change of 143 kcal/mol calculated from a nearest-neighbor analysis (5). van't Hoff enthalpies derived from optical melting curves and DSC shape analysis (8) are also in good agreement with the calorimetric enthalpy. These agreements in the enthalpy data provide further confirmation of our assignment of the second transition to duplex disruption. Further inspection of the data in Table 6.1 reveals that the transition enthalpy, per base interaction, for triplex dissociation, is much lower than for duplex dissociation. Specifically, at pH 6.5 we measure 2.0 ± 0.1 kcal/(mol of base triplets) for the transition from triplex to duplex plus single strand and 6.3 ± 0.3 kcal/(mol of base pairs) for the transition from duplex to single strand. It is interesting to note that the triplex transition enthalpy data we measure is similar to the value that has been measured for a polymeric RNA triplex (41,42). Specifically, Krakauer and Sturtevant (41) report an enthalpy of about 1.3 kcal/(mol of base triplets) for the dissociation of poly(A)·poly(U)$_2$ to poly(A)·poly(U) + poly(U). Thus, for the polymeric RNA triplex and the oligomeric DNA triplex calorimetrically studied to date, the third strand dissociation enthalpies are small and positive. By con-

trast, van't Hoff enthalpies derived from optical melting curves and DSC shape analysis (8) are about 3 times the calorimetric enthalpy. This disparity suggests that caution should be exercised when attempting to extract thermodynamic data for triplexes from optical studies (51).

Table 6.1. Calorimetrically Determined Thermodynamic Parameters for the Melting Transitions of the y15·u21·y21 Triplex

Transition	T_m °C	$\Delta H°$ kcal/mol	$\Delta S°$ cal/mol·K	$\Delta G°$ kcal/mol
First triplex				
y15·u21·y21 → y15 + u21·y21	30.0	30.4 ± 2	97.6 ± 7	1.3 ± 2
Second triplex				
y15 + u21·y21 → y15 + u21 + y21	65.7	128 ± 8	370 ± 24	17.2 ± 1.6
Duplex				
u21·y21 → u21 + y21	64.7	133 ± 6	388 ± 15	17.2 ± 1.2

Triplex and duplex data represent the averages of four and three measurements, respectively. Errors shown represent ± 1 standard deviation.

The entropic contribution to the triplex and to the duplex transitions may be calculated from the area under the secondary $\Delta C_p/T$ vs. T curves which can be derived from the experimental ΔC_p vs T curves (8). These data also are reported in Table 6.1. Inspection of these data reveals the expected large and positive entropy changes favoring structure disruption. Note, however, that the entropy contributes less to triplex than to duplex disruption, as one might anticipate considering the order that still remains in the products after the initial disruption of the triplex.

Given the enthalpic and entropic contributions we have measured calorimetrically at the melting temperature, and assuming a negligible difference in heat capacity between the initial and final states, the free energy at any reference temperature can be calculated using the standard thermodynamic relationship $\Delta G° = \Delta H° - T\Delta S°$. Table 6.1 includes the free energy data at 25°C calculated in this manner. Inspection of these $\Delta G°$ data reveals that at 25°C and 200 mM Na^+, the triplex structure is only marginally stable ($\Delta G°$ = 1.3 kcal/mol) relative to the significant stability of its host duplex ($\Delta G°$ = 17.2 kcal/mol). Nevertheless, this relatively small $\Delta G°$ of 1.3 kcal/mol corresponds to an association constant, K, of ~9.0, which means that for this system, under the equilibrium conditions employed, the vast majority of third strands are bound to the target duplex sequence. Interestingly, this marginal stability of the fully-bonded triplex may be of practical importance since it will reduce the probability of probe hybridization to secondary duplex sites containing sequences which are similar but not fully "complementary" to a third strand probe (e.g., mismatches). Thus, the application of high ionic strengths and low pH to enhance triplex stability may not always be desirable, since one seeks conditions that optimize, but do not necessarily maximize, third strand binding.

The Triplex Melts as if Monomolecular. To design optimal hybridization conditions for triplex formation, it is necessary to evaluate the influence of strand concentration on triplex thermal stability. To this end, we have measured how the triplex melting temperature, T_m, varies with total strand concentration, C_{tot}.

Triplex formation via the association of a third strand (e.g., y15) with a host duplex (e.g. u21·y21) formally is a bimolecular process. Consequently, the position of this equilibrium and therefore the melting temperature, T_m, should, in principle, depend on the strand concentration, C_{tot}. Traditionally, this depen-

dence is assessed by measuring the slope of a plot of T_m^{-1} versus $\ln C_{tot}$ (35). Such a plot should be linear with a slope that depends on both the enthalpy of complex formation, ΔH°, and the molecularity (8). We have measured the dependence of T_m^{-1} on $\ln C_{tot}$ for the triplex and find the slope to be nearly a factor of 30 less than expected based on the calorimetrically measured triplex transition enthalpy of 30.4 kcal/mol and the formal molecularity of two (8). Thus, we conclude that, for the sequences studied here, triplex formation is approaching the limit of pseudo-first order behavior, as is observed for association/dissociation equilibria in *polymeric* DNA systems. This limiting behavior occurs when the monomolecular triplex elongation ("growth") steps dominate the bimolecular nucleation ("initiation") steps (46). In such cases, even an equilibrium with a formal molecularity greater than unity will functionally behave as if it were monomolecular.

On a practical level, the results described above mean that for the triplex studied here, one cannot significantly increase the thermal stability simply by using higher concentrations of the third strand and that the extraction of triplex thermodynamic parameters from concentration-dependent melting studies will be in error.

Methylation of the Third Strand Cytosines Enhances the Stability of the Triplex. To examine the effects of base modification on triplex stability, we also have studied the analogue of the y15 single strand in which all cytosines are C5-methylated [Me-y15]. This modified single strand exhibits an ε value of 1.069 x 10^5 $M^{-1}cm^{-1}$ at 260 nm and 25°C. We find that triplexes also form readily when the u21·y21 duplex is mixed with the Me-y15 single strand. The mixing curves and the uv monitored thermal melting profiles we measure for the Me-y15 containing triplex are qualitatively similar to those of the triplex produced from the

unmodified third strand, y15. Furthermore, the circular dichroism spectra are similar to those for the unmodified triplex, thereby suggesting that the modificaton does not alter the global triplex structure. As with the unmodified system, we find the transition from triplex to duplex plus single strand (Me-y15) to be nearly independent of oligomer concentration.

The pH dependence of T_m for the transition from triplex to duplex plus Me-y15 single strand is the same as that which we measured for the triplex with the unmodified single strand (y15). In other words, the slopes of the T_m vs. pH curves are the same for the modified and unmodified triplex systems. Significantly, however, we find the Me-y15 containing triplex to be thermally more stable by 10°C at any given pH. For example, at pH 6.5 the y15 triplex melts at 29.9°C whereas the Me-y15 triplex melts at 40.0°C. Such enhancement of triplex stability by cytosine methylation has been observed in polymeric systems (52) and in oligonucleotide-directed site-specific binding of large duplex DNA (13).

It is of interest to compare the difference in the known pK_a's of the unmodified and modified cytosine monomers with the apparent pK_a difference we deduce from our melting measurements. The pK_a of 2'-deoxyribocytidine 5'-monophosphate (dCMP) is 4.6 while the pK_a of 5-methyl-2'-deoxycytidine 5'-monophosphate (5-Me-dCMP) is 4.4, thereby reflecting a methylation-induced ΔpK_a of -0.2 units (53). By contrast, based on the observed triplex melting behavior, we calculate an apparent methylation-induced ΔpK_a of +0.5 units. This result is consistent with that obtained on essentially the same sequences using the affinity cleaving method (13). Thus, it is possible that increased hydrophobicity imparted to the third strand by methylation, rather than just a change in the cytosine pK, is responsible for part, if not all, of the enhancement of triplex stability that we observe with the modified third strand (13).

The practical consequence of this study on the modified triplex is that cytosine methylation at C5 provides one approach for altering a third strand oligonucleotide in a manner that increases the thermal stability of the resultant triplex (13). This approach may prove important for modulating the biological activities controlled by triplex formation.

Concluding Remarks

The development and expanded use of oligonucleotides (or their analogs) for modulating biochemical activities through site-specific triplex formation demands a complete understanding of how triplex stability depends on chain-length, pH, salt, base sequence, and base modifications. Compiling the requisite body of thermodynamic data is a minimum first step required to make these assessments.

Addendum to Chapter Six

In addition to reporting the first direct calorimetric measurement of the enthalpy of DNA triple helix formation, the significance of this research is two-fold. First, we observe that the triplex melting behavior approaches the pseudo monomolecular limit. Second, we observe a disparity between the enthalpy value determined by direct integration of the calorimetric excess heat capacity profile (2.0 kcal per mole of base triplets; see Table 6.2) and the values extracted from van't Hoff analyses of optical or DSC melting curve shapes (6.2 kcal per mole of base triplets). Measurements of the enthalpy of triplex formation for a number of different sequences under a wide range of conditions using the concentration-dependence of the T_m or van't Hoff analysis of melting curve shapes have yielded a broad range of values greater than 3 kcal per mole of base triplets (see Chapter Seven). Importantly, if the observed calorimetric enthalpy represents the "true" enthalpy for the y15·u21·y21 triplex, our data suggest the placement of limitations on the interpretation of these thermodynamic parameters extracted from optical studies. Because of the potential impact of these data on the field, a critical discussion of them is presented below.

Monomolecular Triplex Melting Behavior

A common method for the determination of thermodynamic parameters for a nucleic acid order-disorder transition involves measuring the dependence of the transition T_m on the nucleic acid concentration. A relationship between the T_m and $\Delta H°$, $\Delta S°$, the concentration of nucleic acids (C_{tot}), and the molecularity of the transition (n) has been derived (7,8,47):

$$\frac{1}{T_m} = \frac{(n-1)R}{\Delta H^\circ} \cdot \ln C_{tot} + \frac{\Delta S^\circ - (n-1)R\ln(2n)}{\Delta H^\circ} \quad (9)$$

Using eq 9 and an assumption regarding the molecularity of the reaction, ΔH° can be extracted from the slope of a plot of T_m^{-1} versus $\ln C_T$. Although the formal molecularity of the transition can normally be used (*e.g.*, $n = 2$ for the triplex → duplex + single strand reaction), equilibria involving polymeric nucleic acids often approach the monomolecular limit as the pseudo-intramolecular helix elongation ("growth") steps dominate the intermolecular helix nucleation ("initiation") step. These systems behave as if $n \approx 1$ and the slopes of the corresponding T_m^{-1} versus $\ln C_T$ plots approach zero.

Table 6.2. Calorimetric and van't Hoff Thermodynamic Parameters for Order-Disorder Transitions of the y15·u21·y21 Triplex.[a]

Transition	ΔH_{cal} (kcal/mol)	ΔS_{cal} (cal/mol·K)	$\Delta H_{cal}/T_m$ (cal/mol·K)	$\Delta H_{vH,dsc}$[b] (kcal/mol)	$\Delta H_{vH,uv}$[c] (kcal/mol)
First triplex y15·u21·y21 → y15 + u21·y21	30.4	97.6	100	88	97
Second triplex y15 + u21·y21 → y15 + u21 + y21	128	370	377	112	119
Duplex u21·y21 → u21 + y21	133	388	393	113	116

[a] ΔH_{cal} and ΔS_{cal} are repeated from Table 6.1. A "mol" indicates one mole of the corresponding structure, either triplex or duplex. [b] This parameter extracted from shape analysis of the differential scanning calorimetric denaturation profile. [c] This parameter was extracted from shape analysis of the UV denaturation profile.

For the 15mer investigated at pH 6.5 in 200 mM Na^+, we expected the slope of the T_m^{-1} versus $\ln C_T$ plot to be about -6.5×10^{-5} K^{-1} or -2.2×10^{-5} K^{-1}, depending on whether the calorimetric or van't Hoff enthalpy was correct (*vide infra*). The actual slope of the plot (Figure 6.8) was measured to be much smaller (-2.4×10^{-6} K^{-1}). Hence, we concluded that this triplex approaches monomolecular melting behavior. This conclusion is supported by the observation that the $\Delta S°$ value estimated from $\Delta H_{cal}/T_m$ is in excellent agreement with that determined calorimetrically from $\Delta S° = \int(C_p/T)dT$ (Table 6.2). The identity $\Delta S° = \Delta H_{cal}/T_m$ holds only for monomolecular processes since $\Delta G° = 0 = \Delta H° - T_m\Delta S°$ at the T_m. If this behavior is a general characteristic of triplexes of varying sequence in solutions of varying salt and pH conditions, then enthalpies determined using eq 9 with $n = 2$ will be larger than the true enthalpies. To date, this behavior has not been systematically investigated with different triplexes or under conditions different from those reported above.

Disparity between van't Hoff and Calorimetric Enthalpies

The analysis of the shapes of thermal denaturation profiles using the van't Hoff equations allows the determination of thermodynamic parameters using a UV-Vis spectrophotometer equipped with a temperature-control device. In contrast, microcalorimetric measurements require equipment that is more sophisticated, less readily available, and more expensive. Moreover, calorimetric techniques generally require much larger amounts of the nucleic acid molecules. For these reasons, optical measurements have often been employed to measure the enthalpies of nucleic acid transitions. Carefully conducted calorimetric and van't Hoff measurements can usually be expected to yield parameters that agree

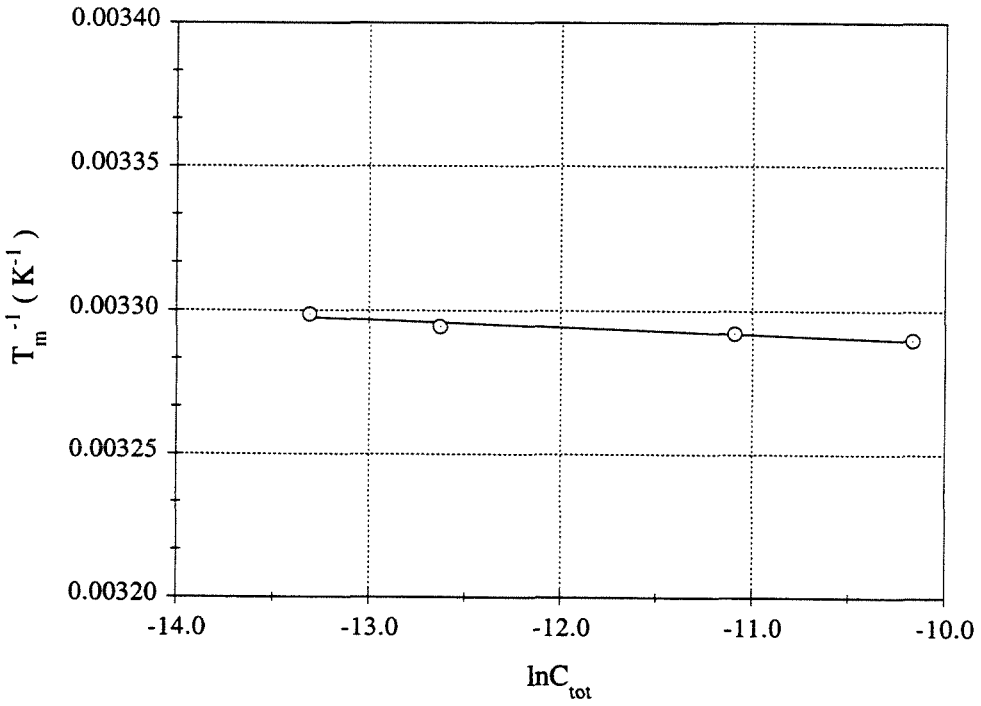


Figure 6.8. A plot of the reciprocal of the melting temperature (T_m) of the y15·u21·y21 triplex as a function of the DNA concentration. The concentrations spanned the range from 1.7 μM to 38 μM, and the T_m increased from 30.0°C to 30.8°C. The line shows the best linear fit to the data points and has a slope of -2.4×10^{-6} K^{-1}.

within experimental uncertainties. However, in the case of the triple helix system described above, the enthalpies derived from van't Hoff and calorimetric measurements (ΔH_{vH} and ΔH_{cal}, respectively) differ by a factor of three (Table 6.2). Among the phenomena that can account for a small calorimetric enthalpy are the presence of competitive equilibria involving the formation of aggregates or alternative complexes by triplex components, and the fractional formation of triplex at the start of the experiment as a result of kinetic limitations. Although these effects could, in principle, affect the calorimetric data, two recent investigations have revealed significant differences between ΔH_{vH} and ΔH_{cal}. Moreover, as discussed below, hysteresis effects can affect the observed ΔH_{vH}.

Aggregation Phenomena. The ratio $\Delta H_{vH}/\Delta H_{cal}$ indicates the size of the "cooperative unit" in nucleic acid order-disorder transitions, and has generally been used in the interpretation of differences between ΔH_{vH} and ΔH_{cal} (7,8). In the case of denaturation of a polymeric nucleic acid structure, ΔH_{vH} is often less than ΔH_{cal} (an enthalpy ratio less than unity). This occurs when one optically observes the melting of structural units which are smaller than the full size of the polymeric structure, and the size of the cooperative unit — the segment that melts cooperatively during the optical experiment — is obtained by multiplying the enthalpy ratio by the full length of the structure in monomer units (*e.g.*, base pairs). On the other hand, an enthalpy ratio greater than unity has traditionally been interpreted as indicating the presence of aggregates, whereby the cooperatively melting unit is larger than a single structural unit. This effect can be magnified in a differential scanning calorimetry sample for which the concentrations of nucleic acid species must be higher than for optical measurements.

Several features of the data presented in Table 6.2 that pertain to the question of aggregation in the y15·u21·y21 triplex should be noted. For the triplex to

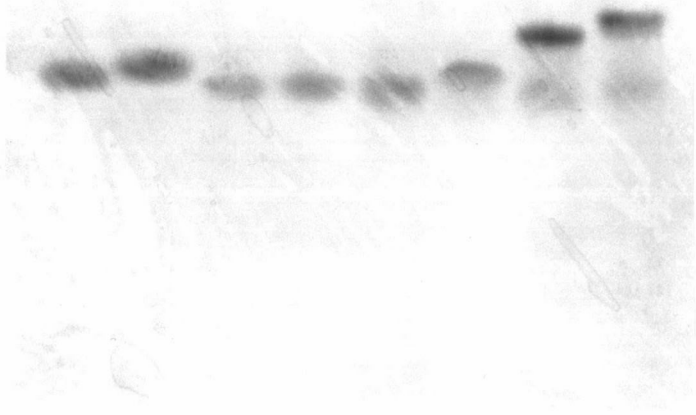
duplex plus single strand transition, we found the value for $\Delta H_{vH,uv}$ to be in good agreement with the van't Hoff enthalpy determined from the shape of the DSC melting transition, $\Delta H_{vH,dsc}$. These two values which we measured at 1.6 μM and 58 μM DNA concentrations, respectively, are three-fold higher than that measured directly by integration of the excess heat capacity profiles. In contrast, the model-dependent van't Hoff and model-independent calorimetric enthalpies for the duplex to single strands transition are in excellent agreement, regardless of the presence of the y15 single strand. Therefore, none of these data provide independent evidence for aggregation.

The most compelling evidence against aggregation was obtained by examining the electrophoretic behaviors of the y15·u21·y21 triplex, the u21·y21 duplex, and their constituent single-stranded oligonucleotides (Figure 6.9). The experiments involved incubation of solutions of the oligonucleotides under conditions which approximated those of the prior calorimetric measurements — 60 μM in each oligo, 350 mM Na^+, pH 6.5 — at 4°C, followed by electrophoretic separation at 4°C in buffer to maintain the appropriate salt concentrations and pH throughout the course of the experiment. The bands visualized using Stain-All™ dye provide no indication of aggregation under these solution conditions.

Other Competitive Equilibria. The formation of an alternative structure by any of the components of the y15·u21·y21 triple helix, and the concomitant reduction in the concentration of triplex present, could reduce the observed transition enthalpy. One alternative structure is an "acid helix" which is stabilized by the formation of C^+·C self-pairs (Figure 6.10) in a parallel-stranded double-helical complex (54-56). Such a competitive equilibrium involving the y15 strand is suggested by the observation of self-complexes of an 11mer containing two pairs of adjacent cytosines during the characterization of a triplex at pH 5.0 (51). It is

Figure 6.9. Gel migration pattern for components of the y15·u21·y21 triplex. Each sample was prepared by dissolving the oligonucleotide(s) in 340 mM NaCl buffered at pH 6.5 by 10 mM NaH_2PO_4 (350 mM total Na^+ concentration) to a final concentration of 60 μM (each). These solutions were allowed to stand at 4°C for 30 min unless otherwise indicated. The samples were fractionated electrophoretically on a 17.5% polyacrylamide gel (29:1 acrylamide:bis-acrylamide) prepared with 89 mM Tris-acetate (pH 6.5), 1 mM EDTA, and 350 mM NaCl (TEAS buffer). The electrophoresis was performed at 4 V/cm (100 mA) at 4°C in TEAS buffer. The buffer was recirculated between the top and bottom reservoirs, maintaining a constant pH 6.5 throughout the electrophoretic run. DNA bands were visualized using Stains-All™ dye and photographed. Lane 1: $d(T)_{14}$ obtained from Pharmacia LKB and heat-denatured at 60°C for 30 min prior to electrophoresis; lane 2: $d(T)_{16}$ obtained from Pharmacia LKB and heat-denatured at 60°C for 30 min prior to electrophoresis; lane 3: y15 heat-denatured at 60°C for 30 min prior to electrophoresis; lane 4: y15; lane 5: y21; lane 6: u21; lane 7: y21 and u 21; and lane 8: y15, u21, and y21.

	1	2	3	4	5	6	7	8
T14:	+	–	–	–	–	–	–	–
T16:	–	+	–	–	–	–	–	–
CT15:	–	–	+	+	–	–	–	+
YT21:	–	–	–	–	+	–	+	+
RA21:	–	–	–	–	–	+	+	+
60°C:	+	+	+	–	–	–	–	–



important to note that the y15 sequence contains no adjacent cytosine residues and that our study was conducted at pH 6.5, two factors that would work together to reduce the stability of $C^+ \cdot C$ base pairs (57). Indeed, we observe no evidence for self-association of the y15 oligonucleotide. In particular, there is no difference in the gel migration patterns of heated and non-heated samples of y15, both of which migrate similarly to dT_{14} and dT_{16} oligomers (Figure 6.9), no transition is observed in the y15 melting curve (Figure 6.11A), and the y15 CD spectrum (Figure 6.11B) is similar to that of poly[d(CT)] at high pH rather than the low-pH acid helix spectrum (57).

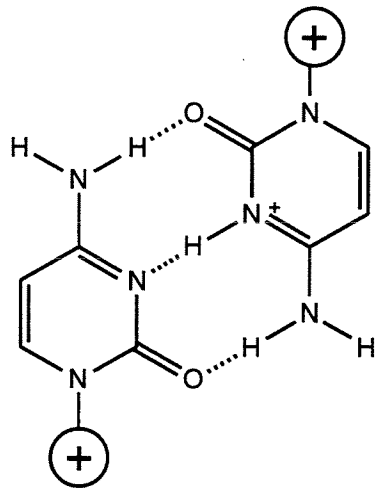


Figure 6.10. Two-dimensional schematic of the hydrogen-bonding in a $C^+ \cdot C$ homo base pair. Such homo pairing is proposed to stabilize a parallel double-helical complex at low pH.

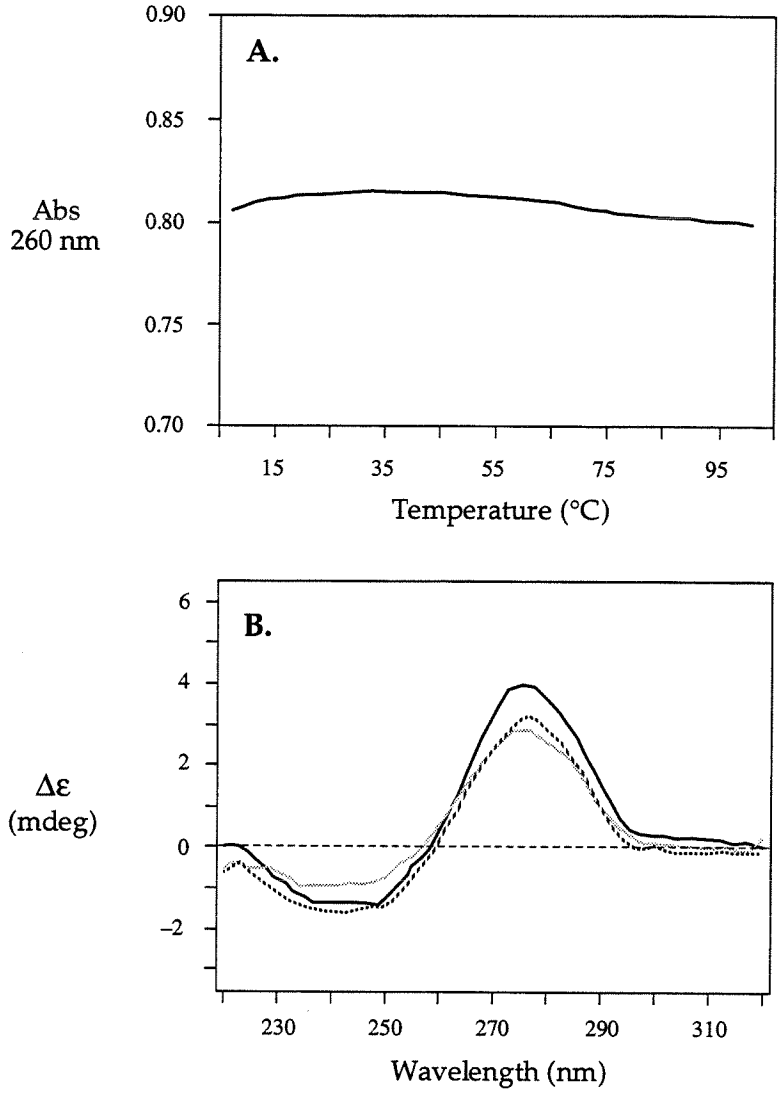


Figure 6.11. (A) The absorbance versus temperature melting profile of y15 alone obtained by observation at 260 nm using a scan-rate of 0.5°C/min. y15 is present at a concentration of 1.65 μM in 200 mM Na^+, 1 mM EDTA, 10 mM phosphate buffer at pH 6.5. (B) Circular dichroism spectra for y15 alone under the same conditions as above. CD spectra were recorded at 20°C (black curve), 45°C (dotted curve), and 80°C (gray curve).

Kinetic Limitations on Triplex Formation. As with the competing equilibria discussed above, a kinetic limit can prevent full formation of the triple helix, resulting in a smaller observed transition enthalpy. Reports from the laboratories of both Dervan (58) and Hélène (59) have demonstrated that the bimolecular rate of triple helix formation is slow relative to that of duplex formation. In fact, the kinetic association rate constant (k_{on}; see eq 10) has been measured to be 300 $M^{-1}s^{-1}$ at 15°C in 200 mM NaCl buffered by 10 mM cacodylate at pH 6.8.

$$\mathbf{O} + \mathbf{D} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathbf{T} \tag{10}$$

For the equilibrium described by eq 10, the integrated rate equation can be written (see eqs 6 and 7 in Chapter One) as

$$\alpha_t / \alpha_{eq} = 1 - \exp\left\{-\left[k_{off} + k_{on} O_0 (1 - \alpha_t)\right] t\right\} \tag{11}$$

where the meanings of the parameters and subscripts are the same as defined above (eqs 3–5). By extrapolation of k_{on} to 4°C using the measured activation energy and the Arrhenius equation as described (59), the time required for equilibration of a triplex can be calculated using eq 11. In this manner, we estimate that ~ 10 min are sufficient for equilibration of the y15·u21·y21 triplex to within 5% of an equilibrium position where 99% of the strands are bound in the triple helix (*i.e.*, $\alpha_{eq} = 0.99$ and $\alpha_t/\alpha_{eq} = 0.95$) under the conditions used at the start of a DSC experiment ($O_0 = D_0 = 58$ μM in 200 mM Na^+ phosphate-buffered at pH 6.5 at a temperature of 4°C). For the experiments described here, equilibration times were always much longer than 10 min. The expectation of fully formed

y15·u21·y21 complex is consistent with the absence of detectable endothermicity or hypochromicity that would accompany triple helix formation at the initiation of a calorimetric or spectrophotometric experiment, respectively.

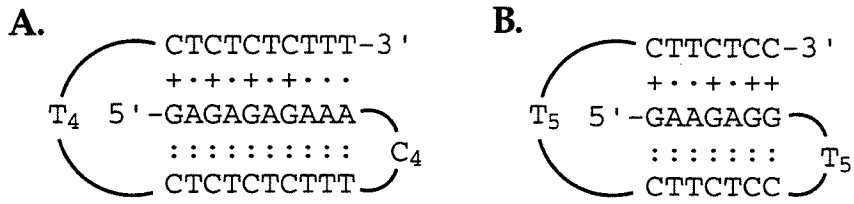


Figure 6.12. Secondary structural models of intramolecular triplexes studied by (A) Völker et al. (60) and (B) Breslauer and Plum (61). Watson-Crick hydrogen-bonding between vertical pairs of bases is indicated by two dots (:), while Hoogsteen-type interactions involving thymidine and cytosine are indicated by single dots (·) and plusses (+), respectively.

Other Relevant Data. For over two years, the report described herein was the only published systematic comparison between ΔH_{vH} and ΔH_{cal} using identical complexes and solution conditions. However, Klump and co-workers recently reported $\Delta H_{vH} \gg \Delta H_{cal}$ for a 10mer intramolecular triplex formed by folding of a 38mer oligonucleotide (Figure 6.12A) (60). The use of an intramolecular triplex eliminates the requirement of a bimolecular association step. Presumably, this change minimizes any kinetic limit on triplex formation and reduces the potential effects of competitive equilibria by enhancing the triplex stability relative to

intermolecular aggregates. In 100 mM NaCl buffered at pH 6.7 by 20 mM cacodylate, ΔH_{vH} was determined to be about 6 kcal per mole of base triplets, while ΔH_{cal} was measured to be 4 kcal per mole of base triplets by DSC. As was the case for the y15·u21·y21 intermolecular system, there is no evidence for aggregation and the duplex-to-coil melting transition displays no enthalpy disparity. Importantly, the difference between the transition enthalpies measured using two different methods is not observed as the stability of the triplex is increased by lowering the pH to 4.5. Xodo et al. have also obtained good agreement between enthalpies obtained by DSC and van't Hoff analysis at pH 5.0 (62).

Breslauer and co-workers have also investigated an intramolecular triplex by DSC (61). In this system (Figure 6.12B), the triplex is composed on seven base triplets formed by folding of a 31mer, and was characterized in 1 M Na^+ buffered at pH 6.3 with 10 mM phosphate. The preliminary data suggests ΔH_{vH} to be two-fold larger than ΔH_{cal}, whether the van't Hoff shape analysis is conducted on spectroscopic or calorimetric melting profiles. Moreover, the calorimetric value has proven useful for predicting the triplex stability under a range of conditions, while the van't Hoff value is required to reproduce the shapes of the transitions.

Potential Problems with Optical Techniques. There are at least two potential difficulties in using optical techniques that may provide alternative explanations for the disparity between the van't Hoff and calorimetric enthalpies. One alternative is that the triplex melting transition is not a true two-state process (60). In order to analyze thermal denaturation profiles, one assumes that the transition is a true two-state process, *i.e.*, that the observed transition involves only a single initial state and a single final state, with intermediate states being only transiently populated at low levels of occupancy. For the melting of double-helical

nucleic acids, limitations in this approximation can lead to relative errors of up to 40% in the van't Hoff enthalpy (63-65). To date, there have been no reported attempts to characterize intermediates in the triplex melting process by either experimental or theoretical means. It is worthy of note that molten globule intermediates in the folding pathways of proteins are characterized by $\Delta H_{vH} > \Delta H_{cal}$ without aggregation (66-70).

An experimental difficulty in extracting thermodynamic data from optical studies has recently been reported by Rougee et al., who found marked hysteresis when comparing heating and cooling profiles for a 22mer triplex in 100 mM NaCl and 10 mM cacodylate (pH 6.8) (59). Remarkably, their data suggest T_m shifts of several degrees from the equilibrium T_m even at temperature scanning rates as low as 0.002°C/s. Typical scan-rates are five to ten-fold higher. This effect is proposed to result from the slow kinetics of triple helix formation and denaturation. Heating profiles were observed to yield T_m values higher than the equilibrium melting temperature, and the transitions were markedly steeper than the equilibrium transition. For a typical analysis, these two effects would combine to result in an apparent enthalpy significantly larger than the true transition enthalpy (see eq 8).

Conclusions

Based on the four independent investigations discussed above (4,6,60,61), it appears likely that the calorimetric measurements have not been grossly influenced by competitive equilibria or kinetic limitations. Recall also that the temperature-dependence of the equilibrium constant for triple helix formation at pH 7.0 by the y15 oligonucleotide has provided an independent corroboration of the transition enthalpy measured by DSC (see Chapter Four) (4).

Thus, values of ΔH_{cal} determined by DSC near neutral pH appear to be the best estimates of "true" transition enthalpies because they provide predictive model-independent determinants of triplex stability. Van't Hoff estimates of transition enthalpies under these conditions may be affected by problems with the two-state model as well as kinetic factors. It seems equally clear, however, that some characteristics of the triplex melting transition can only be adequately described by ΔH_{vH}. Perhaps this is related to the problem of fast temperature scanning relative to the kinetics of triplex dissociation (59).

At low pH, the association rates are higher (58,71) and the kinetic limitation on optical experiments is removed. Indeed, the results of Völker et al. at pH 4.5 and those of Xodo et al. at pH 5.0 suggest that van't Hoff enthalpies are equally reliable estimates of the true enthalpy at low pH. A full understanding of the origins of the observed enthalpy disparity will require carefully designed and executed investigations.

References

(1) Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114**, 6957-65.

(2) Singleton, S. F. & Dervan, P. B. (1992) *Biochem.* **31**, 10995-1003.

(3) Singleton, S. F. & Dervan, P. B. (1993) *Biochem.* **32**, 13171-9.

(4) Singleton, S. F. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* , submitted.

(5) Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3746-3750.

(6) Plum, G. E., Park, Y. W., Singleton, S. F., Dervan, P. B. & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 9436-40.

(7) Breslauer, K. J. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology*, ed. Hinz, H.-J. (Springer-Verlag, Berlin), pp 402-427.

(8) Marky, L. A. & Breslauer, K. J. (1987) *Biopolymers* **26**, 1601-1620.

(9) Sturtevant, J. M. (1987) *Annu. Rev.* **38**, 463.

(10) Privalov, P. L. & Potekhin, S. A. (19XX) *Meth. Enzymol.* **131**, 4.

(11) Breslauer, K. J., Freire, E. & Straume, M. (1992) *Meth. Enzymol.* **211**, 533-567.

(12) Moser, H. E. & Dervan, P. B. (1987) *Science* **238**, 645-50.

(13) Povsic, T. J. & Dervan, P. B. (1989) *J. Am. Chem. Soc.* **111**, 3059-61.

(14) Wilson, W. D., Hopkins, H. P., Mizan, S., Hamilton, D. D. & Zon, G. (1994) *J. Am. Chem. Soc.* **116**, 3607-3608.

(15) Felsenfeld, G., Davies, D. R. & Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023-2024.

(16) Arnott, S., Bond, P. J., Selsing, E. & Smith, P. J. C. (1976) *Nucleic Acids Res.* **3**, 2459-2470.

(17) Stevens, C. L. & Felsenfeld, G. (1964) *Biopolymers* **2**, 293-314.

(18) Riley, M., Maling, B. & Chamberlin, M. J. (1966) *J. Mol. Biol.* **20**, 359-389.

(19) Blake, R. D., Massoulie, J. & Fresco, J. R. (1967) *J. Mol. Biol.* **30**, 291-308.

(20) Lipsett, M. N. (1964) *J. Biol. Chem.* **239**, 1256-1260.

(21) Inman, R. B. (1964) *J. Mol. Biol.* **10**, 137-146.

(22) Broitman, S. L., Im, D. D. & Fresco, J. R. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5120-5124.

(23) Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V. & Fresco, J. R. (1988) *Biochem.* **27**, 9108-9112.

(24) Arnott, S. & Bond, P. J. (1973) *Nature New Biol.* **244**, 99-101.

(25) Arnott, S. & Selsing, E. (1974) *J. Mol. Biol.* **88**, 509-521.

(26) Lee, J. S., Johnson, D. A. & Morgan, A. R. (1979) *Nucleic Acids Res.* **6**.

(27) Rajagopal, P. & Feigon, J. (1989) *Biochemsitry* **28**, 7859-7870.

(28) Rajagopal, P. & Feigon, J. (1989) *Nature* **339**, 637-40.

(29) De los Santos, C., Rosen, M. & Patel, D. (1989) *Biochem.* **28**, 7282-9.

(30) Pilch, D. S., Levenson, C. & Shafer, R. H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1942-1946.

(31) Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987) *Nature* **330**, 495-497.

(32) Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M. & Wohlrab, F. (1988) *FASEB J* **2**, 2939-2949.

(33) Htun, H. & Dahlberg, J. E. (1989) *Science* **243**, 1571-1576.

(34) Strobel, S. A., Moser, H. E. & Dervan, P. B. (1988) *J. Am. Chem. Soc.* **110**, 7927-7929.

(35) François, J.-C., Saison-Behmoaras, T., Barbier, C., Chassignol, M., Thuong, N. T. & Helene, C. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9702-9706.

(36) Praseuth, D., Perrouault, L., LeDoan, T., Chassignol, M., Thuong, N. & Helene, C. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1349-1353.

(37) Sun, J. S., Francois, J.-C., Montenay-Garcotier, T., Saison-Behmoras, T., Roig, V., Thuong, N. T. & Helene, C. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9198-9202.

(38) Maher, L. J., III, Wold, B. & Dervan, P. B. (1989) *Science* **245**, 725-30.

(39) Hanvey, J. C., Shimizu, M. & Wells, R. D. (1990) *Nucleic Acids Res.* **18**, 157-161.

(40) Ross, P. D. & Scruggs, R. L. (1965) *Biopolymers* **3**, 491-496.

(41) Krakauer, H. & Sturtevant, J. M. (1968) *Biopolymers* **6**, 491-512.

(42) Neumann, E. & Ackermann, T. (1969) *J. Phys. Chem.* **73**, 2170-2178.

(43) Sinha, N. D., Biernat, J. & Köster, M. (1983) *Tetrahedron Lett.* **24**, 5843-5846.

(44) Gaffney, B. L. & Jones, R. A. (1989) *Biochem.* **28**, 5881-5889.

(45) Snell, F. D. & Snell, C. T. (1949) *Colorimetric Methods of Analysis* (Van Nostrand, New York), 3rd Ed., Vol. 2, pp 671.

(46) Cantor, C. R. & Schimmel, P. R. (1980) *Biophysical Chemistry* (Freeman, San Francisco).

(47) Gralla, J. & Crothers, D. M. (1973) *J. Mol. Biol.* **73**, 301-319.

(48) Job, P. (1928) *Ann. Chim. (Paris)* **9**, 113-134.

(49) Felsenfeld, G. & Rich, A. (1957) *Biochim. Biophys. Acta* **26**, 457-468.

(50) Bush, C. A. (1974) in *Basic Principles in Nucleic Acid Chemistry*, ed. T'so, P. O. P. (Academic Press, New York), Vol. 2, pp 91-169.

(51) Manzini, G., Xodo, L. E., Gasparotto, D., Quadrifoglio, F., Van der Marel, G. A. & Van Boom, J. H. (1990) *J. Mol. Biol.* **213**, 833-43.

(52) Lee, J. S., Woodsworth, M. L., Latimer, L. J. P. & Morgan, A. R. (1984) *Nucleic Acids. Res.* **12**, 6603-6614.

(53) Dunn, D. B. & Hall, R. H. (1975) in *Handbook of Biochemistry and Molecular Biology*, ed. Fasman, G. P. (CRC, Clevelan, OH), 3rd Ed., Vol. 1, pp 198-199.

(54) Pilch, D. S. & Shafer, R. H. (1993) *J. Am. Chem. Soc.* **115**, 2565-2571.

(55) Robinson, H., van der Marel, G. A., van Boom, J. H. & Wang, A. H. J. (1992) *Biochem.* **31**, 10510-10517.

(56) Robinson, H. & Wang, A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5224-5228.

(57) Gray, D. M., Ratliff, R. L., Antao, V. P. & Gray, C. W. (1987) in *Structure and Expression, Vol. 2: DNA and its Drug Complexes*, ed. Sarma, R. H. & Sarma, M. H. (Adenine Press, Inc., Guilderland, NY).

(58) Maher, L. J., III, Dervan, P. B. & Wold, B. J. (1990) *Biochem.* **29**, 8820-6.

(59) Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T. & Helene, C. (1992) *Biochem.* **31**, 9269-78.

(60) Völker, J., Botes, D. P., Lindsey, G. G. & Klump, H. H. (1993) *J. Mol. Biol.* **230**, 1278-90.

(61) Plum, G. E. & Breslauer, K. J. (1994) *Biochem.*, in preparation.

(62) Xodo, L. E., Manzini, G. & Quadrifoglio, F. (1990) *Nucleic Acids Res* **18**, 3557-64.

(63) Pörschke, D. & Eigen, M. (1971) *J. Mol. Biol.* **62**, 291-308.

(64) Craig, M. E., Crothers, D. M. & Doty, P. (1971) *J. Mol. Biol.* **62**, 383-401.

(65) Anshelevich, V. V., Vologodskii, A. V., Lukashin, A. V. & Frank-Kamenetskii, M. D. (1984) *Biopolymers* **23**, 39-58.

(66) Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E. & Razgulyaev, O. I. (1990) *FEBS Lett.* **262**, 20-24.

(67) Xie, D., Bhakuni, V. & Freire, E. (1991) *Biochem.* **30**, 673-678.

(68) Filimonov, V. V., Prieto, J., Martinez, J. C., Bruix, M., Mateo, P. L. & Serrano, L. (1993) *Biochem.* **32**, 12906-12921.

(69) Haynie, D. T. & Freire, E. (1993) *Proteins* **16**, 115-140.

(70) Xie, D., Bhakuni, V. & Freire, E. (1993) *J. Mol. Biol.* **232**, 5-8.

(71) Lyamichev, V. I., Mirkein, S. M., Frank-Kamenetskii, M.D. & Cantor, C. R. (1988) *Nucleic Acids Res.* **16**, 2165-2178.