LBL-15125

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

CHEMICAL BIODYNAMICS DIVISION

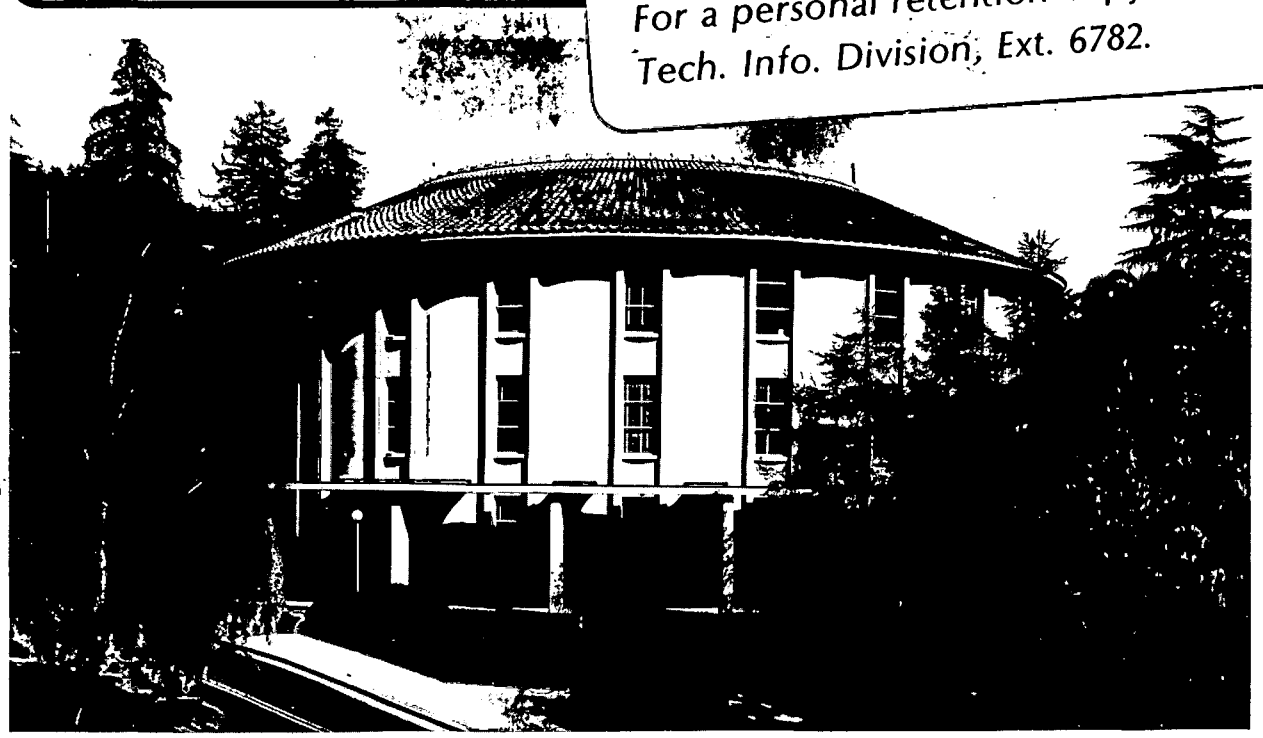
RECEIVED
LAWRENCE
BERKELEY LABORATORY
DEC 7 1982
LIBRARY AND
DOCUMENTS SECTION

THERMODYNAMICS AND KINETICS OF DNA, RNA AND HYBRID OLIGONUCLEOTIDE DOUBLE-STRAND FORMATION

Jeffrey Wayne Nelson
(Ph.D. thesis)

September 1982

TWO-WEEK LOAN COPY
This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782.



LBL-15125 c.2

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-15125

THERMODYNAMICS AND KINETICS OF DNA, RNA AND HYBRID OLIGONUCLEOTIDE DOUBLE-STRAND FORMATION*

Jeffrey Wayne Nelson
(Ph.D. thesis)

Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

September 1982

*This work was supported by the U. S. Department of Energy under Contract No. DE-AC03-76SF00098.

Thermodynamics and Kinetics of DNA, RNA and Hybrid Oligonucleotide Double-Strand Formation

by

Jeffrey Wayne Nelson

ABSTRACT

The double strands formed by the RNA, DNA and RNA·DNA hybrid oligonucleotides rCA_5G + rCU_5G, dCA_5G + dCT_5G and rCA_5G + dCT_5G were studied in order to determine the differences in the stability and dynamics of RNA and DNA. The thermodynamics of double-strand formation were determined by measuring the absorbance vs. temperature at 260 nm for different strand concentrations. The deoxyribo-oligonucleotides were found to be more stable, due to a more favorable enthalpy, than the ribo-oligonucleotides. The double strands were found to aggregate, and the extent of aggregation was determined by analytical untracentrifugation. Mixtures of oligonucleotides capable of forming bulges, such as rCA_6G + rCU_5G, were studied. It was found that the double strands formed dangling ends by breaking the terminal C·G base pairs and forming all of the internal A·U base pairs.

The kinetics of double-strand formation of these oligonucleotides were studied using temperature-jump kinetic techniques. The greater stability of these deoxyribo-oligonucleotides relative to the ribo-oligonucleotides was found to be due to both a faster recombination rate and a slower dissociation rate for the deoxyribo-oligonucleotides. The hybrid oligonucleotide kinetic properties

were similar to the ribo-oligonucleotides, suggesting that the properties of the hybrid are closer to those of the ribo-oligonucleotides.

The intercalation of ethidium ion into these oligonucleotide double strands was measured by monitoring the absorbance vs. temperature at two wavelengths, 260 and 283 nm. Measurements at 260 nm monitor mostly the double-strand to single-strand transition, whereas measurements at 283 nm monitor mostly the binding of ethidium. The ribo-oligonucleotides were fit equally well by two models, one which assumes that the two terminal binding sites are stronger than the internal binding sites by a factor of 140, and another which assumes that the binding of ethidium is cooperative, with a cooperativity parameter $\omega = 0.1$. The deoxyribo-oligonucleotides were fit best by a model assuming all of the binding sites are equal, with no cooperativity. The binding was weaker in the deoxyribo-oligonucleotides relative to the ribo-oligonucleotides.

ACKNOWLEDGEMENTS

Nacho Tinoco receives my thanks and appreciation for his supervision of this project. He put me in charge of the project with little interference, but was always available to discuss problems as they arose. Most confusions were quickly straightened out with a brief discussion of the basic physical chemistry involved. I also thank him for his patience during my periods of less-than-optimal productivity.

Frank Martin taught me most of what I know about oligonucleotides. Much of this work was done in collaboration with him. He was also responsible for synthesizing the deoxyribo-oligonucleotides used in this study.

Arthur Pardi, Ken Dahl, Steve Winkle and Bruce Johnson all had a large influence during my early days in the lab, acquainting me with the procedures used with nucleic acids. We had numerous valuable discussions about our work, and how the different techniques complemented each other. Carlos Bustamante was always a contagious source of enthusiasm towards science.

Kathy Morden and Joe Kao were responsible for bringing much joy and laughter into the lab. Without their cheery outlook, and cherished friendships, life in the lab would have been very dull indeed.

Barbara Dengler and David Koh form the nucleus, around which the lab revolves. With the constant flow of students into and out of the lab, they insure the continuity of the lab, as well as acting

as social directors and story tellers of past personalities. They were also largely responsible for the synthesis of the ribo-oligonucleotides used in this study.

Karen Elliott is responsible for the expert typing of this thesis, as well as being responsible for keeping me calm during the hectic period of writing.

The electronics genius of Gary Smith of the Laboratory of Chemical Biodynamics deserves mention. He was always most happy to attack electronics problems immediately, and the results were always of superlative quality.

Of all of the contributions made by these people, and many of whom escaped mention, the most important was their friendship and comradery, without which much of the enjoyment of earning a Ph.D. would have been absent.

Above all else, the most important factor in completing my graduate studies was the love and support of my parents, James and Dorothy Nelson, to whom this thesis is lovingly dedicated.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

TABLE OF CONTENTS iii

CHPATER I. Introduction 1

1. Thermodynamics of Double-Strand Formation 3

2. Kinetics of Double-Strand Formation 10

3. Ethidium Bromide Binding to Nucleic Acids 11

4. Scope of This Study 14

CHAPTER II. Thermodynamics of Oligonucleotide Double-Strand Formation 17

1. Synopsis 17

2. Introduction 17

3. Experimental Methods 20

A) Synthesis of Oligonucleotides 20

B) Obtaining Melting Curves 23

C) Analysis of Melting Curves 24

D) Equilibrium Ultracentrifugation 26

4. Results 27

A) Melting Curves 27

B) Ultracentrifugation of $rCA_7G + rCU_7G$ 39

5. Discussion 40

A) Comparing $dCA_5G + dCT_5G$ and $rCA_5G + rCU_5G$ 45

B) Chain-Length Dependence of $rCA_nG + rCU_nG$ Thermodynamics 46

C) Low-Temperature Behavior of the Double Strands ... 48

D) Aggregation of Double Strands 49

E) Mismatched Double Helices 59

CHAPTER III. Kinetics of Oligonucleotide Double-Strand Formation 68

1. Synopsis 68

2. Introduction 68

3. Methods and Materials 69

A) Sample Preparation 69

B) The Temperature-Jump Instrument 70

C) Analysis of the Temperature-Jump Kinetic Data 71

4. Results 72

A) Comparison of rCA_5G + rCU_5G, dCA_5G + dCT_5G and rCA_5G + dCT_5G 80

B) The Second Exponential Component 87

5. Discussion 89

A) Comparison of the Ribo-, Deoxyribo- and Hybrid Oligonucleotides 89

B) The Second Exponential Component 90

C) The Hybrid Behaves Similarly to the Ribo-Oligonucleotide 92

D) Comparison with Previous Results 93

E) The Mechanism of Double-Strand Formation 97

CHAPTER IV. Ethidium Ion Binding to Oligonucleotides 100

1. Synopsis 100

2. Introduction 100

3. Experimental 104

4. Theory 106

A) Analysis of Melting Curves 106

B) Description of the Model 109

C) Fitting the Model Parameters to Experimental Results 115

5. Results 116

A) Optical Properties of Ethidium Bromide in the UV 116

B) Oligonucleotide Optical Properties 123

C) Interaction of Ethidium Bromide with rCA_5G + rCU_5G 126

D) Qualitative Interpretation of rCA_5G + rCU_5G + Ethidium 139

E) Interaction of Ethidium Bromide with dCA_5G + dCT_5G 145

6. Discussion 150

A) rCA_5G + rCU_5G + Ethidium at Low Ethidium:Strand Ratios 159

B) rCA_5G + rCU_5G + Ethidium at High Ethidium:Strand Ratios 169

C) dCA_5G + dCT_5G + Ethidium 176

D) Comparison of the Models 184

E) Comparisons with Previous Results 187

7. Conclusion 193

APPENDIX A. Using the Gilford Spectrophotometer 196

APPENDIX B. Using the Temperature-Jump Instrument 221

APPENDIX C. Computer Programs to Analyze Ethidium Melting Data 262

REFERENCES 300

CHAPTER I

Introduction

There are two major classes of nucleic acids, ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). The chemical differences between these two are slight - the DNA lacks the hydroxyl group on the C2' of the ribose sugar, and RNA lacks the methyl group on the C5 of the uracil base, which is on the thymine base of DNA. However, these two differences are enough to account for the dramatic differences in the biological and physical properties of DNA and RNA.

The structure of RNA is much more rigid than DNA. RNA is almost always found in the A-form geometry, no matter what the conditions. However, DNA structure varies from the normal B-form to A-form, C-form, Z-form, etc., by changing the salt concentration or by adding ethanol or other organic compounds to the solvent. Biologically, DNA is generally found as fully base-paired double-helices, whereas RNA is usually single-stranded and forms base-paired secondary structure with numerous loops and bulges.

The stability of the double helix of DNA and RNA depends strongly on the relative amounts of G-C and A-T or A-U bases (Marmur & Doty, 1962; Kallenbach, 1968). G•C base pairs are more stable than A•T or A•U base pairs, although the sequence is also important for the stability (Borer et al., 1974).

Polymer studies using naturally occuring DNA and RNA or synthetic polynucleotides have been used to determine the physical

properties of nucleic acids. While polymer studies are good for measuring the average properties of the DNA, any sequence-specific effects are unobtainable either because of the random sequence of DNA or the limited sequences available from synthetic polymers. These problems can be overcome to a large extent by using oligonucleotides as models to determine the sequence-dependent properties of DNA. Ribo-oligonucleotides of defined sequence may be made enzymatically using polynucleotide phosphorylase and T4 RNA ligase. Deoxyribo-oligonucleotide synthetic procedures have generally lagged behind those of the ribo-oligonucleotide because there are no analogous enzymes for DNA synthesis. However, with the development of chemical DNA synthetic techniques, this trend is being reversed. Hence, deoxyribo-oligonucleotide studies may soon become more prevalent.

Besides the advantage of the availability of defined sequences, oligonucleotides have another important advantage over polymers. Oligonucleotides can be treated as simple molecules, as far as the physical chemist is concerned, and hence the problems associated with polymer studies are eliminated. Hence, the thermodynamic and kinetic studies can be carried out assuming simple bimolecular processes for the single-strand to double-strand transition.

One of the important areas of research on the biological functions of DNA centers around the question of the chemical induction of mutations of DNA. It is thought that frameshift mutations are important in the induction of cancer. One theory is that a chemical mutagen interacts with the DNA by intercalating between the base

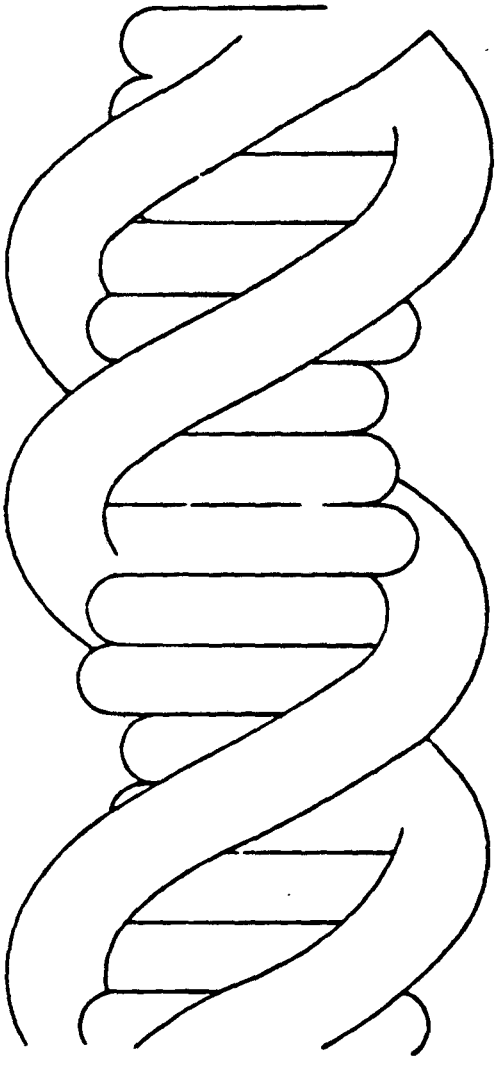
pairs. Figure 1 shows a picture of what this intercalation complex looks like. After intercalating, the mutagen can covalently react with the DNA, causing a distortion of the DNA structure. This perturbation is recognized by repair systems in the cell, which try to fix the DNA by cutting out the region and filling in the gap. During this process, an extra base may be inserted by mistake, or a base may be left out. After replication, the newly synthesized double helix will contain one more or one less base pair. This frameshift mutation usually destroys the function of the protein that is coded by that section of the DNA. Figure 2 shows the scheme for this model.

A theory has been proposed by Streisinger et al. (1966) wherein the incidence of frameshift mutations is increased if the intercalating molecule stabilizes the bulges which can transiently form. If this bulge persists while the section is "repaired", the bulge will be locked into the DNA. Thus, an important question deals with whether an intercalator stabilizes a bulged structure.

1. Thermodynamics of Double-Strand Formation

The thermodynamics of double-strand formation for polynucleotides have been measured by calorimetry. For a review, see chapter 6 of Bloomfield et al. (1974). As mentioned before, studies on polymers are not well suited for determining sequence effects, since average properties are determined for natural DNA, and synthetic polymers are limited as to sequences available. Statistical theories of DNA melting have been developed, and have been quite successful in predicting differential melting profiles for natural

Figure 1. A schamatic representation of the double helix of DNA. Left: the normal helix. Right: the double helix with an intercalated molecule.



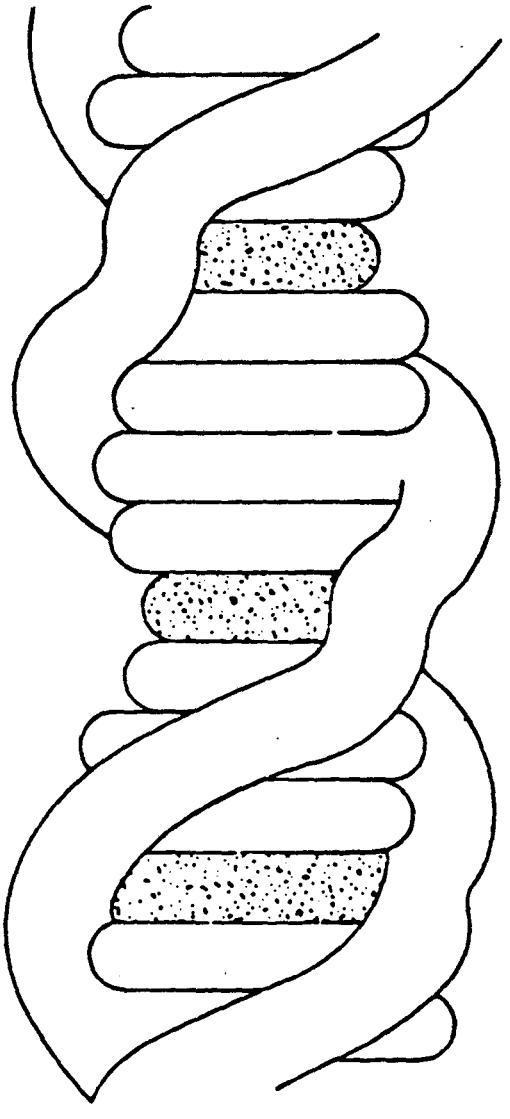
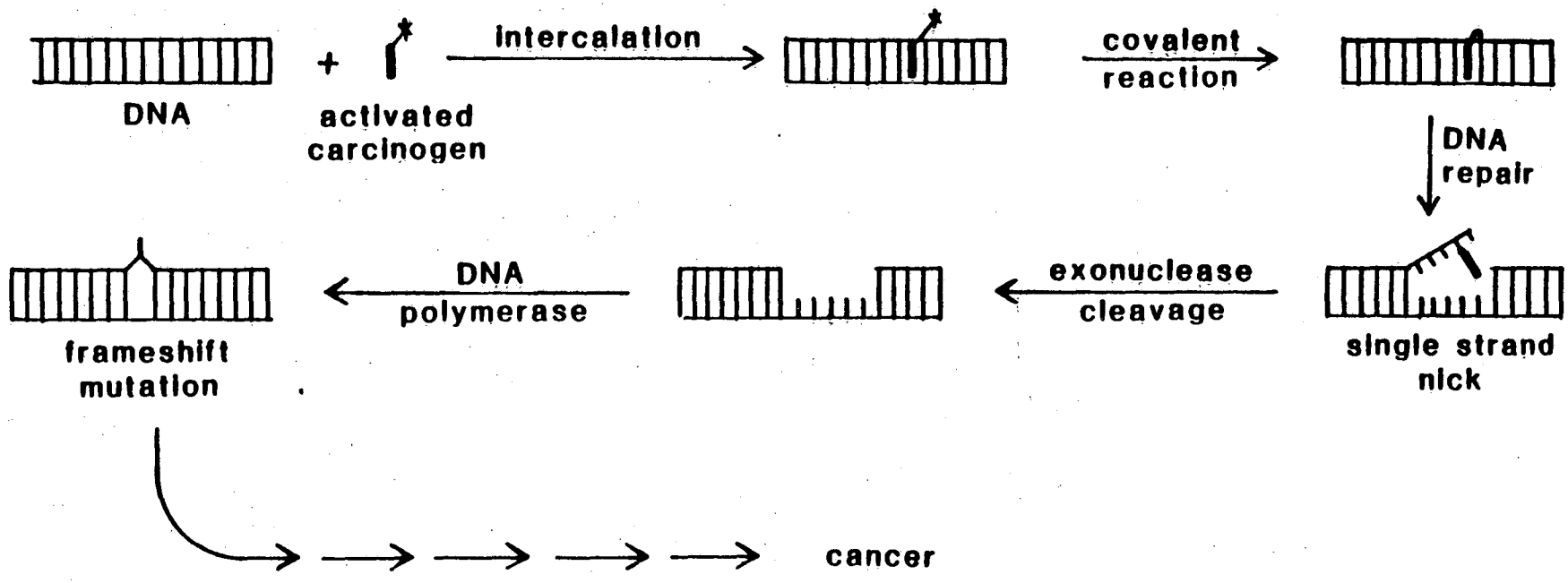


Figure 2. A proposed mechanism for frameshift mutagenesis.

Proposed Mechanism for Frameshift Mutagenesis



DNA's of known sequence (Azbel, 1973, 1975, 1979; Poland, 1974; Fixman and Freire, 1977; Lyubchenko et al., 1978; Vologodskii and Frank-Kamenetskii, 1978).

The thermodynamics of double-strand formation have also been measured for a number of oligonucleotides. Studies on ribo-oligonucleotides are rather extensive. Most studies have analyzed the melting data assuming a two-state equilibrium, in which only single strands and fully base-paired double-strands exist in solution. Comparing the results using many different sequences, the stability of a base pair was found to depend on the bases on either side of it, and not only on whether it is a G•C or an A•U base pair. The melting temperatures of ribo-oligonucleotides could be predicted by assigning thermodynamic values for each of the 10 nearest-neighbor double-strand sequences (Borer et al., 1974). The data base for deoxyribo-oligonucleotides is far less complete, and hence no comparable study has been made.

A statistical theory for oligonucleotide double-strand formation has been worked out, which allows double strand to have partially un-paired bases on the ends (Applequist and Damle, 1963, 1965; Pörschke et al., 1973; Pörschke, 1971; Levine, 1974). The effect of single-strand stacking has also been incorporated into the theory (Appleby & Kallenbach, 1973). The application of the statistical theory on the melting curves of the ribo-oligonucleotides rA_nU_n revealed the possibility that a significant fraction of the double helices have their terminal base pairs broken, indicating the

need to be careful in the application of the two-state model (Levine, 1974).

In order to better understand what forces are important in stabilizing the double strands, studies have been carried out on the deoxyribo-oligonucleotide dG-C-G-C-G-C in different mixtures of aqueous and organic solvents (Albergo et al., 1981; Albergo & Turner, 1981). The results showed that the enthalpy of double-strand formation was different if the solvent was D_2O instead of H_2O, indicating that hydrogen bonding might contribute to the stability of the double strands (Albergo et al., 1981).

The stabilities of the double strands for DNA and RNA both depend on the relative composition of G•C and A•T or A•U base pairs. However, the dependence is different for the two. For DNA, the dependence of the melting temperature, T_m, on G-C content in 0.2M NaCl was found to be $T_m(°C) = 69.3 + 41f_{GC}$ where f_{GC} is the mole fraction of G•C base pairs (Marmur & Doty, 1962). For RNA, the relation was found to be (Kallenbach, 1968): $T_m(°C) = 62 + 78f_{GC}$. Thus, DNA A•T base pairs are more stable than RNA A•U base pairs, whereas RNA G•C base pairs are more stable than DNA G•C base pairs. Comparisons of DNA and RNA oligonucleotides of identical sequences should help in understanding what differences between base pairs in DNA and RNA account for this.

The oligonucleotides dC-G-C-G-A-A-T-T-C-G-C-G, dC-G-T-G-A-A-T-T-C-G-C-G, and C-G-C-A-G-A-A-T-T-C-G-C-G have been studied extensively by NMR and calorimetry to understand the structural and stabilizing effects of a G•T wobble base pair and a bulged adenine

on double strand formation (Patel et al., 1982a,b,c; Pardi et al., 1982). From chemical shifts, it was determined that the "bulged" adenine was stacked in the helix. The enthalpies of double-strand formation were near -103 kcal/mol for both the normal helix and the helix with the bulged adenine (Patel et al., 1982c). However, these oligonucleotide single strands form hairpin loops, which interfere with the comparison of the stabilities (Kenneth J. Breslauer, personal communication). From measurements of the exchange rates of the exchangable imino base-pairing protons, it was determined that the G•T base pair causes only local perturbations, whereas the bulged adenine causes perturbations throughout the double helix (Pardi et al., 1982).

2. Kinetics of Double-Strand Formation

Temperature-jump relaxation studies have been carried out on a number of oligonuceotides. See Bernasconi (1976) for a discussion of relaxation kinetics techniques. As is the case in the thermodynamics, ribo-oligonucleotides have been studied far more than deoxyribo-oligonucleotides. The references are given in Chapter III for many of the studies.

Models for double-strand formation in oligonucleotides have been developed in which the mechanism is assumed be a sequential base-pairing of the helix, often called a "zipper" model (Craig et al., 1971; Pörschke, 1974). From activation energies determined from oligonucleotide studies, the mechanism for double-strand formation is proposed to be the formation of a nucleus of a few base pairs. The rate-determining-step is then the formation of the next

base pair, after which the double strands zipper up quickly to form the fully base-paired double strands.

Because of the lack of data for deoxyribo-oligonucleotides, there are no good comparisons between ribo- and deoxyribo-oligonucleotide double-strand kinetics. Such comparisons would help determine the dynamic differences between DNA and RNA to complement the knowledge of thermodynamic differences.

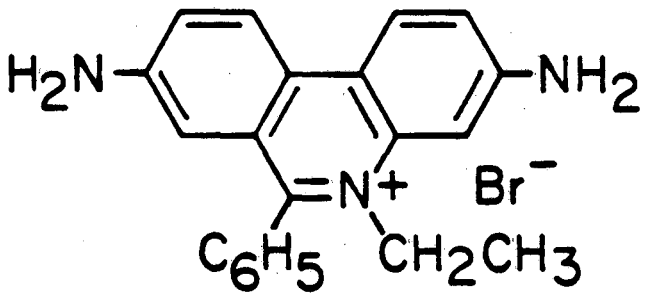
3. Ethidium Bromide Binding to Nucleic Acids

Ethidium bromide is an excellent probe to use to understand what features are important in stabilizing the intercalated complex. It has a large equilibrium constant for intercalation, and forms dimers with itself to a negligible extent. It also exhibits large optical changes upon intercalation. Studies usually involve either binding to DNA polymers or to dinucleotides. Chapter IV lists the references for many of these studies. Figure 3 shows the structure of ethidium bromide.

The studies involving DNA polymers showed quite clearly that ethidium bromide binds with nearest-neighbor exclusion, meaning that an ethidium intercalated between two base pairs prohibits the intercalation to the next base pair. Also, the binding is non-cooperative: ethidium binding to one site does not affect the next binding site two base pairs away. A statistical model accounting for the nearest-neighbor exclusion (Crothers, 1968; McGhee & von Hippel, 1974) fits the polymer data very well (Bresloff & Crothers, 1975).

From binding studies using dinucleotides, it was determined that ethidium bromide binds much more strongly to pyrimidine-purine

Figure 3. The structure of ethidium bromide.



Ethidium bromide

sequences than to purine-pyrimidine sequences (Krugh & Reinhardt, 1975; Dahl et al., 1982). However, the magnitude of this difference could not be determined, since the dinucleotides form very unstable double strands in the absence of ethidium bromide.

Kinetic studies of ethidium binding to calf thymus DNA exhibited a very surprising effect: ethidium bromide can be transferred from one binding site in the DNA to another, without ever going into solution (Bresloff & Crothers, 1975). This direct-transfer mechanism allows the ethidium to move very quickly along the DNA double helix.

The kinetics of ethidium binding to dinucleotides indicated that the ethidium bound to a single-stranded dinucleotide (Davanloo & Crothers, 1976). This complex then had to re-arrange in some manner, which was the rate-determining step at high concentrations. A second dinucleotide then base-pairs to this complex, forming the double strand with an intercalated ethidium bromide.

No extensive study has been made of ethidium bromide binding to oligonucleotides in an effort to determine the magnitude of the sequence specificity for binding, or to test the models for ethidium binding. Considering the advantages of using oligonucleotides as models for DNA properties, it would seem to be the next logical step in the characterization of ethidium bromide intercalation into nucleic acids.

4. Scope of This Study

There were four major goals for this project. First, I wanted to determine the thermodynamic differences between RNA and DNA

double-strand formation. Second, I wanted to determine the magnitude of the destabilizing effect of a bulged base on the double strand. Third, I wanted to determine the kinetic differences between RNA and DNA oligonucleotides. And last, I wanted to find out how much information I could obtain about the seqence specificity of ethidium intercalation into double-stranded oligonucleotides.

Towards these ends, I have studied the properties of oligonucleotides of the form rCA_nG + rCU_nG and dCA_nG + dCT_nG. By characterizing the thermodynamics of these analogous sequences, I have determined what differences are important in determining the relative stabilities of these oligonucleotides. It was also hoped that I could study the destabilizing effect of a bulge when strands were mixed with unequal numbers of A's and U's, for example rCA_6G + rCU_5G. However, instead of forming a bulged A, the structure had a dangling end on one side or the other, allowing all the A•U base pairs to form in the interior of the helix. During the course of these studies, the problem of aggregation of the double strands was recognized, and the effect of this on the thermodynamics was estimated. These studies are described in Chapter II.

Temperature-jump kinetic techniques were applied to these oligonucleotides, and on the hybrid formed by rCA_5G + dCT_5G, in order to determine the kinetic differences between these oligonucletoides. A comparison of the DNA, RNA and hybrid kinetics showed that the hybrid double strand behaved more like the RNA than the DNA. Also, the kinetics of recombination of the deoxyribo-oligonucleotides was found to be faster than the ribo-oligonucleotides,

which is consistent with the same observation of the kinetics of single-strand stacking (Dewey & Turner, 1979). The results of these studies are reported in Chapter III.

The final part of this study involved the interaction of ethidium bromide with these oligonucleotides. A statistical model was developed which allows the ethidium ion to bind to the six potential binding sites in any combination consistent with nearest-neighbor exclusion. From melting curves on the mixture of strands and ethidium, the fraction ethidium bound can be fairly directly measured. Comparisons between the statistical model and the experimental results allows the determination of the thermodynamics of ethidium binding to these oligonucleotides, and the validity of different models can be compared. The binding to the deoxyribo-oligonucleotides was found to occur with very little sequence specificity, with a binding constant and enthalpy consistent with work done on ethidium binding to DNA polymers. However, the binding of ethidium to the ribo-oligonucleotides was found to occur with significant cooperativity. One possible explaination is that the ethidium binds preferentially to the ends of the helix, where the distortions of the double-helical structure might be less than when binding to the interior. These studies are reported in Chapter IV.

CHAPTER II

Thermodynamics of Oligonucleotide Double-Strand Formation

1. Synopsis

The thermodynamic parameters for the double-strand formation of the molecules rCA_mG + rCU_nG, m,n = 5-7, and dCA_mG + dCT_nG, m,n = 5,6, were measured from optical melting curves. Normal helices are formed when m = n. The deoxyribo-oligonucleotides are more stable than the ribo-oligonucleotides, characterized by a more favorable enthalpy. Double helices with mismatched bases can be formed by mixing oligonucleotides with m ≠ n. Such helices may form several possible structures. A structure with a dangling base is favored over a structure with a bulged base. The destabilization of the double strands by the formation of a bulged base was determined to be greater than 1.6 kcal/mol of free energy at 10° C. The extent of aggregation in the oligonucleotide double strand $rCA_7G \cdot rCU_7G$ was determined using ultracentrifugation equilibrium. The possible effects of aggregation on the determination of the thermodynamic parameters for double-strand formation are discussed.

2. Introduction

Stability of nucleic acid secondary structure is a sequence-dependent property which helps determine the three-dimensional folding of single-stranded DNA and RNA and is likely to affect enzymatic copying of DNA and RNA templates (Gilbert, 1976; Sims et al., 1978; Schaller, 1978; Huang & Hearst, 1980). Estimates of RNA

secondary structure stability are usually based on procedures utilizing data obtained from optical melting studies of ribo-oligonucleotides (Tinoco, Jr., et al., 1973; Borer et al., 1974). The deoxyribo-oligonucleotide data base for a similar attempt at DNA stability prediction is more limited (Patel & Canual, 1979; Wells et al., 1977; Haasnoot et al., 1979). Sequence dependence of helix stability is evident from the differences in melting temperature of repeating polymers of identical base composition (Wells et al., 1970; Felsenfeld & Miles, 1967), as well as the well-known dependence of T_m on G•C content of DNA (Marmur & Doty, 1962).

Prediction of secondary structure stabilities has been made taking into account both short- and long-range interactions. An attempt has been made to predict the high-resolution melting profiles of DNA molecules without taking into account nearest-neighbor (or longer range) effects (Lyubchenko et al., 1978; Vizard et al., 1978). However, the melting behavior of the duplexes of the block oligonucleotides $d(C_mA_n)\cdot d(T_nG_m)$ have been interpreted as indicating the importance of long-range interactions (Wells et al., 1977).

The stabilities of perturbed DNA structures might be important in evaluating mechanisms of mutations. In particular, the stability of a bulged base might be important in frame-shift mutations by intercalating agents (Streisinger et al., 1966). By chemically modifying a small fraction of adenine bases in poly(A), the destabilizing influence of a bulged modified adenine on the formation of poly(A)•poly(U) was determined to be 2.8 kcal/mol (Fink & Crothers, 1972b). It should be possible to determine more

precisely the nature of this destabilization by studying oligonucleotide models which lend themselves to more direct physical study.

I have studied the thermal helix-coil transitions of a series of oligonucleotide duplexes of DNA and RNA of related sequence:

$$rCA_mG + rCU_nG, \quad m,n = 5\text{-}7$$

$$dCA_mG + dCT_nG, \quad m,n = 5,6$$

These sequences were chosen for study because the terminal G·C base pairs are expected to minimize fraying of the termini, a complicating factor in previous oligonucleotide studies (Martin et al., 1971; Uhlenbeck et al., 1971; Breslauer et al., 1975). Further, these sequences closely model both the frame-shift mutation "hot spots" of Streisinger et al. (1966) and the sequences of rho-independent transcriptional termination sites (Gilbert, 1976; Adhya & Gottesmann, 1978). I have investigated the possibility that imperfect duplexes of the type $rCA_nG + rCU_{n\pm1}G$ might form bulged double helices such as those which Streisinger has proposed as intermediates in frame-shift mutagenesis (Streisinger et al., 1966).

The relative stabilities of DNA, RNA, and DNA·RNA hybrid duplexes for a few of these oligonucleotides have been reported previously and have been related to the process of termination of transcription (Martin & Tinoco, Jr., 1980).

The thermal stability of base-paired complexes can be measured by monitoring any of several physical properties as a function of temperature and oligonucleotide concentration; absorbance of UV

radiation at 260 nm was monitored in the experiments reported here. Proton magnetic resonance studies and calorimetry can give complementary information and do have particular advantages, but absorbance methods require less material and allow measurements over a wider range of oligonucleotide concentration. This advantage can become crucial when aggregation may be a complicating factor. Romaniuk et al. (1979) have studied the thermal transitions of several ribo-oligonucleotides, including rC-A-U-G, using the chemical shifts of nonexchangeable protons at oligonucleotide concentrations on the order of 10^{-2}M. The correlation of these results with those of optical methods at much lower concentrations is problematic without a more complete understanding of aggregation of oligonucleotide duplexes.

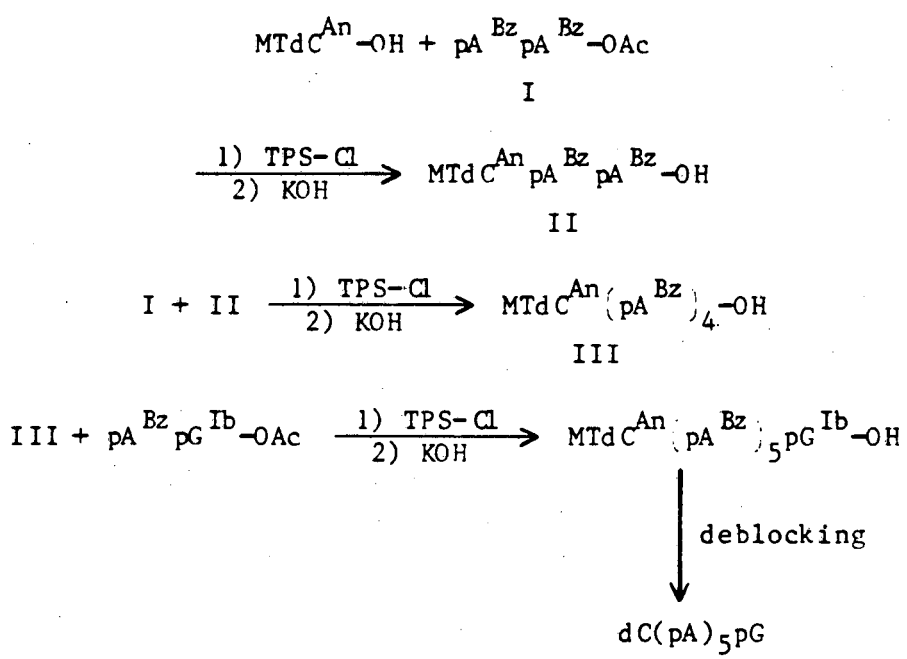
The relationship between optical and NMR melting transition measurements is investigated in more detail in a paper (Pardi et al., 1981) reporting the NMR studies on some of the duplexes discussed here. The results of this chapter have been published (Nelson et al., 1981).

3. Experimental Methods

A) Synthesis of Oligonucleotides

The deoxyribo-oligonucleotides were synthesized by Dr. Francis H. Martin using the classical diester approach developed in the laboratory of Khorana (Khorana, 1968; Goeddel et al., 1977). Reagents employed for blocking of base amino groups were benzoyl (Bz) chloride for adenine, isobutyric (Ib) anhydride for guanine, and anisoyl (An) chloride for cytosine. 5'-Hydroxyls were blocked when

necessary with a monomethoxytrityl (MT) group, 3'-hydroxyls with acetyl (Ac) groups, and 5'-phosphates with cyanoethyl (CE) groups. 2,4,6-Triisopropylbenzenesulfonyl chloride (TPS-Cl) was the activating agent for condensations. Following each condensation, the 3'-hydroxyl-blocking group was removed with 1M KOH in 50% aqueous pyridine (5 min at 0°C). Oligonucleotides were synthesized by sequential condensation of dinucleotide or trinucleotide blocks to the 3'-hydroxyl of the growing chain:



and similarly for the synthesis of the $dC(pT)_n pG$ oligonucleotides. Blocked oligonucleotides were separated by DEAE-cellulose chromatography using triethylammonium bicarbonate (TEAB) and/or ethanol gradients (Van de Sande et al., 1976), or by use of reverse-phase (C_{18}) columns (Fritz et al., 1978). Deblocked oligonucleotides were rinsed well at low ionic strength after adsorbance on DEAE-cellulose columns, eluted with a solution of high ionic strength (NaCl or

TEAB), and repurified by RPC-5 column chromatography using NaCl gradients at neutral pH. The NMR spectra showed no detectable contamination, and all oligonucleotides formed helical complexes specifically with their complements. Ribo-oligonucleotides were synthesized by enzymatic procedures using primer-dependent polynucleotide phosphorylase (PNPase) at high ionic strength (0.4 - 0.8M) and appropriate nuclease treatment (Martin et al., 1971; Uhlenbeck et al., 1971):

$$rCpA + rADP \xrightarrow{PNPase} rC(pA)_n, \quad n = 0,1,2,3,\ldots(\text{mixture})$$

$$\xrightarrow{\text{RPC-5 chromatography}} \text{pure } rC(pA)_5 \text{ and other pure } rC(pA)_n\text{'s}$$

$$rC(pA)_5 + rGDP \xrightarrow{PNPase,\ T_1\ RNase} rC(pA)_5pGp$$

$$rC(pA)_5pGp \xrightarrow[\text{phosphatase; RPC-5 column}]{\text{alkaline}} rC(pA)_5pG$$

Extinction coefficients of rCA_5G, rCA_8G, rCU_5G, and rCU_8G were determined by measurement of absorbance before and after alkaline hydrolysis to nucleotides (Warshaw, 1965). Extinction coefficients of other RNA compounds were estimated from the ones measured. Extinction coefficients of the deoxyribo-oligonucleotides were estimated from the extinction coefficients of mononucleotides and dinucleoside phosphates (Handbook of Biochemisry and Molecular Biology, Nucleic Acids, 1975). All are tabulated below.

Oligonucleotide	Extinction Coefficient at 260 nm ($\times 10^{-4}$ mol ℓ^{-1} cm $^{-1}$) 25° C	50° C
rCA_5G	7.10	7.64
rCA_6G	8.06	8.70
rCA_7G	9.00	9.74
rCA_8G	9.92	10.77
rCU_5G	6.31	6.34
rCU_6G	7.22	7.26
rCU_7G	8.13	8.19
rCU_8G	9.04	9.13
dCA_5G	7.88	8.28
dCT_5G	5.79	5.81

B) Obtaining Melting Curves

The melting curves were measured on a Gilford model 250 spectrophotometer equipped with a Gilford model 2527 thermoelectric temperature programmer. The data were collected either by recording the absorbances and temperatures manually or automatically by a Commodore PET model 2001 microcomputer *via* an interface to the Gilford spectrophotometer. The readings were at 260 nm, with a temperature scan rate of 1° C/min. At this rate, the instrument exhibited no lag between the temperature in the cell and that indicated by the thermoprogrammer. Details on the use of the instrument are given in Appendix A.

The cells for the Gilford have dimensions of 1.2 (ℓ) x 0.6 (w) x 1.9 (h) cm. The path lengths were 1.0, 0.5, 0.2, or 0.1 cm. Path

lengths of 0.01 and 0.02 cm were obtained by using a 0.2 cm cell fitted with a quartz spacer of 0.19 or 0.18 cm, respectively. The actual path lengths were determined from absorbances of sodium dichromate solutions of known concentrations. The path lengths of 0.1 cm and longer were essentially correct; those of the 0.01 and 0.02 cm varied somewhat between different cells and spacers. The path lengths of all the combinations were determined.

The cells were soaked in concentrated nitric acid before using. The samples were degassed either by purging the buffer with helium prior to mixing or by heating the samples briefly to about 60° C and shaking to eliminate bubbles prior to filling the cells.

Evaporation was controlled in cells without stoppers by floating silicon oil (Dow Corning 200 Fluid, 20cS viscosity) over the sample. This had no effect on the absorbance of the solution and was very effective in reducing evaporation. In cells with Teflon stoppers, no oil was used. In most cases, the samples were returned to 0° C after the melting curve was completed. The evaporation, as indicated by an increase in absorbance, was usually less than 1%. Samples with evaporation greater than this were not used in the analysis. Additional information on the use of the Gilford cells is given in Appendix A.

The buffers used were either 0.2 or 1M NaCl in 0.01M phosphate buffer, pH = 7, and 0.1 mM EDTA.

C) Analysis of Melting Curves

The thermodynamic parameters were obtained from melting curves using the method of Martin et al. (1971). The relative absorbance

vs. temperature was plotted by normalizing the absorbance of single strands to 1.0 at 50°C. This allows us to compare the melting at different concentrations directly. In some cases, the double-strand to single-strand transition is not complete at 50°C. In these cases, the absorbance of the single strands at 50°C is obtained by extrapolating from higher temperatures (see figure 1).

If we know the upper (single-strand) and lower (double-strand) base lines, we can determine f, the fraction of strands in double helices at any temperature, from the formula

$$1 - f = \frac{A(T) - A_d(T)}{A_s(T) - A_d(T)} \qquad (1)$$

where $A(T)$, $A_d(T)$, and $A_s(T)$ are the relative absorbances of the experimental curve, the lower base line, and the upper base line, respectively. We have assumed a two-state model. Since the oligonucleotides in this study are not self-complementary, we can directly measure $A_s(T)$, the single-strand base line. However, $A_d(T)$, the double-strand base line, can only be estimated from the behavior of the melting curve at low temperatures.

We can relate f to the equilibrium constant using equation 3 by setting $C_a = C_b$, the total concentrations of oligonucleotides A and B, respectively:

$$A + B \rightleftharpoons DS$$

$$f = \frac{(DS)}{(DS) + (A)} = \frac{(DS)}{C_a} \qquad (2)$$

$$K = \frac{(DS)}{(A)\ (B)} = \frac{f}{(1 - f)^2 C_a} \qquad (3)$$

The melting temperature, T_m, is defined as the temperature at which half of the strands are in double strands (f = 0.5). From the concentration dependence of the T_m, we can calculate the $\Delta H°$ for the transition from the slope of a plot of $1/T_m$ vs. the log of the concentration:

$$\Delta H^0 = 2.303R \frac{d \log(C_a)}{d(1/T_m)} \quad (4)$$

where C_a is the total concentration of each of the oligonucleotide strands.

In addition, assuming a two-state model, we can determine $\Delta H°$ from the slope of a melting curve at the T_m using the formula:

$$\Delta H^0 = 6R(T_m)^2 \left(\frac{df}{dT}\right)_{T_m} \quad (5)$$

The $\Delta H°$ calculated by both equations should agree if the system is two-state, and if we have drawn the base lines correctly.

D) Equilibrium Ultracentrifugation

Equilibrium ultracentrifuge experiments were carried out on the single strand rA_7G and the double strands formed from rCA_7G + rCU_7G using a Beckman model E analytical ultracentrifuge equipped with UV scanning optics. The speed was controlled and monitored electronically.

For rA_7G, a sample with a concentration of 180 μM was prepared in a 1M NaCl solution using a centrifuge cell with a 1-mm double-sector titanium center piece. The centrifuge was run at 34,900 rpm (3660 rad/s) at a temperature of 2°C. Scans were taken several hours apart until successive scans were the same, indicating that

equilibrium had been attained. The total time of the experiment was 4-1/2 days. The absorbance at 260 nm was too high to measure; therefore, scans were taken at 265, 270, 275, 280, and 285 nm. The final concentrations ranged from 18 μM at the top of the cell to 490 μM at the bottom.

The double-strand experiment was performed using the same 1-mm cell, using an initial concentration of 220 μM in both rCA_7G and rCU_7G. A speed of 20,400 rpm was used (2140 rad/s); the temperature was controlled at 3°C. The attainment of equilibrium was determined in the same manner as described above. The final concentrations were 27 μM at the top and 1200 μM at the bottom. The equilibrium constant at 3°C for rCA_7G + rCU_7G double-strand formation was measured to be 1.0×10^9 (see results), so at a concentration of 27 μM, greater than 99% of the strands are in double helices. The scans were taken at 280, 285, 290, and 295 nm. The total time of the experiment was again 4-1/2 days.

4. Results

A) Melting Curves

The melting curves for seven concentrations of dCA_5G + dCT_5G in 1M NaCl are shown in figure 1. The curves are all normalized to an absorbance of 1.0 at 50°C. The upper line is the experimental melting curve for the single strands. The melting curves do not superimpose at low temperatures; the hypochromicity increases with higher concentrations. This effect is due to aggregation of the double strands and will be discussed later. Because of the concentration-dependent hypochromicities, I assign a different base line

Figure 1. Melting curves for seven concentrations of dCA_5G + dCT_5G in 1M NaCl. The absorbances were normalized to 1.0 at 50° C (see text). The concentrations and melting temperatures are listed in Table I. The upper curve is the experimental single-strand melting curve.

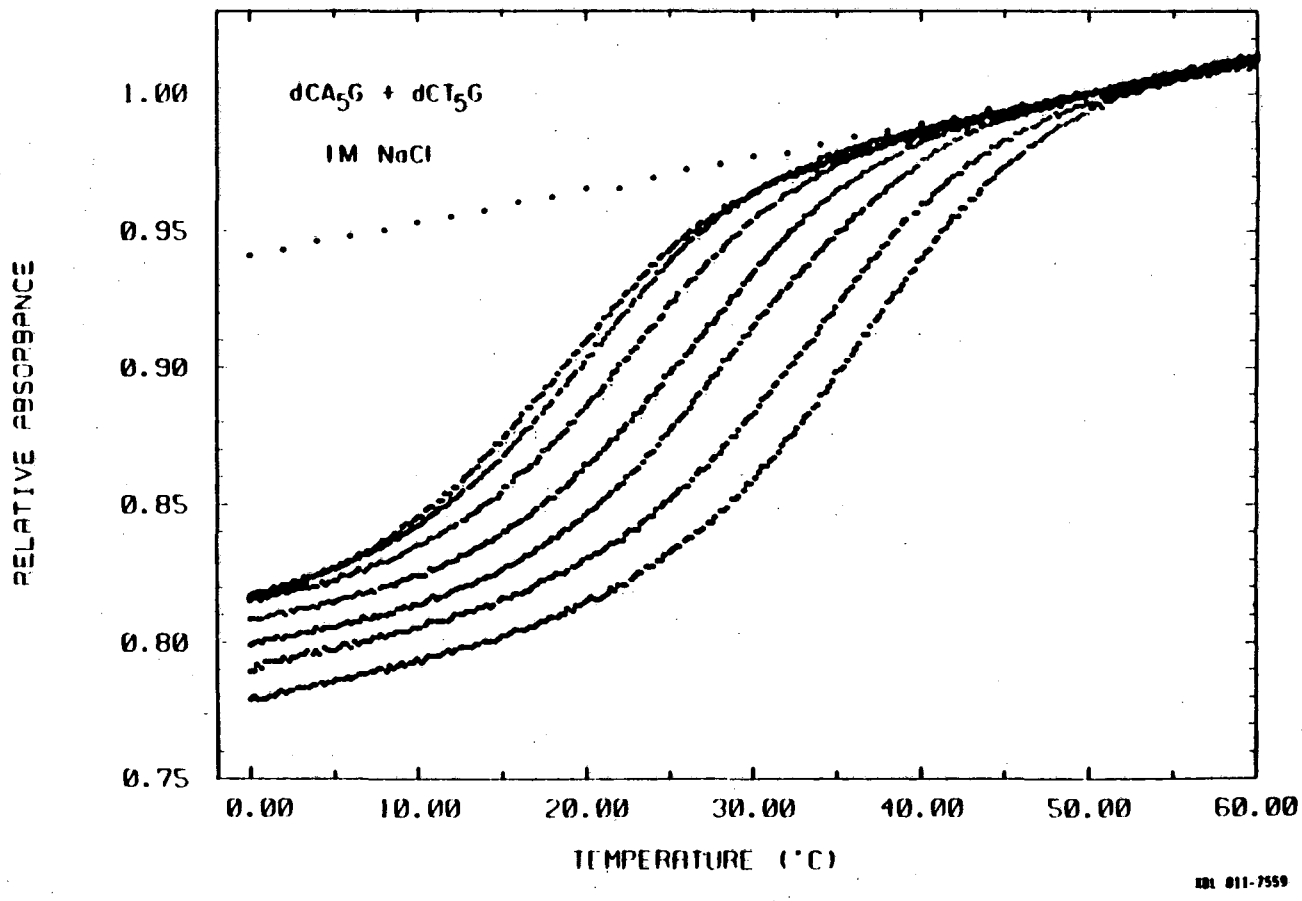

dCA5G + dCT5G
1M NaCl
1.00
0.95
0.90
0.85
0.80
0.75
RELATIVE ABSORBANCE
0.00
10.00
20.00
30.00
40.00
50.00
60.00
TEMPERATURE (°C)
BBL 811-7559

TABLE I

Percent Hypochromicities and T_m's Calculated for dCA_5G + dCT_5G Using Flat and Sloping Double-Strand Base Lines, 1M NaCl

	Flat Base Lines		Sloping Base Lines[a]	
Conc. (μM)	Hypo-chromicity (%)	T_m (°C)	Hypo-chromicity (%)	T_m (°C)
606	22	32.9	22.1	36.1
273	21	30.0	21.0	33.2
91.0	20	26.3	19.9	29.2
44.4	19.5	23.6	19.2	26.6
17.6	19	20.3	18.5	23.0
9.79	19	17.8	18.5	20.2
5.86	19	16.7	18.5	19.1

[a]Slope = 1.4×10^{-3}/°C.

Figure 2. The plots of $1/T_m$ vs. log(concentration) for dCA_5G + dCT_5G in 1M NaCl: ●, flat double-strand base lines; ■, sloping double-strand base lines.

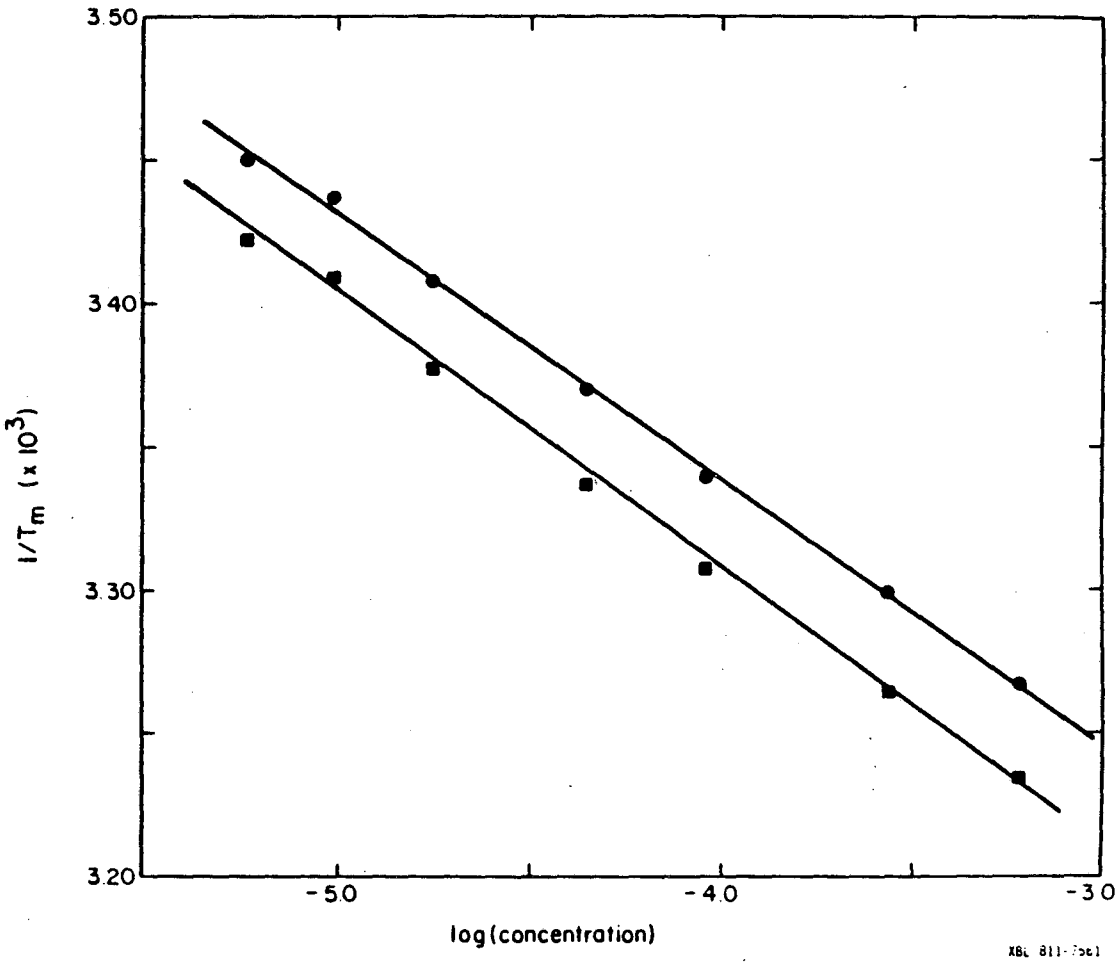

3.50
3.40
3.30
3.20
$1/T_m$ (x 10^3)
-5.0
-4.0
-3.0
log (concentration)
XBL 811-7561

TABLE II

Thermodynamic Parameters for $dCA_5G + dCT_5G$, 1M NaCl

	Flat Base Line		Sloping Base Line[a]	
	Conc. Dependence	$\left(\frac{df}{dT}\right)_{T_m}$	Conc. Dependence	$\left(\frac{df}{dT}\right)_{T_m}$
$\Delta H°$ (kcal/mol)	-49 ± 1	-42 ± 2	-47 ± 2	-51 ± 3
$\Delta S°$ (e.u.)	-145 ± 5		-136 ± 5	
$\Delta G°$ (kcal/mol,25°C)	-6.1		-6.6	
T_m (200μM,°C)	28.8		32.0	

[a]Slope = 1.4×10^{-3}/°C.

to each curve. If one base line is chosen for all of the curves, the results are significantly different.

The melting curves do not have zero slope at low temperatures. This effect could be due to a temperature-dependent extinction coefficient due to a small conformational change or perhaps differential melting of the ends. I did the analyses using both flat and sloping lower base lines. The resulting hypochromicities and melting temperatures (T_m) for each concentration are shown in Table I for dCA_5G + dCT_5G. We can determine $\Delta H°$ from a plot of $1/T_m$ vs. $\log(C_a)$ using equation 4. The resulting plots are essentially straight lines, as shown in figure 2.

Table II summarizes the results of the analyses, including the $\Delta H°$ calculated from the slope of the melting curve using equation 5. In addition to $\Delta H°$, $\Delta S°$ and $\Delta G°$, I tabulate the T_m for a solution 200μM in each strand (400μM total strand concentration) as a reference to compare stabilities of different oligomers. This corresponds to a concentration of 100μM for a self-complementary oligonucleotide. The $\Delta H°$ calculated from the concentration dependence is essentially the same, whether flat or sloping base lines are used. However, the value for $\Delta H°$ calculated from the slope of the curve at the T_m using equation 5 increases when using a sloping base line. The stability, as indicated by the $\Delta G°$ or the T_m (200μM), increases when a sloping base line is used. This is simply because the T_m's are shifted to higher temperatures. I will use flat base lines unless specifically noted. This will allow us to make direct comparisons to earlier work, all of which assumed flat lower base

lines. I will also use values for $\Delta H°$ derived from the concentration dependence of the T_m.

Table III compares the thermodynamics of dCA_5G + dCT_5G with rCA_5G + rCU_5G in 0.2 and 1M NaCl. The $\Delta H°$ for the deoxyribo-oligonucleotides is about 7-8 kcal greater than for the ribo-oligonucleotides. In both cases, the $\Delta H°$ does not change significantly when the salt concentration is increased from 0.2 to 1M NaCl, although the stabilities increase somewhat: 2.4°C for the deoxyribo- and 4.7°C for the ribo-oligomers.

The effects of chain length on the thermodynamic stability are tabulated in Table IV for the ribo-oligomers rCA_nG + rCU_nG, where n = 5-7, in 1M NaCl. The values calculated using the parameters from Ebrer et al. (1974) are shown below the experimental parameters. The calculated stabilities are lower than those observed experimentally, although the agreement is better for the longer oligomers. It is also seen in Table IV that the values for $\Delta H°$ calculated from the concentration dependence of T_m and from the slope of the melting curve at the T_m agree fairly well for rCA_5G + rCU_5G and rCA_6G + rCU_6G but differ significantly for rCA_7G + rCU_7G.

When oligomers are mixed with unequal numbers of A's and U's (or T's), the double helices can form bulged structures. One example is mixing rCA_6G with rCU_5G. The results of such mismatches are shown in Table V. The mixtures rCA_7 + rCU_5G and rCA_5G + rCU_7 are included to compare the "bulged" structure to a helix which must dangle bases off one end. The thermodynamics for the normal double helices are included for comparison. Since the mismatched double

TABLE III

Thermodynamic Parameters for
$dCA_5G + dCT_5G$ and $rCA_5G + rCU_5G$ Using
Flat Lower Base Lines

Oligomer	NaCl Conc. (M)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (e.u.)	$\Delta G°$ (kcal/mol,25°C)	T_m (200μM,°C)
$dCA_5G + dCT_5G$	0.2[a]	-50	-149	-5.7	26.5
	1.0	-49	-145	-6.1	28.9
$rCA_5G + rCU_5G$	0.2	-43	-130	-4.6	19.2
	1.0	-41	-120	-5.3	23.9

[a]These values differ from those previously published (Martin & Tinoco, Jr., 1980). An experimental error in preparing the buffer was detected and corrected. The values reported therein are: $\Delta H°$ = -41 kcal/mol and T_m(200μM) = 27.5°C.

TABLE IV

Thermodynamic Parameters for $rCA_nG + rCU_nG$, 1M NaCl

Oligomer	ΔH° (kcal/mol) Conc. Dependence	ΔH° (kcal/mol) $\left(\frac{df}{dT}\right)_{T_m}$	ΔS° (e.u.)	ΔG° (kcal/mol, 25° C)	T_m (200μM, °C)
$rCA_5G + rCU_5G$	-41	-45	-120	-5.3	23.9
Calculated[a]	-44.6		-136	-4.0	15.7
$rCA_6G + rCU_6G$	-50	-46	-148	-6.2	29.2
Calculated	-52.8		-160	-5.2	23.6
$rCA_7G + rCU_7G$	-63	-53	-187	-7.3	33.8
Calculated	-61.0		-183	-6.4	29.7

[a]Calculated using the parameters of Borer et al. (1974).

TABLE V

Effect of Mismatched Bases on Double-Strand Stability, 0.2M NaCl

Oligomer	$\Delta H°$ (kcal/mol)	$\Delta G°$ (kcal/mol, 10°C)	T_m (200μM, °C)
$rCA_5G + rCU_5G$	-43	-6.5	19.2
$rCA_6G + rCU_5G$	-39	-5.3	10.8
$rCA_7 + rCU_5G$	-35	-4.9	7.7
$rCA_5G + rCU_6G$	-33	-5.5	12.9
$rCA_5G + rCU_7$	-37	-4.9	8.1
$rCA_5G + dCT_5G$	-50	-8.1	26.5
$dCA_5G + dCT_6G$	-45	-6.3	17.6

strands are not stable at 25°C, the values for $\Delta G°$ are tabulated at 10°C. The "bulged" mismatches destabilize the double helices by about 1.0-1.2 kcal/mol for the ribo-oligomers. The destabilization is about 1.8 kcal for a mismatched thymine in the deoxyribo-oligomers. By comparing rCA_7 + rCU_5G with rCA_6G + rCU_5G, we will see that the mismatched double helices probably form structures with dangling ends rather than bulges (see Discussion).

B) Ultracentrifugation of rCA_7G + rCU_7G

Ultracentrifugation equilibrium studies were carried out on the single-strand oligomer rA_7G and the rCA_7G + rCU_7G double strand in 1M NaCl. The slope of a plot of log(conc) vs. r^2, where r is the distance from the axis of rotation, is given by equation 6:

$$\frac{d\ \log(\text{conc})}{dr^2} = \frac{M_w(1 - \bar{v}\rho)\omega^2}{4.606RT} \qquad (6)$$

where M_w is the weight-average molecular weight (hereafter called apparent molecular weight), ω is the rotational velocity in rad/s, $\bar{v}$ is the specific volume of the molecules, ρ is the density of the solution, R is the gas constant, and T is the temperature. If the molecules do not aggregate or dissociate, the plot of log(conc) vs. r^2 should be a straight line. If there is aggregation, the slope will increase with increasing concentration.

The results of the ultracentrifugation for the single-strand rA_7G in 1M NaCl at 2°C is essentially a straight line over the concentration range of 18-490μM, indicating that this single strand does not aggregate to a significant degree. We can use the molecular weight of the oligonucleotide (2742 daltons for the Na^+

salt) to obtain a value for $(1 - \bar{v}\rho)$. The resulting value of 0.328, with the density of 1M NaCl of 1.04 g/ml, yields a value of 0.646 ml/g for the specific volume of the oligomer. This specific volume corresponds to the hydrated molecule. The corresponding value for double-stranded Na-DNA in 1M NaCl was measured to be 0.563 ml/g (Cohen & Eisenberg, 1968). The higher value for the oligomer might be caused by a lowering of density due to a lower percentage of phosphates (there were no terminal phosphates) and to a different degree of hydration in the oligomer vs. the polymer. The buoyant density of single-stranded DNA is only about 3% larger than for double strands (Wiesehahn et al., 1976). Thus, the value for $(1 - \bar{v}\rho)$ determined above was used in the determination of the apparent molecular weight for the double-stranded oligomer.

A plot of log(conc) vs. r^2 for the double strands $rCA_7G + rCU_7G$ is shown in figure 3. The slope increases with increasing radius (increasing concentration), clearly indicating that aggregation is occurring. From the slopes of the curve at different values of log(conc), we can determine the apparent molecular weight as a function of concentration. The results are shown in Table VI.

5. Discussion

It is important that we know the nature of the molecules in solution, especially since oligo(A) + oligo(U) form triple-stranded structures in high salt concentrations (Pörschke, 1971). Also, we want to know whether the double helices are fraying significantly at the ends.

Figure 3. Log(concentration) vs. r^2 for the equilibrium ultracentrifugation of rCA_7G + rCU_7G in 1M NaCl at 3°C: ■, 295 nm; ●, 290 nm; ▲, 285 nm.

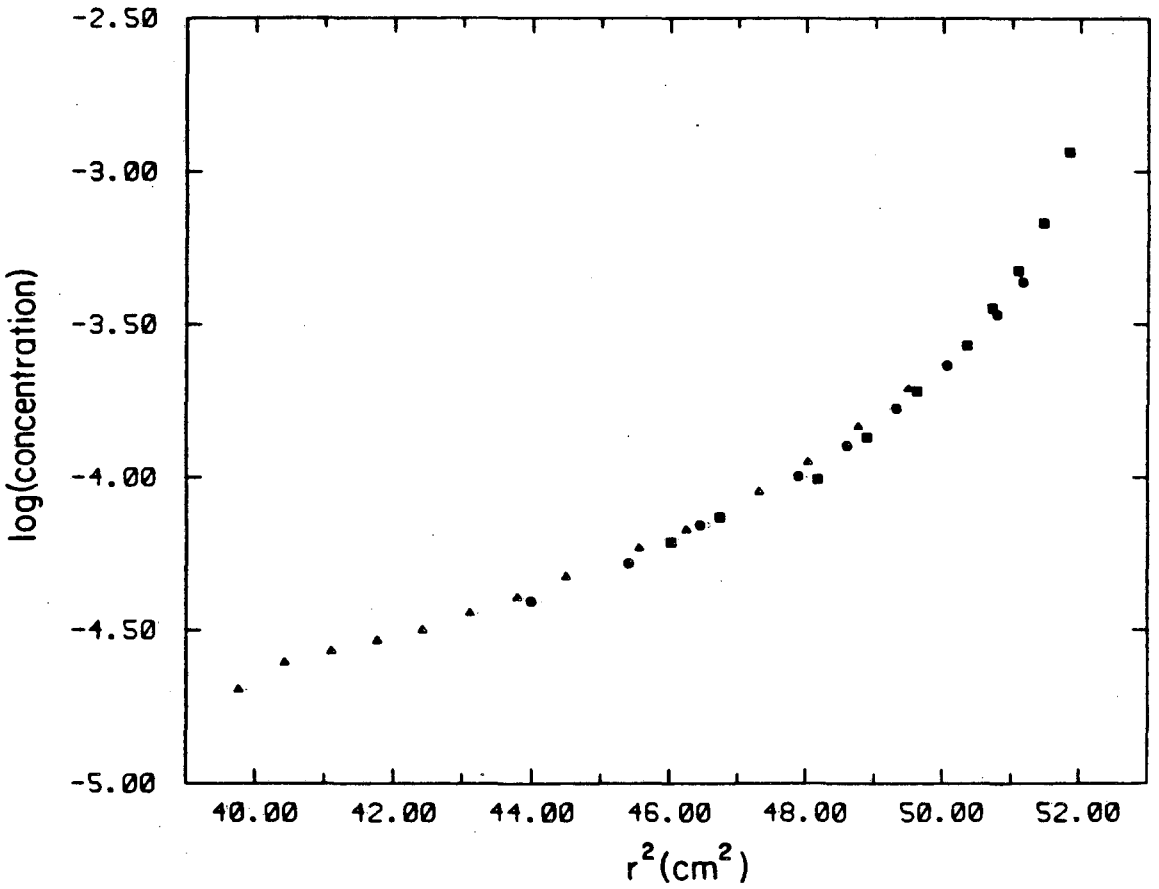

-2.50
-3.00
-3.50
-4.00
-4.50
-5.00
log(concentration)
40.00
42.00
44.00
46.00
48.00
50.00
52.00
r²(cm²)

TABLE VI

Apparent Molecular Weight for rCA_7G + rCU_7G in 1M NaCl

Conc. (μM)	Apparent Molecular Weight	Degree of Polymerization[a]
30	5070	.85
75	8250	1.6
150	13,000	2.6
410	24,100	4.7
910	42,900	8.5

[a]Apparent molecular weight divided by 5977, the molecular weight of the double helix.

Several pieces of evidence indicate that the oligomers are not forming triple strands. Job plots (Uhlenbeck et al., 1971) for rCA_5G + rCU_5G in 0.05 and 1M NaCl at several wavelengths show no significant concentrations of triple strands. Also, melting studies performed by mixing the strands at a ratio of 1 rCA_5G:2 rCU_5G gave results consistent with double-strand formation, even though formation of triple strands would be encouraged. Finally, NMR studies on dCA_5G + dCT_5G and rCA_5G + rCU_5G in 0.2M NaCl indicate the absence of any significant amount of triple-strand formation (Pardi et al., 1981).

The double-stranded helices do not melt differentially at the ends; that is, there is not a detectable concentration of double-stranded complexes in which terminal base pairs are broken. For dCA_5G + dCT_5G and rCA_5G + rCU_5G in 1.0M NaCl, the values for $\Delta H°$, calculated by equations 4 and 5, agree fairly well (Tables II and IV). Apparent broadening of the melting curves, due to partially melted duplexes leading to small $(df/dT)_{T_m}$ values, has been observed with oligomers such as A_nU_n, where the value of $\Delta H°$ obtained from equation 5 is about 30% lower than that from equation 4 (Martin et al., 1971). The fact that the two $\Delta H°$'s agree rather well in the present study might indicate that the single-strand to double-strand transition is essentially behaving in a two-state manner. From NMR spectroscopy on the nonexchangeable base protons, it was determined that the terminal base pairs melt at the same temperature, within experimental error, as the internal base pairs for dCA_5G + dCT_5G in

0.2M NaCl (Pardi et al, 1981), thus further indicating the double helix melts in a two-state manner.

From statistical considerations, one would expect the terminal base pairs would not fray significantly in these oligonucleotides, since they end in G•C base pairs. These have a greater stability constant than do A•T (or A•U) base pairs; so at temperatures low enough to form double strands, the stability constant for the terminal base pairs is large enough to ensure completely base-paired double helices.

A) Comparing dCA_5G + dCT_5G and rCA_5G + rCU_5G

Table II gives the thermodynamic parameters determined for dCA_5G + dCT_5G in 1M NaCl. The analysis was performed assuming both flat and sloping lower base lines. The values for $\Delta H°$, calculated using equation 4, were the same within experimental error: -49 kcal/mol assuming a flat base line, and -47 kcal/mol assuming a sloping base line. The calculated $\Delta H°$ did not depend on whether flat or sloping baselines were used for all of the oligonucleotides in this study. The same result was found for dG-C-G-C-G-C in 1.0M NaCl (Albergo et al., 1981). The significance of a sloping lower base line will be discussed later.

The value for $\Delta H°$ calculated from the shape of the melting curve, using equation 5, does depend on the lower base line. Since equation 5 assumes a two-state system, the calculation is fairly model-dependent, and one would expect changing the lower base line to affect the value of $\Delta H°$. In contrast, the calculation of $\Delta H°$ using the concentration dependence of the T_m (equation 4) is far

less model-dependent. In order to calculate the correct $\Delta H°$, you only need to select the melting temperatures such that they all correspond to the same position of the equilibrium. Thus, if the melting temperatures all corresponded to f = 0.4 instead of f = 0.5, the $\Delta H°$ calculated using equation 4 would not be affected.

The slope of the $1/T_m$ vs. $\log(C_a)$ plot is also more accurately determined than $(df/dT)_{T_m}$. For these reasons, I use values of $\Delta H°$, determined using flat base lines and equation 4, unless otherwise specified. This will allow us to make direct comparisons with earlier work.

A comparison between the thermodynamics of dCA_5G + dCT_5G and rCA_5G + rCU_5G in 0.2 and 1M NaCl are shown in Table III. The deoxyribo-oligomers are more stable than the ribo-oligomers. The increase in stabilization is enthalpic: -50 kcal/mol for dCA_5G + dCT_5G vs. -43 kcal/mol for rCA_5G + rCU_5G in 0.2M NaCl. In the case of both deoxyribo- and ribo-oligomers, increasing the salt concentration from 0.2 to 1M NaCl has little effect on the $\Delta H°$. The stabilities of both are increased somewhat, as indicated by more favorable entropies at the higher salt concentrations. The T_m's of the deoxyribo-oligomers increase less than the ribo-oligomers, 2.4 and 4.7°C, respectively.

B) Chain-Length Dependence of rCA_nG + rCU_nG Thermodynamics

The thermodynamic results for rCA_nG + rCU_nG for n = 5-7 in 1M NaCl are shown in Table IV. Also included are the values predicted by Borer et al. (1974).

As n becomes larger, the oligonucleotides would be expected to behave more like A_n + U_n, with a larger tendency toward triple-strand formation. A slope discontinuity at a ratio of 2U:1A strands is detectable in the Job plots of rCA_8G + rCU_8G in 1M NaCl at low temperatures (1°C). The shapes of the melting curves of these longer oligonucleotides were noticeably different from those for n = 5-7, as seen by a significant curvature in the low temperature base line. Finally, the ΔH° calculated from the concentration dependence was anomalously high, whereas the apparent melting temperatures were too low, when compared to the shorter oligonucleotides. As a result, rCA_8G + rCU_8G and longer oligonucleotides in this series have not been included in this analysis.

A small amount of triple-helix formation in mixtures of rCA_7G + rCU_7G may be indicated by the discrepancy between the values of ΔH° calculated from equations 4 and 5, but the amount of triple helix was too small to detect in Job plots or by increased curvature of the melting curves at low temperature.

From the chain-length dependence of the thermodynamics, we can calculate the contribution of an internal $\begin{smallmatrix} A-A \\ \vdots \\ U-U \end{smallmatrix}$ base pair to the stability of the double helix. Using the differences for n = 5 and n = 6, we obtain ΔH° (addition of $\begin{smallmatrix} A-A \\ \vdots \\ U-U \end{smallmatrix}$ base pair stack) = -9 kcal/mol, ΔS° = -28 e.u., and ΔG° = -0.9 kcal/mol. The comparable numbers obtained by Borer et al. (1974) are -8.2 kcal/mol, -24 e.u. and -1.2 kcal/mol, respectively. This agreement is reasonable, as the present study compares only two oligonucleotides, whereas the Borer study used data for several oligonucleotides.

From Table IV, we see that the predicted stabilities of the oligonucleotides by Borer et al. (1974) are lower than are found experimentally, although the calculations improve with increasing chain length. The stability parameters were obtained primarily from oligomers of the type A_nXYU_n where XY was A•U, C•G, or G•C. As mentioned earlier, double helices with melted ends seem to contribute significantly in oligomers of the type A_nU_n. From statistical mechanical calculations, it was postulated that the terminal base pairs in A_nU_n were less stable than the internal base pairs (Levine, 1974). The procedure used by Borer et al. (1974) would underestimate the stability of internal $\begin{smallmatrix}A-A\\ \bullet\ \ \bullet\\ U-U\end{smallmatrix}$ base-pair stacks. This is because they essentially determine the average stability of all the $\begin{smallmatrix}A-A\\ \bullet\ \ \bullet\\ U-U\end{smallmatrix}$ stacks and assign this value to the (more stable) internal $\begin{smallmatrix}A-A\\ \bullet\ \ \bullet\\ U-U\end{smallmatrix}$ base pairs. Since the oligomers in the present study do not terminate in A•U base pairs, the result is to underestimate the stability of these double helices.

Another possible reason for underestimating the double-helix stability is that none of the oligonucleotides used in the Borer study contained the $\begin{smallmatrix}C-A\\ \bullet\ \ \bullet\\ G-U\end{smallmatrix}$ stacking interaction. They report the average stability for the $\begin{smallmatrix}C-U\\ \bullet\ \ \bullet\\ G-A\end{smallmatrix}$, $\begin{smallmatrix}G-A\\ \bullet\ \ \bullet\\ C-U\end{smallmatrix}$, and $\begin{smallmatrix}G-U\\ \bullet\ \ \bullet\\ C-A\end{smallmatrix}$ stacks. If the $\begin{smallmatrix}C-A\\ \bullet\ \ \bullet\\ G-U\end{smallmatrix}$ stack is more stable than this average value, the calculated stability will be too low.

C) <u>Low-Temperature Behavior of the Double Strands</u>

The slope of the melting curves at low temperatures is surprisingly large; in fact, it is about as large as the slope of the single-strand melting. Since the low-temperature base lines of ab-

sorbance melting curves for polynucleotides are flat, the temperature dependence of the absorbances of oligonucleotide double helices is thought to be an end effect. In the case of oligonucleotides terminating in A•U base pairs, the slope can at least in part be attributed to changes in hypochromicity due to changes in the degree of base-pair melting at the ends, but I do not think this occurs for the oligonucleotides studied here. It is possible that the conformation of the ends is not as rigid as it is in the interior of the double helix. Evidence supporting conformational changes comes from NMR spectra. Although the oligonucleotides in this study are not stable enough to measure the low-temperature base lines from NMR spectra, for the more stable double strand formed by the oligomer dG-G-A-A-T-T-C-C, some of the base protons on each of the base pairs (including the interior ones) continue to exhibit a change in chemical shift with temperature down to 0° C (Patel & Canuel, 1979). We cannot say how large this change in conformation is, nor why it would occur only near the ends of double helices.

D) Aggregation of Double Strands

The problem of aggregation giving rise to the concentration-dependent hypochromicities was presented earlier. Aggregation effects are also observed in NMR studies on dG-G-A-A-T-T-C-C (Patel & Canuel, 1979) and rA-A-G-C-U-U (Ebrer et al., 1975), as evidenced by excessive line widths of the nonexchangeable protons. The NMR spectra were run at a much higher concentration (10mM) than was used in this study.

To determine if the problem occurred optically for the single strands, the oligomers rCA_7G and rCU_7G were melted separately in 1M NaCl over a broad concentration range. For rCA_7G, the range was from 7.9 to 720μM; for rCU_7G, 8.7 to 800μM. The relative melting curves superimposed over this range, indicating that the effect was not occurring in the single strands. This also rules out the possibility that the effect of hypochromicities might be due to instrumental artifacts.

In an attempt to understand the nature of the double-strand aggregation, I used equilibrium ultracentrifugation. This allows one to determine directly the molecular weight of a molecule, and the extent to which the molecules are aggregating.

The results of the ultracentrifugation of the single-strand rA_7G show no aggregation. The plot of log(conc) vs. r^2 was a straight line from 18 to 490μM, indicating a constant apparent molecular weight vs. concentration.

The results of the ultracentrifugation for rCA_7G + rCU_7G are shown in Table VI. As can clearly be seen, the oligomers are aggregating to a significant extent over the concentration ranges commonly used for optical studies. The aggregation is significant even down to 75μM.

The third column in Table VI shows the ratio of the apparent (weight average) molecular weight to the molecular weight calculated for the rCA_7G + rCU_7G double helix. At the low concentration range, 30μM, the ratio is 0.85. The double-strand equilibrium constant at this temperature, determined from the optical measurements, suggests

that the oligomers are greater than 99% in double strands at this concentration. Thus, there is some discrepancy in the data. There are several possibilities that explain this. The errors in determining the slopes are rather large. The calculation of the apparent molecular weight is very sensitive to the value of $(1 - \bar{v})$. The value used was that obtained from rA_7G. There might be some error in this assumption, although the specific volumes of double-stranded and single-stranded DNA are nearly the same. Also, degradation of the oligomer would result in a lowering of the apparent molecular weight. The experiment took 4-1/2 days, so degradation is a definite possibility.

The simplest model to describe the aggregation would involve assigning the same equilibrium constant, K_p, for the addition of each double strand (H) to an aggregate of n double strands:

$$H_n + H \rightleftharpoons H_{n+1}, \qquad K_p = (H_{n+1})/(H)(H_n) \tag{7}$$

The weight-average molecular weight is given by the expression:

$$M_w = \frac{M \sum_{n=1}^{\infty} (n^2)(S)^{n-1}}{\sum_{n=1}^{\infty} (n)(S)^{n-1}} = \frac{M(1 + S)}{(1 - S)} \tag{8}$$

$$S = K_p(H) = \frac{1 + 2K_pC_t - \sqrt{1 + 4K_pC_t}}{2K_pC_t} \tag{9}$$

The best fit to the ultracentrifuge data is $K_p = 12{,}000 \pm 6000$, which corresponds to a free energy of aggregation of -5.2 ± 0.4 kcal/mol at 3°C. The fit to the data is not especially good. This

could be due in part to degradation of the strands, as mentioned earlier. This would result in the experimental data being too low at low oligomer concentrations, as was observed. It is also possible that the model is not valid; the equilibrium constant may differ with the number of double helices in the aggregate. Thus, the magnitude of the aggregation energy should be considered only as an approximate value. However, it allows us to estimate the possible effects aggregation might have in the determination of the thermodynamic values for double-strand formation.

Of course, from these data alone we cannot determine what type of aggregation is occurring. One reasonable possibility is end-to-end aggregation, where the G•C base pair of one helix stacks on that of another helix, forming a sort of double-helix polymer. This could have a favorable enthalpy from the stacking interactions. It would also be accompanied by a hypochromicity, giving rise to the observed concentration-dependent hypochromicities.

An important quesion is whether aggregation affects significantly the thermodynamic parameters we measure. To try to answer this, I tested a model composed of the single-strand to double-strand transition linked to the aggregation of double strands. The four parameters for such a model are the $\Delta H°$ and $\Delta S°$ for the double-helix formation and for aggregation. The experimental parameters I fit are the melting temperatures at different concentrations.

We can simplify the model by assuming that the melting curves directly measure α, the fraction of strands that are single stranded. The T_m is the temperature where $\alpha = 0.5$. I also ignore the

concentration-dependent hypochromicity by assigning a separate low-temperature base line for each curve:

$$\alpha = 1 - f = (A)/C_a \tag{10}$$

where (A) is the concentration of single strands, and C_a is the total concentration of each of the strands. Since $C_a = (A) + (H) + 2(H_2) + 3(H_3) + \cdots + n(H_n) + \cdots$, this equation can be rewritten:

$$\alpha = \frac{(A)}{(A) + (H) + (2H_2) + 3(H_3) + \cdots + n(H_n) + \cdots} \tag{11}$$

I used a value of -5.2 kcal at 3°C for the free energy of aggregation, the value determined from the ultracentrifugation. Since we do not know how to distribute this between $\Delta H°$ and $\Delta S°$, the calculations were performed by allowing the stabilization to be entirely enthalpic or entirely entropic. The results of these model calculations can then be used as a guide to determine the potential effect of the aggregation.

The fits were determined using two sets of aggregation parameters, $\Delta H°_p = -5.2$ kcal, $\Delta S°_p = 0$, and $\Delta H°_p = 0$, $\Delta S°_p = -5.2/276$ kcal/deg. The $\Delta H°$ and $\Delta S°$ for double-helix formation were chosen to best fit the experimental T_m's. The resulting values are shown in Table VII. Calculated plots of $1/T_m$ vs. $\log(C_a)$ using both sets of $\Delta H°$ and $\Delta S°$ were essentially identical.

The effect is calculated to be quite large. Assuming the aggregation is totally enthalpy-stabilized, the $\Delta H°$ for double-helix formation is calculated to be -82 kcal/mol, as compared to the value of -63 kcal/mol in the absence of aggregation. The aggregation

TABLE VII

Effect of Aggregation on Calculated Double-Helix Stability for $rCA_7G + rCU_7G$

$\Delta H^\circ_{aggreg.}$ (kcal/mol)	$\Delta S^\circ_{aggreg.}$ (e.u.)	ΔH°_{helix} (kcal/mol)	ΔS°_{helix} (e.u.)
0	0	-63	-187
-5.2	0	-82	-250
0	18.8	-94	-290

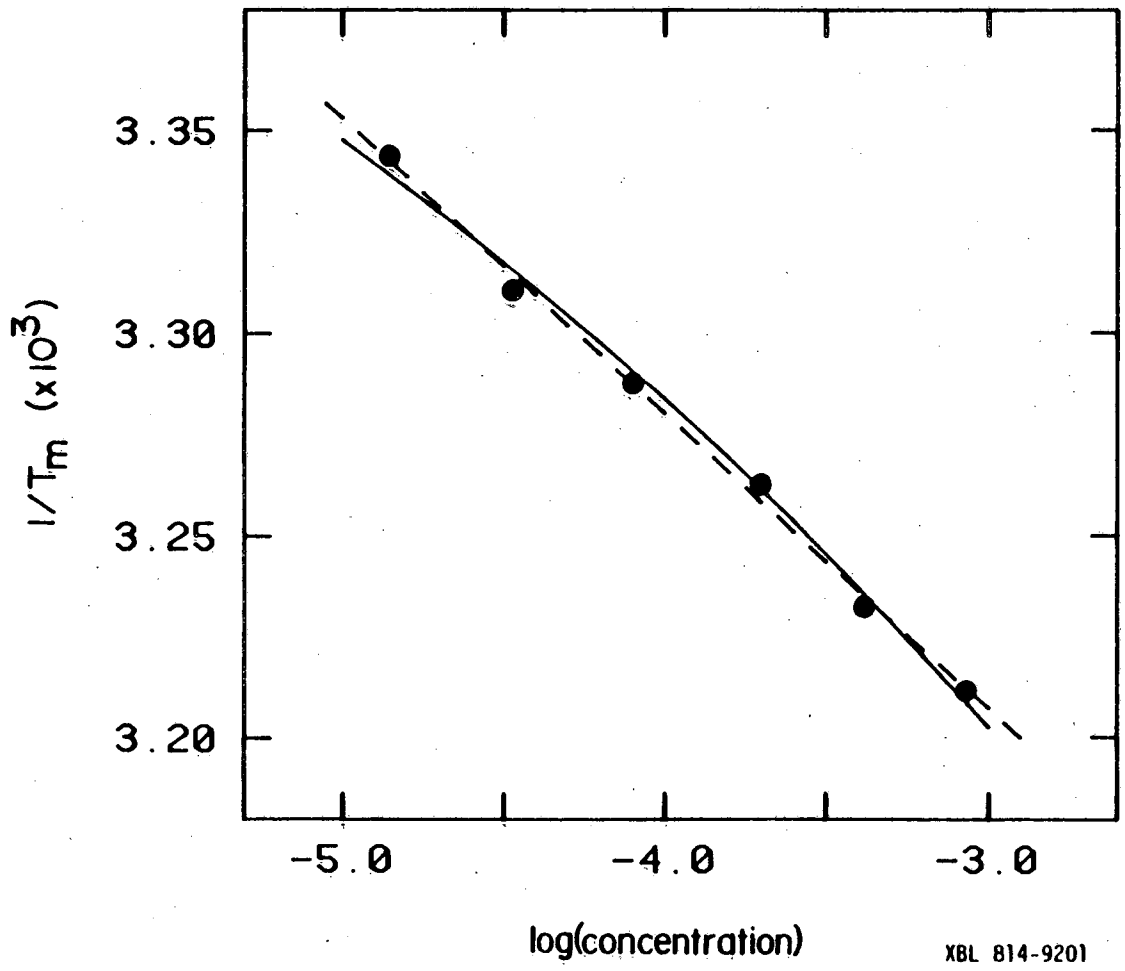
model tends to increase the value calculated for the double-helix formation. This is because aggregation tends to increase the T_m at higher concentrations. This results in a steeper slope in the $1/T_m$ vs. log(conc) plot, and a lower apparent $\Delta H°$. To compensate for this, the $\Delta H°$ for double-helix formation must be made larger to decrease the slope back to that of the experimental curve. This result indicates that spectroscopic determinations will underestimate the $\Delta H°$ if aggregation occurs.

The results, assuming enthalpic aggregation stabilization, are shown in figure 4, along with the experimental T_m's. The aggregation model calculations exhibit a curvature. Since the slope of such a plot is related to the $\Delta H°$, the $\Delta H°$ for double-strand formation is concentration-dependent. This is an expected result, as the double strands are, of course, stabilized relative to the single strands when aggregates are formed. At higher concentrations, the extent of aggregation, and hence the double-strand stability, increases.

There is no noticeable curvature in any of the experimental $1/T_m$ vs. $\log(C_a)$ plots I measure. The model might overestimate the effect of aggregation on the double-strand formation. However, the effect of the aggregation model on the thermodynamics is large enough so that a significant effect may be present and not cause any apparent curvature in the $1/T_m$ vs. $\log(C_a)$ plots.

The extent of the effect of aggregation on the double-strand thermodynamics is not apparent from spectroscopic analysis, aside from the effect of the concentration-dependent hypochromicity.

Figure 4. The plot of $1/T_m$ vs. log(concentration) for rCA_7G + rCU_7G in 1M NaCl: ●, experimental data; ———, best fit from simple aggregation model, assuming enthalpic stabilization of the aggregation; - - - - -, straight line fit assuming no aggregation. See Table VII for the corresponding thermodynamic values.


3.35
3.30
3.25
3.20
$1/T_m$ ($\times 10^3$)
-5.0
-4.0
-3.0
log(concentration)
XBL 814-9201

Unfortunately, the effect would also be difficult to detect in microcalorimetric measurements, which directly measure the amount of heat absorbed by the molecules between the low and high temperatures. If the amount of aggregation is large throughout the melting transition, the transition may be described as

$$\text{double strands (aggregated)} \rightleftharpoons \text{single strands}$$

The aggregation is concentration-dependent; thus, in principle, the effect could be determined from the concentration dependence of the enthalpy. However, in scanning differential microcalorimetry on oligonucleotides, the experimenters usually work at concentrations in the range of 1mM strand concentration and vary the concentration by a factor of about 2. Since the extent of aggregation is predicted to be large in this concentration range, the concentration effect may be too small to observe.

We have seen that the effect of aggregation on the determination of single-strand to double-strand energetics can be significant. A direct estimation of the extent of aggregation is obtained only from the equilibrium ultracentrifuge analysis. Thus, when studying oligonucleotides at concentrations in the range of 100-1000μM, care must be taken to be sure that aggregation will not significantly interfere with the results. If the aggregation is comparable in all the oligonucleotide studies, the comparison of energetics from one set of oligomers to another set of closely related oligomers should be basically valid.

E) Mismatched Double Helices

A bulged structure must occur whenever a frame-shift mutation occurs, and hence the stability of such a structure might be important in understanding the molecular mechanism of frame-shift mutagenesis. We can potentially form bulges by mixing oligonucleotides with different numbers of A and U bases, for example, rCA_6G + rCU_5G. The stability of a perturbed double helix can thus be compared to that of the normal double helix.

The best estimate to date of the destabilization of a bulge is from the work of Fink and Crothers (1972b). They studied poly(U) + poly(A,A*), where A* represents adenine residues modified by monoperphthalic acid to adenine-N-1-oxide. This modified base forms a bulge. By monitoring the T_m at different mole fractions of modified adenine, they were able to determine that a bulge of one nucleotide destabilizes the double helix by 2.8 kcal/mol and the triple strands by 2.3 kcal/mol. From CD measurements, they determined that the modified adenine was probably stacked in the single-strand helix but was bulged out of the double helix (Fink & Crothers, 1972a).

I studied the thermodynamics of the mismatched double helices formed by the three mixtures: rCA_6G + rCU_5G, rCA_5G + rCU_6G, and dCA_5G + dCT_6G, in 0.2M NaCl. The results of these studies are given in Table V, along with the normal double helices for comparison. I also studied the stabilities of the mixtures rCA_7 + rCU_5G and rCA_5G + rCU_7 in order to help determine the structure of the "bulges" in the mismatched double helices. The structures of the "bulged" double helices are ambiguous. In contrast, the structure of a

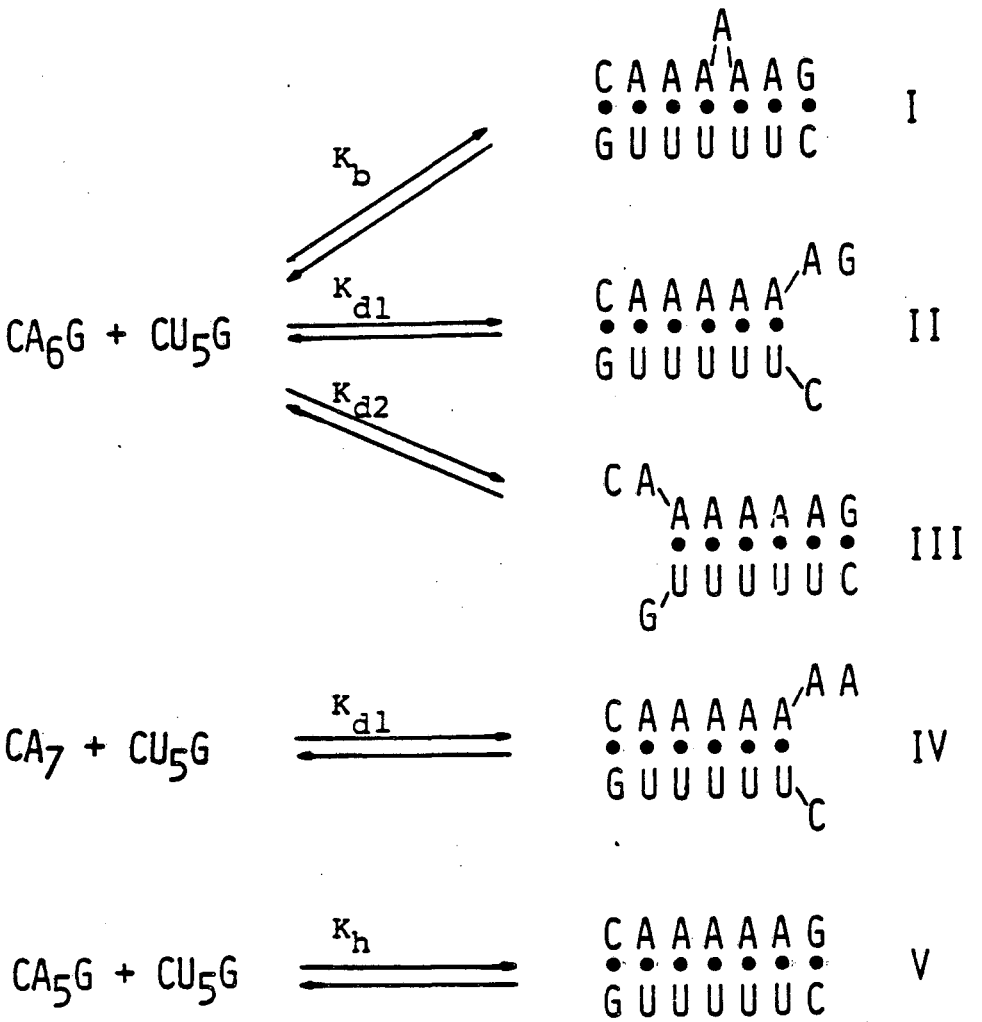
double helix like $rCA_7 \cdot rCU_5G$ is constrained to have a dangling end (see below).

The values for $\Delta G°$ are reported at 10° C, since the mismatched double helices are not stable enough to study at 25°C. An extra rA or rU destabilizes the double helix by about 1.2 and 1.0 kcal/mol, respectively. The extra dT destabilizes by about 1.8 kcal/mol. In this instance, the perturbed structures are more destabilizing for deoxyribo-oligonucleotides than for ribo-oligonucleotides; however, it would be inappropriate to generalize from such limited information.

We can use the comparison between the stabilities of the double helices formed by the mixtures rCA_6G + rCU_5G with rCA_7 + rCU_5G. The possible structures for the former are represented by structures (I), (II), and (III) in figure 5; the most likely structure for the latter is structure (IV). Equilibrium constants for each of these possible structures are defined in figure 5.

Structures (II) and (IV) both have dangling ends containing an adenine followed by another purine, either adenine or guanine. It was found that a dangling end in oligomers of the kind $A_{n+1}U_n$ stabilized the double helix relative to A_nU_n. A second dangling base, $A_{n+2}U_n$, stabilized even more, but the effect was less than the first (Martin et al., 1971). Since the second dangling base contributes only a little, and in both structures (II) and (IV) the second dangling base is a purine, they have been assigned the same equilibrium constant, K_{d1}.

Figure 5. Plausible structures and equilibrium constants for the double strands formed by mixing rCA_6G + rCU_5G. Structure (I) should be considered as schematic only - the bulged bases could be in several positions, and could be inside or outside the double helix.


CA6G + CU5G
Kb
Kd1
Kd2
CA7 + CU5G
Kd1
CA5G + CU5G
Kh
I
II
III
IV
V

The other possible dangling end structure, (III), is assigned a different equilibrium constant, K_{d2}, since it is of a different nature. The bulged structure, (I), is assigned an equilibrium constant, K_b, which is an apparent equilibrium constant for all possible bulged structures. Structure (I) should be considered a schematic structure only. We cannot say whether the bulged base is inside the double helix or pushed out into solution. In these oligonucleotides, there exists a number of possible sites for the bulged A. Since we do not know what the structure of the bulge is, or what the stability is at different sites, we can only assign an apparent equilibrium constant.

The equilibrium constant for structure (IV), K_{d1}, can be determined experimentally, as can be the sum of the equilibrium constants $K_b + K_{d1} + K_{d2}$. [In this case, we measure the apparent equilibrium constant for formation of all double-stranded species K_{app}(double helix) = $K_b + K_{d1} + K_{d2}$.] We want to make a comparison with the corresponding normal double helix, structure (V), which is assigned the equilibrium constant K_h.

The data in Table V allow us to calculate the ratio of the apparent equilibrium constants for double-strand formation of structures (I), (II), and (III) to the equilibrium constant for structure (IV), since $\Delta G^{\circ}_{app} = -RT \cdot \ln(K_{app})$:

$$\frac{(K_b + K_{d1} + K_{d2})}{K_{d1}} = \frac{\exp(\Delta G^0_{app}/RT)}{\exp(-\Delta G^0/RT)} = \exp(\Delta G^0 - \Delta G^0_{app})/RT \qquad (12)$$

Table V tells us that the "bulged" structures (I), (II), and (III) are more stable than the dangling-end structure (IV) by 0.4 kcal/mol at 10° C. I estimate the error in this to be at most 0.2 kcal/mol. Using the extreme value of 0.6 kcal/mol, the value of the ratio in equation 12 is 2.9. This allows us to calculate an estimate for the lower bound for the amount of destabilization produced by a bulged adenine.

We do not know the relationship between K_{d1} and K_{d2}. If we assume $K_{d1} = K_{d2}$, then $K_b = 0.91K_{d1}$. This assumption is probably not strictly valid, since the stability of the dangling end probably depends on sequence. However, if we set $K_{d2} = 0.5K_{d1}$, the resulting free energy changes by only 0.2 kcal/mol.

Keeping in mind that 0.6 kcal/mol was an upper limit, the corresponding lower limit on the bulge destabilization can be calculated by comparing $-RT \cdot \ell n(K_b)$ with $-RT \cdot \ell n(K_{d1})$ and $-RT \cdot \ell n(K_h)$. The last two values are directly measured. Using $K_b = 0.9K_{d1}$, the $\Delta G°$ for forming structure (I) is calculated to be -4.9 kcal/mol. Comparing these $\Delta G°$'s of double-strand formation with the value for the normal double helix, -6.5 kcal/mol, we see that the bulge destabilizes the double helix by at least 1.6 kcal/mol.

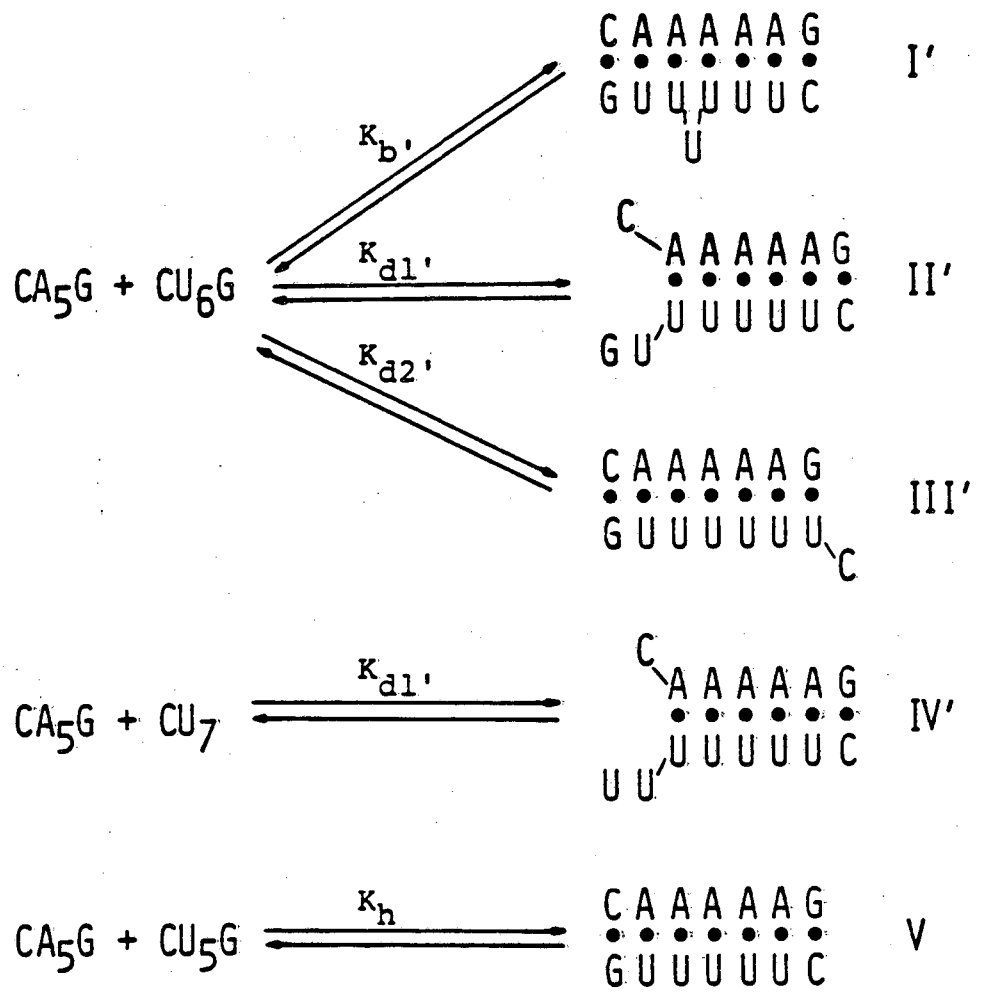
We must view these numbers with some caution, however. If we set $K_{d1} = K_{d2}$, and assume $K_b << K_{d1}$, we would expect the mixture of rCA_6G + rCU_5G to be $RT \cdot \ell n(2)$ more stable than rCA_7 + rCU_5G on the basis of statistical considerations. (The former can form dangling ends on either end; the latter on only one.) $RT \cdot \ell n(2) = 0.39$ kcal/mol at 10° C. The difference we measured was 0.4 ± 0.2

kcal/mol. Thus, these data are consistent with no significant amount of bulged species, and hence the destabilizing effect of the bulge could be much greater than we just calculated. Also, since the bulge could exist in a number of places, the stability of any one of these structures would be lower than we calculate because of statistical considerations.

I also investigated the effect of an extra rU or dT base by studying the thermodynamics of rCA_5G + rCU_6G, rCA_5G + rCU_7, and dCA_5G + dCT_6G. These results are included in Table V. The possible structures for the ribo-oligomers are shown in figure 6. In this case, we cannot follow a similar argument as above, because now the dangling-end structure (III') can form a G•U base pair; thus, structure (III') is expected to be more stable than structure (II'). The difference in stability between the "bulged" structures (I'), (II'), and (III') and the dangling-end structure (IV') should now be larger than we measured above. The observed value for a mismached rU was 0.6 ± 0.2 kcal/mol.

In summary, I attempted to study the destabilizing effect of a bulge on double-helix formation. The system I chose has the ability to form either a bulge or a dangling end. These results indicate that the dangling end is more stable than the bulge, which indicates that a bulge destabilizes the double helix more than other perturbations. To favor the formation of a bulge, it would be better to use a molecule such as dC-A-A-A-C-A-A-A-G, which when mixed with dCT_6G should form structures with a bulged C.

Figure 6. Plausible structures and equilibrium constants for the double strands formed by mixing rCA_5G + rCU_6G.


CA5G + CU6G
Kb'
Kd1'
Kd2'
C A A A A A G
G U U U U U C
U
I'
C
A A A A A G
U U U U U C
G U
II'
C A A A A A G
G U U U U U U
C
III'
CA5G + CU7
Kd1'
C
A A A A A G
U U U U U C
U U
IV'
CA5G + CU5G
Kh
C A A A A A G
G U U U U U C
V

CHAPTER III

Kinetics of Oligonucleotide Double-Strand Formation

1. Synopsis

The kinetics of double-strand formation were measured using temperature-jump kinetics techniques for the DNA oligonucleotides dCA_5G + dCT_5G, the analogous RNA oligonucleotides rCA_5G + rCU_5G, and the hybrid rCA_5G + dCT_5G. The DNA oligonucleotides have a faster rate of recombination and a slower rate of dissociation at 12.0° C than the RNA oligonucleotides; the hybrid has about the same recombination rate and a slightly faster dissociation rate than the RNA oligonucleotides. The activation energy for recombination for the DNA and RNA oligonucleotides are both near zero kcal/mol. The difference in dissociation and recombination activation energies are consistent with the thermodynamic results obtained earlier.

The relaxation process is composed of two exponential components for the RNA and hybrid oligonucleotides at temperatures of 12.0°C and lower. One exponential component is observed for these oligonucleotides above 12.0° C and for the DNA oligonucleotides at all temperatures.

2. Introduction

The kinetics of double-strand formation for a number of ribo-oligonucleotides have been studied using the temperature-jump technique (Craig et al., 1971; Podder, 1971; Pörschke & Eigen, 1971; Pörschke et al., 1973; Ravetch et al., 1974; Breslauer & Bina-Stein,

1977). Studies on deoxyribo-oligonucleotides are more limited (Drobnies, 1979; Freier et al., 1982). One RNA-DNA hybrid oligonucleotide double helix has been studied (Hoggett & Maass, 1971). However, no thorough study has been made comparing the kinetics of the same sequence for RNA, DNA and hybrid oligonucleotides. Such knowledge is important in determining how chemical differences between DNA and RNA manifest themselves in the dynamic properties of the double strands.

In this report I present results of temperature-jump kinetics studies of double-strand formation of the RNA oligonucleotides rCA_5G + rCU_5G, the analogous DNA oligonucleotides dCA_5G + dCT_5G, and a hybrid double helix composed of rCA_5G + dCT_5G. The other hybrid dCA_5G + rCU_5G does not form stable double strands (Martin & Tinoco, Jr., 1980). The results of this chapter have been published (Nelson & Tinoco, Jr., 1982).

3. Methods and Materials

A) Sample Preparation

The synthesis and characterization of the oligonucleotides have already been discussed in Chapter II. Concentrations of all mixtures capable of forming double strands were determined by measuring absorbances at 50°C, where the strands exist as single strands. The oligonucleotides are subject to degradation due to the high intensity of the UV lamp used in temperature-jump experiments. Therefore, concentrations were determined before and after the temperature-jump measurements, and were always within 3%. Additionally, absorbances were recorded at 0°C before and after, to check the

hypochromicity. The hypochromicity, $(A_{50°}-A_{0°})/A_{50°}$, generally decreased 0.5 to 2% after the measurements. Samples with larger discrepancies were not used in the analysis.

The buffer used throughout these studies, 1M NaCl, 0.01M sodium phosphate buffer, pH=7, 0.1mM EDTA, was filtered using Uni-Pore Polycarbonate Membranes (Bio-Rad) with a pore size of 3 μm. Samples were degassed prior to temperature-jump measurements by purging the buffer with helium for 3 to 5 minutes before mixing with a small volume of a stock solution of oligonucleotides.

Measurements were taken at the mercury line near 267 nm for the ribo- and deoxyribo-oligonucleotides. Due to problems developing from the sample cell, the hybrid oligonucleotides were more conveniently measured at 250 nm (see Appendix B). The temperature was measured to an accuracy of 0.2°C using a copper-constantin digital thermocouple (Fluke model 2100A) in contact with the upper electrode.

B) The Temperature-Jump Instrument

The temperature-jump instrument was manufactured by DiaLog (West Germany) and has been previously described (Rigler et al., 1974; Drobnies, 1979). The data were collected digitally by a Biomation model 805 transient recorder, using 2048 data points. A PET micro-computer was used to transfer the data from the Biomation to a VAX 11/780 computer, where the data were analyzed.

The program DISCRETE written by S. W. Provencher (Provencher, 1976a,b) was used to analyze the data. The program determines, from the data, what number of exponential components results in the best

fit. The program was allowed to search for three exponential components. Two exponentials were reproducibly found for the ribo- and hybrid oligonucleotides at 6.7°C and 12.0°C. One exponential was found above this temperature and for the deoxyribo-oligonucleotides.

Any further description of the set-up and use of the instrument will be deferred until Appendix B, which describes the instrumental and experimental procedures in detail.

C) Analysis of the Temperature-Jump Kinetic Data

For the reaction scheme given by equation 1, k_r corresponds to

$$A + B \underset{k_d}{\overset{k_r}{\rightleftharpoons}} AB \qquad (1)$$

the recombination rate constant for the formation of double strands AB from the non-selfcomplementary strands A and B, whereas k_d corresponds to the dissociation rate constant. When the system at equilibrium is perturbed by producing a fast jump in temperature, the system relaxes exponentially to its equilibrium at the final temperature according to equation 2.

$$1/\tau = 2k_r C_s + k_d \qquad (2)$$

C_s is the equilibrium concentration of A at the final temperature, and the total concentrations of A and B are assumed to be equal. τ is the relaxation time characterizing the process.

I have considered three methods to determine k_r and k_d by measuring $1/\tau$ at several concentrations of strands. The first is to calculate C_s at each total concentration using the equilibrium con-

stant determined thermodynamically, and plotting $1/\tau$ vs. C_s. The slope gives k_r; the intercept gives k_d.

The second method is to use the kinetic equilibrium constant determined by the first method, $K_{kinetic} = k_r/k_d$, and calculate new values for C_s, k_r and k_d. Iteration will result in the determination of a kinetically determined equilibrium constant consistent with the data.

The third method is to square equation 2 and eliminate C_s from the equation to get equation 3:

$$1/\tau^2 = 4k_r k_d C_{tot} + k_d^2 \qquad (3)$$

A plot of $1/\tau^2$ vs. C_{tot}, the total concentration, gives an intercept $= k_d^2$ and slope $= 4k_r k_d$.

The second and third analysis methods have the advantage that previous thermodynamic knowledge is not used in the final determination of the kinetics. A comparison of the results of the kinetics and thermodynamics allows a check of the consistency of the reaction scheme. I used all three methods of analysis.

4. Results

Figure 1 shows a trace of voltage vs. time for the ribo-oligonucleotides rCA_5G + rCU_5G at 14.2μM concentration and a final temperature of 6.7°C. The size of the temperature jump was 1.8°C; the initial signal was 5 volts. Figure 2 shows a semi-log plot of the trace in figure 1. The small amount of curvature indicates that there are two exponential components. The two lines indicate the best fits assuming one exponential (- - - -) and two exponentials

Figure 1. A trace of a temperature-jump experiment on 14.2μM rCA_5G + rCU_5G in 1M NaCl, 0.01M phosphate buffer, pH=7, 0.1mM EDTA, at a final temperature of 6.7°C. The temperature jump was 1.8°C; λ = 267nm. The initial signal was 5 volts. The first 5 msec shows the signal prior to the temperature jump.

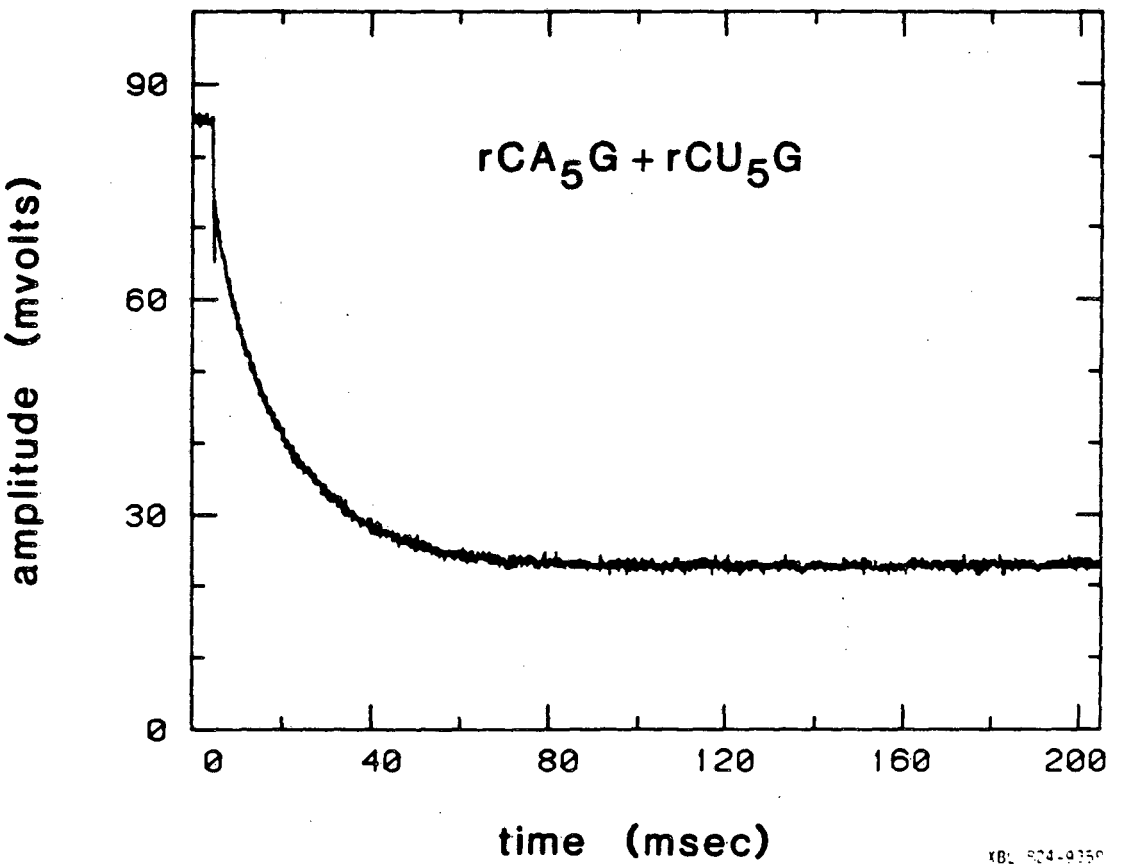
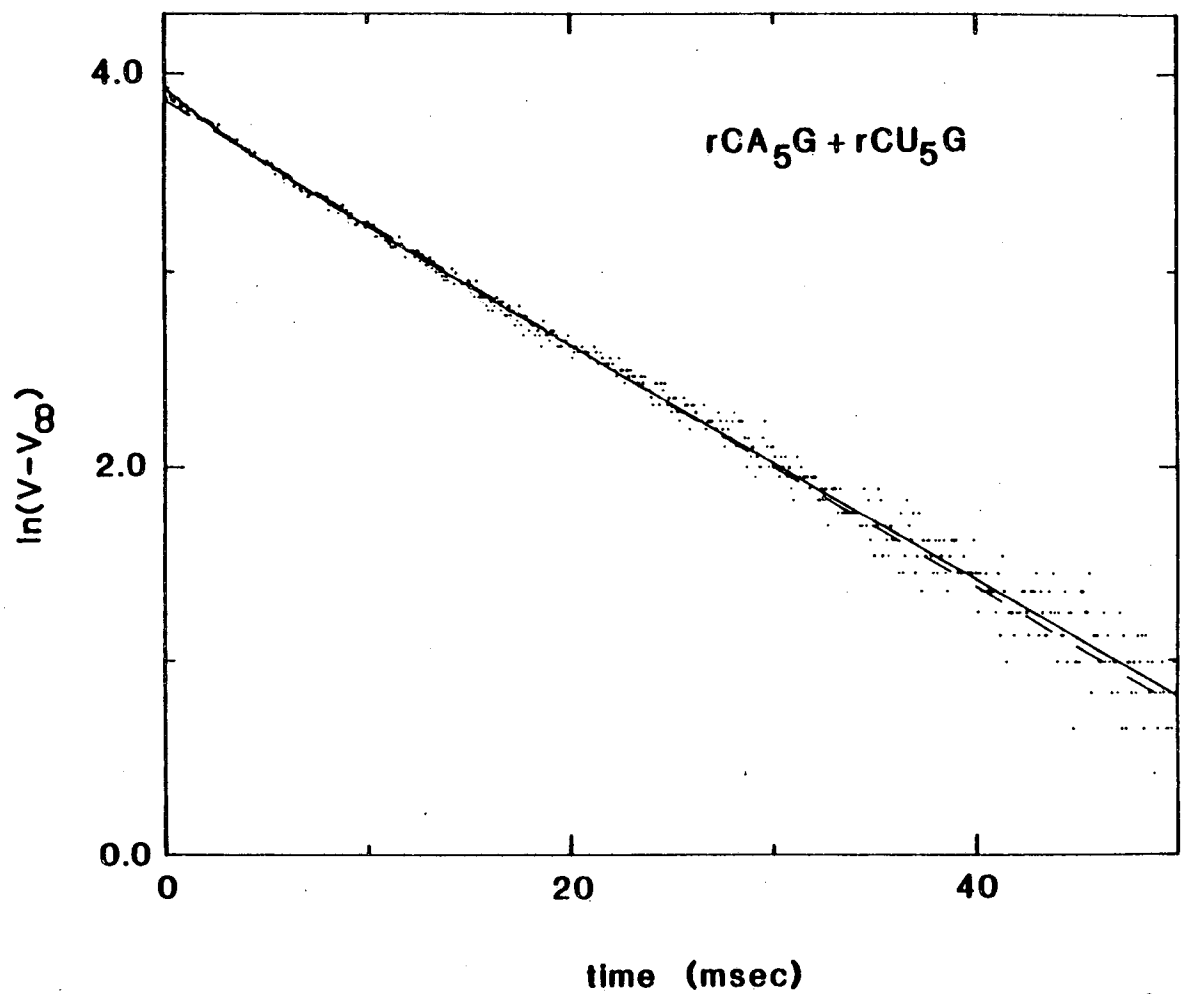

rCA5G + rCU5G
amplitude (mvolts)
90
60
30
0
0
40
80
120
160
200
time (msec)
XBL 904-9359

Figure 2. Semi-log plot of the trace in Figure 1. The amplitude at long times, V_∞, was determined from the two component fit of program DISCRETE. The two lines show the two component fit (———) (τ_1 = 16.6msec, a_1 = 45.8mV; τ_2 = 2.87msec, a_2 = 4.92mV) and the one component fit (- - - -) (τ_1 = 16.1msec, a_1 = 47.8mV).



XBL 824-9360

(———). Due to the small amplitude of the second component, the relaxation time for the predominant relaxation for both fits are within 3%. Since the second component is seen only at low temperature, the one-component fit is used to characterize the bimolecular reaction. Using the two exponentials, when observed, does not change the results. The second component is seen only for the ribo- and hybrid oligonucleotides at temperatures at or below 12°C. It is not observed at higher temperatures or for the deoxyribo-oligonucleotides at all temperatures.

The relaxation times for the predominant relaxation were averages of 5 to 10 relaxations for each sample. The standard deviation of the relaxation times was generally 2-6%. These standard deviations were used in linear regression to give the errors for the recombination and dissociation rate constants reported. Using these errors in the linear regression to determine the activation energies yielded the errors reported for the activation energies.

The temperature range of the instrument is limited by the sample cell to initial temperatures no lower than 5°C. The upper end of the temperature range is reached when the concentration of double strands becomes too low and the signal becomes too small to measure. Thus, the temperature ranges studied were 6.7 to 21.1°C for the ribo-oligonucleotides and 8.0 to 24.8°C for the deoxyribo-oligonucleotides.

Table I lists the rate constants and activation energies determined for the ribo-oligonucleotides rCA_5G + rCU_5G at four tempera-

tures, using the three methods of analysis discussed above. The first uses thermodynamic equilibrium constants to calculate the equilibrium single strand concentrations, C_s, in equation 2; the recombination rate constant, k_r, and the dissociation rate constant, k_d, are determined from the slope and intercept of a plot of $1/\tau$ vs. C_s. This procedure has the disadvantage of requiring previous thermodynamic knowledge, which might bias the kinetic results. The second method avoids this problem by calculating the equilibrium constant determined from kinetics, $K_{kinetic} = k_r/k_d$. This equilibrium constant is then used to calculate the next set of rate constants. Iteration produces rate constants which give a consistent equilibrium constant determined kinetically; previous thermodynamic parameters do not bias the results. The third procedure uses equation 3, from which a plot of $1/\tau^2$ vs. C_{tot}, the total concentration, gives directly k_d^2 from the intercept, and $4k_rk_d$ from the slope. For this data the intercept of such a plot is small, and hence k_d has a large uncertainty. This also contributes to the error of the recombination rate constant, k_r. Thus, although this procedure requires no thermodynamic knowledge, it is less accurate than the second procedure.

Of course, if the mechanism is correct, and the kinetics and thermodynamics are consistent, the three methods should give consistent results. The data in Table I show that the three analytical methods give consistent results within experimental error. The second method is preferred, since it requires no previous thermodynamics results, and is more accurate than the third method. It

TABLE I

Kinetic Results for $rCA_5G + rCU_5G$ Using Different Analytical Methods

Temperature	Recombination Rate Constant ($\times 10^{-6}$) ($\ell\ mol^{-1}\ sec^{-1}$)			Dissociation Rate Constant (sec^{-1})		
	$1/\tau$ vs. C_s	iterating $1/\tau$ vs. C_s	$1/\tau^2$ vs. C_{tot}	$1/\tau$ vs. C_s	iterating $1/\tau$ vs. C_s	$1/\tau^2$ vs. C_{tot}
6.7°C	6.2±0.7	5.2±0.6	5.6±6.4	11±4	12±4	11±1
12.0°C	6.4±0.4	6.8±0.4	6.8±1.9	38±4	37±4	37±8
16.4°C	5.9±0.6	6.0±0.6	6.0±1.3	108±9	107±9	107±12
21.1°C	4.6±1.1	4.4±1.0	4.5±1.3	330±20	330±20	330±20
k (12.0°C)	6.2×10^6	6.2×10^6	6.7×10^6	38	38	37
E_a (kcal/mol)	−2±2	0±2	−6±6	39±2	39±2	40±3
$\Delta S^{\ddagger}$ (e.u.)	−36±7	−29±7	−50±20	84±6	84±6	87±10

will be used for the remainder of the kinetic results reported in this paper. Figure 3 shows plots of $1/\tau$ vs. C_s for the ribo-oligonucleotides, using the second method of analysis.

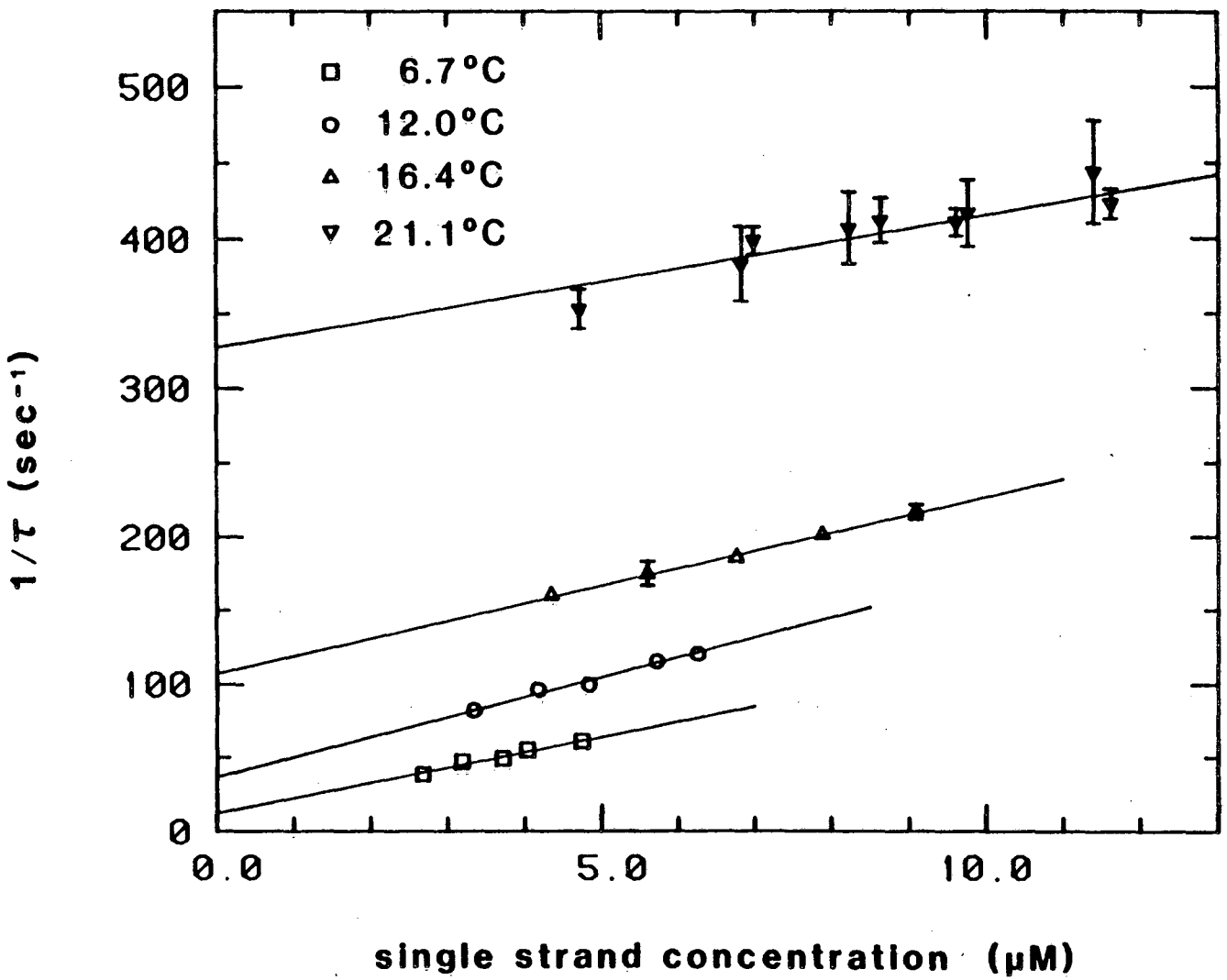
A) Comparison of $rCA_5G + rCU_5G$, $dCA_5G + dCT_5G$ and $rCA_5G + dCT_5G$

Table II shows the kinetic results for the ribo-, deoxyribo- and hybrid oligonucleotides. Activation energies were determined for the first two; rate constants at 12.0° C were determined for the third. The rate constants are all compared at 12.0°C. The recombination rate for the deoxyribo-oligonucleotides (8.3×10^6 ℓ mol^{-1} sec^{-1}) is 34% faster than for the ribo-oligonucleotides (6.2×10^6 ℓ mol^{-1} sec^{-1}); the hybrid is essentially the same as the ribo-oligonucleotides, (6.6×10^6 ℓ mol^{-1} sec^{-1}).

The activation energy, E_a, and the activation entropy, $\Delta S^{\ddagger}$, are obtained from the temperature dependence of the rate constants. Figure 4 shows an Arrhenius plot for the recombination and dissociation rate constants for the ribo- and deoxyribo-oligonucleotides. The recombination activation energies for the ribo- and deoxyribo-oligonucleotides are small and the same within experimental error: 0 ± 2 and -0.5 ± 2 kcal/mol, respectively. The activation entropies are negative and nearly equal, -29 and -30 e.u., respectively. The activation parameters were not measured for the hybrid oligonucleotides; they are assumed to be consistent with the ribo- and deoxyribo-oligonucleotide data since the recombination activation energies are the same in the ribo- and deoxyribo-oligonucleotides.

The dissociation rate constant at 12.0° C for the deoxyribo-oligonucleotides (9.6 sec^{-1}) is several times smaller than the ribo-

Figure 3. $1/\tau$ as a function of equilibrium single strand concentration of rCA_5G + rCU_5G, determined by the second method (see text). The four temperatures are: □ = 6.7°C, ○ = 12.0°C, △ = 16.4°C, ▽ = 21.1°C. The error bars not shown are on the order of the symbol size. The results are tabulated in Table I.


6.7°C
12.0°C
16.4°C
21.1°C
500
400
300
200
100
0
0.0
5.0
10.0
1/τ (sec⁻¹)
single strand concentration (μM)

XBL 824-9359

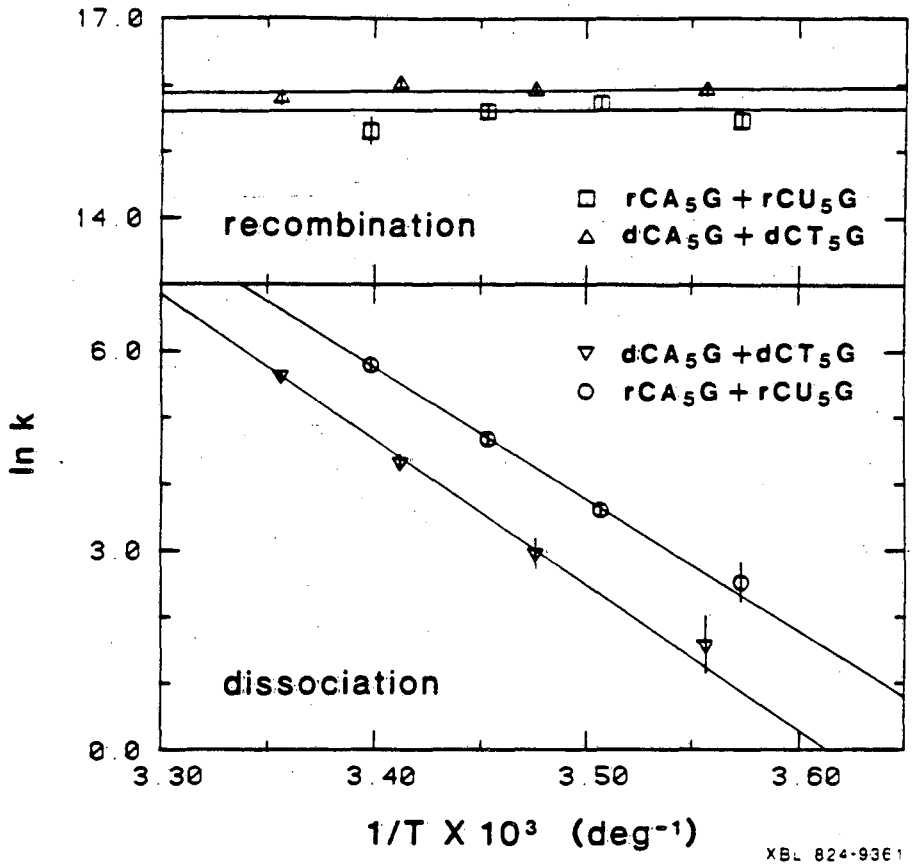
TABLE II

Kinetic Results for rCA_5G + rCU_5G, dCA_5G + dCT_5G and rCA_5G + dCT_5G

	rCA_5G + rCU_5G	dCA_5G + dCT_5G	rCA_5G + dCT_5G
k_r (12.0° C, ℓ mol^{-1} sec^{-1})	6.2×10^6	8.3×10^6	6.6×10^6
$E_{a,r}$ (kcal/mol)	0±2	-0.5±2	---
$\Delta S^{\ddagger}_r$ (e.u.)	-29±7	-30±4	---
k_d (12.0° C, sec^{-1})	38	9.6	50
$E_{a,d}$ (kcal/mol)	39±2	43±3	---
$\Delta S^{\ddagger}_d$ (e.u.)	84±6	95±9	---
$\Delta H°$ (kcal/mol)[a]	-41	-49	
$\Delta S°$ (e.u.)[a]	-120	-145	

[a]Thermodynamic results from Chapter II.

Figure 4. Arrhenius plot of the natural logarithm of the rate constants as a function of the inverse temperature for the recombination and dissociation rate constants. The slope of such a plot is $-E_a/R$. Δ = dCA_5G + dCT_5G recombination rate; $\square$ = rCA_5G + rCU_5G recombination rate; ∇ = dCA_5G + dCT_5G dissociation rate; $\bigcirc$ = rCA_5G + rCU_5G dissociation rate. The results are given in Table II.


17.0
14.0
6.0
3.0
0.0
ln k
recombination
rCA5G + rCU5G
dCA5G + dCT5G
dCA5G + dCT5G
rCA5G + rCU5G
dissociation
3.30
3.40
3.50
3.60
1/T X 10³ (deg⁻¹)
XBL 824-9361

oligonucleotides (38 sec^{-1}). The hybrid dissociation rate constant (50 sec^{-1}) is slightly faster than that of the ribo-oligonucleotides. The dissociation activation energies are large. The deoxyribo-oligonucleotide activation energy (43 kcal/mol) is somewhat larger than that for the ribo-oligonucleotide (39 kcal/mol). The activation entropies for the deoxyribo- and ribo-oligonucleotides (95 e.u. and 84 e.u., respectively) are both large and positive.

The activation energies determined from kinetics can be used to calculate the enthalpy and entropy for the reaction scheme. Comparison of these values with those determined thermodynamically provides a useful check for consistency of the reaction scheme. The thermodynamic parameters for these oligonucleotides have been previously determined from melting curve analysis (Chapter II; Nelson et al., 1981) and are included in Table II. The kinetically determined enthalpy is given by $\Delta H_{kinetic} = E_{a,r} - E_{a,d}$; the entropy is given by $\Delta S_{kinetic} = \Delta S_r^{\ddagger} - \Delta S_d^{\ddagger}$. For the ribo-oligonucleotides, the values of the enthalpy determined from thermodynamics and kinetics are -41 and -39 kcal/mol, respectively, in excellent agreement. The corresponding values for the entropy are -120 and -113 e.u., respectively. For the deoxyribo-oligonucleotides, the comparisons for the enthalpy are -49 and -44 kcal/mol; for the entropy, -145 and -125 e.u. The consistency of these values derived from independent techniques lends confidence to the procedures used to characterize the double-strand formation of these oligonucleotides.

B) The Second Exponential Component

Table III summarizes the relaxation times found for the faster second component for the ribo- and hybrid oligonucleotides at 6.7°C and 12.0°C. For these molecules the second component was not observed above 12°C; no second component was observed for the deoxy-ribo-oligonucleotides at any temperature. There was no trend of the relaxation time with concentration. The data shown are averages of 16 to 23 relaxations at several concentrations for each temperature. Due to the small amplitude of the effect (5-10% of the total amplitude), the errors in determining the relaxation times are large. This relaxation was 7-12 times faster at 6.7°C and 15-25 times faster at 12.0°C than the predominant bimolecular component. No trend in the amplitude with concentration or temperature was determinable.

The relaxation times for the ribo- and hybrid oligonucleotides are the same within experimental error at both temperatures: 440 and 470 sec^{-1}, respectively, at 6.7°C; and 2000 and 1900 sec^{-1} at 12.0°C. The apparent activation energies determined were large and nearly equal for the ribo- and hybrid oligonucleotides: 45 and 42 kcal/mol, respectively. Since I do not know what the process is, the apparent rate cannot be resolved into forward and reverse components.

The predominant relaxation is due to the bimolecular single-strand to double-strand transition. The fast minor component may be a rearrangement between two double helical species (Freier et al., 1982) (see Discussion).

TABLE III

Apparent Rate Constants and Activation Energies for the Second Exponential Component

	$k_{app} = 1/\tau_2$ (sec^{-1})	
	$rCA_5G + rCU_5G$	$rCA_5G + dCT_5G$
6.7°C	440 ± 230	470 ± 150
12.0°C	2000 ± 1500	1900 ± 1100
$E_{a(app)}$	45 kcal/mol	41 kcal/mol

5. Discussion

A) Comparison of the Ribo-, Deoxyribo- and Hybrid Oligonucleotides

These studies allow us to make direct comparisons of the kinetics of ribo-, deoxyribo- and hybrid oligonucleotides of equivalent sequences under identical conditions. Thermodynamic results on these same oligonucleotides showed the order of stability to be $dCA_5G + dCT_5G > rCA_5G + dCT_5G > rCA_5G + rCU_5G$ (Martin & Tinoco, Jr., 1980). The hybrid $dCA_5G + rCU_5G$ was very unstable, and formed triple strands at low temperatures (Martin & Tinoco, Jr., 1980; Pardi et al., 1981); it was not studied here.

The recombination rate constant for the deoxyribo-oligonucleotides at 12.0°C is faster than that for the ribo-oligonucleotides (Table II). The hybrid recombination rate constant is essentially the same as that for the ribo-oligonucleotides. The dissociation rate constant for the deoxyribo-oligonucleotides at 12.0°C is much smaller than for the ribo-oligonucleotides; the hybrid dissociation rate constant was slightly larger than the ribo-oligonucleotides. From this, we can conclude that the deoxyribo-oligonucleotide double helix is more stable than the ribo-oligonucleotide double helix because of both a greater recombination rate and a smaller dissociation rate. However, the difference in dissociation rates is the more important factor. The data for the hybrid double helix is only at one temperature, and thus comparisons with it are not as meaningful.

The activation energy for recombination of the deoxyribo- and ribo-oligonucleotides were both small; 0 and -0.5 kcal/mol, respec-

tively. This suggests that although the rates are different, the mechanism is nearly the same for both oligonucleotides. The dissociation activation energies were both large, with the deoxyribo-oligonucleotide activation energy being greater than that of the ribo-oligonucleotides. The values of the activation energies are consistent with the greater enthalpy of double strand formation for the deoxyribo-oligonucleotides (Chapter II; Nelson et al., 1981). The activation energies for the hybrid oligonucleotides were not determined, but are presumably consistent with the deoxyribo- and ribo-oligonucleotides.

The kinetics of exchange with water for the base-pairing imino protons of the three helices in this work were studied by NMR (Pardi & Tinoco, Jr., 1982). For dCA_5G + dCT_5G in 0.18M NaCl, the rate constant and activation energy for exchange at 12.0°C were 7 sec^{-1} and 47 ± 7 kcal/mol, respectively, for the interior A•T base pairs. The similarity to the values in the present work, 9.6 sec^{-1} and 43 ± 3 kcal/mol, led the authors to conclude that the rate limiting step for exchange of the imino protons was the dissociation of the double strands to single strands. The exchange from the single strands was fast. The agreement of kinetic studies using techniques as diverse as temperature-jump and NMR lends valuable support to the interpretations of the results from both techniques.

B) <u>The Second Exponential Component</u>

As reported in the results section, the relaxation data are best fit using a two-exponential fit for the ribo- and hybrid oligonucleotides at temperatures of 12.0°C and below. One-exponential

fits are observed above 12.0°C and for the deoxyribo-oligonucleotides at all temperatures.

Freier et al. (1982) obtained two-exponential fits for the deoxyribo-oligonucleotide dG-C-G-C-G-C at temperatures below the melting temperature (T_m). They attribute the second relaxation to a reaction between two different double helical species. Since the oligonucleotides in this study are less stable than dG-C-G-C-G-C, we obtained less data below the T_m. They found that the relaxation time of the second component was independent of concentration. The amplitude of their signal was also independent of concentration and temperature. My data also indicate no trend of relaxation time with concentration; the data do not allow a conclusion to be drawn concerning the behavior of the amplitude.

The second component found by Freier et al. had an amplitude of 10-15% of the maximum signal; it contributed a larger fraction of the signal at low temperatures, where the signal from the bimolecular process was small. The second component observed in this study was never more than 5 to 10% of the total amplitude. The second component of Freier et al. was 4 to 10 times faster than the predominant component; in this study, it was 7 to 25 times faster, depending on the temperature (see Results).

The rates for the fast process in dG-C-G-C-G-C are calculated to be 9.6 and 24 sec^{-1} at 6.7°C and 12.0°C, respectively, with an activation energy of 23 kcal/mol (Freier et al., 1982). For the ribo-oligonucleotides in this study, the corresponding rates are 440 and 2000 sec^{-1} at 6.7°C and 12.0°C, respectively, with an activation

energy of 45 kcal/mol. The process observed in this study behaves very differently from that observed by Freier et al. They observe the effect in the deoxyribo-oligonucleotide; here it is observed for the ribo- and hybrid oligonucleotides, but not for the deoxyribo-oligonucleotide. It is not clear that this process is similar to the one observed by Freier et al. Further studies on this effect will be necessary before the process can be characterized to any extent.

C) The Hybrid Behaves Similarly to the Ribo-Oligonucleotide

The rate constants of the hybrid oligonucleotides at 12.0° C are closer to the ribo-oligonucleotides than the deoxyribo-oligonucleotides. This suggests that the properties of the hybrid more closely resemble the RNA than the DNA. More convincing evidence comes from the behavior of the fast second component. The rates and activation energies were essentially the same for the ribo- and hybrid oligonucleotides, and the effect was not observed in the deoxyribo-oligonucleotides (Table III).

NMR studies of the chemical shifts of the base-pairing imino protons and the non-exchangable base and sugar protons of the helices used in this study (Pardi et al., 1981) indicate that the structure of the hybrid is close to the structure of the ribo-oligonucleotide. The NMR study on exchange rates of base-pairing imino protons mentioned earlier (Pardi & Tinoco, Jr., 1982) showed that at 5° C the dissociation rate of the hybrid and ribo-oligonucleotides are nearly equal, and an order of magnitude faster than for the deoxyribo-oligonucleotide.

The evidence seems to indicate that the properties of the hybrid double helices are governed mainly by the ribo-oligonucleotide strand. RNA generally exhibits less structural diversity than DNA. This lower freedom of flexibility of the RNA might dictate the properties of the more flexible DNA strand in the hybrid double helix.

D) Comparison with Previous Results

The kinetics of double-strand formation have been studied for a number of ribo-oligonucleotide sequences containing only A•U base pairs (Craig et al., 1971; Pörschke & Eigen, 1971; Breslauer & Bina-Stein, 1977), both A•U and G•C base pairs (Pörschke et al., 1973; Ravetch et al., 1974), and one sequence containing only G•C base pairs (Podder, 1971). The data for deoxyribo-oligonucleotides are much more limited (Drobnies, 1979; Freier et al., 1982). One hybrid has been studied (Hoggett & Maass, 1971).

The rate of recombination increases considerably as the ionic strength is increased. The dissociation rate is roughly independent of ionic strength (Pörschke et al., 1973). Both recombination and dissociation activation energies are independent of salt concentrations between 0.05 and 1M NaCl (Pörschke et al., 1973). Thus we can compare activation energies and dissociation rate constants from experiments done at different ionic strengths; recombination rate constants can be compared only when measured at the same ionic strength. Table IV summarizes the results for the activation energy for recombination for several oligonucleotides.

TABLE IV

Recombination Activation Energies
for Several Oligonucleotides

Oligonucleotide	(NaCl)	E_a (kcal/mol)	Reference
$rA_n + rU_n$, n=8-18	0.05M	-9	1
rA_nU_n, n=4-7	0.25 to 1M	-4 to -6	2,3
rA_nGCU_n, n=2-4 $rA_nCG + rCGU_n$, n=6,8 $rA_nG_2 + rC_2U_n$, n=4,5	0.05 to 1M	+6 to +9	4
rCG_3+rG_3C	0.18M	+4.5	5
dG-C-G-C-G-C	1M	+1	6
dA-T-G-C-A-T	1M	-4	7
$dA_8 + dT_8$	0.05M	-2	7
$rA_n + dT_n$, n=7-9	1M	-1	8

1. Pörschke & Eigen (1971)
2. Breslauer & Bina-Stein (1977)
3. Craig et al. (1971)
4. Pörschke et al. (1973)
5. Podder (1971)
6. Freier et al. (1982)
7. Drobnies (1979)
8. Hoggett & Maass (1971)

Ribo-oligonucleotides containing only A•U base pairs recombine with a negative activation energy: -9 kcal/mol for $rA_n + rU_n$ and -4 to -6 kcal/mol for rA_nU_n (Table IV). Rate constants and activation energies were essentially independent of chainlength for both systems. The negative activation energy suggests that the mechanism of recombination requires the formation of a stable nucleus with two or three base pairs; the rate-determining step is the formation of the next base pair (Craig et al., 1971; Pörschke & Eigen, 1971) (see below).

Ribo-oligonucleotides containing both G•C and A•U base pairs have positive activation energies for recombination: +6 to +9 kcal/mol for sequences like rA_nGCU_n, $rA_nCG + rCGU_n$ and $rA_nG_2 + rC_2U_n$ (Table IV). The activation energy was roughly independent of length or sequence. However, the rate constants decreased with increasing chainlength. The rates were faster if the G•C base pairs were at the end rather than the middle of the sequence. It appears recombination is faster when formation of the stable nucleus involves more stable G•C base pairs. Adding more A•U base pairs slows the rate due to increased steric hindrance; G•C bases at the end could be more accessible than in the middle of the sequence (Pörschke et al., 1973). The ribo-oligonucleotides rG-G-G-C + rG-C-C-C, which contain only G•C base pairs, have a recombination activation energy of +4.5 kcal/mol. The positive activation energy of oligonucleotides with G•C base pairs probably means the nucleus involved in the rate-determining step requires one or two G•C base pairs (Pörschke et al., 1973).

Much less work has been done on deoxyribo-oligonucleotide kinetics. The recombination activation energy for dG-C-G-C-G-C is +1 kcal/mol (Freier et al., 1982). For dA_8 + dT_8 and dA-T-G-C-A-T, the corresponding values are -2 and -4 kcal/mol, respectively (Drobnies, 1979). The hybrid oligonucleotides rA_n + dT_n have an activation energy of -1 kcal/mol (Hoggett & Maass, 1971).

The trend in activation energies for the ribo-oligonucleotides is from negative values when only A•U base pairs are present (-4 to -9 kcal/mol), zero when isolated G•C base pairs are present, and positive when two or more G•C base pairs are adjacent (+5 to +9 kcal/mol). Due to the greater stability of G•C base pairs, fewer base pairs are required to form a stable nucleus, thus the larger activation energy when G•C base pairs are present. This trend is not as apparent in the deoxyribo-oligonucleotides. The activation energies vary from -2, -0.5, -4, +1 kcal/mol for the series dA_8 + dT_8, dCA_5G + dCT_5G, dA-T-G-C-A-T, and dG-C-G-C-G-C. This might mean the number of bases required for the stable nucleus does not depend on the presence of G•C base pairs. The difference in stability between A•T and G•C base pairs in DNA is not as great as for RNA; thus, in DNA the stability of the nucleus will not depend as strongly on the presence of G•C base pairs, and the number of bases in the nucleus and thus the activation energies vary less than for RNA.

Because of the restriction that recombination rate constants must be compared at the same salt concentration, there are fewer results to compare. Values determined in 1M NaCl at 12.0°C from earlier work (in terms of ℓ mol^{-1} sec^{-1}) are: rA_7U_7, 2 x 10^6

(Breslauer & Bina-Stein, 1977); rA_2GCU_2, 5 x 10^6 (Pörschke et al., 1973); dA-T-G-C-A-T, 8 x 10^6 (Drobnies, 1979); dG-C-G-C-G-C, 11 x 10^6 (Freier et al., 1982). No systematic comparisons may be made from the data available.

E) The Mechanism of Double-Strand Formation

The data presented in Tables II and III allow us to investigate the differences between the kinetics of ribo-, deoxyribo- and hybrid oligonucleotides.

The negative activation energy found for the recombination of ribo-oligonucleotides rA_n + rU_n means there is a pre-equilibrium step involved in double strand formation. The rate-determining-step of the recombination is hypothesized to be the addition of the next base pair to the nucleus composed of a few base pairs (Craig et al., 1971). The nucleus is the species which adds the next base pair faster than it dissociates. Thereafter the double helix quickly zippers up to the fully base-paired double helix. The nucleus is in a fast equilibrium with the single strands. Thus the forward rate is given by k_f(nucleus), where k_f is the rate of forming the next base pair. Dissociation of the strands occurs by breaking enough base pairs to get to the nucleus, which then quickly dissociates. The forward rate is then characterized by a small activation energy, positive or negative depending on the number of base pairs in the nucleus. The dissociation rate is characterized by a large activation energy required for breaking several base pairs to get to the nucleus.

Differences in the forward rate will depend on two factors: the concentration of the nucleus and the rate of adding base pairs to an existing nucleus.

It was stated earlier that the recombination of dCA_5G + dCT_5G was faster than rCA_5G + rCU_5G. The hybrid rCA_5G + dCT_5G recombines at essentially the same rate as rCA_5G + rCU_5G. The dissociation rates went as hybrid $\gtrsim$ ribo > deoxyribo. The behavior of the second exponential component also suggests that the behavior of the hybrid is similar to the ribo-oligonucleotide.

The similarity of the recombination activation energies for the deoxyribo- and ribo-oligonucleotides might suggest that the nuclei for double strand formation are of similar stability. Thus the increased rate for the deoxyribo-oligonucleotides might be manifested in a larger k_f than for the ribo-oligonucleotides. Freier et al. suggest that the zippering rate, k_f, might be limited by the diffusion of the single stranded bases into a helical conformation (Freier et al., 1982). The faster rate of recombination of the deoxyribo-oligonucleotides might be a result of faster single strand stacking. From laser temperature-jump studies, it was determined that single strand stacking in poly(dA) is about 4 times faster than in poly(rA): k(single strand stacking) = 2.7×10^7 and 0.7×10^7 sec^{-1}, respectively, at 25°C in 0.05M sodium cacodylate (Dewey & Turner, 1979). This reasoning suggests that the hybrid recombination rate would be similar to the ribo-oligonucleotides, since the helix zippering would be limited by the stacking of the slower rCA_5G strand.

The data presented in this chapter are insufficient to fully justify the explanations for the differences between deoxyribo- and ribo-oligonucleotide kinetics. More studies comparing deoxyribo- and ribo-oligonucleotides would greatly clarify the situation. Ribo-G•C base pairs are more stable than deoxyribo-G•C base pairs. It would be very informative to compare the kinetics of double strand formation for two analogous G-C-containing deoxyribo- and ribo-oligonucleotides. This would determine if the differences in stability are due to the ribo-oligonucleotide having a greater rate of recombination, a slower rate of dissociation, or both. Also comparing recombination activation energies and single strand stacking rates would determine whether the differences in recombination rates are due to a more stable nucleus as manifested by a more negative activation energy, or a faster zippering rate.

CHAPTER IV

Ethidium Ion Binding to Oligonucleotides

1. Synopsis

The thermodynamics of ethidium ion binding to the double strands formed by the ribo-oligonucleotides rCA_5G + rCU_5G and the analogous deoxyribo-oligonucleotides dCA_5G + dCT_5G were determined by monitoring the absorbance vs. temperature at 260 and 283 nm at several concentrations of oligonucleotides and ethidium bromide. A statistical model is described which allows up to three ethidium ions to bind per double strand. The data from the ribo-oligonucleotides were equally well fit by models which assumed negative cooperativity between next-nearest-neighbor binding sites with a cooperativity parameter $\omega = 0.1$, and a model which assumed the terminal binding sites on both ends of the helix are stronger than the interior sites by a factor of $\tau = 140$. The data from the deoxyribo-oligonucleotides were fit best by a model which assumed no cooperativity and identical binding sites. The enthalpy and equilibriuim constant at 25° C for the ribo-oligonucleotides were determined to be about -11 to -14 kcal/mol and 25×10^4 to 60×10^4 ℓ mol^{-1}, respectively, depending on the model. For the deoxyribo-oligonucleotides, the values were -9 kcal/mol and 2.5×10^4 ℓ mol^{-1}, respectively.

2. Introduction

A large number of molecules that cause frameshift mutations intercalate between the base pairs of nucleic acids. Streisinger et

al. (1966) proposed a model in which the mutagen promotes frameshift mutations via the stabilization of a bulge after strand breakage in DNA. This bulge is then locked into the sequence when the break is repaired. A strong correlation was found between frameshift mutagenicity and chemical carcinogenicity (McCann et al., 1975). Ethidium bromide has been used extensively as a probe to help characterize the factors that are important in stabilizing the intercalation complex. Ethidium bromide has been shown to be a frameshift mutagen in the Ames test (McCann et al., 1975).

Several factors make ethidium bromide an ideal probe. It intercalates with a large binding constant, making it possible to prepare samples in which essentially all of the ethidium is bound. Intercalation is accompanied by a large shift in the visible absorption band at 480 nm to longer wavelengths (Waring, 1965; Waring, 1966). Also, the fluorescence is enhanced greatly upon intercalation (LePecq & Paoletti, 1967). Ethidium bromide forms dimers with itself in solution; however the extent of aggregation is small compared to intercalators such as the acridines (Thomas & Roques, 1972; Reinhardt & Krugh, 1978). Ethidium bromide binds to single-stranded nucleic acids (Kreishman & Chan, 1971), but the binding is very much weaker than intercalation in double-stranded nucleic acids (Waring, 1966).

Studies of ethidium bromide binding to DNA have demonstrated that binding occurs with nearest-neighbor exclusion, and with very little cooperativity (Bresloff & Crothers, 1975). Overall binding constants may be obtained by this procedure; however sequence-

specific properties are inaccessible due to the randomness of the DNA sequence.

Studies carried out on dinucleotides have shown that there is a preference for ethidium binding to double-stranded sequences in the order pyrimidine-purine > purine-purine > purine-pyrimidine (Krugh & Reinhardt, 1975; Krugh et al., 1975; Reinhardt & Krugh, 1978; Pardi, 1980; Dahl, 1981; Dahl et al., 1982). The purine bases are adenine and guanine, whereas the pyrimidine bases are cytosine, uracil and thymine. However, the dinucleotides form very unstable mini-double helices by themselves, as evidenced by the equilibrium constants for double strand formation from the self-complementary dinucleotide pdG-dC being on the order of 10 ℓ mol^{-1} (Young & Krugh, 1975; Krugh et al., 1976). Dinucleotides composed of only A•T or A•U base pairs have equilibrium constants which are too small to measure. Thus the sequence specificity found in these studies are complicated by the fact that the stability of the double strand in the absence of ethidium bromide is not well-known.

Ethidium bromide forms complexes with the double helix formed by the tri-nucleotide rCpUpG by intercalating between the two C•G base pairs, bulging the two uracils into solution (Lee & Tinoco, Jr., 1978). There is also a complex formed between a mixture of rGpUpG, rCpC, and ethidium, wherein the two C•G base pairs are formed with the uracil bulged into solution (Lee & Tinoco, Jr., 1978). The equilibrium constants for the complex of ethidium with rCpG and with rCpUpG were measured at 0°C to be 100 x 10^6 and 1 x 10^6 ℓ^2 mol^{-2}, respectively (Pardi, 1980). Thus, the bulged uracils

destabilize the structure significantly. The equilibrium constant for the complex formed by rGpUpG, rCpC, and ethidium was less than 10^5 (Pardi, 1980). Polymer studies on the double strands poly(I)•poly(C,A_x), where x denotes the mole fraction of A residues on the C strand, showed that the equilibrium constant of ethidium binding to the site with the A•I mismatch is about 20 times greater than that for binding to the normal base pairs (Helfgott & Kallenbach, 1979). This indicates that ethidium bromide might relieve some of the destabilizing effect of the mismatched bases.

Using oligonucleotides, it is possible to determine sequence effects by studying different sequences. Oligonucleotide studies have the advantage that the properties of the double strands in the absence of ethidium are known (Chapter II, Nelson et al., 1981). Potentially, oligonucleotide studies can be used to determine the destabilizing effect of a bulged base, and the extent to which ethidium binding relieves this strain.

In this chapter, I report the results on the binding of ethidium bromide to the double strands formed by the oligonucleotides rCA_5G + rCU_5G and dCA_5G + dCT_5G. The extent of ethidium binding is measured by monitoring the large change in absorbance in the UV band of ethidium at 283 nm. Fortuitously, at this wavelength, the double-stranded oligonucleotides absorb to the same extent as the single strands. A statistical model is described wherein the ethidium cation can bind between any combination of base pairs within the nearest-neighbor limit.

3. Experimental

The synthesis and characterization of the oligonucleotides rCA_5G, rCU_5G, dCA_5G and dCT_5G were described in Chapter II. Ethidium bromide was purchased from Sigma. Ethidium bromide often contains ethanol impurity (Bresloff & Crothers, 1975). To remove any ethanol present, the ethidium bromide was lyophilized twice with double-distilled water prior to use.

The buffer used throughout this study consisted of 0.2M NaCl, 0.01M sodium phosphate buffer, pH = 7, and 0.1mM EDTA.

Samples were prepared by adding small amounts of concentrated stock solutions of the oligonucleotides and ethidium bromide to the buffer. The buffer was degassed by purging with helium for three to four minutes prior to preparing the samples. The cuvettes were teflon-stoppered with path lengths of 0.1, 0.2, 0.5 or 1 cm. Melting curves were returned to 0°C after going to high temperature to check for evaporation, which was always less than 1%.

Samples of the ribo-oligonucleotides were made up with nominal concentrations of 50μM, 25μM, and 12μM, with ethidium:strand ratios of approximately 0, 0.1, 0.2, 0.5, 1, 2, and 3. One sample contained 50μM strands and a 4:1 ratio of ethidium:strands. The Eppendorf 1.5ml polypropylene micro centrifuge tubes were used to prepare the samples. Ethidium bromide is adsorbed irreversibly to these tubes when filled the first time. However, after rinsing a tube thus treated, no significant additional ethidium is adsorbed. Therefore, tubes used for sample preparations were pre-treated by rinsing with an ethidium bromide solution. The actual concentra-

tions were determined using the absorbances at 50° C, where the oligonucleotides are single-stranded, and the ethidium is unbound. When the melting curve was not finished by 50° C, the absorbances of single strands and free ethidium at 50°C were determined by extrapolating the absorbances at higher temperatures.

The deoxyribo-oligonucleotide samples were made at a concentration of roughly 40μM, with ethidium:strand ratios of roughly 0.4, 0.6, and 0.8. Because ethidium does not bind as strongly to the deoxyribo-oligonucleotides, studies were impractical at ethidium:strand ratios greater than 1 (see results).

Melting curves were obtained using a Gilford Model 250 UV-VIS spectrophotometer, with a Gilford Model 2527 thermoprogrammer. Data were obtained concurrently at 260 and 283 nm using a Gilford Model 2530 wavelength scanner. The data were collected by a Commodore PET Model 2001 microcomputer interfaced to the instrument and were later transmitted to a VAX 11/780 computer, where the analysis was done. The melting data were interpolated to every 1° C, since analysis of the data requires knowing A_{260} and A_{283} at the same temperature (see below). The temperature range was generally 0° C to 70° C. Spectra were taken on the Gilford Model 250 spectrophotometer modified to allow the PET computer to control the wavelength.

Samples for melting curves of ethidium bromide in buffer were prepared in unstoppered cuvettes, which were covered with silicon oil (Dow Corning 200 Fluid, 20cS viscosity) to ensure that evaporation was negligible. No detectable amount of ethidium went into the

oil. The melting curves were taken from 0°C to 90°C. Spectra of ethidium were taken at 0°C, 25°C, and 50°C in stoppered cuvettes.

The computer programs used to determine the fraction ethidium bound from melting curves, and to carry out the model calculations, are listed in Appendix C.

4. Theory

A) Analysis of Melting Curves

From the melting curves, we want to measure f_b, the fraction of ethidium ions bound, and f_h, the fraction of strands in double helices (with or without ethidium ions bound). We can obtain this information from the melting curves obtained at two wavelengths (260 and 283 nm) as follows.

We can write the mass balance equations for the total concentration of strands, C_t, and the total concentration of ethidium, C_d, as:

$$C_t = C_s + C_h \quad (1)$$

$$C_h = C_0 + C_1 + C_2 + C_3 + \ldots + C_n \quad (2)$$

$$C_d = C_f + C_b \quad (3)$$

$$C_b = C_1 + 2C_2 + 3C_3 + \ldots + nC_n \quad (4)$$

where C_s = concentation of single strands, C_h = concentration of all double helices, with or without ethidium ions bound, C_0 = concentration of double helices with no ethidium bound, C_1 = concentration of double helices with one ethidium bound, etc., and where C_f = concentration of free ethidium and C_b = total concentration of bound ethidium. The maximum number of ethidium ions that can bind per double

helix is n (3 in our case).

If ℓ is the pathlength in cm, and the extinction coefficients are labelled as ε, then we can write the expression for the absorbance at 260 nm as:

$$A_{260}/\ell = \varepsilon_{s,260} C_s + \varepsilon_{f,260} C_f + \sum_{i=0}^{n} \varepsilon_{i,260} C_i \qquad (5)$$

with a similar expression for the absorbance at 283 nm.

We have made the assumption that the extinction coefficients of the species with ethidium bound do not depend on which sites the ethidium ions bind to. For example, we assign the same $\varepsilon_{1,260}$ when the ethidium is bound on the left side as we do when it is bound in the middle of the double helix.

Define the quantities $\Delta\varepsilon_i$ such that $\varepsilon_i = \varepsilon_0 + \Delta\varepsilon_i$. (We have excluded the wavelength subscript for clarity.) Thus, $\Delta\varepsilon_i$ is the change in the extinction coefficient of the double strands due to the binding of the i ethidium ions. By assuming that each ethidium ion makes the same contribution to the extinction coefficient, we can write

$$\varepsilon_i = \varepsilon_0 + i\Delta\varepsilon_1 \qquad (6)$$

We can assign $\Delta\varepsilon_1$ as the extinction coefficient for an ethidium ion bound to the double helix, namely $\varepsilon_b = \Delta\varepsilon_1$. By making these substitutions, and by assigning $\varepsilon_h = \varepsilon_0$, the extinction coefficient for the double helix with no ethidium bound, we can re-write equation 5 as:

$$A_{260}/\ell = \varepsilon_{s,260}\, C_s + \varepsilon_{f,260}\, C_f + \varepsilon_{h,260}\, C + \varepsilon_{b,260}\, C_b \qquad (7)$$

(again with a similar equation for the absorbance at 283 nm).

This approach assigns all of the change in the extinction coefficients upon binding ethidium ions to the ethidium chromophore. Actually, this change is probably due to changes in both the nucleic acid and the ethidium chromophores. However, this makes no difference in analyzing the data in this manner, using the two assumptions stated above.

We can define the fraction of strands in double helices, $f_h = C_h/C_t$, and the fraction of ethidium ions bound, $f_b = C_b/C_d$, and write:

$$A_{260}/\ell = C_t\left[f_h\varepsilon_{h,260} + (1-f_h)\varepsilon_{s,260}\right] + C_d\left[f_b\varepsilon_{b,260} + (1-f_b)\varepsilon_{f,260}\right] \qquad (8)$$

$$A_{283}/\ell = C_t\left[f_h\varepsilon_{h,283} + (1-f_h)\varepsilon_{s,283}\right] + C_d\left[f_b\varepsilon_{b,283} + (1-f_b)\varepsilon_{f,283}\right] \qquad (9)$$

Thus, we need 8 extinction coefficients to determine the values for f_h and f_b from the melting cuves at two wavelengths. The extinction coefficients generally will have a temperature dependence, for example, $\varepsilon(T) = \varepsilon(T_o) + m(T-T_o)$ describes an extinction coefficient that is linear with temperature. The method used to determine the extinction coefficients is explained in the results section.

Equations 8 and 9 can be solved simultaneously to determine values for f_h and f_b. Alternatively, since the melting of these

double helices to single strands is essentially isosbestic, namely $\varepsilon_{h,283} \approx \varepsilon_{s,283}$, the fraction dye bound, f_b may be directly determined from equation 9 alone. This approach requires the determination of only three extinction coefficients, $\varepsilon_{s,283}$, $\varepsilon_{f,283}$ and $\varepsilon_{b,283}$.

The above approach allows us to determine rather directly the fraction of ethidium ions bound from melting curves, provided we have good estimates for the extinction coefficients.

B) Description of the Model

Three different statistical models were used to describe the binding of ethidium ions to the double helices formed by rCA_5G + rCU_5G or dCA_5G + dCT_5G. One model assumed that the pyrimidine-purine site was stronger than the purine-purine sites by a factor σ. Of course, $\sigma = 1$ corresponds to the model with all binding sites equal. A second model assumed that the two terminal binding sites, $\begin{smallmatrix} C-A \\ G-U \end{smallmatrix}$ and $\begin{smallmatrix} A-G \\ U-C \end{smallmatrix}$, were stronger than the internal $\begin{smallmatrix} A-A \\ U-U \end{smallmatrix}$ binding sites by a factor τ. A third model assumed that there was cooperativity between binding sites, where the parameter ω describes the effect of one bound ethidium ion on the adjacent next-nearest-neighbor binding site. Two assumptions will be made for all of the models:

1) Binding is only via intercalation between base pairs in the double helix. No binding is allowed to the ends of the helix, or to single strands.

2) Binding occurs with nearest-neighbor exclusion: binding an ethidium ion between two base pairs makes the adjacent site unavailable for binding another ethidium ion.

I will explain the development of the model using the assumption that the pyrimidine-purine site is σ times stronger than the purine-purine sites, with no cooperativity. As will be explained, the theory is very easily modified to accomodate the alternative models.

We denote the equilibrium constant for double-helix formation K_h:

$$rCA_5G + rCU_5G \underset{}{\overset{K_h}{\rightleftharpoons}} \begin{array}{c} rC\text{-}A\text{-}A\text{-}A\text{-}A\text{-}A\text{-}G \\ G\text{-}U\text{-}U\text{-}U\text{-}U\text{-}U\text{-}Cr \end{array} \qquad (10)$$

The equilibrium constant for binding ethidium to any of the purine-purine sites is denoted K_d:

$$\text{ethidium}(\mid) + \begin{array}{c} rC\text{-}A\text{-}A\text{-}A\text{-}A\text{-}A\text{-}G \\ G\text{-}U\text{-}U\text{-}U\text{-}U\text{-}U\text{-}Cr \end{array} \overset{K_d}{\rightleftharpoons} \begin{array}{c} rC\text{-}A\text{-}A \\ G\text{-}U\text{-}U \end{array} \mid \begin{array}{c} A\text{-}A\text{-}A\text{-}G \\ U\text{-}U\text{-}U\text{-}Cr \end{array} \qquad (11)$$

The equilibrium constant for binding ethidium to the pyrimidine-purine site is denoted σK_d. Hence the pyrimidine-purine site is σ times stronger than the purine-purine sites. Nearest-neighbor exclusion allows only three ethidium ions to bind per double helix (n = 3 in equations 3 and 4). Table I shows some of the possible arrangements of binding one, two, or three ethidium ions. There are 6 ways to bind one ethidium ion, 10 ways to bind two, and 4 ways to bind three.

TABLE I

Statistical Weights of Double Helices
with Ethidium Ions Bound

no. of bound ethidium ions	C–A / C–U site filled			C–A / C–U site vacant			
	example	statistical weight*	no. of species	example	statistical weight*	no. of species	g_i
0		---	-		K_h	1	1
1		$\sigma K_h S$	1		$K_h S$	5	$\sigma + 5$
2		$\sigma K_h S^2$	4		$K_h S^2$	6	$4\sigma + g$
3		$\sigma K_h S^3$	3		$K_h S^3$	1	$3\sigma + 1$

* Statistical weight of single strands = 1. $S = C_f K_d$.

The concentration of double helices with one ethidium ion bound to a specific purine-purine site is $K_h K_d C_f C_s^2$. The concentration of double helices with one ethidium bound to the pyrimidine-purine site is $\sigma K_h K_d C_f C_s^2$. If we assign $S = K_d C_f$, we can write the concentration of double helices with i ethidium ions bound in a specific arrangement as $K_h S^i C_s^2$ if the pyrimidine-purine site is vacant, and $\sigma K_h S^i C_s^2$ if it is filled.

To develop the statistical approach, I define the statistical weight of the species as shown in Table I, namely $K_h S^i$ or $\sigma K_h S^i$ for double strands with i ethidium ions bound, depending on whether the pyrimidine-purine site is filled. The single strands are defined as the reference state, and hence have a statistical weight of 1.

The partition function describing the double-helical species with or without ethidium ions bound is defined as Q:

$$Q = \sum_{i=0}^{n} \begin{array}{c}\text{(statistical weight of double helices} \\ \text{with i ethidium ions bound)}\end{array} \qquad (12)$$

In order to simplify the calculation of the sum of statistical weights, I will define statistical factors g_i for the double helices with i ethidium ions bound, with i = 0 to n. The factor g_i describes the number of ways i ethidium ions can bind, and the relative strengths. For example, there are six ways one ethidium ion can bind to these double strands. One way to the pyrimidine-purine binding site and five ways to a purine-purine site. Thus, $g_1 = \sigma + 5$. Table I shows the different ways to bind 0 to 3 ethidium ions, the statistical weights, and the statistical factors g_i. As shown,

$g_0 = 1$, $g_1 = \sigma + 5$, $g_2 = 4\sigma + 6$, and $g_3 = 3\sigma + 1$. Thus, equation 12 can be written:

$$Q = \sum_{i=0}^{n} g_i K_h S^i \tag{13}$$

The fraction of all double helices which have i ethidium ions bound is denoted P_i:

$$P_i = C_i/C_h = g_i K_h S^i/Q \tag{14}$$

If the total concentrations of the non-selfcomplementary single strands are equal, the total concentration of double helical species, C_h, is obtained from the partition function by:

$$\frac{C_h}{C_s^{\,2}} = \frac{C_h}{(C_t - C_h)^2} = Q \tag{15}$$

which we can solve to get:

$$f_h = C_h/C_t = \frac{1 + 2QC_t - \sqrt{1 + 4QC_t}}{2QC_t} \tag{16}$$

The expression for the total amount of ethidium bound is:

$$C_b = \sum_{i=1}^{n} i \cdot C_i = C_h \cdot \sum_{i=1}^{n} i \cdot P_i = f_h C_t \sum_{i=1}^{n} i \cdot P_i \tag{17}$$

which can be written as the fraction of ethidium bound, $f_b = C_b/C_d$:

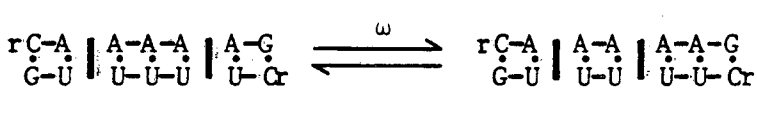
$$f_b = \frac{f_h C_t}{C_d} \cdot \sum_{i=1}^{n} i \cdot P_i \tag{18}$$

We need a procedure to calculate the quantities f_h and f_b from the parameters K_h, K_d, C_t, C_d and σ. The right-hand side of equation 18 depends upon K_h, K_d, C_t, C_d, σ, and f_b. Thus, we want to determine f_b such that:

$$f_b - \frac{f_h C_t}{C_d} \cdot \sum_{i=1}^{n} i \cdot P_i = 0 \qquad (19)$$

where the value of f_b determines the second term. This is readily done by standard numerical solution techniques; I utilized Newton's method.

The previous development may be performed for a number of different models. The only difference is the form of the statistical factors g_i. We can include in the model cooperativity in binding ethidium by including a cooperativity parameter ω which is defined as the equilibrium:



(20)

Thus, $\omega > 1$ means a bound ethidium enhances binding at the next available site, and $\omega < 1$ means a bound ethidium reduces binding at the next available site. The statistical weights of the species with two or more ethidium ions bound are then multiplied by ω for each time that two bound ethidium ions are two base pairs apart. The values for the statistical factors for this model are easily shown to be: $g_0 = 1$, $g_1 = \sigma + 5$, $g_2 = \sigma(\omega + 3) + 3\omega + 3$, and $g_3 = \sigma(\omega^2 + 2\omega) + \omega^2$.

C) Fitting the Model Parameters to Experimental Results

After experimentally determining f_b, the fraction of ethidium bound to double helices, we need a procedure to determine the best values of K_d and σ to fit the data. I have developed the following method:

The first step is to pick an arbitrary value for σ. In general, $\sigma = \exp(\Delta S_\sigma^\circ/T - \Delta H_\sigma^\circ/RT)$, where ΔH_σ° and ΔS_σ° are the differences between the enthalpy and entropy between the pyrimidine-purine and purine-purine sites. Usually, the difference will be attributed entirely to the entropy, making σ independent of temperature.

Using this fixed value for sigma, equation 19 may be solved numerically for K_d, using the experimental value of f_b at any particular temperature. Several melting curves at different strand concentrations and ethidium:strand ratios are fit simultaneously. A plot of $\ln(K_d)$ vs. $1/T$ should give a linear plot whose slope will be $\Delta H_d^\circ/R$, and whose intercept will be $\Delta S_d^\circ/R$ by the van't Hoff relation. Linear regression is used to determine the best values for ΔH_d° and ΔS_d°. Having thus determined an estimate for K_d at each temperature, we may now calculate theoretical values for f_b for each point in the melting curves. The goodness of fit can be evaluated both by the linear correlation coefficient, r, from the linear regression analysis, or by the total reduced chi-squared determined for all the melting curves, $\chi_\nu^2 = (f_{d,calc} - f_{d,expt})^2/(N - 2)$ where there are N data points from all the melting curves (Bevington, 1969). Hopefully, by repeating the process for a range of σ's, the best value will become apparent.

5. Results

A) Optical Properties of Ethidium Bromide in the UV

Most of the work on the binding of ethidium bromide to nucleic acids have been done by monitoring the absorbance of ethidium bromide in the visible band near 480 nm. This has the advantage that the absorbance of the nucleic acids does not interfere with the measurement, allowing the use of a large excess of nucleic acids. However, in this work, I monitored the behavior of both the ethidium bromide and the oligonucleotide, and hence carried out the studies in the UV.

A range of extinction coefficients for ethidium bromide at 480 nm have been reported (in terms of $\ell\ mol^{-1}\ cm^{-1}$): 5.6×10^3 (Waring, 1965), 5.45×10^3 (LePecq, 1971), 5.9×10^3 (Wakelin & Waring, 1974), 5.85×10^3 (Bresloff & Crothers, 1975), and 5.86×10^3 (Reinhardt & Krugh, 1978). I used an average value of $(5.8 \pm 0.2) \times 10^3\ \ell\ mol^{-1}\ cm^{-1}$.

Ethidium bromide forms dimers at moderate concentrations, which causes a shift in the maximum absorbance to wavelengths greater than 480 nm (Thomas & Roques, 1972). The equilibrium constant at 25°C for this dimerization was found to be 70 M^{-1} in 0.1M NaCl by NMR (Reinhardt & Krugh, 1978). At the most concentrated solution of ethidium bromide used in this study, 0.2mM, less than 3% of the ethidium was dimerized, and hence was considered negligible. Indeed, there was no significant departure from Beer's law at 260 or 283 nm for ethidium bromide up to 1mM concentration.

The absorption spectrum in the range of 220 to 380 nm of ethidium bromide depends on temperature. The maximum absorbance was found to be 284 nm at 0°C, 285 nm at 25°C, and 286 nm at 50°C. The absorption spectra at 0°C and 50°C are shown in figure 1. The absorption spectra were superimposable between 260 and 380 nm over the concentration range of 0.01mM to 0.2mM. The spectra deviated with concentration below 250 nm, with higher concentrations having larger absorbances.

The extinction coefficients at 283 nm and 260 nm at 25°C were obtained by comparing the absorbance of a dilute solution (20μM) at 480, 283 and 260 nm: $\varepsilon_{283} = 5.6 \times 10^4$ and $\varepsilon_{260} = 1.74 \times 10^4$.

The extinction coefficient at 260 nm varied linearly with temperature and did not depend on concentration. It fit well to the equation $\varepsilon_{260}(T) = 1.7 \times 10^4 - 25.5(T - 50°C)$ ($\ell\ mol^{-1}\ cm^{-1}$).

At 283 nm, the melting curve depended on concentration. Melting curves taken from 0°C to 90°C are shown in figure 2 for three ethidium concentrations. The effect with concentration is not large, only about 3% at 0°C, and negligible above 50°C. The melting curves were fit very well by the empirical equation $\varepsilon_{283}(T)/\varepsilon_{283}(90°C) = 1 - 1.45 \times 10^{-3}(T - 90) + a_3(T - 90)^3$ where $a_3 = 4.9 \times 10^{-6} \cdot \sqrt{\text{concentration}} + 8.6 \times 10^{-9}$. The dependence of these equations on the square root of the concentration is an empirical relation to fit the data; no theoretical motivation was involved. The curves calculated using this equation are also shown in figure 2. $\varepsilon_{283}(90°C)$ was found to be 5.1×10^4. All of the extinction coefficients are shown in Table II.

Figure 1. Absorption spectra of ethidium bromide in 0.2M NaCl, 0.01M phosphate buffer, pH = 7, 0.1mM EDTA. The ethidium bromide concentration was 17μM; the path length was 1 cm. o = 0°C; x = 50°C.

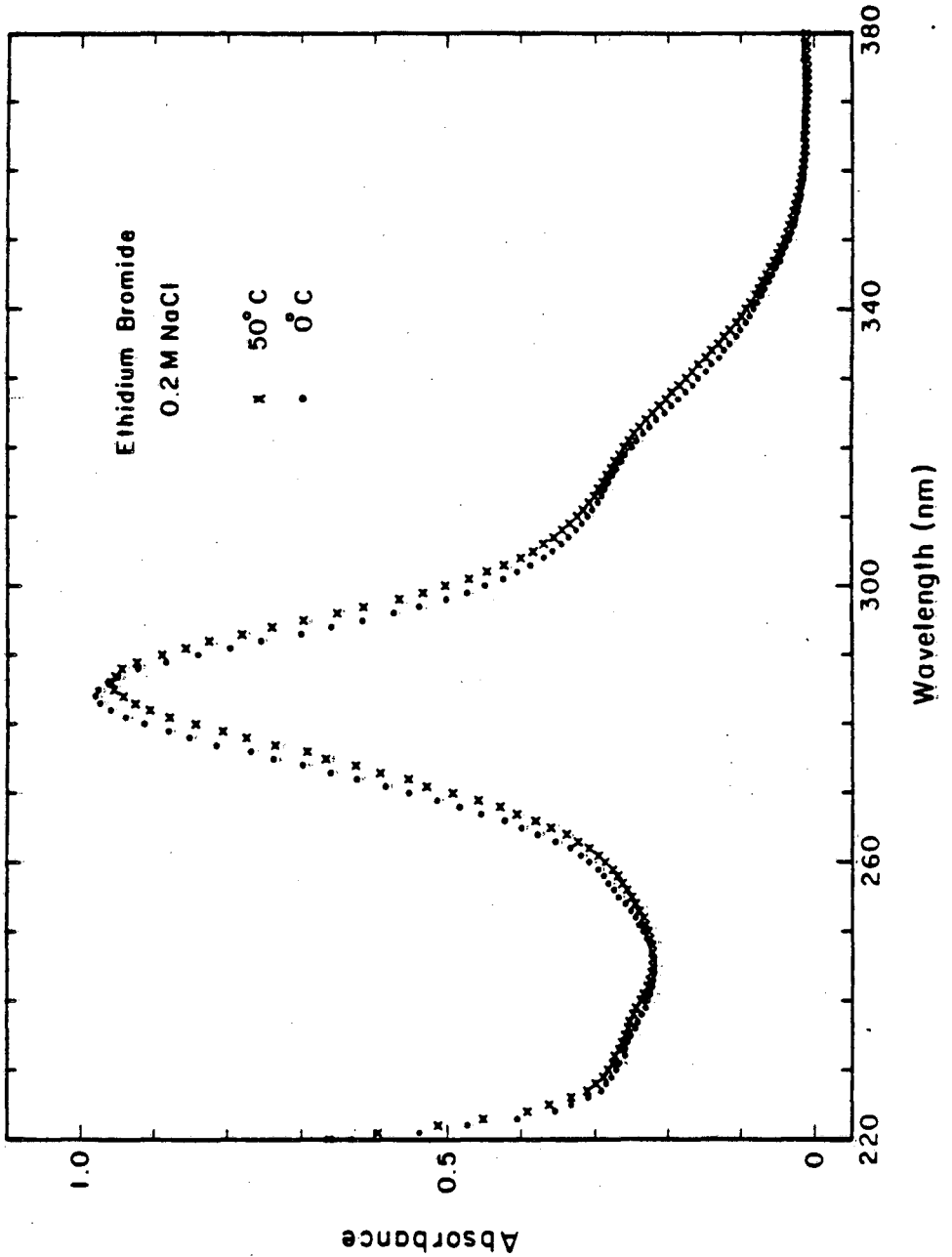

Ethidium Bromide
0.2 M NaCl
50°C
0°C
Absorbance
1.0
0.5
0
220
260
300
340
380
Wavelength (nm)

Figure 2. Melting curves of ethidium bromide in 0.2M NaCl, 0.01M phosphate buffer, pH = 7, 0.01mM EDTA. The curves are normalized by dividing by the absorbance at 90°C. Hence the parameter plotted is $\varepsilon_{283}(T)/\varepsilon_{283}(90°C)$. The concentrations were: Δ, 19μM; O, 96μM and □, 184μM. Th lines show the fit using the equation $\varepsilon_{283}(T)/\varepsilon_{283}(90°C) = 1 - 1.45 \times 10^{3} \cdot (T - 90°C) + a_3(T - 90°C)^3$ where $a_3 = 4.9 \times 10^{-6} \cdot \sqrt{\text{concentration}} + 8.6 \times 10^{-9}$ (see text).

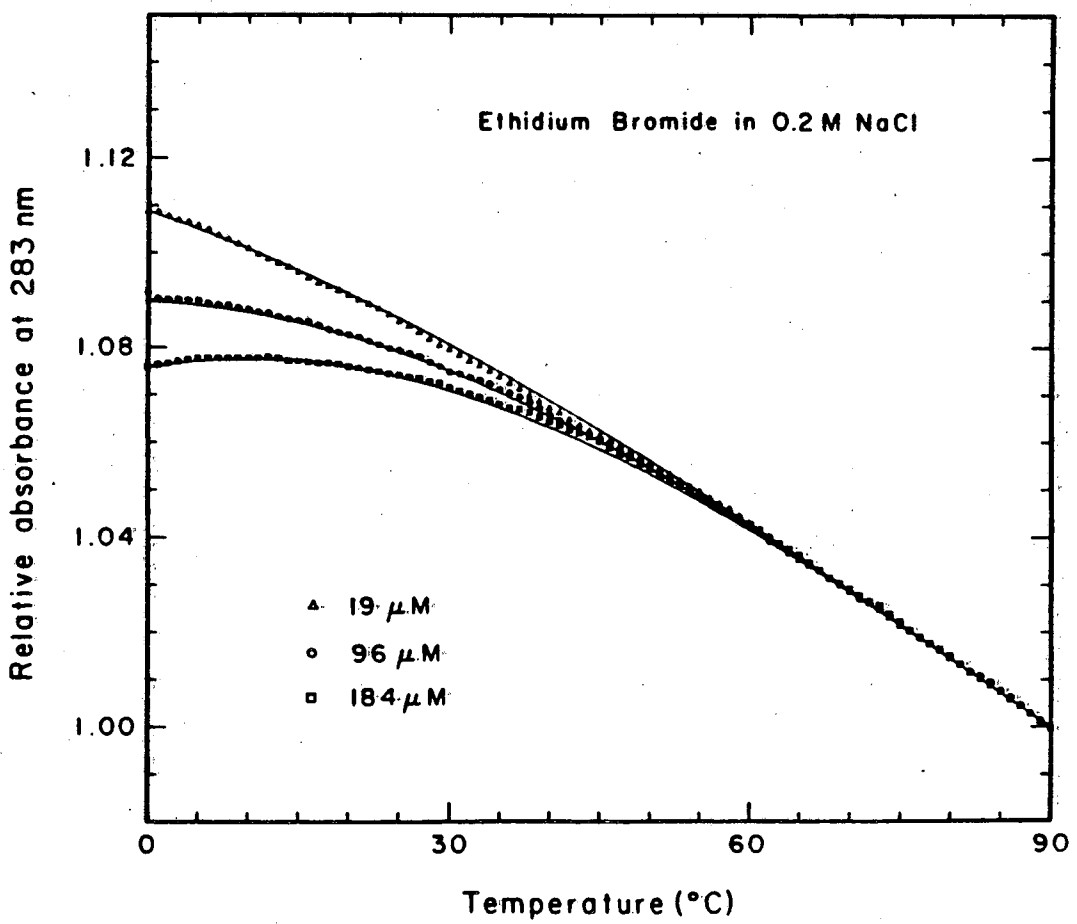

Ethidium Bromide in 0.2 M NaCl
Relative absorbance at 283 nm
1.12
1.08
1.04
1.00
19 μM
96 μM
184 μM
0
30
60
90
Temperature (°C)

TABLE II

Extinction Coefficients at 260 and 283 nm
in 0.2M NaCl, 0.01M Phosphate Buffer, pH=7, 0.1mM EDTA[a]

Ethidium bromide:

$\varepsilon_{260} = 1.7 \times 10^4 - 25.5(T - 50)$

$\varepsilon_{283} = 5.1 \times 10^4 - 74(T - 90) + 5.1 \times 10^4 \cdot a_3 \cdot (T - 90)^3$

$a_3 = 4.86 \times 10^{-6} \cdot \sqrt{\text{concentration}} + 8.64 \times 10^{-9}$

$rCA_5G + rCU_5G$ single strands:

$\varepsilon_{260} = 1.40 \times 10^5 + 190(T-50) - 2.35(T-50)^2 - 0.0237(T-50)^3$

$\varepsilon_{283} = 4.17 \times 10^4 - 10.4(T - 50)$

$rCA_5G + rCU_5G$ double strands:

$\varepsilon_{260} = 1.14 \times 10^{5*} + 240(T)$

$\varepsilon_{283} = 4.21 \times 10^4$

Ethidium bromide bound to $rCA_5G \cdot rCU_5G$

$\varepsilon_{260} = 9.1 \times 10^3 - 15(T)$

$\varepsilon_{283} = 2.0 \times 10^{4*} + 50(T)$

$dCA_5G + dCT_5G$ single strands:

$\varepsilon_{260} = 1.41 \times 10^5 + 170(T - 50)$

$\varepsilon_{283} = 5.4 \times 10^4 + 40(T - 50)$

$dCA_5G + dCT_5G$ double strands:

$\varepsilon_{260} = 1.14 \times 10^5 + 150(T)$

$\varepsilon_{283} = 5.25 \times 10^4$

Ethidium bromide bound to $dCA_5G + dCT_5G$:

$\varepsilon_{260} = 1.4 \times 10^4 - 15(T)$

$\varepsilon_{283} = 2.2 \times 10^4 + 75(T)$

---

[a]Temperatures are in °C.

*Value varies slightly with concentration of strands.

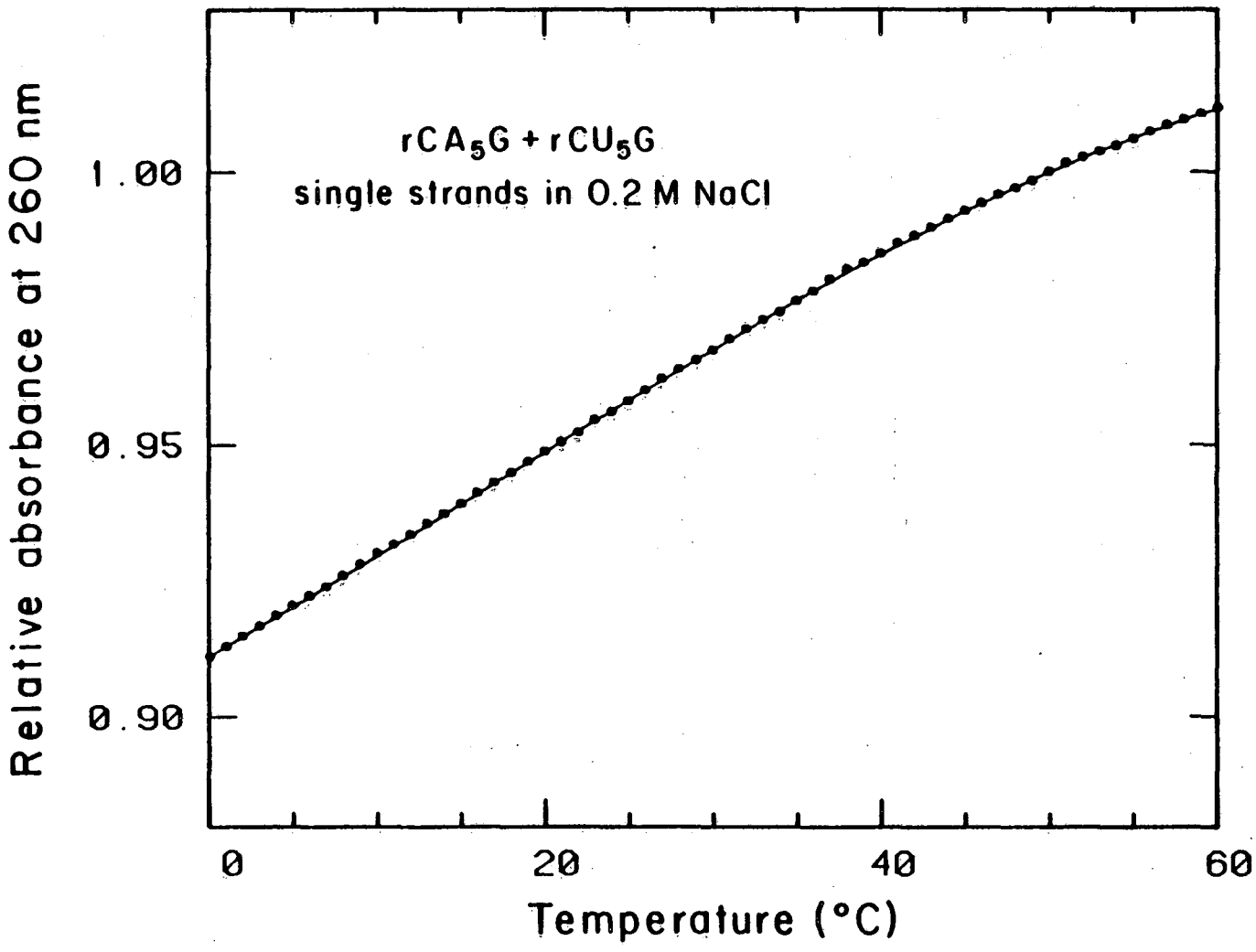
B) Oligonucleotide Optical Properties

The extinction coefficients for the single-strand mixtures of rCA_5G + rCU_5G and dCA_5G + dCT_5G are readily determined by melting the single strands separately. At 283 nm, the extinction coefficients of the ribo-oligonucleotides were linear with temperature: $\varepsilon_{283}(T) = 4.17 \times 10^4 - 10.4(T - 50°C)$. The slope contributes only a 1% change between 0°C and 50°C. At 260 nm, the extinction coefficient of the ribo-oligonucleotides shows significant curvature, and is fit well to the third order equation: $\varepsilon_{260}(T) = 1.40 \times 10^5 + 190(T - 50) - 2.35(T - 50)^2 - 0.0237(T - 50)^3$. The fit by a second order expression was significantly worse. The melting curve and fit at 260 nm are shown in figure 3.

The deoxyribo-oligonucleotide single strands are characterized by a linear extinction coefficient at 283 nm, $\varepsilon_{283} = 5.4 \times 10^4 + 40(T - 50)$. At 260 nm, the curve is also linear, $\varepsilon_{260}(T) = 1.41 \times 10^5 + 170(T - 50)$.

The fitting of the extinction coefficients of the double strands is problematic because the oligonucleotides form only moderately stable double strands. The problem is complicated by the fact that the double strands generally have a non-zero slope (Chapter II; Nelson et al., 1981; Breslauer et al., 1975; Albergo et al., 1981), and the fact that the double strands aggregate, which causes the extinction coefficient of the double strands to depend on concentration (Chapter II; Nelson et al., 1981). However, since the absorbance at 283 nm changes very little for both the ribo- and deoxyribo-oligonucleotides studied here, the errors in the extinction co-

Figure 3. The melting curve of the single strands of rCA_5G + rCU_5G obtained by melting the strands separately. The curves are normalized by dividing by the absorbance at 50°C. The line shows the fit using the equation $\varepsilon_{283}(T)/\varepsilon_{260}(50°C) = 1 + 1.36 \times 10^{-3}\cdot(T - 50) - 1.67 \times 10^{-5}\cdot(T - 50)^2 - 1.69 \times 10^{-7}\cdot(T - 50)^3$ (see text).


rCA5G + rCU5G
single strands in 0.2 M NaCl
Relative absorbance at 260 nm
1.00
0.95
0.90
0
20
40
60
Temperature (°C)

efficients of the double strands at 260 nm have little effect on the determination of the fraction ethidium bound.

The spectral effects of the melting from double strands to single strands are shown by the spectra at 0°C, 20°C, and 50°C for rCA_5G + rCU_5G in figure 4, and for dCA_5G + dCT_5G in figure 5. It can be seen that both oligonucleotides change in absorbance very little at 283 nm upon melting to single strands. Thus we find that for the double-stranded ribo-oligonucleotides, $\varepsilon_{283} = 4.21 \times 10^4$, and for the deoxyribo-oligonucleotides, $\varepsilon_{283} = 5.25 \times 10^4$, both assumed independent of temperature.

The slope of the extinction coefficient at 260 nm of the double strands formed by rCA_5G + rCU_5G was determined from melting curves at 1mM concentration. The resulting expresson is: $\varepsilon_{260}(T) = \varepsilon_{260}(0°C) + 240(T)$, with $\varepsilon_{260}(0°C)$ depending slightly upon concentration: 1.13×10^5 at 50μM, 1.135×10^5 at 25μM, and 1.15×10^5 at 12μM.

The deoxyribo-oligonucleotides were studied at only one concentration. The expression for the double strands at 40μM was determined to be $\varepsilon_{260}(T) = 1.14 \times 10^5 + 150(T)$. The extinction coefficients are included in Table II.

C) <u>Interaction of Ethidium Bromide with rCA_5G + rCU_5G</u>

The spectral effects of ethidium binding to the ribo-oligonucleotides are shown in figure 6. At high temperatures, the spectrum resembles the sum of the free ethidium and the oligonucleotide single strand spectra. However, at low temperature, the spectrum of the complex of the oligonucleotides with bound ethidium changes

Figure 4. Spectra of the mixture rCA_5G + rCU_5G at a concentration of each strand of 55μM. The path length was 0.1 cm. Δ = 50° C (mostly single stranded); ○ = 20° C (roughly equal double and single strands); □ = 0° C (mostly double strands).

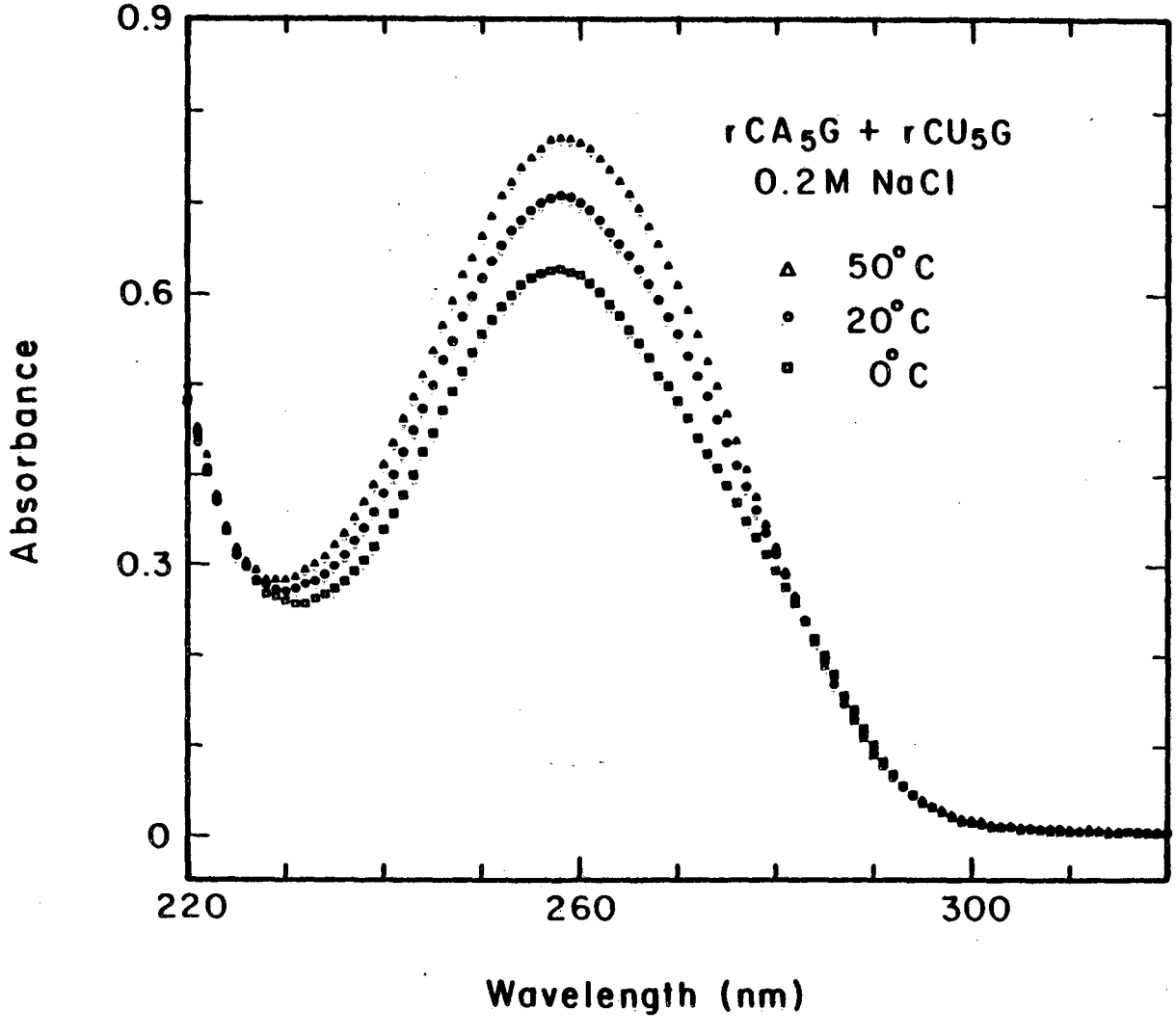

0.9
0.6
0.3
0
Absorbance
220
260
300
Wavelength (nm)
rCA5G + rCU5G
0.2M NaCl
50°C
20°C
0°C

Figure 5. Spectra of the mixture dCA_5G + dCT_5G at a concentration of each strand of 47μM. The path length was 0.2 cm. Δ = 50° C (mostly single strands); O = 20° C (roughly equal single and double strands); □ = 0° C (mostly double strands).

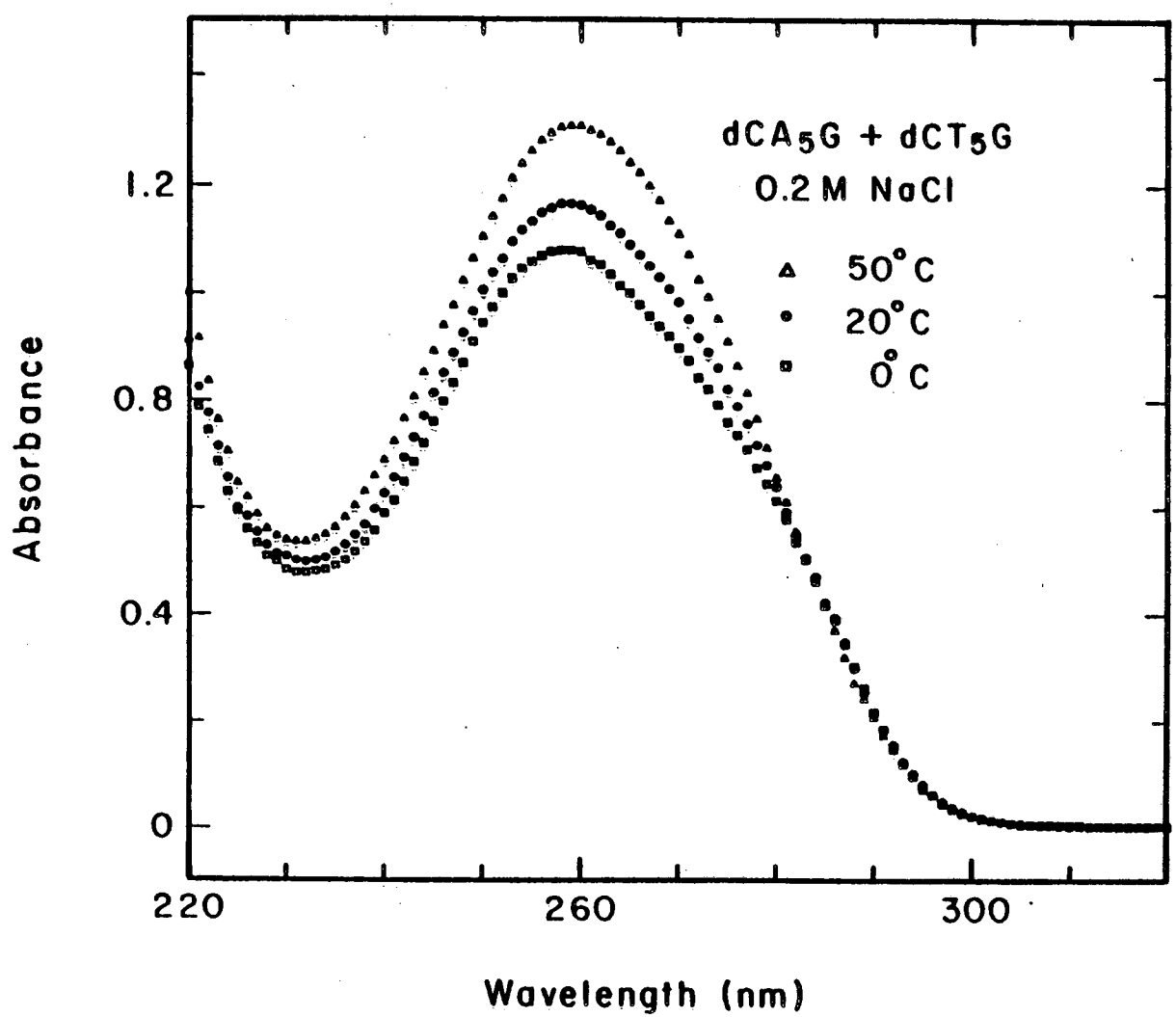
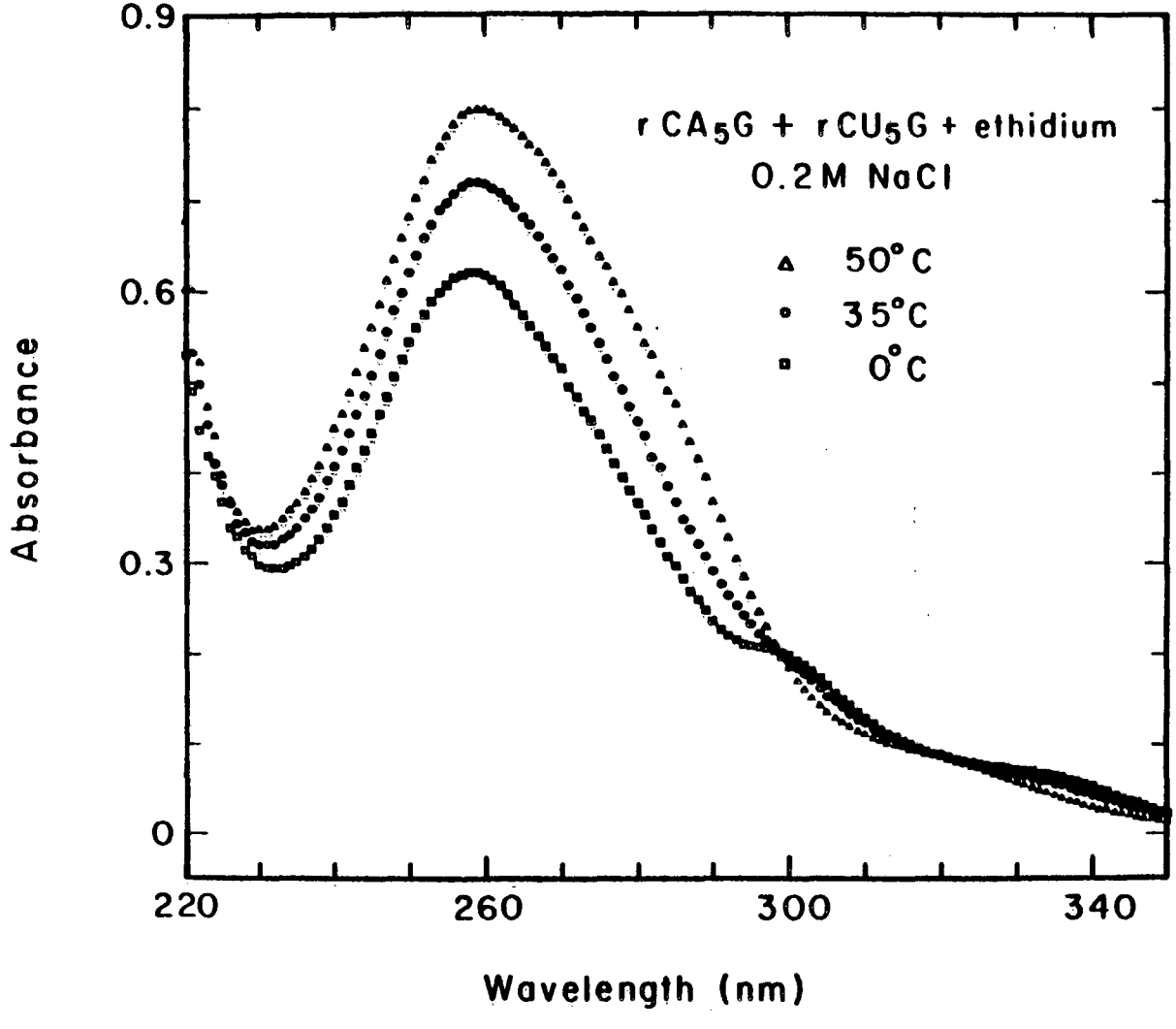

dCA5G + dCT5G
0.2 M NaCl
50°C
20°C
0°C
Absorbance
1.2
0.8
0.4
0
220
260
300
Wavelength (nm)

Figure 6. Spectra of the mixture rCA_5G + rCU_5G + ethidium. The concentrations of the strands were 50μM; the concentration of the ethidium was 55μM. The path length was 0.1 cm. △ = 50° C; ○ = 35° C; □ = 0° C.

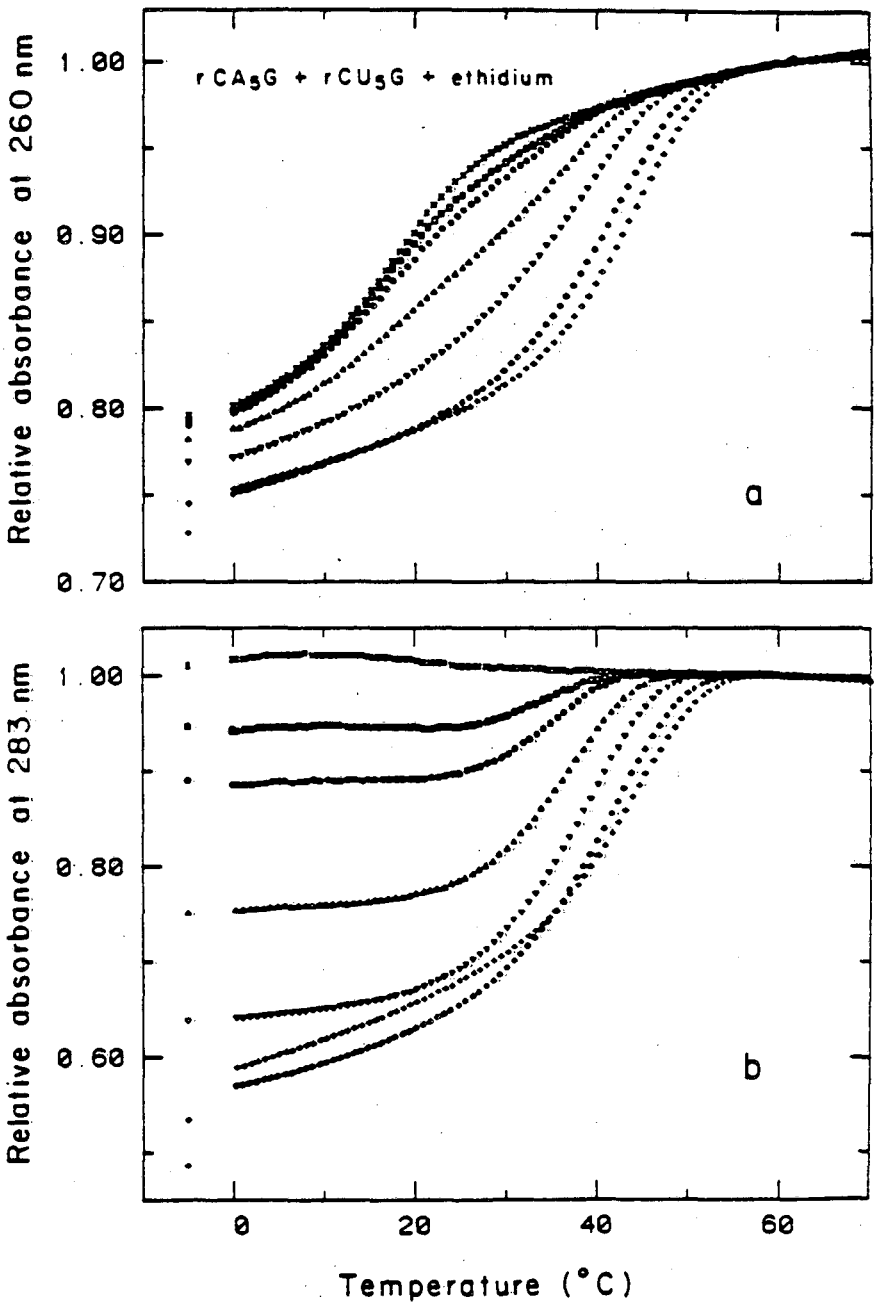

0.9
0.6
0.3
0
Absorbance
220
260
300
340
Wavelength (nm)
r CA5G + r CU5G + ethidium
0.2 M NaCl
50°C
35°C
0°C

markedly. The isosbestic point at 283 nm that characterized the oligonucleotides now exhibits significant hypochromicity upon ethidium binding. A shoulder also appears near 300 nm, which is probably due to changes in the ethidium chromophore in the complex.

By monitoring the absorbance at 260 and 283 nm from low to high temperatures, we can monitor the fraction double strands and the fraction ethidium bound, as described in the theory section. Since the melting of the double strands is nearly isosbestic at 283 nm, this wavelength monitors the state of ethidium binding. Conversely, 260 nm monitors mostly the double-strand to single-strand transition. The melting curves at 260 and 283 nm at a strand concentration of about 50μM at a range of ethidium:strand ratios are shown in figure 7. The curves at 283 nm exhibit quite clearly the sigmoidal behavior characteristic of a cooperative transition at low ratios of ethidium:strand.

Since the leveling off of the curves at low temperature and low ethidium:strand ratios clearly indicates that the ethidium is fully bound, we can use the curve at low temperature to determine the extinction coefficient of the bound ethidium. The extinction coefficient thus determined is $\varepsilon_{283}(T) = 2.0 \times 10^4 + 50(T°C)$. The extinction coefficient of bound ethidium is only 35% of that for free ethidium at 283 nm. The best fit for the curve with low ethidium:strand ratios varied for individual melting curves by about 10%. However, this is a small effect, considering the magnitude of the change.

Figure 7. (a) Melting curves at 260 nm of rCA_5G + rCU_5G + ethidium, holding the concentration of strands roughly equal and varying the ratio of ethidium:strands. The data are all normalized at 60°C. The concentrations of ethidium:strands were (μM): ×, 0:62; □, 4.4:52; ○, 9.4:51; △, 25:50; ▽, 50:49; ◇, 106:50; +, 156:49. The symbols to the left of the curves are the expected absorbances at 0°C if the double strands are fully bound and all of the ethidium is bound. (b) Melting curves at 283 nm.


rCA5G + rCU5G + ethidium
Relative absorbance at 260 nm
1.00
0.90
0.80
0.70
a
Relative absorbance at 283 nm
1.00
0.80
0.60
b
0
20
40
60
Temperature (°C)

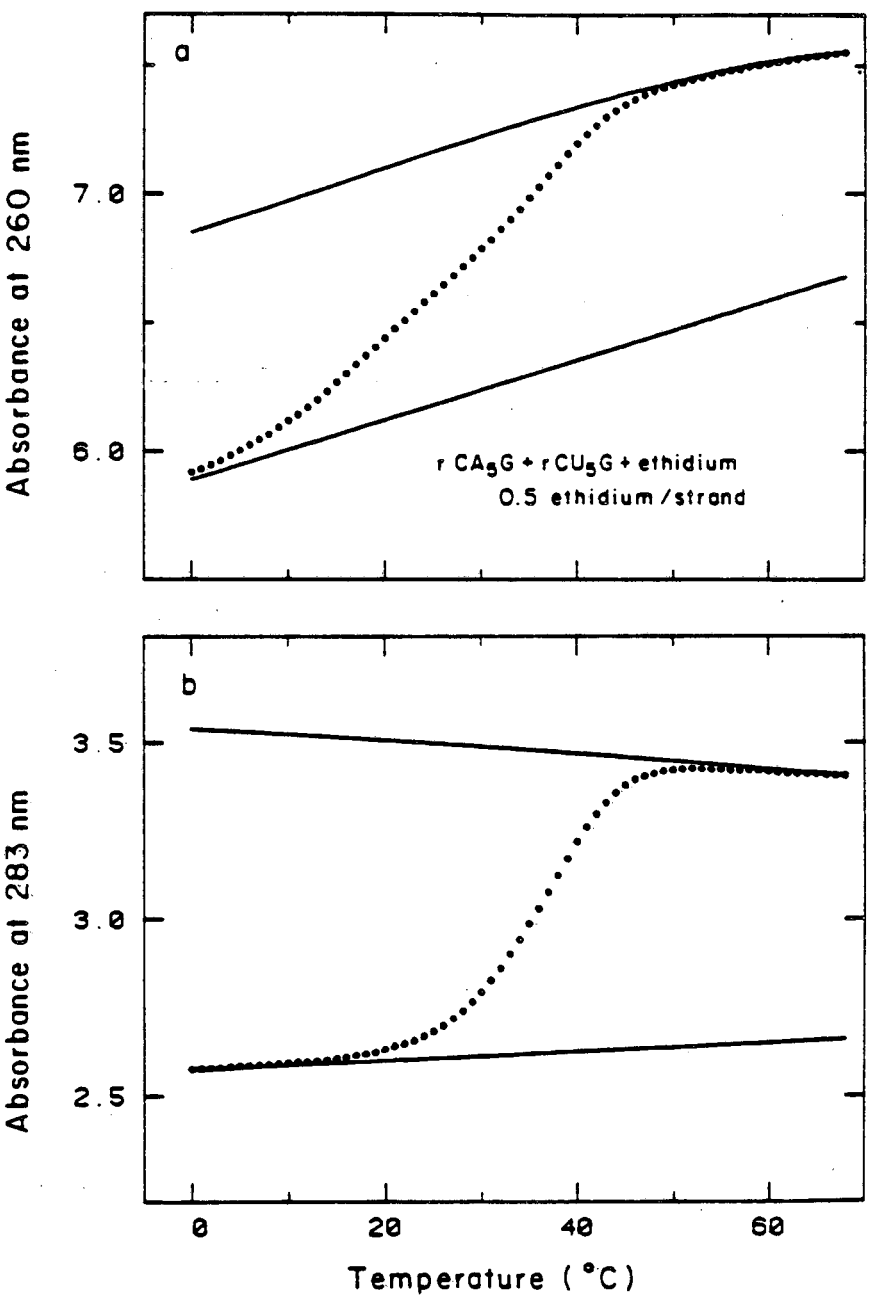
The melting curves at 260 nm are more difficult to interpret, due to the fact that for these oligonucleotides, the single-strand to double-strand transition is not fully over at 0°C in the absence of ethidium. Thus the determination of the extinction coefficient of bound ethidium is only an estimate: $\varepsilon_{260}(T) = 9.1 \times 10^3 - 15(T)$. Thus, the extinction coefficient of bound ethidium changes by about 50% at 260 nm. However, since the extinction coefficient of the oligonucleotides at 260 nm is much larger than that of ethidium, the effect of ethidium binding on the absorbance is not as large as it is at 283 nm.

As mentioned in the theory section, all of the changes in absorbance upon ethidium binding are attributed to the ethidium. This is a simplification, since the optical properties of both the oligonucleotides and the ethidium probably change upon binding ethidium. However, this arbitrary designation does not affect the determination of the extent of ethidium binding.

The melting curves and baselines discussed above are shown for a strand concentration of 50μM and a 0.5 ethidium:strand ratio at 260 and 283 nm in figure 8. These extinction coefficients are included in Table II.

Since the melting of the double strands to single strands is not completely isosbestic at 283 nm, the fraction double strands and ethidium bound were calculated by using equations 8 and 9. Assuming the melting of the strands was isobestic, and using the absorbance at 283 nm only, calculation of the fraction of ethidium bound changed by less than 1%.

Figure 8. (a) The melting curve, low temperature and high temperature baselines at 260 nm for rCA_5G + rCU_5G + ethidium. The strand concentration was 50μM; the ethidium concentration was 25μM. The path length was 0.2 cm. (b) The melting curves and baselines at 283 nm.


a
Absorbance at 260 nm
7.0
6.0
r CA5G + r CU5G + ethidium
0.5 ethidium /strand
b
Absorbance at 283 nm
3.5
3.0
2.5
0
20
40
60
Temperature (°C)

D) Qualitative Interpretation of rCA_5G + rCU_5G + Ethidium

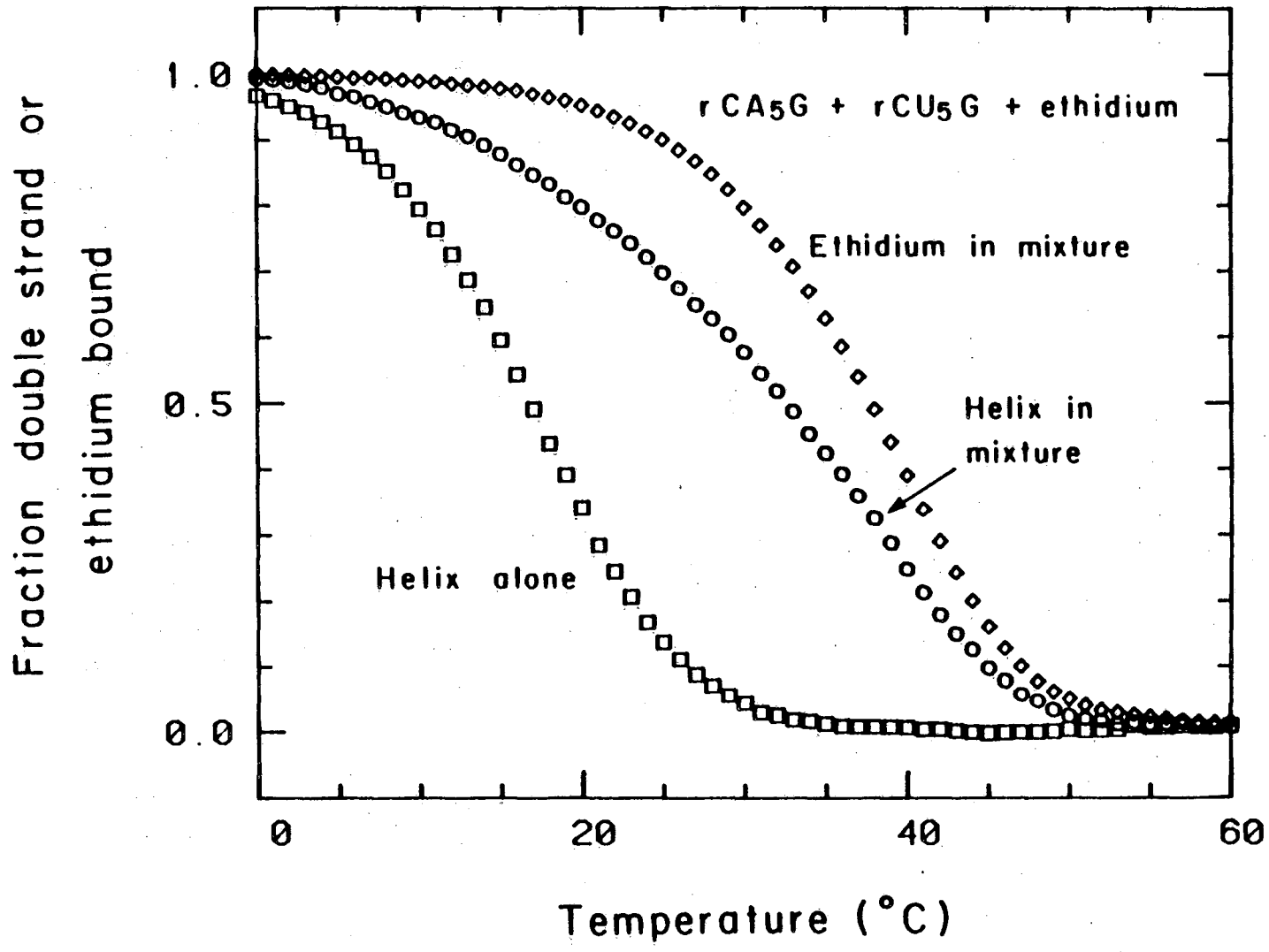
The stabilization of the double strands by ethidium binding is shown in figure 9, which shows the fraction double strands in the absence and presence of a 1:1 ratio of ethidium bromide. Also shown is the fraction ethidium bound. Several features are apparent from this figure.

1) The melting of the double strands is shifted to higher temperature and is broadened significantly when ethidium bromide is present.
2) The ethidium melting curve occurs at higher temperature than the strands, and is a sharper transition.
3) The strands in the presence of ethidium are significantly double-stranded at temperatures where essentially no double strands form in the absence of ethidium bromide.

We can describe the process shown in the melting curves going from high to low temperature as follows:

At high temperatures, where the double strands in the presence of ethidium are just forming, essentially every double strand has at least one bound ethidium. At this point, there is a large excess of ethidium relative to the double strands, and more than one ethidium binds. For example, when the ethidium is approximately 50% bound, only roughly 33% of the strands are double-stranded. Thus, every double strand has an average of approximately 1.5 ethidium ions

Figure 9. The fraction double strands formed by rCA_5G + rCU_5G alone (□), the fraction double strands in a mixture of strands and ethidium (○), and the fraction ethidium bound in the mixture (◇). The concentration of strands alone was 62μM. The concentration of strands in the mixture was 49μM, the concentration of ethidium was 50μM.

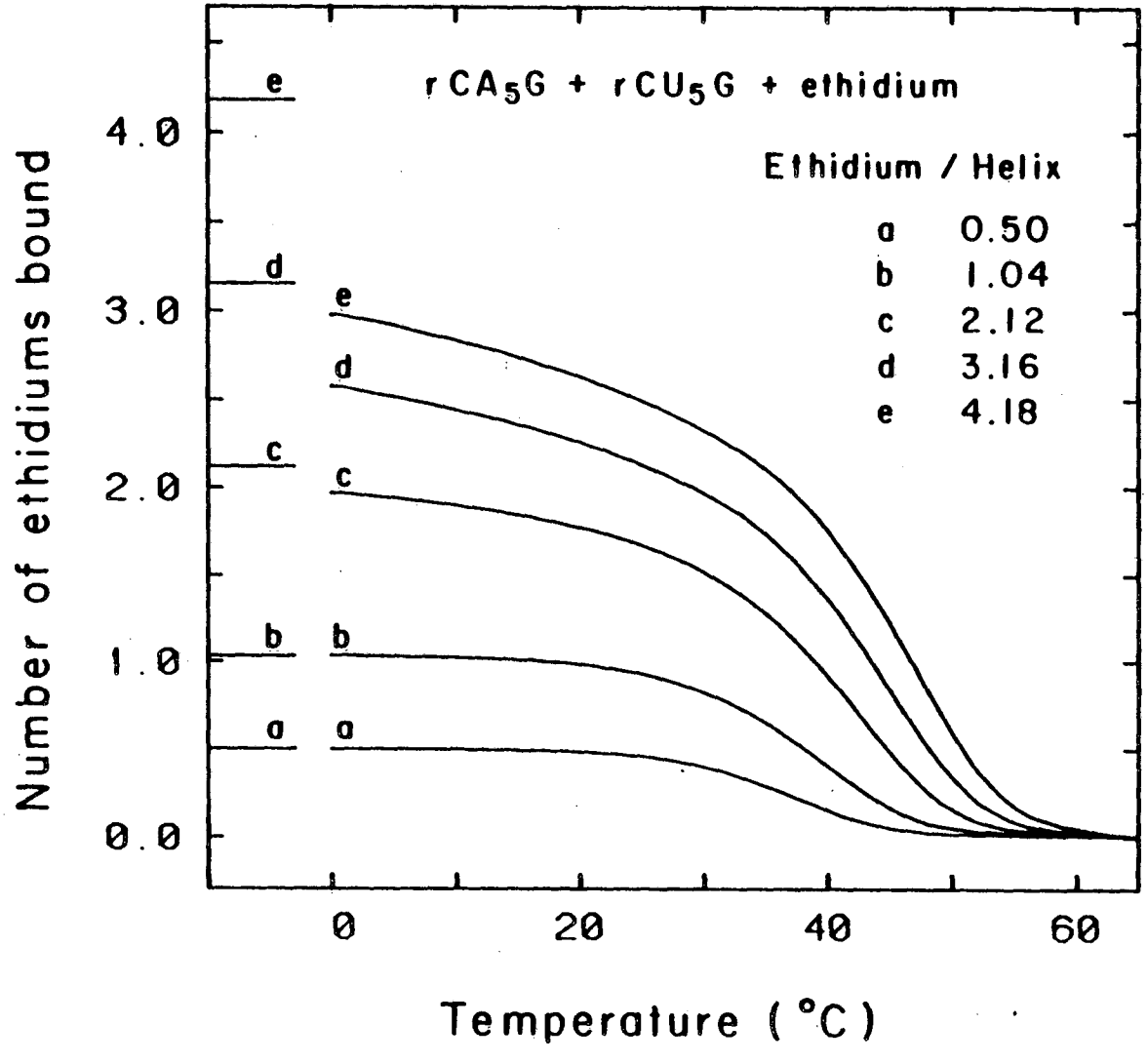

Fraction double strand or ethidium bound
1.0
0.5
0.0
0
20
40
60
Temperature (°C)
rCA5G + rCU5G + ethidium
Ethidium in mixture
Helix in mixture
Helix alone

bound. When half of the strands are double-stranded, approximately 73% of the ethidium is bound, hence approximately 1.5 ethidium ions bind per double strand. This behavior continues, until the point where 90% of the ethidium is bound, and approximately 70% of the strands are double-stranded. Binding more than one ethidium ion per double strand explains why the ethidium melting curve is sharper than the oligonucleotide melting curve.

Further qualitative results may be derived by looking at the behavior of ethidium binding at different ratios of ethidium:strand. Figure 10 shows the melting curves at ethidium:strand ratios of approximately 0.5, 1, 2, 3, and 4, at a constant strand concentration of 50μM. The binding curves at ethidium:strand ratios of 0.5 or 1 show all of the ethidium is bound between 0°C and about 15°C. The second ethidium ion does not bind as strongly, indicated by the fact that all of the ethidium is not bound at 0°C at an ethidium:strand ratio of 2.12. The third ethidium ion binds even less strongly, since at an ethidium:strand ratio of 3.16 at 0°C, an average of only about 2.6 ethidium ions are bound per double strand. Figure 10 also shows very clearly that when 4 ethidium ions are present, only three bind. This confirms the assertion that ethidium binds with nearest-neighbor exclusion.

Thus, we can conclude that the first ethidium binds very strongly; this binding is coupled to the formation of a double strand from single strands, which is a very cooperative process. The second ethidium binds to an already-formed double strand, and does so readily but not as strongly. The third ethidium binds much

Figure 10. The number of ethidium ions bound at a roughly constant strand concentration of 49-53μM at ethidium:strand ratios of (a) 0.50; (b) 1.04; (c) 2.12; (d) 3.16; and (e) 4.18. The lines to the left of the curves indicate the input ethidium:strand ratios.


rCA5G + rCU5G + ethidium
Ethidium / Helix
a 0.50
b 1.04
c 2.12
d 3.16
e 4.18
Number of ethidiums bound
4.0
3.0
2.0
1.0
0.0
0
20
40
60
Temperature (°C)

more weakly, and a fourth ethidium is excluded from binding altogether.

Quantitative analysis of the results to determine the magnitude of the binding constants will be presented in the discussion.

E) Interaction of Ethidium Bromide with $dCA_5G + dCT_5G$

The spectral effects of ethidium binding to the deoxyribo-oligonucleotides are shown in figure 11. The features are similar to the spectra of ethidium bound to $rCA_5G + rCU_5G$, except that the shoulder at about 300 nm is not as pronounced with the deoxyribo-oligonucleotides. Figure 12 shows the melting curves at 260 and 283 nm, respectively, at a strand concentration of about 40μM, with ethidium:ratios of about 0.4, 0.6, and 0.8.

The melting curves at 283 nm do not level off at low temperatures, which indicates that ethidium does not bind as strongly to the deoxyribo-oligonucleotides as it does to the ribo-oligonucleotides. This makes it much more difficult to determine the extinction coefficient for the bound ethidium, since all of the ethidium is not bound at 0°C. The values of the extinction coefficients for bound ethidium at 260 and 283 nm were determined by an iterative process whereby a value was estimated, and the fraction double strands and ethidium bound were calculated. This was fit to the model, and the calculated and measured curves of the fraction of ethidium bound were compared. The value for the extinction coefficients were varied until the agreement was good. This procedure to determine extinction coefficients is not as direct as that used for the ribo-oligonucleotides and could potentially bias the results

Figure 11. Spectra of the mixture dCA_5G + dCT_5G + ethidium. The concentration of strands was 46μM; the concentration of the ethidium was 48μM. The path length was 0.2 cm. △ = 50° C; ○ = 30°; □ = 0° C.

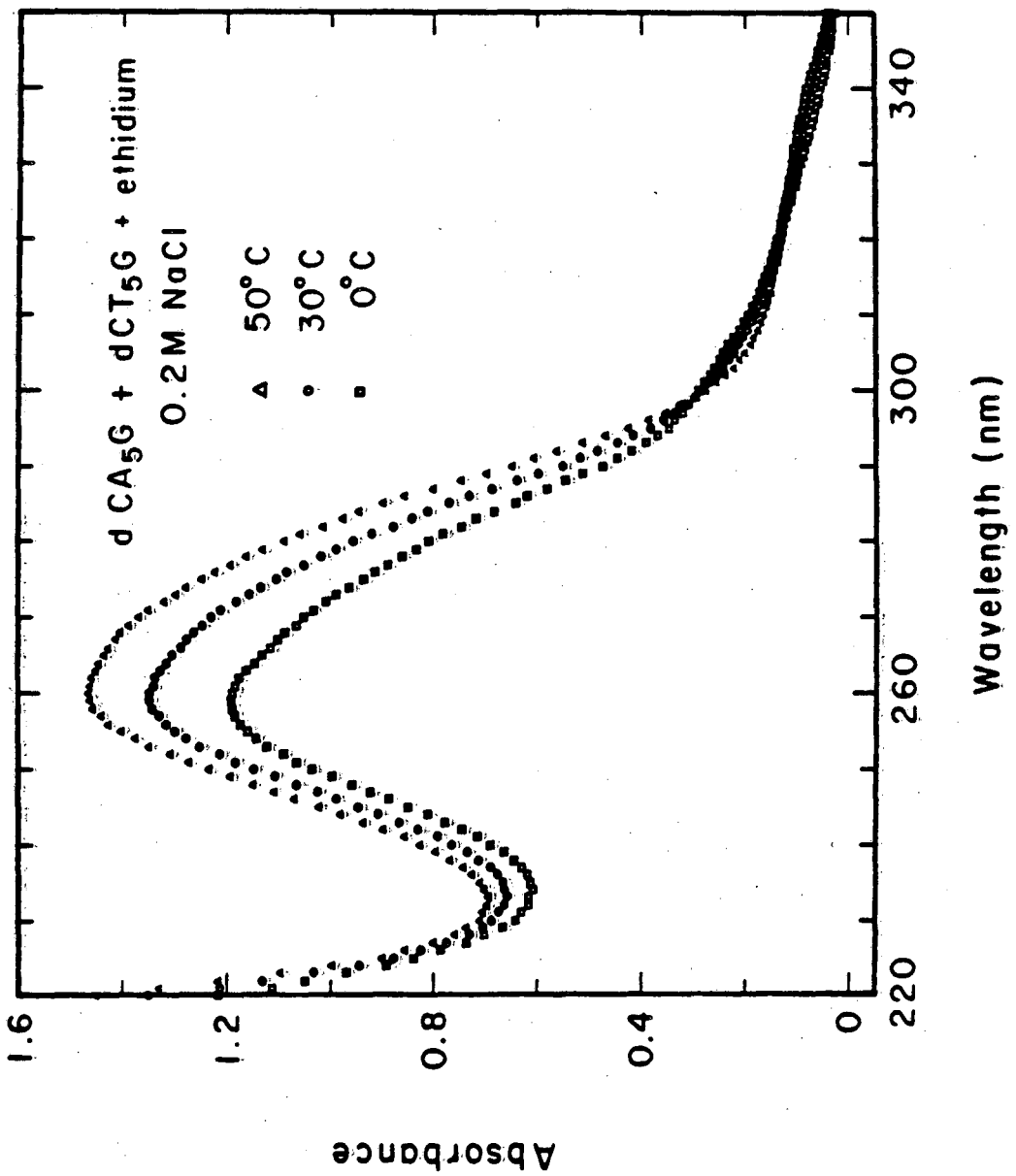
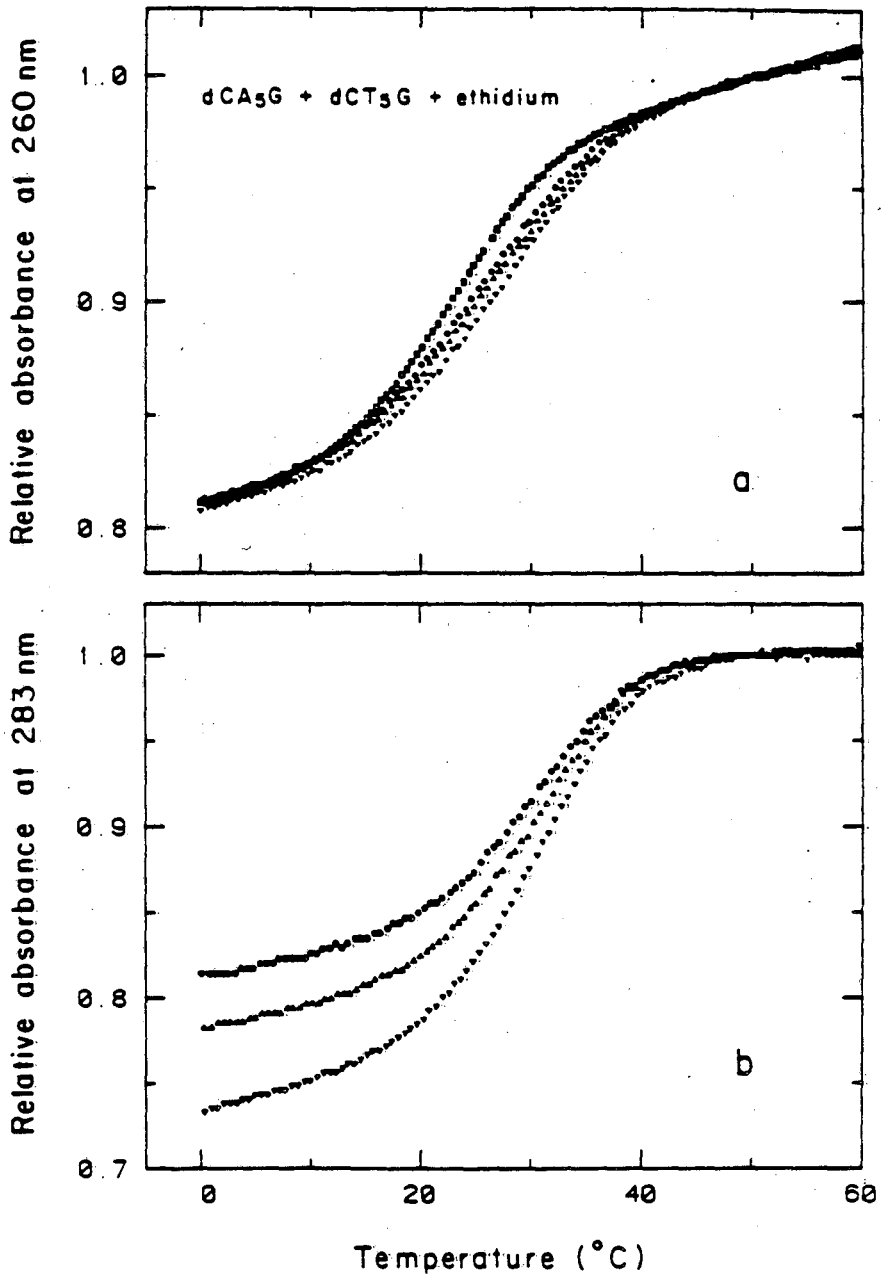

d CA5G + dCT5G + ethidium
0.2M NaCl
50°C
30°C
0°C
Absorbance
1.6
1.2
0.8
0.4
0
220
260
300
340
Wavelength (nm)

Figure 12. (a) Melting curves at 260 nm of dCA_5G + dCT_5G + ethidium, holding the concentration of strands roughly equal and changing the ratio of ethidium:strands. The data are all normalized at 50°C. The concentrations of ethidium:strands (μM) were: □, 0:44; ○,19:44; △, 24:42; ▽,31:40. (b) Melting curves at 283 nm.


dCA5G + dCT5G + ethidium
Relative absorbance at 260 nm
1.0
0.9
0.8
a
Relative absorbance at 283 nm
1.0
0.9
0.8
0.7
b
0
20
40
60
Temperature (°C)

to fit the model (see Discussion). The resulting extinction coefficients of bound ethidium are $\varepsilon_{260} = 1.4 \times 10^4 - 15(T)$ and $\varepsilon_{283} = 2.2 \times 10^4 + 75(T)$, and are included in Table II. Figure 13 shows the melting curves and base lines at 260 and 283 nm at a strand concentration of 42μM and an ethidium:strand ratio of 0.6.

Figure 14 shows the fraction double-strands and fraction ethidium bound at an ethidium:strand ratio of 0.8. Comparing this with figure 9, it is clear that ethidium binding does not stabilize the double strands of the deoxyribo-oligonucleotides nearly as much as it stabilizes the ribo-oligonucleotides. The comparison of the strength of ethidium binding to the deoxyribo-oligonucleotides is shown more clearly in figure 15, which shows the binding at different ethidium:strand ratios at comparable strand concentrations, roughly 40μM for the deoxyribo- and 50μM for the ribo-oligonucleotides. Clearly, the binding is weaker for the deoxyribo-oligonucleotides. The double strands formed by these deoxyribo-oligonucleotides are more stable than the ribo-oligonucleotides (Chapter II; Nelson et al., 1981), so this lower binding to the deoxyribo-oligonucleotides is not explained by differences in double strand stability. The determination of the binding constants is explained in the Discussion.

6. Discussion

From the results presented so far, we have determined that the extent of ethidium binding can be determined rather directly by monitoring the melting curves of the oligonucleotides in the presence of ethidium bromide. Two features make this possible. The

Figure 13. (a) The melting curve, low temperature and high temperature baselines at 260 nm for dCA_5G + dCT_5G + ethidium. The strand concentration was 42μM; the ethidium concentration was 24μM. The path length was 0.1 cm. (b) The melting curves and baselines at 283 nm.

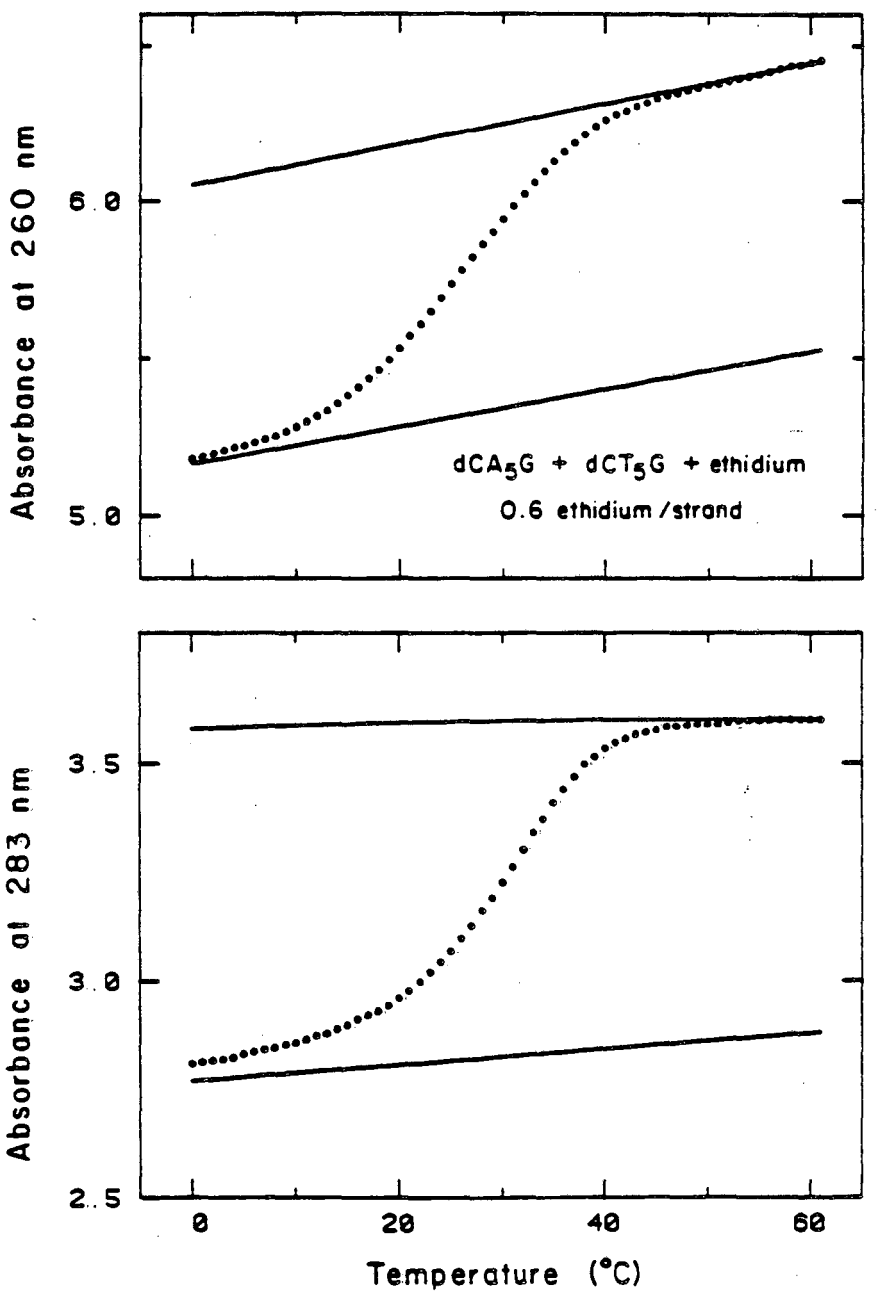

Absorbance at 260 nm
6.0
5.0
dCA5G + dCT5G + ethidium
0.6 ethidium /strand
Absorbance at 283 nm
3.5
3.0
2.5
0
20
40
60
Temperature (°C)

Figure 14. The fraction double strands formed by dCA_5G + dCT_5G alone (□), the fraction double strands in a mixture of strands and ethidium (O), and the fraction ethidium bound in the mixture (◇). The concentration of the strands alone was 44μM. The concentration of the strands in the mixture was 40μM; the ethidium concentration in the mixture was 31μM.

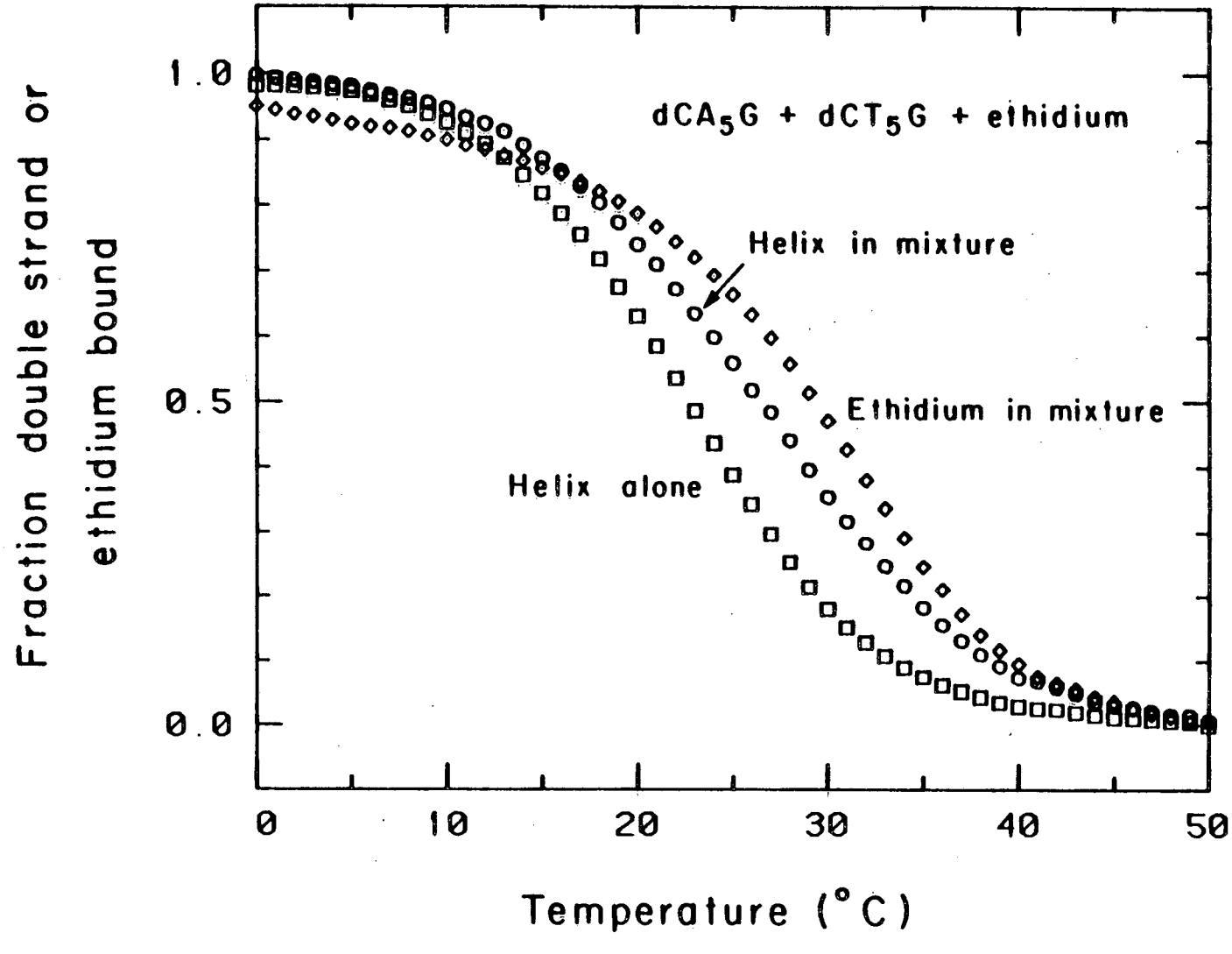
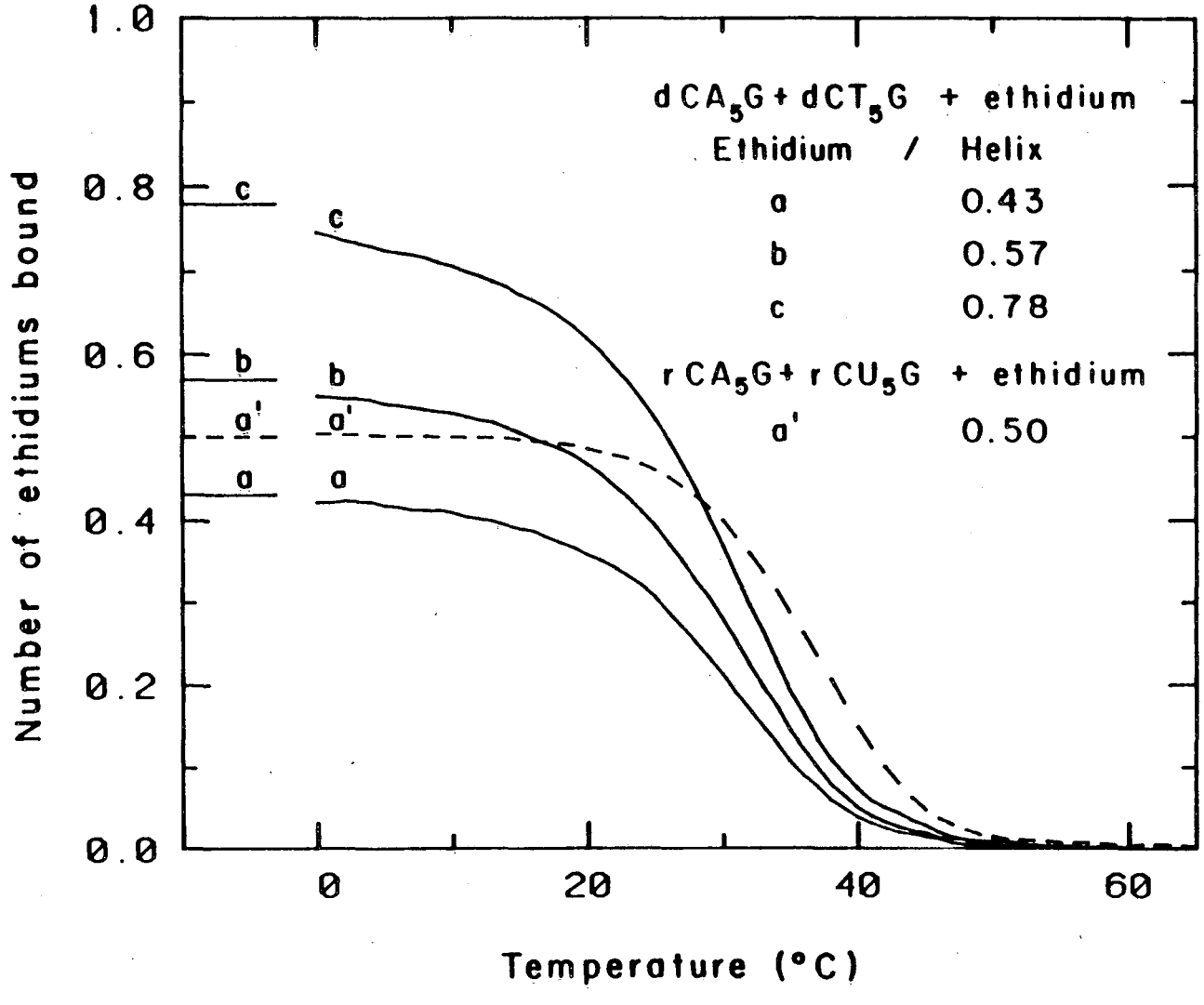

dCA5G + dCT5G + ethidium
Helix in mixture
Ethidium in mixture
Helix alone
Fraction double strand or ethidium bound
1.0
0.5
0.0
0
10
20
30
40
50
Temperature (°C)

Figure 15. The number of ethidium ions bound *vs.* temperature at a roughly constant strand concentration of dCA_5G + dCT_5G (40μM, ———) at ethidium:strand ratios of (a) 0.43; (b) 0.57; and (c) 0.78. The corresponding plot for rCA_5G + rCU_5G (50μM, - - - -) at an ethidium:strand ratio of 0.50. The lines to the left of the curves indicate the input ethidium:strand ratios.


1.0
0.8
0.6
0.4
0.2
0.0
Number of ethidiums bound
0
20
40
60
Temperature (°C)
c
b
a'
a
dCA5G + dCT5G + ethidium
Ethidium / Helix
a 0.43
b 0.57
c 0.78
rCA5G + rCU5G + ethidium
a' 0.50

double-strand to single-strand transitions of these oligonucleotides exhibit essentially no change in absorbance at 283 nm. Also, this is the wavelength of the maximum absorbance in the UV for ethidium bromide. The absorbance decreases dramatically upon intercalating into the double strands.

In the model presented above, the assumption was made that ethidium bromide binds only to double helices by intercalation. It is well known that the binding of ethidium to single strands is very much weaker than binding to double strands. Ethidium binding studies on homopolymers showed very clearly the dramatic increase in ethidium binding to the double-stranded poly(A)•poly(U) relative to the binding to either single-stranded poly(A) or poly(U) (Waring, 1966). The same result was obtained for the deoxyribo-dinucleotides dCpA and dTpG, wherein binding of the ethidium bromide as measured by fluorescence increased greatly when the non-selfcomplementary dinucleotides were mixed, relative to the separate dinucleotides (Reinhardt and Krugh, 1978).

In studies on the binding of ethidium bromide to the tetranucleotide dC-G-C-G, Kastrup et al. (1978) determined from circular dichroism measurements that two ethidium ions can bind to the ends of the double helices. This binding was much weaker than intercalation, and occurred to a significant degree only if the ratio of ethidium:strand became large. Since the ethidium:strand ratio for all of the melting curves analyzed in the present study were always 3 or less, outside binding probably contributes very little to the binding of ethidium. Further justification is seen in figure 10,

where three ethidium ions clearly bind when an excess of ethidium is present, using the extinction coefficients of bound ethidium determined at low ethidium:strand ratios.

The assumption that the extinction coefficient of the bound ethidium does not depend on which site is filled, and that two ethidium ions bound to a double strand absorb twice as much as one ethidium bound, are more difficult to verify. However, nothing in the analysis of the data suggests that this assumption is not valid.

The question I want to answer in this section is: what information can we obtain, using the experimental data and the theory presented, about the magnitudes of the binding constants and the sequence specificity of ethidium binding?

The procedure to do this was set out in the Theory section. Briefly reiterating, we will assume a value for σ, the factor which describes the relative strength of the pyrimidine-purine site to the purine-purine sites. Using the experimentally-determined value for f_b, the fraction of ethidium bound, the total concentrations, and K_h, the equilibrium constant for forming double strands from single strands, the theory allows us to calculate a value of K_d, the equilibrium constant for ethidium binding to a purine-purine site. The best value for all of the data points at several temperatures and several concentrations of strands and ethidium will be determined by doing linear least-squares regression on $\ln(K_d)$ vs $1/T$. The slope of such a plot is $\Delta H°/R$, using the van't Hoff relation, and the intercept is $\Delta S°/R$.

The criterium for the goodness-of-fit will be r, the linear correlation coefficient, and the reduced chi-squared for the fit, $\chi_\nu^2 = [f_b(calc) - f_b(expt)]^2 / (N - 2)$ where N is the number of data points for all of the melting curves. Both criteria predict the same model parameters, however χ_ν^2 is more sensitive than r.

A) <u>rCA_5G + rCU_5G + Ethidium at Low Ethidium:Strand Ratios</u>

The results obtained for several values of σ are shown in Table III for ethidium binding to rCA_5G + rCU_5G. Nine melting curves were used in this analysis, all with ethidium:strand ratios between about 0.08 and 1. The concentrations of ethidium:strands (μM) were: 1) 4.4:52; 2) 9.4:51; 3) 4.3:25; 4) 25:50; 5) 12.2:25; 6) 5.8:12.4; 7) 50:49; 8) 25:25; and 9) 12.2:12.1. The data were analyzed between f_b = 0.2 and 0.8, where the accuracy is the greatest. The values used for the thermodynamics of double-strand formation were $\Delta H°$ = -43 kcal/mol and $\Delta S°$ = -128 e.u. (Chapter II; Nelson et al., 1981).

It can be seen from Table III that the enthalpy for ethidium binding, $\Delta H_d°$, varies very little, from -14.3 to -14.6 kcal/mol for values of σ ranging from 1 to 80. However, the equilibrium constant, K_d, varies from 18 x 10^4 with σ = 1, to 2.9 x 10^4 with σ = 80. This is because the strength of the purine-purine sites can decrease as the pyrimidine-purine site increases, and still bind the ethidium to the same extent. However, by comparing r and χ_ν^2, it is clear that the best fit occurs with σ = 20. The fit is noticably worse for σ = 10 and σ = 40.

TABLE III

Ethidium Binding to $rCA_5G + rCU_5G$
at Low Ethidium:Strand Ratios
One Strong Site Model

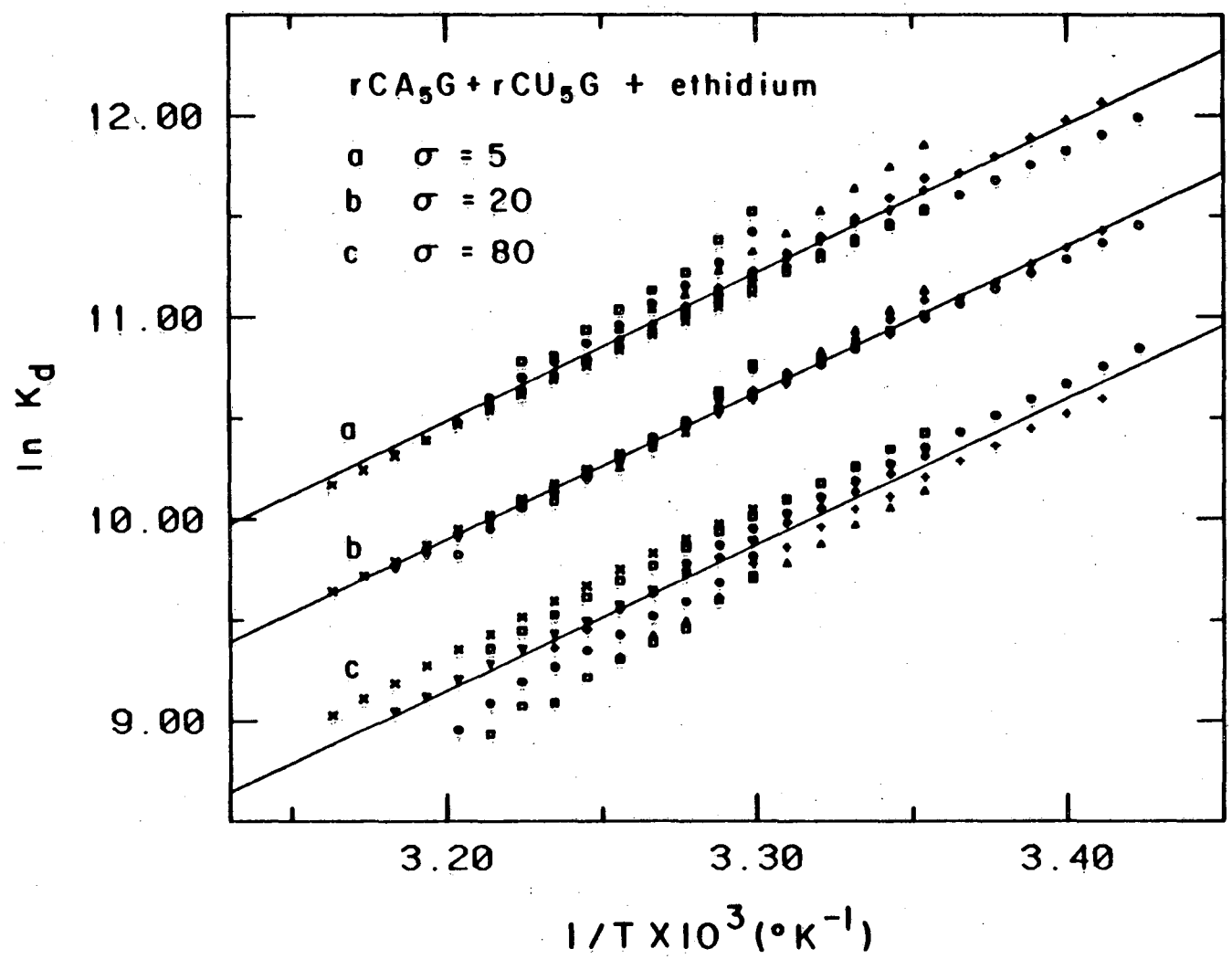
σ	ΔH_d° (kcal/mol)	ΔS_d° (e.u.)	K_d(25°C) ($\times 10^4$ ℓ mol^{-1})	σK_d	r	χ_ν^2 ($\times 10^4$)
1	-14.6	-24.9	18	18	0.977	5.0
5	-14.6	-25.9	11.2	56	0.979	4.2
10	-14.5	-26.2	8.4	84	0.988	2.1
20	-14.4	-26.6	6.1	122	0.996	0.7
40	-14.4	-27.2	4.3	172	0.988	2.6
80	-14.4	-27.8	2.9	232	0.950	10.5

The effect of σ on the plot of $\ln(K_d)$ vs. $1/T$ is shown in figure 16. Each melting curve uses a different symbol, with some repetitions. However, it is possible to follow several of the melting curves. The fit is visibly best using a value of $\sigma = 20$. Also apparent is the observation that the enthalpy changes very little with different values of σ, whereas the equilibrium constant changes noticeably. This figure gives us visual justification for using the behavior of $\ln(K_d)$ vs. $1/T$ to determine how good the fit is.

The behavior of the plot of $\ln(K_d)$ vs. $1/T$ with different values of σ can be explained by comparing the data at different ratios of ethidium:strand. The ratio of the concentrations of double strands with two ethidium ions bound to those with one bound is just the ratio of the statistical weights, namely $P_2/P_1 = S(g_2/g_1) = K_d C_f(4\sigma + 6)/(\sigma + 5)$. At very low ethidium concentrations, the value of S is small, and hence $P_2 \approx 0$. Thus, we measure the apparent equilibrium constant for binding one ethidium ion to the double strands, $K'_{app} = (\sigma + 5)K'_d$. (The prime indicates a determination at low ethidium concentration.) Since the value of K'_{app} does not change, the value calculated for K'_d will vary inversely as $\sigma + 5$. Thus increasing σ will decrease the value determined for K'_d.

However, if the concentration of ethidium is increased, the concentration of double helices with two bound ethidiuim ions increases, and the ratio $C_f K_d(4\sigma + 6)/(\sigma + 5)$ becomes important. As σ increases, the ratio $(4\sigma + 6)/(\sigma + 5)$ increases, however the value of K'_d calculated at the low ethidium concentration decreasses. At

Figure 16. Plots of $\ell n(K_d)$ vs. 1/T for rCA_5G + rCU_5G + ethidium at low ethidium:strand ratios, assuming various values of σ. The concentrations are given in the text. (a) $\sigma = 5$; (b) $\sigma = 20$; (c) $\sigma = 80$. The results are tabulated in Table III.

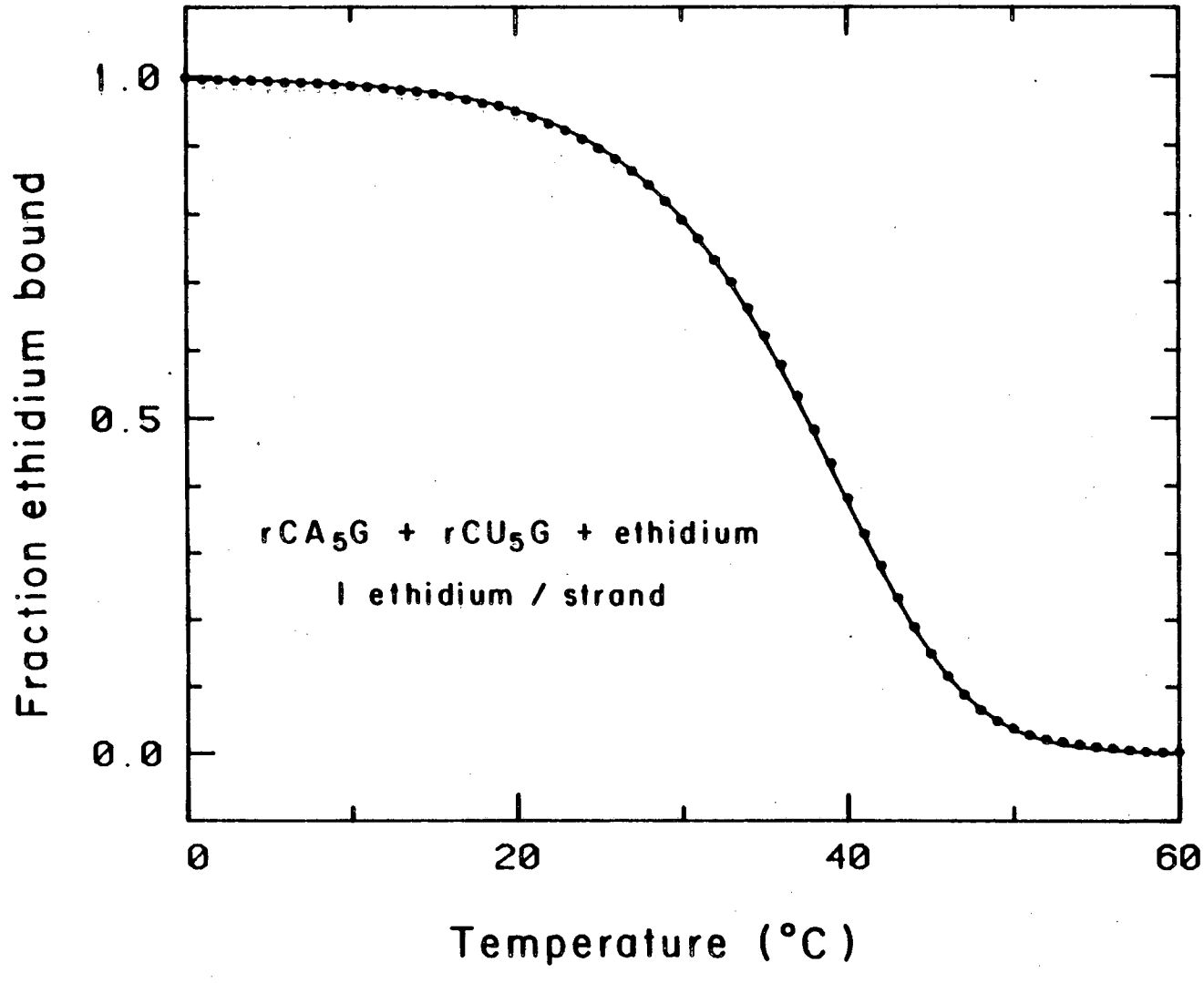

rCA5G + rCU5G + ethidium
a σ = 5
b σ = 20
c σ = 80
a
b
c
12.00
11.00
10.00
9.00
ln Kd
3.20
3.30
3.40
1/T X10^3 (°K^-1)

values of $\sigma = 1$ and $\sigma = 10$, the values of $K_d'(4\sigma + 6)/(\sigma + 5)$ are $0.28K_{app}'$ and $0.20K_{app}'$, respectively. Thus, at the larger value of σ, the binding of the second ethidium becomes less favorable. As σ increases, the value of K_d must be increased relative to K_d' to fit the high ethidium concentration data. If the guessed value of σ is higher than the "true" value, the plot of $\ln(K_d)$ vs. 1/T for the high ethidium concentration will lie above that for low concentrations. The best fit for σ results in the plot of $\ln(K_d)$ vs. 1/T in which the curves at different ethidium concentrations are co-linear.

Figure 17 shows the experimental and calculated curves for the fraction of ethidium bound using the thermodynamic parameters corresponding to $\sigma = 20$ (Table III), with ethidium and strand concentrations of 50μM and 49μM, respectively. The fit is seen to be extremely good.

In this, I have assumed that the value of σ is independent of temperature. Thus the contribution of σ was assumed to be entirely entropic, $\sigma = \exp(\Delta S_o^\circ/R)$. We can also make the assumption that the contribution was entirely enthalpic, $\sigma = \exp(\Delta H_o^\circ/RT)$. This assumption makes no changes in the determination of the value of σ; however it increases the calculated enthalpy, by about 0.8 kcal, and leaves the magnitude of K_d unchanged. This is a reasonable result, because if binding to the stronger pyrimidine-purine site contributes additional enthalpy to the complex, the purine-purine sites can make a proportionately lower contribution. However, the fit to the data cannot distinguish the value of the enthalpic contribution to σ, and hence we will continue to assume it is zero.

Figure 17. The experimental and calculated curves for rCA_5G + rCU_5G + ethidium, using a value of $\sigma = 20$ (see Table III). The concentration of the strands was 49μM; the concentration of the ethidium was 50μM.


1.0
0.5
0.0
Fraction ethidium bound
rCA5G + rCU5G + ethidium
1 ethidium / strand
0
20
40
60
Temperature (°C)

Errors in the thermodynamics of the single-strand to double-strand transition can also contribute errors to the determination of σ. However, by using values of -47 kcal/mol and -141 e.u. for ΔH_h° and ΔS_h° instead of -43 kcal/mol and -128 e.u., respectively, the value for the enthalpy of ethidium binding changes by about 2 kcal/mol, and the equilibrium constant changes by about 10%. The best value for σ remains unchanged. Since the errors in the value of ΔH_h° were estimated to be 10% in chapter II, this gives us a reasonable estimate of the error in the determination of the binding parameters for ethidium binding.

We can postulate different models for ethidium binding to these oligonucleotides, and attempt to verify their validity by comparing the fits between models. One possiblility we may address is whether the last purine-purine site, $\begin{smallmatrix} C-A \\ \bullet\ \ \bullet \\ G-U \end{smallmatrix}$, has a different binding constant than the other purine-purine sites. Thus, we can test the model that both terminal binding sites, $\begin{smallmatrix} C-A \\ \bullet\ \ \bullet \\ G-U \end{smallmatrix}$ and $\begin{smallmatrix} A-G \\ \bullet\ \ \bullet \\ U-C \end{smallmatrix}$, are stronger than the interior binding sites by the factor τ. The statistical factors for this model are $g_0 = 1$, $g_1 = 2\tau + 4$, $g_3 = \tau^2 + 6\tau + 3$, and $g_3 = 2\tau^2 + 2\tau$. Table IV shows the results of the calculations for different values of τ. The best fit to the data are for τ between about 40 and 200, with the equilibrium constant of the strong sites, $\tau K_d(25°C)$ between 56 and 64 x 10^4. As the value of τ approaches ∞, the model reduces to one with two independent binding sites. In this case, the binding constant at 25°C is 66 x 10^4. If $\tau = \infty$, $\tau K_d(25°C) = 66 \times 10^4$. Thus the data at low ethidium:strand ratios are consistent with two strong binding sites on the ends of the

TABLE IV

Ethidium Binding to $rCA_5G + rCU_5G$
at Low Ethidium:Strand Ratios
Two Strong Sites Model

τ	ΔH_d° (kcal/mol)	ΔS_d° (e.u.)	K_d(25°C) ($\times 10^4$ ℓ mol^{-1})	τK_d	r	χ_ν^2 ($\times 10^4$)
1	-14.6	-24.9	18.0	18	0.977	5.0
5	-14.6	-26.8	7.4	37	0.973	5.9
10	-14.5	-27.4	4.5	45	0.980	4.1
20	-14.4	-28.1	2.6	52	0.987	2.3
40	-14.2	-28.8	1.4	56	0.992	1.4
80	-14.1	-29.5	0.77	62	0.993	1.3
200	-14.0	-30.8	0.32	64	0.991	1.6
∞*	-13.8	---	---	66	0.988	2.3

*Corresponds to two independent binding sites.

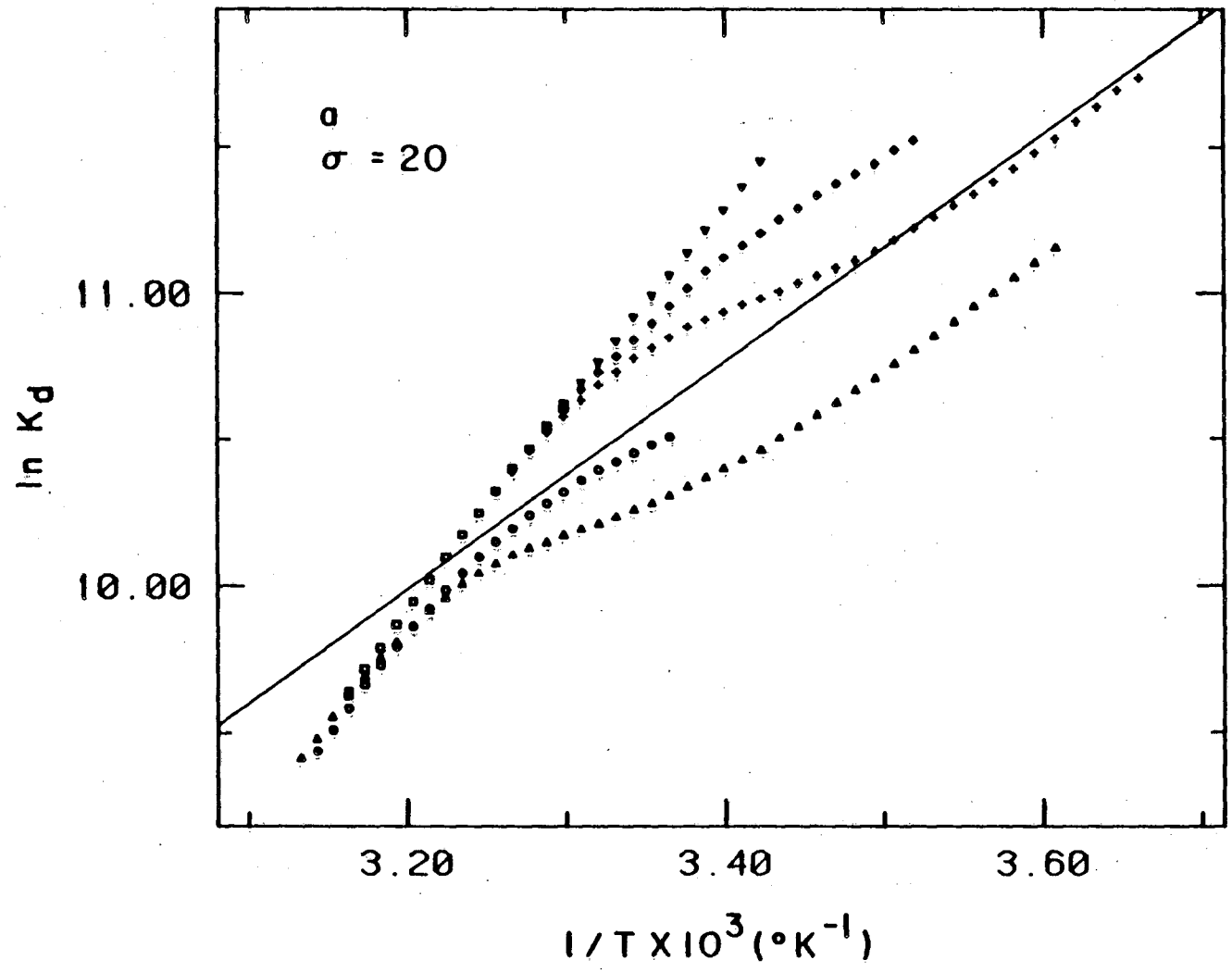
helix. However, comparing Tables III and IV, the calculated enthalpy for ethidium binding is -14 kcal/mol in both cases. Comparing the values of r and χ^2_ν, both models fit the data equally well.

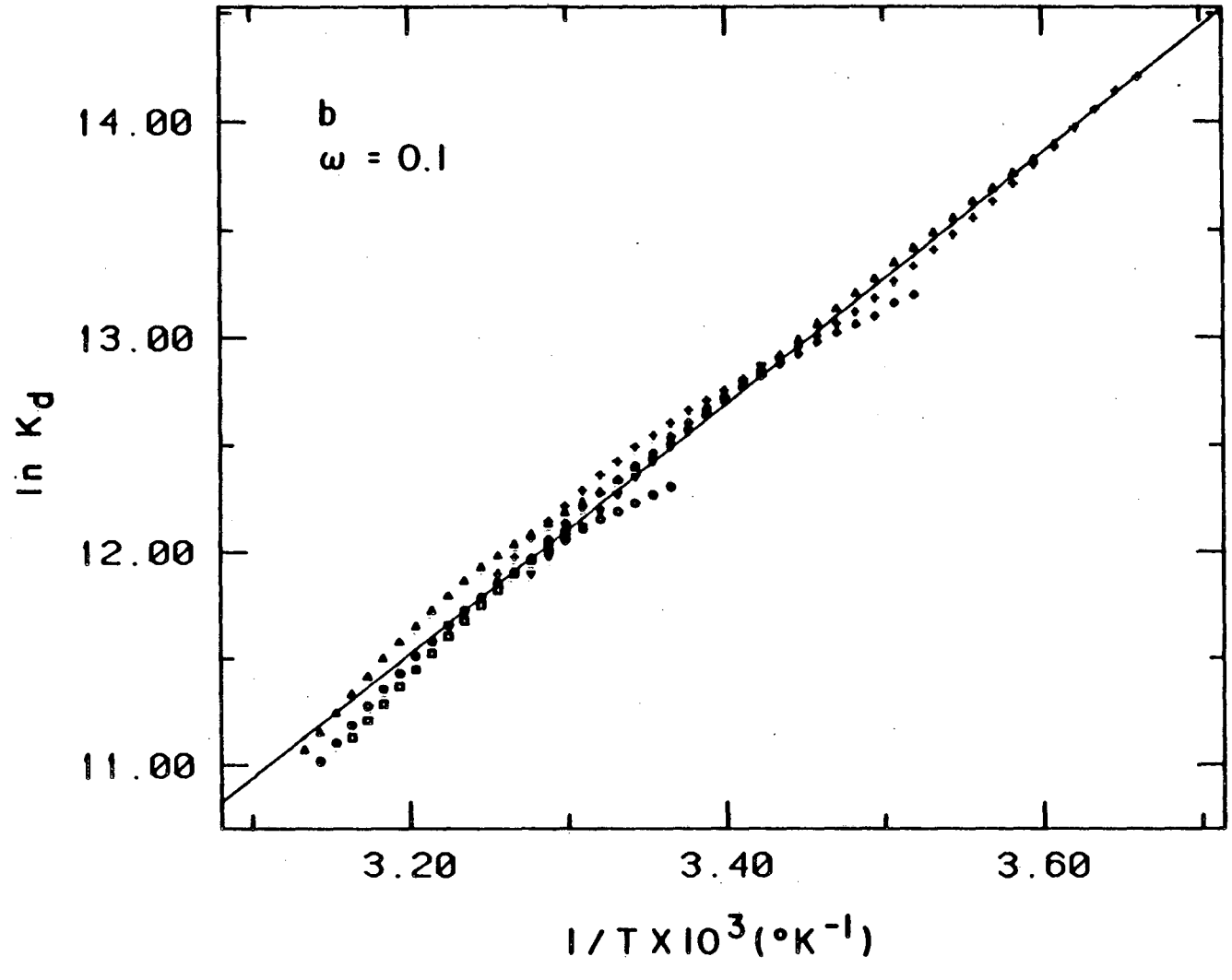
B) <u>rCA_5G + rCU_5G + Ethidium at High Ethidium:Strand Ratios</u>

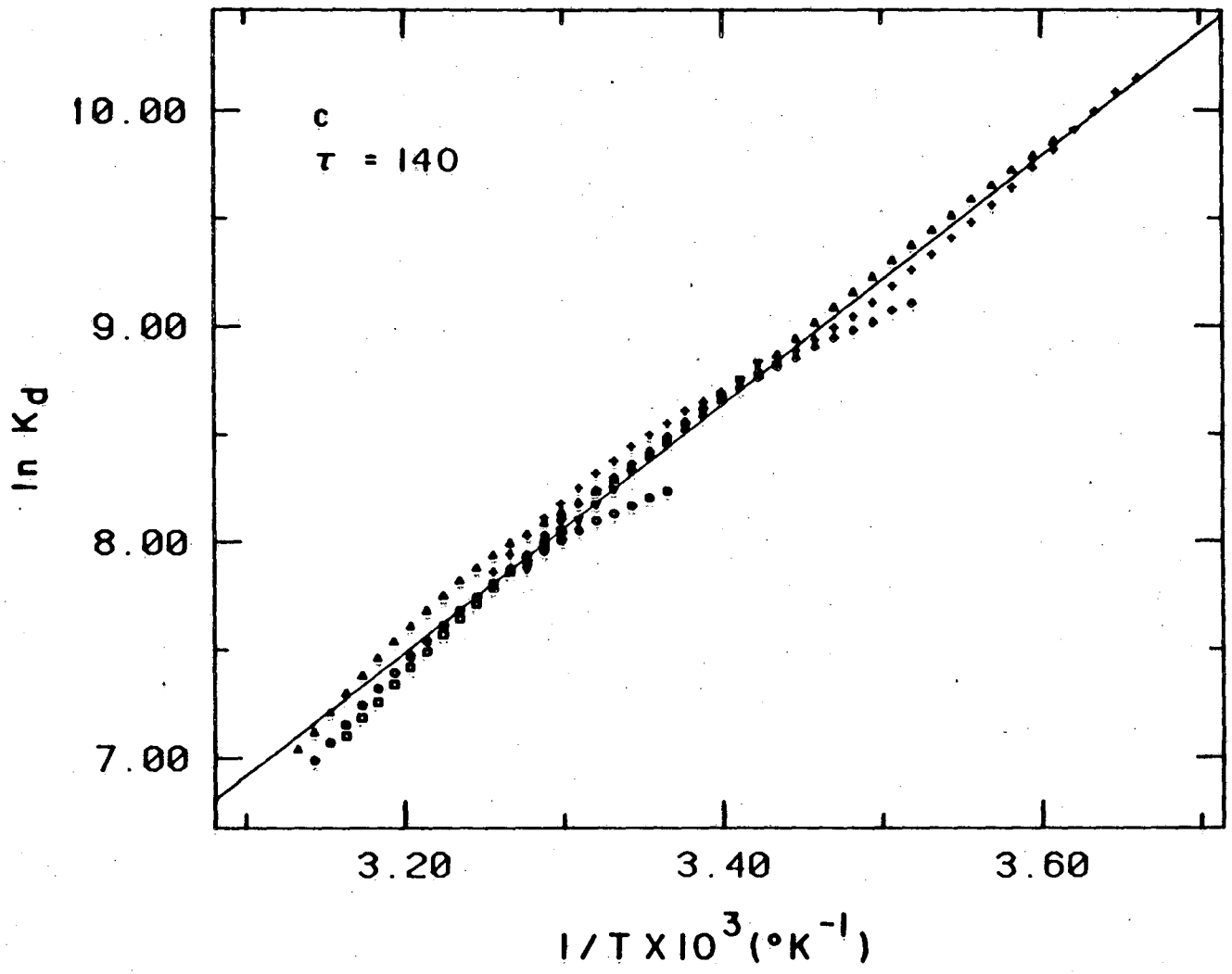
Figure 18a shows a plot of $\ell n(K_d)$ <u>vs</u>. 1/T for the melting curves at ratios of ethidium:strand of between 1 and 3, assuming σ = 20. The concentrations of six ethidium:strand melting curves (μM) are: 1) 50:49; 2) 106:50; 3) 156:53; 4) 12.1:12.1; 5) 24.2:11.9; and 6) 36.7:11.9. It is immediately clear that the plots of the several melting curves are not co-linear, as was the case at low ethidium:strand ratios. The curves at higher ethidium:strand ratios curve downward significantly relative to those at low ratios. This effect was described qualitatively using figure 10, where we observed that the second ethidium binds less cooperatively than the first, and the third binds only reluctantly. The downward deviation shown in figure 18a is explained by the fact that the model over-estimates the ability of the successive ethidium ions to bind. This must be compensated for by a lowering of the apparent equilibrium constant for ethidium binding, causing the downward curvature. Thus, the model that worked very well at low binding ratios must be modified to explain the data at high ratios.

One possibility is that the binding of ethidium is associated with a negative cooperativity, namely $\omega < 1$ in equation 20. The binding of the second and third ethidium ions would be hindered.

Figure 18. (a) The plot of $\ell n(K_d)$ vs. 1/T for rCA_5G + rCU_5G + ethidium at high ethidium:strand ratios, assuming one strong site with σ = 20 and ω = 1. The concentrations of ethidium:strands were (μM): □, 50:49; ○, 106:50; △, 156:53; ▽, 12.1:12.1; ◇, 24.2:11.0; +, 36.7:11.9. (b) The plot of $\ell n(K_d)$ vs. 1/T assuming two strong sites with τ = 140. (c) The plot of $\ell n(K_d)$ vs. 1/T assuming cooperativity, with σ = 1 and ω = 0.1. The results are shown in Table V.


a
σ = 20
11.00
10.00
ln Kd
3.20
3.40
3.60
1/T X10³(°K⁻¹)


b
ω = 0.1
14.00
13.00
12.00
11.00
ln Kd
3.20
3.40
3.60
1/T X 10^3 (°K^-1)


c
τ = 140
10.00
9.00
8.00
7.00
ln Kd
3.20
3.40
3.60
1/T X10^3 (°K^-1)

Two ethidium ions could bind six ways without the cooperative interaction, and four ways with such an interaction. Three ethidium ions could bind two ways with one cooperative interaction, and two ways with two cooperative interactions. Thus, the reluctance of the binding of the third ethidium is clearly explained. Table V and figure 18b show the results of the calculations using the model including cooperativity (ω) and a strong pyrimidine-purine site (σ). Clearly, the model fits the data well with $\sigma = 1$ and $\omega = 0.1$. An equally good fit is calculated with $\sigma = 20$ and $\omega = 0.2$ (Table V). The resulting enthalpies calculated with either set of values for σ and ω are about 3 kcal/mol more positive than in the calculations using low ethidium:strand ratios; the equilibrium constant increases by about 25 - 40%. Clearly, no unique set of σ and ω will best fit the data.

If we return to the data at low ethidium:strand ratios, and use the model of cooperativity and the strong pyrimidine-purine site, the best fit assuming $\omega = 0.1$ is $\sigma = 1$, $\Delta H_d^\circ = -14.0$ kcal/mol, and $K_d(25°C) = 25 \times 10^4\ \ell\ mol^{-1}$ ($r = 0.991$, $\chi_\nu^2 = 2.0 \times 10^{-4}$). Assuming $\omega = 0.2$, the best fit is again with $\sigma = 1$, $\Delta H_d^\circ = -14.2$ kcal/mol, and $K_d(25°C) = 24 \times 10^4\ \ell\ mol^{-1}$ ($r = 0.994$, $\chi_\nu^2 = 1.2 \times 10^{-4}$). Thus, the data at low ethidium:strand ratios is consistent with either cooperativity or a strong pyrimidine-purine binding site, but not both. The data at low and high ethidium:strand ratios are both consistent only with $\sigma = 1$ and $\omega = 0.1$.

Another possible model which can explain the reluctance of the third ethidium ion to bind is the model considered earlier with two

TABLE V

Ethidium Binding to $rCA_5G + rCU_5G$
at High Ethidium:Strand Ratios
One Strong Site with Cooperativity Model

σ	ω	ΔH_d° (kcal/mol)	ΔS_d° (e.u.)	$K_d(25°C)$ ($x10^4$ ℓ mol^{-1})	σK_d	r	χ_ν^2 ($x10^4$)
1	1.0	-6.4	+1.2	9.5	9.5	0.784	41.6
1	0.3	-8.8	-5.4	16.6	16.6	0.949	11.4
1	0.2	-9.7	-8.4	19.5	19.5	0.979	5.6
1	0.1	-11.6	-14.2	25	25	0.995	1.7
1	0.05	-13.7	-21.0	31	31	0.981	5.1
20	1.0	-7.7	-4.9	4.0	80	0.885	23.1
20	0.5	-9.1	-9.1	5.3	106	0.958	9.4
20	0.3	-10.4	-12.7	6.5	130	0.987	3.7
20	0.2	-11.4	-16.0	7.5	150	0.995	1.9
20	0.1	-13.5	-22.5	9.3	196	0.987	4.3

strong sites (τ) on the ends of the helix. The third ethidium ion must bind to a weaker interior site. Table VI shows the fits to the data at high ethidium:strand ratios, assuming several values of τ. The best fit occurs with $\tau = 140$, $K_d(25°C) = 0.44 \times 10^4$ and $\tau K_d = 60 \times 10^4$. The results for this fit are shown in figure 18c. The enthalpy for the fit is about 3 kcal/mol greater than that calculated at low ethidium:strand ratios, as was the case with the cooperativity model. However, the fits to the data at low and high ethidium:strand ratios (Tables IV and VI) result in essentially the same values for τ and τK_d, 140 and 60×10^4, respectively. Thus, this model fits the data more consistently than the model with cooperativity.

C) <u>dCA_5G + dCT_5G + Ethidium Bromide</u>

The binding of ethidium bromide to the double strands formed by dCA_5G + dCT_5G is not as strong as it is to the ribo-oligonucleotides. Because of this, the ethidium is not fully bound to the double strand at 0°C, making it more difficult to determine the extinction coefficients of bound ethidium. Since the oligonucleotides show very little change in absorbance at 283 nm, the fraction ethidium bound is mostly determined at this wavelength. Therefore, as before, errors in estimating the extinction coefficient of bound ethidium at 260 nm will not affect the accuracy of the determination of the fraction of ethidium bound.

The results for the binding of ethidium bromide to the deoxy-ribo-oligonucleotides are shown in Table VII and figure 19 for different values of σ. The values for the enthalpy and entropy for

TABLE VI

Ethidium Binding to $rCA_5G + rCU_5G$
at High Ethidium:Strand Ratios
Two Strong Sites Model

τ	ΔH_d° (kcal/mol)	ΔS_d° (e.u.)	$K_d(25°C)$ ($x10^4$ ℓ mol^{-1})	τK_d	r	χ_ν^2 ($x10^4$)
1	-6.4	+1.2	9.5	9.5	0.784	41.6
20	-7.5	-5.8	1.6	32	0.877	24.6
40	-8.5	-10.3	1.0	40	0.944	12.7
80	-10.0	-16.1	0.64	51	0.986	4.2
140	-11.4	-21.7	0.43	60	0.995	1.7
200	-12.5	-25.6	0.34	68	0.990	2.4
400	-14.7	-34.1	0.21	84	0.968	8.4
∞*	-38	---	---	33	0.805	177

*Corresponds to two independent binding sites.

TABLE VII

Ethidium Binding to $dCA_5G + dCT_5G$
One Strong Site Model

σ	ΔH_d° (kcal/mol)	ΔS_d° (e.u.)	K_d(25° C) ($x10^4$ ℓ mol^{-1})	σK_d	r	χ_ν^2 ($x10^4$)
1	-8.9	-9.7	2.5	2.5	0.994	0.4
3	-8.9	-10.2	1.8	5.4	0.994	0.4
10	-8.7	-10.8	1.1	11	0.988	0.7
30	-8.4	-11.1	0.58	17	0.953	2.1
100	-8.1	-11.7	0.24	24	0.857	5.7

Figure 19. Plots of $\ell n(K_d)$ vs. 1/T for dCA_5G + dCT_5G + ethidium assuming different values of σ. The concentrations are given in the text. (a) $\sigma = 1$. (b) $\sigma = 3$. (c) $\sigma = 10$. The results are tabulated in Table VII.

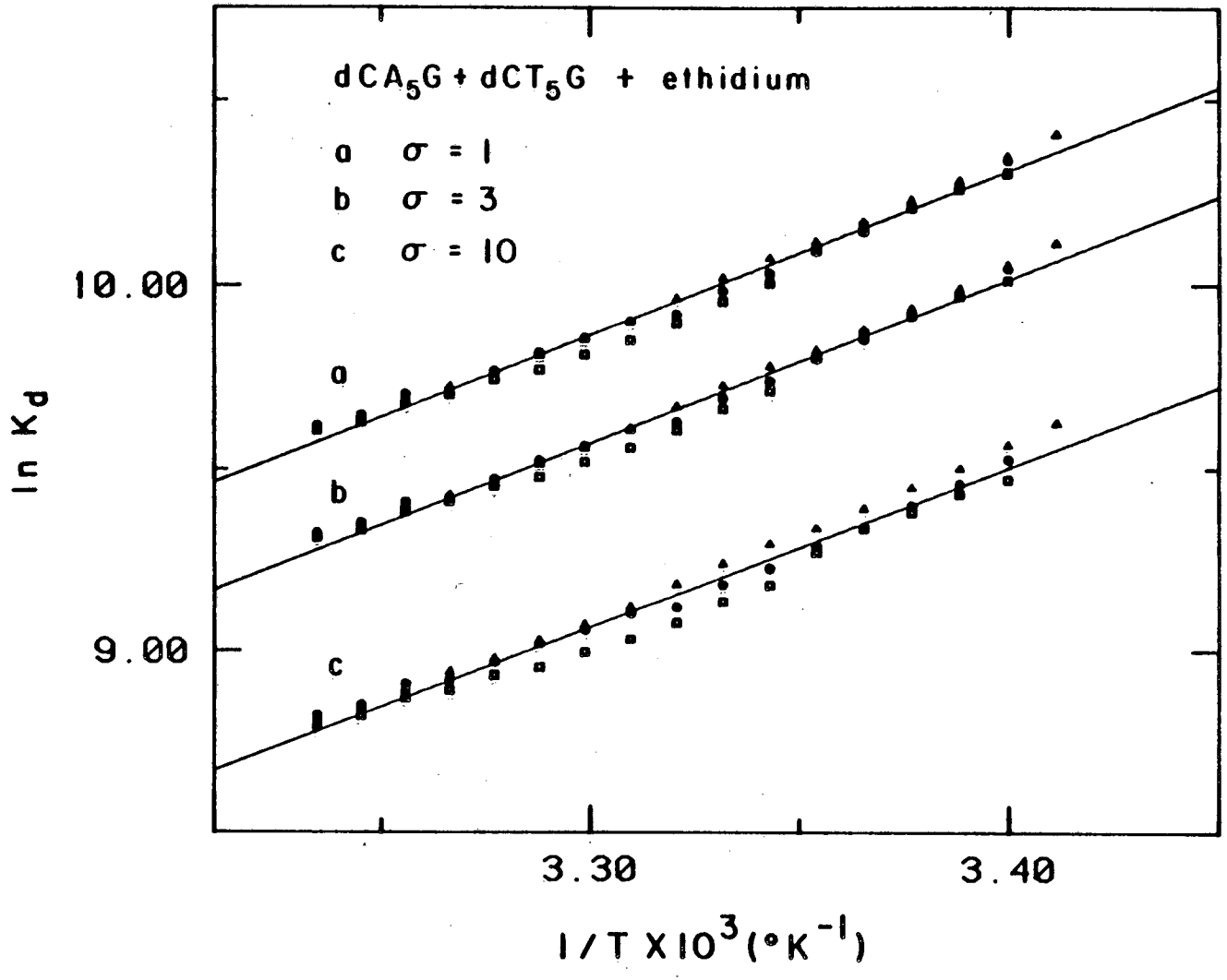
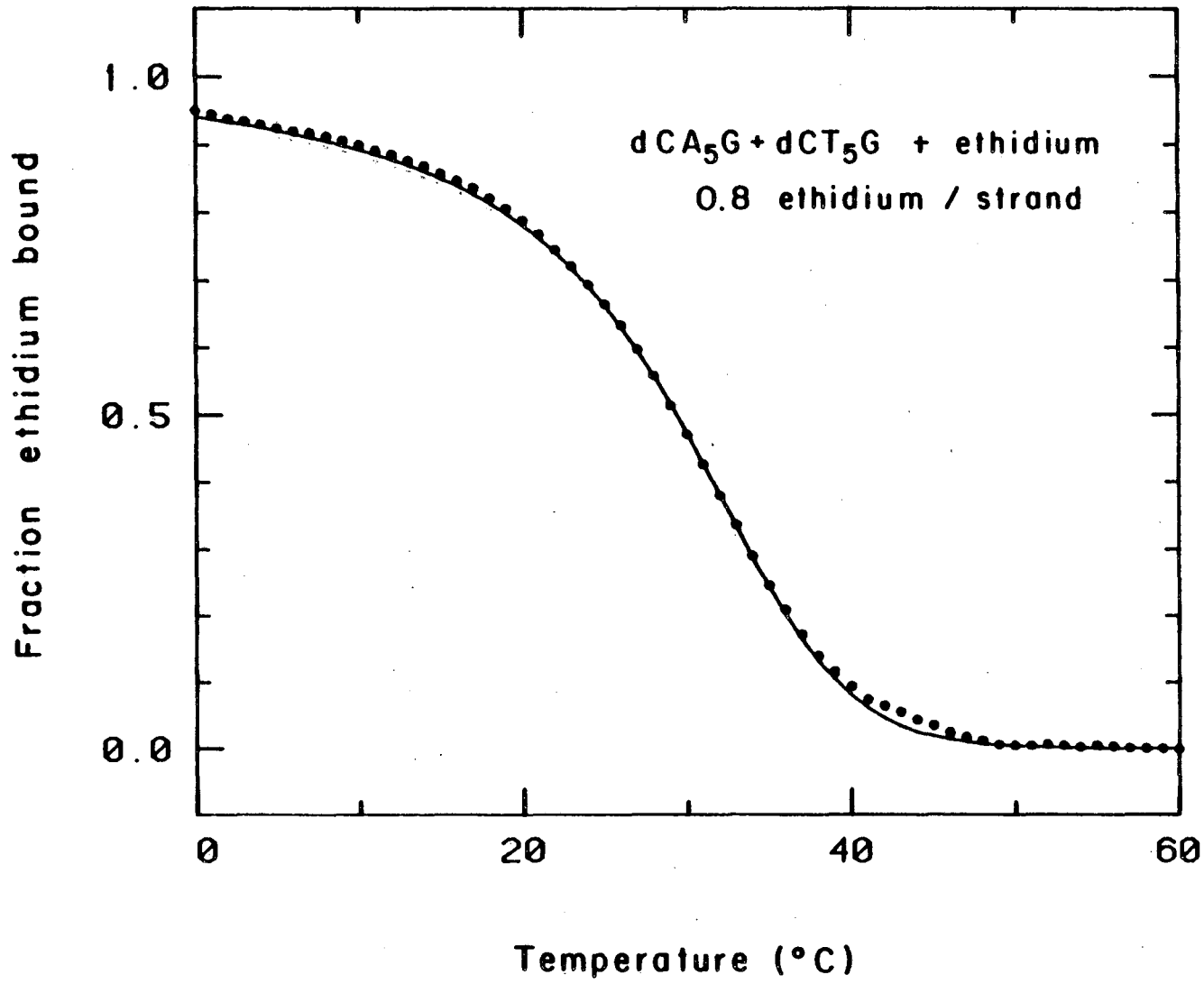

dCA5G + dCT5G + ethidium
a σ = 1
b σ = 3
c σ = 10
10.00
9.00
ln Kd
a
b
c
3.30
3.40
1/T X10^3 (°K^-1)

Figure 20. The experimental and calculated curves for dCA_5G + dCT_5G + ethidium, using a value of $\sigma = 1$ (see Table VII). The concentration of the strands was 40μM; the concentration of the ethidium was 31μM.


dCA5G + dCT5G + ethidium
0.8 ethidium / strand
1.0
0.5
0.0
Fraction ethidium bound
0
20
40
60
Temperature (°C)

double-helix formation were -50 kcal/mol and -148 e.u., respectively (Chapter II; Nelson et al., 1981). The concentrations of ethidium:strands in μM were: 19:44, 24:42 and 31:40. Because of the lower binding constant for ethidium binding, we are constrained to work over a smaller range of ethidium:strand ratios. This makes determination of the value of σ less accurate. From the data shown in Table VII, we see that the best fit occurs with $\sigma < 10$. For $\sigma = 1$, the values for ethidium binding are: $\Delta H_d^\circ = -8.9$ kcal/mol and $K_d(25°C) = 2.5 \times 10^4\ \ell\ mol^{-1}$. Figure 20 shows the experimental and calculated melting curves at a ratio of ethidium:strand of 0.8.

The slope of the extinction coefficient of bound ethidium is not accurately determined. However, the conclusion does not differ if we choose the extinction coefficient of bound ethidium to be constant at 283 nm, namely $\varepsilon_{283}(T) = 2.2 \times 10^4$. Making this change results in a determination of the enthalpy of ethidium binding which differs by 1.2 kcal/mol, and an equilibrium constant at 25°C which differs by about 15%. The best value for σ does not change.

The effect of the thermodynamic parameters for the double-strand formation is also small. Using $\Delta H°$ and $\Delta S°$ values of -45 kcal/mol and -131 e.u. instead of -50 kcal/mol and -148 e.u., respectively, changes the enthalpy for ethidium binding by 3 kcal/mol, and the equilibrium constant at 25° by 10%.

We can test the same models used successfully with the ribo-oligonucleotides. If we assume there are two strong sites, one on each end, the best fit occurs with $\tau = 3$. Assuming a model with cooperative binding, the best fit occurs with $\omega = 1$. Thus, the

deoxyribo-oligonucleotide data are consistent with the model that all binding sites have essentially equal binding constants, with no cooperativity between the binding sites. However, we must keep in mind that the detrermination of the fraction of ethidium bound was not as directly determined as it was for the ribo-oligonucleotides. Also, data could only be collected over a range of ethidium:strand ratios between about 0.4 and 0.8. The comparison of the models in the ribo-oligonucleotides were best done at high ethidium:strand ratios, where the relative strength of the binding sites are more important due to the larger number of ethidium ions bound per double strand.

However, we can clearly conclude that the binding of ethidium to the deoxyribo-oligonucleotides is weaker than to the ribo-oligonucleotides. The enthalpy of binding is comparable: -9 kcal/mol for the deoxyribo- and -11 to -14 kcal/mol for the ribo-oligonucleotides.

D) Comparison of the Models

Table VIII summarizes all of the results for all of the models for the ribo-oligonucleotides and the deoxyribo-oligonucleotides. In the case of the deoxyribo-oligonucleotides, the only model that fit the data well assumed all the binding sites were of equal strength, with no cooperativity between binding sites. For the ribo-oligonucleotides, two models fit the data equally well: either assuming the two terminal binding sites were both stronger than the interior binding sites by a factor of about 140, or assuming all the

TABLE VIII

Summary of Models for Ethidium Binding

Oligomers	Model	Ethidium per strand	$\Delta H°$ (kcal/mol)	$\Delta S°$ (e.u.)	K_d (25°C)[a] ($\times 10^{-4}$ ℓ mol^{-1})
	One strong site, $\sigma = 20$	Low	-14	-27	6.1
		High	not consistent with data		
rCA_5G + rCU_5G	Two strong sites, $\tau = 140$	Low	-14	-30	63
		High	-11	-22	60
	Cooperative, all sites equal, $\omega = 0.1$	Low	-14	-22	25
		High	-12	-14	25
dCA_5G + dCT_5G	All sites equal, not cooperative		-9	-10	2.5

[a] τK_b = strong site equilibrium constant tabulated for the two strong sites model.

binding sites were of equal strength, but there was negative cooperativity between the binding sites with $\omega = 0.1$.

Although the analysis of the data on the deoxyribo-oligonucleotides was problematic due to weaker binding, the results would have detected a cooerativity effect or an end-binding effect that was as large as those found for the ribo-oligonucleotide data. Thus, we can conclude that the binding to the deoxyribo-oligonucleotides shows less end-binding or cooperativity effects than the binding to the ribo-oligonucleotides.

The reason the ribo-oligonucleotides bind ethidium with these large effects might be due to the hypothesis that RNA double helices are more rigid than DNA double helices. X-ray crystalography studies on ethidium complexes with iodoUpA and iodoCpG showed that the iodoU and iodoC sugars have C3' endo conformations, whereas the A and G sugars have C2' endo conformations (Tsai et al., 1977; Jain et al., 1977). From this, they proposed a general model for ethidium binding wherein the sugar conformations become C3' endo - ethidium - C2' endo. DNA B-form has a C2' endo sugar conformation, whereas RNA A-form has a C3' endo conformation. DNA can assume a number of different conformations by changing the solvent conditions by adding high salt concentrations, ethanol, etc., whereas RNA structure remains predominately A-form regardless of the solvent conditions. Thus, the DNA may be better able to adjust the sugar conformation on the 5' side of the bound ethidium to C3' endo, whereas the RNA cannot adjust the sugar on the 3' side of the bound ethidium to C2' endo as readily. Thus, ethidium binds without co-

operativity to the DNA double helices, but the strain on the RNA double helix is large enough to cause cooperativity.

The model assuming two strong terminal binding sites could be explained by the assertion that the sugar conformations are more flexible at the end of the helix, thus the bindng of the ethidium does not cause as much strain as it would in the interior. Alternatively, the model assuming equal binding sites with cooperativity could be explained by the assertion that the binding of the ethidium ion couses distortions which extend further down the helix, making the next available sit less favorable.

This explaination is apparently in contradiction to the observation that ethidium binds more strongly to RNA than DNA (see below). However, the effects are quite separate, and we should not confuse the strength of ethidium binding to a site from the effect of binding to adjacent sites; quite possibly, ethidium both bonds more strongly to RNA, and cuases greater effects to adjacent binding sites, than with DNA.

E) Comparisons with Previous Results

Previous studies of ethidium bromide intercalation into nucleic acids have generally beem carried out on polynucleotides or dinucleotides by monitoring the absorbance or fluorescence of ethidium in the visible near 480 nm, or by NMR techniques.

Studies of ethidium binding to DNA are fit very well by a model that assumes nearest-neighbor exclusion and no cooperativity. McGhee and von Hippel (1974) have developed a model of ligand binding to polynucleotides which allows for cooperativity between

lignads, and the possibility that a ligand occupies more than one base pair. Their equation 15 relates r/C_f to the extent of binding r:

$$r/C_f = K(1-nr)\left(\frac{(2\omega-1)(1-nr) + r + R}{2(\omega-1)(1-nr)}\right)^{n-1}\left(\frac{1 - (n+1)r + R}{2((1-nr)}\right)^{2}$$

$$R = \sqrt{[1 - (n + 1)r]^2 + 4\omega r(1 - nr)}$$

I have changed nomenclature slightly. K is the equilibrium constant for the ligand binding to an isolated site, r is the extent of binding in terms of ligand/base pair, C_f is the free ligand concentration, n is the number of sites covered by the ligand (2 for nearest-neighbor exclusion), and ω is the cooperativity parameter for two ligands bound contiguously. Note an incorrect plus sign in the their paper which I have corrected above. This equation reduces to the equation derived by Crothers (Crothers, 1968; Bresloff & Crothers, 1975) for the binding of a ligand which makes the next site unavailable for binding, namely n = 2 and ω = 1.

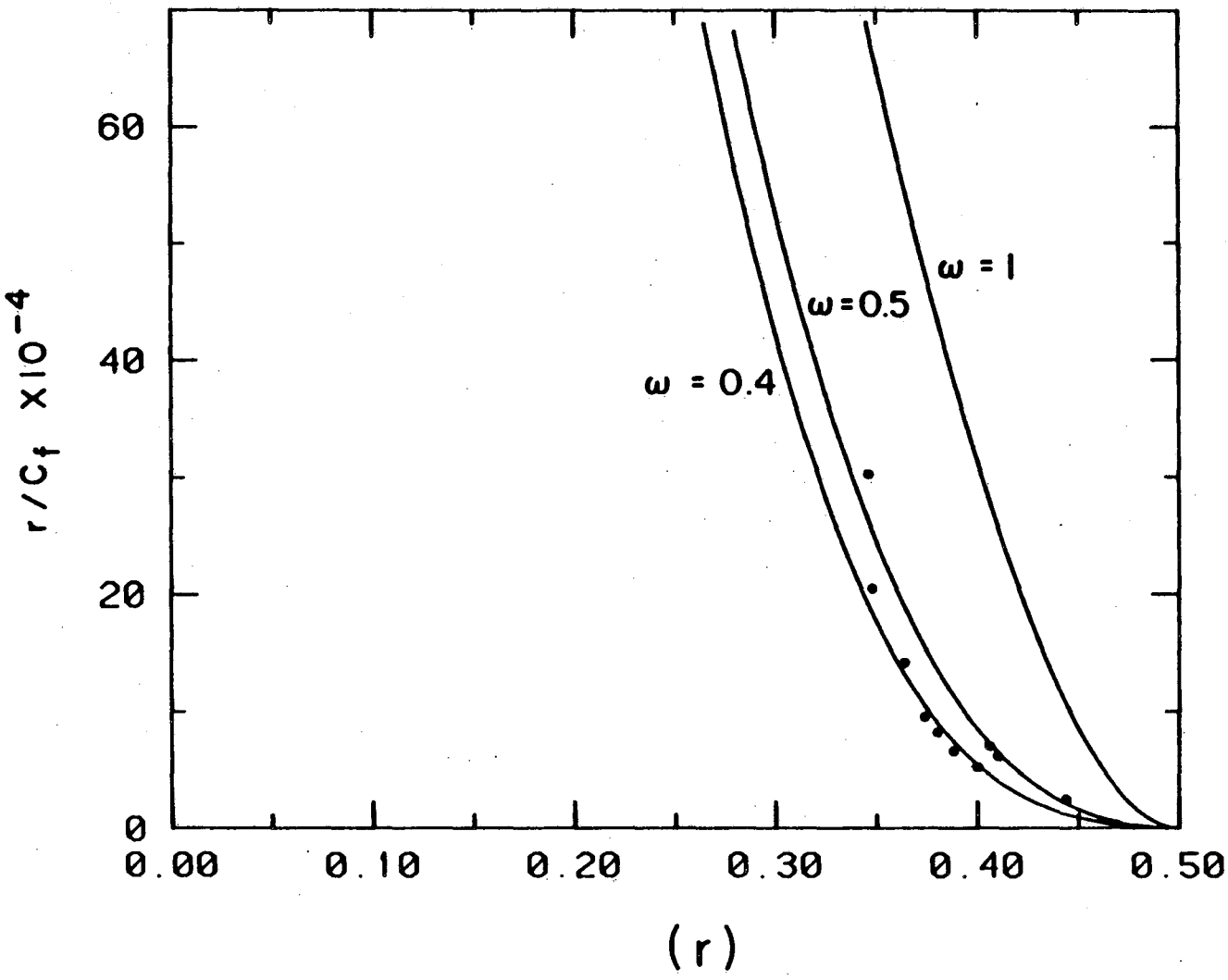
The binding of ethidium bromide to calf thymus DNA in 1M NaCl was found to fit very well to the model of nearest-neighbor exclusion and no cooperativity, with a binding constant at 19°C of 1.8 x 10^4 ℓ mol^{-1} (Bresloff and Crothers, 1975). This binding constant is an average for all the different binding sites in DNA. Also, at the higher salt concentration, ethidium binding is weaker than at 0.2M NaCl. However, their results on DNA compares well with the value determined in this study, 3.5 x 10^4 at 19°C.

The enthalpy of ethidium bromide binding to calf thymus DNA has been measured by batch and flow microcalorimetry (Quadrifoglio et al., 1974). The enthalpy they measured for ethidium binding to DNA in 0.1M KCl at 25°C was -6.7 kcal/mol. The enthalpy was 0.5 kcal/mol more positive in 0.015M salt concentration. This corresponds well with the value measured in this study for the enthalpy of ethidium binding to the deoxyribo-oligonucleotides, -9 kcal/mol in 0.2M salt.

The data on ethidium bromide ninding to RNA is more limited by the unavailability of RNA's which are double-stranded. Douthart et al. (1973) studied ethidium binding to the double-stranded RNA obtained from the mycophage Penicillium chrysogenum. From Scatchard plots at different salt concentrations, they determined that the saturation binding occurred at r = 0.38 ethidium ions bound/base pair in 0.1M sodium cacodylate, 0.32 in 0.01M sodium cacodylate, and 0.18 in 0.001M sodium cacodylate. The binding constant from the slope and the intercept of the Scatchard plot was found to be 4.7 x 10^6 at 25°C. Thus, the binding constant for ethidium bromide is much larger for RNA than for DNA.

In order to see if the data of Douthart et al. (1973) are consistent with a model with cooperativity, I took the data from their figure 2, and re-plotted it in figure 21. Using the equation of McGhee & von Hippel (1974) given above, setting n = 2 and K = 4.7 x 10^6, the fits for ω = 1.0, 0.5 and 0.4 were calculated and are shown in figure 21. Clearly, the data are consistent with a cooperativity parameter ω between about 0.4 and 0.5. It would be inappropriate to

Figure 21. A Scatchard plot of r/C_f vs. r (in terms of ethidium bound/base pair) for ethidium binding to double-stranded RNA. The data are from figure 2 of Douthart et al. (1973). The lines show fits using the McGhee & von Hippel (1974) equation given in the text for $n = 2$, $K = 4.7 \times 10^6$, and $\omega = 0.4$, 0.5 and 1.0.


60
40
20
0
0.00
0.10
0.20
0.30
0.40
0.50
(r)
r/C_f X10^-4
ω = 0.4
ω = 0.5
ω = 1

generalize from the results of one study, but clearly the binding of ethidium bromide to RNA polymers could be associated wtih a co-operativity which is not seen in DNA polymers.

Experiments of ethidium bromide binding to dinucleotides have shown quite clearly that ethidium binds preferentially to pyrimidine-purine sequences compared to purine-pyrimidine sequences (Krugh et al., 1975; Reinhard & Krugh, 1975; Reinhardt and Krugh, 1978; Pardi, 1980; Dahl, 1981; Dahl et al., 1982). For example, the complex rUpA-rUpA-ethidium is about 14 times stronger at 0° C than the complex rApU-rApU-ethidium (Dahl et al., 1982). However, comparisons of the strength of ethidium binding to different sequences is complicated by the fact that the dinucleotides form very unstable double strands in the absence of ethidium, making the determination of the equilibrium constant for double strand formation difficult (Young & Krugh, 1975; Krugh et al., 1976). Also, dinucleotide studies cannot measure cooperative effects, since there is only one binding site. Further verification of the pyrimidine-purine preference was obtained by studies on the tetranucleotides dC-G-C-G, dG-C-G-C, dC-C-G-G and dG-G-C-C using optical (Kastrup et al., 1978) and NMR (Patel & Canuel, 1976) techniques.

The sequence preference for the deoxyribo-oligonucleotides in the present study were small. However, only pyrimidine-purine and purine-purine sites are present in these oligonucleotides, and hence no quantitative difference between the pyrimidine-purine and purine-pyrimidine sites were determined. The technique can definitely be extended using measurements on a series of sequences to help deter-

mine quantitatively the sequence preferences unobtainable by polymer or dinucleotide studies. Also, by comparing ethidium binding to oligonucleotides which can form bulges with the related normal double helix, the strength of ethidium binding to a bulge can be determined. This will help determine whether the interaction of an intercalator with a bulge is important in promoting the incidence of frameshift mutations.

7. Conclusion

In the preceeding discussions, several models were analyzed and compared. It would be useful to summarize here what we found, and the implications to ethidium ion binding to nucleic acids.

In the case of ethidium binding to deoxyribonucleic acids, the polymer data suggests that there is no cooperativity between the binding sites. The data presented here for the deoxyribo-oligonucleotides also indicated no cooperativity. We also found that there is not a strong sequence preference for ethidium binding to the sequence studied here, $dCA_5G + dCT_5G$. However, the conclusions about the correct model are clouded by the fact that the analysis of data at high ethidium:strand ratios was not feasible. This problem might be overcome to some extent by working at strand concentrations much greater than those used here, using path lengths of 0.01 or 0.02 cm.

Ethidium binding to ribonucleic acids is associated with a larger equilibrium constant than to deoxyribonucleic acids. However, the binding is also associated by some sort of negative cooperativity in the ribo-polynucleotides, and either negative co-

operativity or preferential binding to the terminal sites of the ribo-oligonucleotide double strands studied here.

Both the end-binding and the cooperativity models for ethidium binding to the ribo-oligonucleotides are consistent with the assertion that the structureal rigidity of RNA double helices causes the distortion induced by the ethidium ion to affect the base pairs removed from the binding site. It might take a few base pairs before the comformation of the double helix returns to that required for binding the next ethidium ion. Alternatively, binding to the ends occurs because the greater flexibility of the ends of the helix can better accomodate the distortions caused by ethidium binding.

The end-binding model suggetst that ethidium binding to RNA molecules might occur at the ends of double-helical regions, internal loops, or bulge loops in the secondary structure of a natural RNA. This could have implications on the mode of binding of other intercalators to naturally occurring RNA.

The binding of ethidium to both RNA and DNA double helices is stabilized by a favorable enthalpy, with this stabilization being somewhat greater in the ribo-oligonucleotides used in this study (see Table VIII). The binding to RNA might be associated with a greater stacking of the ethidium with the RNA bases; alternatively, the base stacking in the double strands without ethidium bound might be distrubed less in the RNA, compared to the DNA. The ethidium replaces one base-base stack with two base-ethidium stacks. Ethidium binding to RNA double helices results in greater stabilization

than to DNA, however, the comformational distortions are less localized in the RNA.

The questions of the relative stabilities of different sequences will help determine what factors are important in ethidium binding. This would help determine if the greater enthalpic stabilization for the ribo-oligonucleotides found here is general for all sequences, or if the difference between RNA and DNA depends on the sequence. This would also quantitate more completely the difference between pyrimidine-purine and purine-pyrimidine sites, which was not addressed in these studies.

Oligonucleotide studies of ethidium binding are also ideally suited to determine the effect of bulged bases on ethidium binding. The destabilizing effect of a bulged base can be determined by comparing oligonulceotides which form normal double strands vs. the comparable double strand with a bulged base. The question of the degree to which ethidium overcomes this destabilization by bindng more strongly to the bulged site relative to the normal double helix could then be addressed. This would help assess the validity of the model for frameshift mutagenesis which asserts that the mutational rate is enhanced by the stabilization of the bulged base by intercalators.

APPENDIX A

Using the Gilford Spectrophotometer

The Gilford UV-VIS spectrophotometer is set up to collect melting curves at one or two wavelengths, and to take spectra at different temperatures. A PET model 2001 microcomputer acts to collect the data, which is subsequently sent to the PDP-8E or VAX 11/780 for analysis and plotting. The contents of this appendix are:

1. Procedures for Using Gilford Cuvettes
2. Setting Up the Gilford
3. Collecting Melting Curves
4. Taking Spectra on the Gilford
5. Transferring Data to the PDP-8E or the VAX
6. Analyzing Melting Curve Data on the PDP-8E
7. Analyzing Spectra on the PDP-8E
8. Program Listings for the PET

1. Procedures for Using Gilford Cuvettes

There are two types of cuvettes for the Gilford: stoppered and unstoppered. The stoppered cuvettes use teflon stoppers to seal the cuvette, and come in path lengths of 1, 2, 5 and 10 mm. The outside dimensions of these cells are all the same. The unstoppered cuvettes reduce evaporation by floating silicon oil on top of them. The path lengths of these cells are also 1, 2, 5 and 10 mm. The dimensions of the cells are different, depending on the path length.

Generally, the cells are washed by soaking in warm concentrated nitric acid for about 15 minutes, thouroughly rinsed with double-distilled water, and dried. If silicon oil was used to cover a sample, the cuvette is usually rinsed with double-distilled water,

detergent and 95% ethanol a few times to facilitate removal of the oil. If further treatment is required, the cell can be soaked in a solution of KOH in 95% ethanol for a few minutes. It is important not to soak the cell long in this solution, as it etches the cell. Spacers are washed in a similar manner, using lens paper for any wiping of the surfaces.

Samples which are taken to high temperatures should be degassed prior to filling the cell. This can be done by purging the buffer with helium for about 5 minutes, or by heating the sample to high temperature and shaking out the bubbles. Degassing is usually needed only when raising the temperature to 50° C or above. Sometimes bubbles will form along the sides of the cuvettes, which often do not interfere with the light path. Checking the absorbance before and after a measurement, by returning to low temperatures after a melting curve is completed, will verify if the bubbles caused problems.

The stoppered cuvettes are simply filled and stoppered. The recommended volumes for filling the cells of different path lengths are: 1 mm, 40μℓ; 2 mm, 80μℓ; 5 mm, 150μℓ; and 10 mm, 300μℓ. All the cells except for those with a 1 mm path length can be filled and emptied with a Pipetman P-200. Use a teflon needle for the 1 mm cuvettes. After filling the cells with a path length less than 10 mm, a round mound appears above the solution. Evaporation can be reduced if the cell is jarred carefully, causing the solution to wet the corners at the top of the cell. An evaporation of about 1% can

be expected after taking a melting curve to high temperatures, 60° C or above.

The samples in unstoppered cuvettes are covered with silicon oil to prevent evaporation. The oil should cover the sample by about 1 mm in the center. If too much oil is used, it will creep out of the cell, and cover the outside of the cell. The silicon oil is usually very effective in reducing evaporation, which is usually negligible. There is a problem with the oil wetting the corners of the cells, and occasionally the oil creeps down the cell enough to interfere with the light beam, especially at higher temperatures. Checking absorbances at low temperature after the melting curves will determine if this happened, as well as a visual check of the cell. The silicon oil can be separated from the sample after removing the sample from the cell by rolling the sample on a teflon sheet shaped like a watch glass.

Short path lengths of 0.1 and 0.2 mm can be achieved by using 2 mm unstoppered cuvettes with quartz spacers of 0.19 and 0.18 mm, respectively. The cuvette and spacer are carefully washed and dried, being careful to eliminate any dust. The spacer is then carefully inserted into the cuvette. The sample is then carefully added to the cuvette. Usually air is trapped under the sample. This is easily removed by centrifuging the cuvette very carefully in an Eppendorf centrifuge, using an Eppendorf centrifuge tube padded on the bottom with a small piece of tissue. The suggested volume of sample is roughly 20μℓ, but depends slightly on the cuvette and spacer combination. It is important that the solution cover the

spacer completely, as otherwise the silicon oil will creep between the spacer and the window, interfering with the light path. After centrifuging the sample into the cuvette, the sample is covered with oil to a depth of about 1 mm in the center. The sample is recovered by removing the spacer, and using a teflon needle to remove the sample from the spacer and the cell. A teflon sheet is used to separate the oil from the sample. The path length is different for different cuvette and spacer combinations. All combinations are tabulated in the box holding the spacers.

The unstoppered cuvettes require aluminum spacers to fill the remaining space, in order to insure a snug fit. This helps in the thermal contact between the cuvette and the cell block. Different sizes of spacers are available for the different path lengths. By convention, the cuvette is put on the side of the spacer towards the light source. The spacer is inserted first, followed by the cuvette.

The cuvettes are removed from the cell block using the tweezers which are kept by the Gilford. If aluminum spacers were used, the spacer is usually removed first, followed by the cuvette.

2. Setting Up the Gilford

The Gilford is very simple to operate. The following procudure is used to start up the instrument. Two circuit breakers behind the Gilford supply the power. Turn on the lamp, making sure you have selected either the UV, VIS or Both by turning the knob on top of the lamp and the switch immediately in front of it. The lamps require only several minutes to warm up.

Turn on the cuvette positioner, the wavelength and slit control, the thermo-programmer, the interface to the PET, and the PET, as required. If using the thermo-programmer, make sure the distilled water is flowing through the cell block. This cooling water is gravity-fed, and the top reservoir must be replenished about every hour by turning on the pump, using the circuit breaker to the right of the Gilford. If the reservior is empty, and the syphon is broken, re-fill the reservior with the pump, and re-establish the syphon by pulling the air out of the line at the lower end of the tubing with a 50 ml syringe.

If you are planning on going very far from room temperature, turn on the tap water to the right of the Gilford. Only a very slow stream of water is required. If you plan to go to low temperatures, turn on the nitrogen to a flow of about 0.5 on the regulator. Sometimes, when the weather is humid, there is a problem with condensation in the sample compartment. This problem can be eliminated by turning off the distilled water flow to the cell block, and turning the thermo-programmer to about 70°C for about 10 to 15 minutes while purging with nitrogen. Make sure no samples are in the cell block while doing this.

The wavelength is controlled either by the PET or the Gilford, depending on the position of a switch behind the Gilford, where the cable connects the Gilford to the interface. Flip the switch to "OPER" to allow the Gilford to control the Gilford, and to "COMP" to allow the PET to control it.

3. Collecting Melting Curves

The PET is used to collect the data from a melting curve. After inserting the cuvettes into the cell block, turn the thermo-programmer to the desired starting (set point) temperature, and turn the mode knob to "Set Point". Several minutes may be required to reach a low temperature. This process can be speeded up by placing ice in an insulated bucket, and immersing the tubing for the cell block cooling water into it.

Load the Gilford Melt program into the PET by typing "LOAD". Insert the cassette into the cassette player, and press "play". The PET will tell you it is loading the program, after which it says "READY". Type "RUN", and you are off.

The program asks how many cuvettes you are using. You may use between 2 and 4 cuvettes. Cuvette #1 is the reference, which can either be filled with a cuvette containing buffer, or left empty. You are then asked how many data sets you desire. This is usually set to a big number, such as 1000. The melting curve is stopped well short of this manually, after the desired high temperature is reached. You are then asked how many wavelengths - one or two. You then tell the PET what these wavelengths are. If one wavelength is used, turn the mode knob the "Standby" and set the wavelength manually. If using two wavelengths, turn the mode knob to "Set A", and use the wavelength set potentiometer labeled "A" to set the wavelength. Repeat the process for wavelength B. After you have set these, turn the mode knob to "Dual WL". Also turn the reference compensator on.

The computer now asks you for a time (sec) between data sets. This is the time the Gilford waits for the wavelength to change (if two wavelengths are used) and the reference compensator to adjust the slits. Usually, a value of 3 or 4 seconds is used. You are then asked for a time for the absorbance to settle. This is usually 3 seconds, although a shorter value may be used if you want points at very closely-spaced temperature intervals. You are given the option to collect more than one reading per data point and averaging. Usually this option is not used, but if it is, you must specify a delay time between readings, for example 0.1 sec.

You are now asked for a file name, which will be used for the file sent to the cassette or disk. Then, you are asked if you want the output to go to the cassette, disk and teletype. Usually you want to collect the data on the cassette and teletype. Punching a paper tape on the teletype allows a back-up in case the cassette does not read on the PET which sends the data to the PDP-8E. However, the teletype is very good at mis-punching a paper tape, and editing on the PDP-8E is often required after reading the paper tape. Finally, you are asked for an identification header for this run. This should identify the melting curve as completely as required, and appears as the first line of the file.

The program then tells you to turn the reference compensator on, and set the dwell time on the cuvette positioner to 1 sec. Then, push "auto", 1, 2, etc. for all the cuvette positions you are using. The program will now wait until you type "Y", after which it will start collecting data. After the PET has collected one or two

sets at the starting temperature, turn the mode swith on the thermo-programmer to 1°C/minute, or slower if desired. The data collection is now automatic, until you stop the collection by typing "#" when the high temperature is reached. (Make sure you set the heat limit potentiometer on the thermo-programmer to the highest temperature you want to reach.) After typing "#", it is very important to wait until the PET says "DONE WITH DATA COLLECTION" before touching anything. Otherwise, the file on cassette will be messed up. Usually, after reaching the high temperature, I will set the mode knob on the thermo-programmer to "Set-Point" after a melting curve, to check the absorbance after the run, and allow the PET to collect this data. After the low temperature is again reached, I will stop the PET with "#".

The data will subsequently be sent to the PDP-8E or the VAX for analysis and plotting, as will be described later.

4. Taking Spectra on the Gilford

The Gilford is set up the same way as for melting curves, except the switch in back of the Gilford must be switched to "COMP", and the mode knob is set to "Set A". The Gilford Scan program is loaded and run.

The program is broken into three sections: calibrating the wavelength, taking a spectrum, and writing the spectrum to a disk, cassette, teletype or screen. The program is explained in detail in the documentation found with the Gilford, and will not be repeated here. The data are sent to the PDP-8E in the same way as the

melting curve data. Usually, it is required to have a reference cuvette in cell position #1.

5. Transferring Data to the PDP-8E or the VAX

The data is usually sent from the PET to the PDP-8E using a cassette tape. The program "PET TO PDP8" is loaded on the PET next to the PDP-8E, and the directions are followed.

An alternative way to transfer the data is to use the paper tape reader on the PDP-8E. After bootstraping the computer, this is done by typing the following (Underlined text is provided by the PDP-8E):

. R PIP

* FILE.DA<PTR: (the filename can have 6 characters.)

_ (the PDP-8E waits for a character to be typed.)

The paper tape will then read until the end is reached, unless some control character was accidentally punched, in which case the reading will stop prematurely. If the analysis programs will not run on the data, incorrect data were punched, and they must be corrected using the PDP-8E editor prior to analysis. This problem is due to a problem with the interface between the PET and the teletype, or to the teletype itself. It is most desirable to transfer the data using the cassette if at all possible, as very few mistakes are made in the transfer.

For some unknown reason, the temperature is often written incorrectly to the cassette tapes and paper tapes. This might happen once in every other to every fourth melting curve. Usually a number in the temperature is mistyped as a zero. Often an eight is mis-

typed this way. This mistake is usually easily found from the melting curve (the melting curve plot stops when this occurs), and can be corrected with the PDP-8E editor.

The data can be sent from the PDP-8E to the VAX 11/780 by running the program VAX on the PDP-8E. This program allows the LA30 to act as a VAX terminal. Make sure the jumper cables on the teletype communications box connect VAX to PDP8. Log in as usual. Then, enter the editor on the VAX by typing:

. R VAX

VAX:

USERNAME: (enter username and password)

$ EDT DATAFILE.DAT (up to 9 characters are allowed)

FILE DOES NOT EXIST

* SET NONUMBERS (this allows the PDP-8E to interpret the echos from the VAX correctly.)

* I (enter insert mode)

* <ctrl>A (this signals the PDP-8E to ask for a file name.)

FILE NAME: FILE.DA (any disk file may be sent this way.)

VAX: <ctrl>B (this signals the PDP-8E to send the file.)

ECHO TO TTY? N (Y echos the file, taking much longer.)

SENDING TO VAX.

The transmission is very fast, and the PDP-8E will signal the end of the transmission. The PDP-8E verifies what it sends. If an error is detected, the PDP-8E will tell you there was an error in the transmission. You can fix this by:

VAX:

<ctrl>Z (this gets you out of insert mode)

* D %WH (delete the buffer)

XX LINES DELETED

* I (re-enter insert mode)

<ctrl>B (send the file again.)

If you get a transmission error again, the file probably contains some mis-placed control characters, which can be fixed by the editor of the PDP-8E, looking at the file sent to the VAX to determine the position in the file.

If the transmission was correct, complete the process by:

DONE WITH TRANSMISSION.

VAX:

<CTRL>Z (To get out of insert mode.)

* EX (to exit editor and write file.)

DATAFILE.DAT;1 XX LINES

$ LO (to log out)

XYZ LOGGED OUT AA/BB/CC

<ctrl>C (puts you back into the PDP-8E monitor)

There are programs on the VAX to analyze melting curves, similar to those on the PDP-8E. The programs to analyze melting curves of oligonucleotides with intercalating dyes are discussed in Appendix C. In addition, the program on DISK$USERFILE1: [JWN.MELT]MELT is essentially the same as MELT1 on the PDP-8E (see below), and the program [JWN.MELT]PLMELT is nearly identical to the program PLMELT on the PDP-8E. In addition, a program named

[JWN.MELT]MLTCRV calculates the fraction double-strands when you input the lower and upper baseline, and the program [JWN.MELT]LSTFIT performs non-linear least-squares analysis on melting curves. You are referred to these programs for details on how to run them.

6. Analyzing Melting Curve Data on the PDP-8E

The melting curve data are analyzed using the programs MELT1 and PLMELT. Both program listings are in the folder next to the PDP-8E.

The first program to be run is MELT1. This program subtracts the reference cuvette #1 from the others, and subtracts an additional reference blank. It then divides by a high-temperature absorbance, normalizing all the curves at high temperature. To run the program, type "R MELT1" while in the PDP-8E monitor. The program will ask you if you want to output results from an earlier file, which you usually ignore. (In all programs on the PDP-8E, a "yes" response is signaled by typing "1", and a "no" by typing either "0" or merely a carriage return.) You then enter the file name for the input file, which you can have echoed to the teletype. This is usually not useful, as the normalized data are much more interesting.

You are then asked whether you want to manipulate wavelength #1. If so, you can subtract cuvette 1 (the reference) from the others, compensating in errors in zeroing the reference on the Gilford. You then enter the cell blanks, which are generally in the range 0.030 to 0.040 if you used air (no cuvette) in the reference position. You are than asked to input the absorbances at high

temperature, not corrected for blank. Thus, you can read the absorbance directly from the melting curve printed on the teletype while you were collecting data. If you do not want to normalize the data, you must type 1 + blank, as the blank is subtracted before dividing, e.g. 1.030.

You are asked the same questions about wavelength #2, if collected. Finally, if the data are OK, you are asked whether you want the normalized data printed, and for a file name for the output file.

To plot the data, you run the programs DEFAUL, PLOTST, and PLMELT. Program DEFAUL sets up default plotting parameters for a subsequent call to the plotting programs. Afterwards, you can start the plotting programs with the same default parameters by running PLOTST, which reads the defaults from disk and calls the plotting program.

To run defaul, type "R DEFAUL". You are asked for a plotting program, PLMELT in this case. Then, you are asked for titles. The values XMIN, XMAX, and XINCH are, respectively, the X value at the left-hand side of the X-axis, the value at the right-hand side, and the length of the axis in inches. It is set to 7 of the Tektronix is used, otherwise it can be up to 12 inches on the Calcomp. An example is 0., 60., and 7. XST, XTIC and XINC are, respectively, the X value for the first tic mark, the increment between tic marks, and the increment between numeric labels on the X axis. An example is 0., 5., and 20. YMIN,YMAX,YINCH are the corresponding values for the Y-axis. YINCH is 6 for the Tektronix, or up to 9 inches on the

Calcomp. Examples are 0.7, 1.1, and 6. for normalized melting curves. YST, YTIC and YINC might be 0.7, 0.1, and 0.2.

The tic length is usually 8 - 10, and the numeric label and title sizes are usually 2, unless a figure for publication is being made, in which case these sizes are 3. (A label size of 3 will not work on the Tektronix.) Finally, you are asked if you want to call the plotting routine. This is equivalent to running PLOTST.

PLOTST asks if you want to see the default values, and then calls the plotting program you specified in DEFAUL, namely PLMELT. You are given the option to see the long form of the output, which elaborates somewhat on the information it tells you. For routine plotting, you usually don't want to see this information. You then enter the file name, which is the output file from MELT1, followed by the cuvette number (1 to 3, the blank cuvette position is eliminated), the wavelength, and the symbol number. The symbol numbers are: 1 - box; 2 - circle; 3 - triangle; 4 - upside-down triangle; 5 - diamond; 6 - +; 7 - X. They are listed on the side of the PDP-8E. Symbol size is usually 1. You can ignore the first few points in the file, which is useful if you collected several points at the low temperature before starting the scan. A new graph will erase the Tektronix screen, or move the Calcomp paper, and make new axes. New graph parameters will allow you to change any parameters specified in DEFAUL, and finally you have the option of sending the plot to the Calcomp or the Tektronix. If all is OK, the plot is sent; if not, you start over. After the axes are made, you are asked to specify the program to plot every nth point. Usually

this is set to 1, unless for aesthetic reasons, you want fewer points. After a plot is sent, you are given the option to re-plot the data, which allow viewing on the Tektronix before sending to the Calcomp. It also allows you to change plot parameters before re-plotting. By specifying not to re-plot, you return to the program PLMELT.

Several melting curves may be plotted on the same plot, by not specifying a new graph each time. This is generally very useful for plotting relative melting curves at different concentrations.

The programs [JWN.MELT]MELT and PLMELT are very similar routines on the VAX, and can be run without additional information. However, the high temperature absorbances you input must be corrected for the blank, which is different than on the PDP-8E.

7. Analyzing Spectra on the PDP-8E

Spectra are plotted directly using the plot program JNSCAN. The program DEFAUL can also be used for this, by specifying JNSCAN as the plotting program. However, the X-axis title is always 'WAVELENGTH (NM)', and the Y-axis title is always 'ABSORBANCE'.

The first question JNSCAN asks is whether the data are from 2 or more files. This allows you to subtract a spectrum from one file from that in another file, for example at different temperatures. Usually, you don't want to do this. As before, you have the option to have a long form of the output. You must specify the long form to print out the data, but it is not necessary for plotting.

You are asked for the file name that contains the spectra, and the header is echoed. You can then subtract any cell from any

other, but usually you subtract cell 1 (reference) from a sample cell. The exception is if you want to plot difference spectra between two cuvettes in the same file. (For spectra in different files, specify that the data come from 2 or more files above.) You can then subtract an additional baseline, if the sample and reference cuvettes absorb differently. A normalization constant is then input, which is divided into all the absorbance values. This allows normalizing the spectra, or scaling the absorbance to the path length, etc.

You are then asked if you want a new graph, new parameters, the symbol #, the symbol size, as before. If everything is OK, the plot is sent to the Tektronix or Calcomp. Note that in order to change from plotting on the Tektronix or the Calcomp, you must change the plot parameters. Again, as before, several spectra can be plotted on the same graph.

There are no comparable plotting programs on the VAX, as there are for the melting curves.

8. Program Listings

The PET programs Gilford Melt , Gilford Scan and PET TO PDP8 are listed below. For the hardware considerations for the interface and programming techniques, see the folder kept with the Gilford.

Program Gilford Melt. Program to collect melting curves on the Gilford spectrophotometer.

```
10 REM GILFORD MELTING DATA COLLECTION
20 REM VERSION 2, FEB 1981
30 REM WRITTEN BY JEFF NELSON
90 TW=0 BL$="                    "
100 POKE 52,0:POKE 53,35
110 BP=28672
120 TD=BP:TM=BP+1 TD$=""
130 P1=BP+48:P2=BP+49:P3=BP+50:P4=BP+51
140 P5=BP+52:P6=BP+53:P7=BP+54:P8=BP+55
150 POKE P4,146:POKE P8,155
160 DIM A%(4),T%(4)
170 REM
180 REM
500 PRINT"2"
510 PRINT"MELTING CURVE DATA COLLECTION"
520 PRINT
550 INPUT" NUMBER OF CUVETTES";CU%
560 IF (CU%>4) OR (CU%<2) THEN PRINT"":GOTO550
570 INPUT" NUMBER OF DATA SETS";DS%
590 INPUT" NUMBER OF WAVELENGTHS";WL%
600 IF (WL%<1)OR(WL%>2)THEN PRINT"":GOTO590
610 INPUT" WAVELENGTH 1";W1%
620 IF(WL%=2)THEN INPUT" WAVELENGTH 2";W2%
700 PRINT
710 INPUT" TIME (SEC) BETWEEN DATA SETS";DW:DW=DW*60
720 INPUT" TIME (SEC) FOR ABS TO SETTLE";DC:DC=DC*60
730 INPUT" NUMBER OF READINGS PER POINT";NR%
740 DR=0
750 IF NR%<>1 THEN INPUT" TIME BETWEEN READINGS";DR:DR=DR*60
800 PRINT
810 INPUT" FILE NAME";FI$
820 INPUT" OUTPUT TO CASSETTE (Y OR N)";TA$
850 INPUT" OUTPUT TO DISK (Y OR N)";D$
860 IF D$="Y" THEN INPUT" DRIVE #";DR$
900 INPUT" OUTPUT TO TELETYPE";TY$
940 PRINT
950 PRINT" HEADER (80 CHAR):"
960 INPUT ID$
970 INPUT" ALL OK";K$
980 IF K$<>"Y" GOTO 500
985 IF TA$="Y" THEN OPEN1,1,1,FI$
990 IF D$<>"Y" GOTO 1000
995 OPEN8,8,15:PRINT#8,"I"+DR$:CLOSE8
996 OPEN8,8,12,DR$+":"+FI$+",SEQ,WRITE"
1000 REM WRITE FIRST 2 LINES OF FILE
1010 TD$=ID$
1020 GOSUB 4500
1030 TD$=RIGHT$(STR$(CU%),1)
1040 A$=STR$(DS%)
1050 TD$=TD$+RIGHT$(BL$,6-LEN(A$))+A$
1060 A$=STR$(NR%)
1070 TD$=TD$+RIGHT$(BL$,6-LEN(A$))+A$
1080 A$=STR$(WL%)
1100 TD$=TD$+RIGHT$(A$,1)
1110 A$=STR$(W1%)
```

```
1120 TD$=TD$+RIGHT$(BL$,6-LEN(A$))+A$
1130 IF (WL%=1) THEN 1160
1140 A$=STR$(W2%)
1150 TD$=TD$+RIGHT$(BL$,6-LEN(A$))+A$
1160 GOSUB 4500
1180 REM
1190 REM
1200 REM WAIT FOR CUV#1, WL#1
1210 PRINT:PRINT" TURN REF COMPENSATOR ON."
1220 PRINT" SET DWELL=1SEC."
1230 PRINT:PRINT" PUSH AUTO, 1, ..., ";CU%
1240 PRINT:PRINT" TYPE 'Y' TO START."
1250 PRINT" TYPE '#' TO STOP DATA COLLECTION"
1260 GETK$:IFK$<>"Y"GOTO1260
1265 PRINT"Y"
1270 POKE P3,0
1280 IF (PEEK(P7)AND16)<>16 GOTO 1280
1290 IF (PEEK(P7)AND15)<>1 GOTO 1290
1300 POKE P3,2
1310 TW=TI+30
1320 IF TI<TW GOTO 1320
1330 IF(PEEK(P7)AND31)<>17 GOTO 1270
1350 PRINT:PRINT" STARTING COLLECTION."
1360 TW=TI+DW
1370 IF TI<TW GOTO 1370
1380 REM
1390 REM
1500 FOR I=1 TO DS%
1510 FOR J=1 TO WL%
1520 FOR K=1 TO CU%
1530 IF (PEEK(P7)AND15)<>2^(K-1) GOTO 1530
1540 POKE P3,2
1550 TW=TI+DC
1560 IF TI<TW GOTO 1560
1570 A%(K)=0
1580 GOSUB 3000 REM GET TEMP IN TE%
1590 T%(K)=TE%
1600 FOR L=1 TO NR%
1610 GOSUB 3500 REM GET ABS IN AS
1620 A%(K)=A%(K)+AS%
1630 NEXT L
1640 IF NR%<>1 THEN A%(K)=A%(K)/NR%
1650 POKE P3,0
1660 NEXT K
1670 IF(PEEK(P7)AND15)<>1 GOTO 1670
1680 POKE P3,2
1690 TW=TI+DW
1700 GOSUB 4000 REM MAKE STRING TD$
1710 GOSUB 4500 REM PRINT STRING
1720 IF TI<TW GOTO 1720
1730 NEXT J
1740 GET K$:IFK$="#"GOTO 1800
1750 NEXT I
1800 POKE P3,0
1810 TW=TI+30 REM WAIT 1/2 SEC
1820 IF TI<TW THEN 1820
1830 TD$="000":J=1
1840 GOSUB 4500
1850 CLOSE 1:CLOSE 8
1860 PRINT:PRINT" DONE WITH DATA COLLECTION."
1870 STOP
1880 REM
1890 REM
3000 REM GET TEMPERATURE IN TE%
3010 T1%=PEEK(P5):T2%=PEEK(P6)AND15
3020 T3%=PEEK(P5):T4%=PEEK(P6)AND15
```

```
3030 T5%=PEEK(P5):T6%=PEEK(P6)AND15
3040 IF(T1%<>T3%)OR(T1%<>T5%) GOTO3010
3050 IF(T2%<>T4%)OR(T2%<>T6%) GOTO3010
3060 TE%=T2%*100+INT(T1%/16)*10+(T1%AND15)
3070 IF TE%>990 THEN TE%=TE%-1000
3100 RETURN
3110 REM
3120 REM
3500 REM GET ABSORBANCE IN AS%
3510 POKE P3,3
3520 A1%=PEEK(P1):A2%=PEEK(P2)
3530 POKE P3,2
3540 AS%=INT(A2%/16)*1000+(A2%AND15)*100+INT(A1%/16)*10+(A1%AND15)
3550 IF AS%>8000 THEN AS%=AS%=10000
3600 RETURN
3610 REM
3620 REM
4000 REM MAKE STRING TD$
4010 TD$=RIGHT$(STR$(J),1)
4020 FOR M=1 TO CU%
4030 T$=STR$(T%(M))
4040 T$=RIGHT$(BL$,4-LEN(T$))+T$
4050 A$=STR$(A%(M))
4060 A$=RIGHT$(BL$,5-LEN(A$))+A$
4070 TD$=TD$+T$+A$
4080 NEXT M
4090 RETURN
4100 REM
4110 REM
4500 REM PRINT TD$ TO SCREEN, ETC
4510 IF J=2 THEN PRINT"";
4520 PRINTTD$
4530 IF TA$="Y"THEN PRINT#1,TD$
4540 IF D$="Y"THEN PRINT#8,TD$
4550 IF TY$="Y"THEN GOSUB 5000
4600 RETURN
4610 REM
4620 REM
5000 REM PRINT TD$ TO TELETYPE
5010 POKE TM,204 POKE TM,206,POKE TM,206,POKE TM,206,POKE TM,5
5020 FOR TC=1 TO LEN(TD$)
5030 IF(PEEK(TM)AND1)<>1 THEN 5030
5040 POKE TD,ASC(MID$(TD$,TC,1))
5050 NEXT TC
5060 IF(PEEK(TM)AND1)<>1 THEN 5060
5070 POKE TD,141
5080 IF(PEEK(TM)AND1)<>1 THEN 5080
5090 POKE TD,138
5200 RETURN
5210 REM
5220 REM
```

Program Gilford Scan. Program to take spectra on the Gilford spectrophotometer.

```
10 REM GILFORD SCAN PROGRAM
20 REM JEFF NELSON AUG. 1980
25 REL MODIFIED FOR NEW ROMS BY
26 REM JEFF NELSON FEB. 1981
30 REM TAPE RUNS CONTINUOUSLY
100 POKE 52,0:POKE 53,44:REM SAVE MEM
110 PRINT"3"
115 DIM A%(4,201)
120 T=0 I=0 J=0:AV%=0:WL%=0:ST$="":SP$=""
130 SW%=0:EW%=0:IW%=0:NC%=0:NW%=0:NS%=0
140 A1%=0 A2%=0
200 BP=28672:TD=BP:TM=BP+1
210 P1=BP+48:P2=BP+49:P3=BP+50:P4=BP+51
220 P5=BP+52:P6=BP+53:P7=BP+54:P8=BP+55
230 POKE P4,146:POKE P8,155 REM 8255'S
240 INPUT"WAVELENGTH:";WL%
250 IF WL%<200ORWL%>700 THEN PRINT"OUT OF RANGE":GOTO240
260 Q1=BP+192 Q2=BP+193 Q3=BP+194 Q4=BP+195
270 POKE Q4,128:POKE Q3,1
280 GOSUB4000
290 POKE P3,0
500 PRINT
510 PRINT "S2PECTRUM"
520 PRINT "W2RITE"
530 PRINT "L2AMBDA CALIBRATION"
540 PRINT
600 GET K$ IF K$=""GOTO600
610 IF K$="S"THEN PRINT"S":PRINT:GOTO1000
620 IF K$="W"THEN PRINT"W":PRINT:GOTO1500
630 IF K$="L"THEN PRINT"L":PRINT:GOTO5000
640 GOTO500
1000 REM INPUT PARAMETERS FOR SCAN
1010 INPUT"STARTING WL:";SW%
1020 IF SW%<200 OR SW%>700 THEN PRINT"OUT OF RANGE":GOTO1010
1030 WL%=SW% GOSUB4000
1040 INPUT"ENDING WL   :";EW%
1050 IF EW%<200 OR EW%>700 THEN PRINT"OUT OF RANGE":GOTO1040
1060 INPUT"INCREMENT   :";IW%
1070 IF SGN(IW%)<>SGN(EW%-SW%) THEN PRINT"WRONG SIGN":GOTO1060
1071 NW%=(EW%-SW%)/IW%+1
1072 IF NW%>201 THEN PRINT"TNO MANY POINTS, 201 ALLOWED":GOTO1010
1080 NW%=(EW%-SW%)/IW%+1
1090 IF NW%>201 THEN PRINT"TOO MANY POINTS, 201 ALLOWED":GOTO1010
1100 IF SGN(IW%)<>SGN(EW%-SW%) THEN PRINT"WRONG SIGN" GOTO1060
1110 INPUT"# OF CELLS:";NC%
1120 PRINT:PRINT"WAIT":PRINT
1130 FOR I=1TONC% FOR J=1TONW%
1140 A%(I,J)=0
1150 NEXTJ NEXTI
1160 INPUT"# OF SCANS:";NS%
1170 INPUT"TIME FOR ABS TO SETTLE:";TS
1180 INPUT"# OF READS AT EACH WL:";NR%
1190 IF NR%=1 THEN TR=0:GOTO1210
1200 INPUT"TIME BETWEEN READINGS:";TR
1210 PRINT"IDENTIFICATION TITLE:"
```

```
1220 INPUT ID$
1230 TS=TS*60:TR=TR*60
2000 DT=(ABS(IW%)/9+0.5)*60
2010 WL%=SW%:GOSUB4000
2020 PRINT:PRINT"SET DWELL=1 SEC. REF. COMP. OFF"
2030 PRINT"TYPE Y WHEN ALL READY: ";
2040 GET K$.IF K$="" GOTO2040
2050 IF K$<>"Y" GOTO2040
2060 PRINT "Y"
2100 FOR I=1 TO NS%
2105 PRINT"SCAN ";I
2110 POKE P3,0
2120 IF(PEEK(P7)AND15)<>1 GOTO2120
2130 POKE P3,2
2140 IF(PEEK(P7)AND15)<>1 GOTO2110
2150 FOR J=1 TO NC%
2155 PRINT"  CUVETTE ";J
2160 IF J=1 GOTO2215
2170 CC=2^(J-1)
2180 POKE P3,0
2190 IF(PEEK(P7)AND15)<>CC GOTO2190
2200 POKE P3,2
2210 IF(PEEK(P7)AND15)<>CC GOTO2180
2215 T=TI+180
2216 IF TI<T GOTO2216
2220 FOR K=1 TO NW%
2230 WL%=SW%+(K-1)*IW%:GOSUB4000
2240 T=TI+TS+DT
2250 IF TI<T GOTO2250
2260 FOR L=1 TO NR%
2270 POKE P3,3
2280 A1%=PEEK(P1):A2%=PEEK(P2)
2290 POKE P3,2
2300 AS=INT(A2%/16)*1000+(A2%AND15)*100+INT(A1%/16)*10+(A1%AND15)
2310 IF AS>8000 THEN AS=AS-10000
2320 A%(J,K)=A%(J,K)+INT(AS)
2420 T=TI+TR
2430 IF TI<T GOTO2430
2440 NEXT L
2450 NEXT K
2460 WL%=SW%.GOSUB4000
2470 T=TI+(ABS(EW%-SW%)/9+1)*60
2480 IF TI<T GOTO2480
2490 NEXT J
2500 NEXT I
2510 POKE P3,0
2520 AV%=NS%*NR%
2530 IF AV%=1 GOTO2620
2540 FOR I=1 TO NC%
2550 FOR J=1 TO NW%
2560 A%(I,J)=A%(I,J)/AV%
2610 NEXTJ:NEXTI
2620 GOTO500
3500 REM DATA WRITE
3510 PRINT"OUTPUT TO T2APE, D2ISK, S2CREEN OR TTY2? ";
3520 GET K$:IF K$=""GOTO3520
3530 IF K$="S"THEN OPEN10,3:PRINTK$:PRINT:GOTO3600
3540 IF K$="T"THEN PRINTK$:PRINT.INPUT"FILE NAME ";F$:OPEN10,1,1,F$ GOTO3600
3550 IF K$="Y"THEN PRINTK$:PRINT.GOTO3600
3554 IF K$<>"D"GOTO3560
3555 PRINTK$ INPUT"DRIVE #";DN$:INPUT"FILE NAME";FI$
3556 OPEN8,8,15.PRINT#8,"I"+DN$:CLOSE8
3557 OPEN 10,8,2,DN$+":"+FI$+",SEQ,WRITE":GOTO3600
3560 GOTO3400
3600 SP$="          "
3610 IF K$="Y"THENST$=ID$:GOSUB3900:GOTO3630
```

```
3620 PRINT#10,ID$
3630 ST$=STR$(SW%)
3640 NO$=STR$(EW%):GOSUB4500
3650 NO$=STR$(IW%):GOSUB4500
3660 NO$=STR$(NC%):GOSUB4500
3670 NO$=STR$(NS%):GOSUB4500
3680 NO$=STR$(NR%):GOSUB4500
3690 IF K$="Y"THENGOSUB3900:GOTO3710
3700 PRINT#10,ST$:POKE 59411,53
3710 FOR J=1 TO NW%
3720 WL%=SW%+(J-1)*IV%
3730 ST$=STR$(WL%)
3740 FOR I=1 TO NC%
3750 NO$=STR$(A%(I,J)):GOSUB4500
3800 NEXTI
3810 IF K$="Y"THENGOSUB3900:GOTO3830
3820 PRINT#10,ST$ IF K$="T"THEN POKE 59411,53:REM KEEP TAPE MOVING
3830 NEXTJ:CLOSE 10
3840 PRINT"DONE WITH WRITE":GOTO500
3900 REM SEND ST$ TO TTY
3910 POKE TM,204:POKE TM,206:POKE TM,206:POKE TM,206:POKE TM,5
3920 ST$=ST$+CHR$(141)+CHR$(138)
3930 FOR TC=1 TO LEN(ST$)
3940 IF (PEEK(TM)AND1) <> 1 GOTO3940
3950 POKE TD,ASC(MID$(ST$,TC,1))
3960 NEXT:RETURN
4000 DA%=(WL%-200)*8
4010 L%=11264-200+WL%
4020 C%=PEEK(L%)
4030 IF C%>128 THEN C%=C%-256
4040 DA%=DA%+C%
4050 POKE Q2,DA%/256:POKE Q1,DA%AND255
4060 POKE Q3,0:POKE Q3,1
4070 RETURN
4500 REM ADD NO$ TO ST$, 6 CHAR.
4510 ST$=ST$+LEFT$(SP$,6-LEN(NO$))+NO$
4520 RETURN
5000 REM WAVELENGTH CALIBRATION
5010 PRINT "S2TEP WAVELENGTH"
5020 PRINT "C2ONSTANT CORRECTION"
5030 PRINT "G2OTO WAVELENGTH"
5040 PRINT "R2ETURN"
5050 GET K$:IFK$=""GOTO5050
5060 IFK$="S"THENPRINT"S":GOTO5250
5070 IFK$="C"THENPRINT"C":GOTO5500
5080 IFK$="G"THENPRINT"G":GOTO5700
5090 IFK$="R"THENPRINT"R":GOTO500
5100 GOTO5000
5250 PRINT:REM STEP WL FOR CORRECTION
5260 INPUT"STARTING WL =";WL%
5270 IF WL%<200 OR WL%>700 THEN PRINT"OUT OF RANGE":GOTO5260
5280 GOSUB4000
5300 PRINT WL%," NM: H(4), L(6), OK(5),BACKUP(2), OR (R)ETURN?"
5310 GET K$:IFK$=""GOTO5310
5320 IFK$="5"GOTO5410
5330 IFK$="4"THEN POKEL%,(PEEK(L%)+1)AND255:GOTO5360
5340 IFK$="6"THEN POKEL%,(PEEK(L%)-1)AND255:GOTO5360
5343 IF K$="2"THEN WL%=WL%+1:GOTO 5360
5345 IFK$="R"GOTO5000
5350 GOTO5300
5360 WL%=WL%+1:GOSUB4000
5370 T=TI
5380 IF TI<T+9 GOTO5380
5390 WL%=WL%-1:GOSUB4000
5400 GOTO5300
5410 WL%=WL%-1:IF WL%>=200 GOTO5280
```

```
5420 PRINT:PRINT"CALIBRATION DONE".PRINT
5430 GOTO5000
5500 PRINT:PRINT"CONSTANT CORRECTION TO WAVELENGTH":PRINT
5510 INPUT"STARTING WL:";W2%:IFW2%<2000RW2%>700THENPRINT"OUT OF RANGE":GOTO5510
5520 INPUT"ENDING   WL:";W1%:IFW1%<2000RW1%>700THENPRINT"OUT OF RANGE":GOTO5520
5530 INPUT"CORRECTION :";CO%
5540 ZZ=11264-200
5550 FORI=W1%TOW2%
5560 I%=ZZ+I
5570 POKE I%,(PEEK(I%)+CO%)AND255
5580 NEXT
5590 GOTO5000
5700 PRINT:
5710 INPUT"WAVELENGTH.";WL%
5720 IF WL%<200 OR WL%>700 THENPRINT"OUT OF RANGE":GOTO5710
5730 GOSUB4000
5740 GOTO5000
6000 END
```

Program Pet to PDP8. Program to transfer melting curves and spectra from the PET to the PDP-8E.

```
5 REM OCT, 1981 VERSION
10 PRINT"3PROGRAM TO SEND FILES FROM PET"
20 PRINT"   CASSETTE TO PDP-8E."
22 PRINT
30 PRINT"PLEASE FOLLOW THE NEW DIRECTIONS."
40 PRINT
45 PRINT"SWITCHES BELOW PDP-8E SHOULD BE SET"
50 PRINT"   TO 4010 AND 9600."
60 PRINT
70 PRINT"CONNECT CABLES ON SWITCHING BOX"
80 PRINT"   BETWEEN PDP8 AND TEKTRONIX"
82 PRINT
85 PRINT"IF PDP8 DOESN'T RESPOND TO TEKT,"
86 PRINT"   THEN SET SWITCH REGISTER TO 7600"
87 PRINT"   (0-4 UP, 5-11 DOWN), PRESS 'HALT',"
88 PRINT"   'ADDR LOAD', 'CLEAR', 'CONTINUE'"
89 PRINT"   A DOT SHOULD APPEAR ON TEKTRONIX."
90 PRINT
95 R$="":REM SAVE VARIABLE AT TOP OF STACK
100 REM OPEN THE TAPE FILE.
110 PRINT"LOAD CASSETTE.  TYPE <RETURN>"
115 PRINT"   WHEN READY."
120 GET K$:IFK$=""THEN 120
130 PRINT:PRINT"SEARCHING FOR FILE":PRINT
140 OPEN1,1,0
145 REM CHECK FOR END OF FILE
150 IF ST=64 THEN 900
200 INPUT#1,S$
210 PRINT:PRINTS$:PRINT
220 INPUT "RIGHT FILE";K$
230 IF K$<>"Y" THEN 900
300 PRINT:PRINT"ENTER EDITOR ON PDP-8E WITH FILE NAME ":PRINT
310 PRINT"        .R EDIT"
320 PRINT"        *FILE.DA<"
330 PRINT"        @"
335 PRINT
340 PRINT"CONNECT CABLES FROM PDP8 TO PET."
350 PRINT"WHEN THIS IS DONE, TYPE <RETURN>."
360 GETK$:IFK$=""THEN360
365 PRINT
370 INPUT"ECHO THE FILE TO SCREEN";EC$
400 PRINT:PRINT"SENDING FILE TO PDP-8E.":PRINT
410 OPEN4,4
420 PRINT#4,"I"
425 T=TI
430 IF TI<T+30 THEN 430
435 REM SEND 65 LINES TO BUFFER
440 FOR I=1TO65
450 PRINT#4,S$
460 IFEC$="Y" THEN PRINTS$
470 IF FL=64 THEN 800
480 INPUT#1,S$
```

```
485 FL=ST:REM FLAG FOR END OF FILE
490 NEXTI
600 REM WRITE AND CLEAR BUFFER ON PDP-8E
605 PRINT"***":PRINT"*** CLEARING BUFFER":PRINT"***"
610 PRINT#4,CHR$(12);:REM CTRL L
620 T=TI+30:REM WAIT .5 SEC
630 IF TI<T THEN 630
650 PRINT#4,"N"
660 GOSUB 1000:REM WAIT FOR PDP8E
670 GOTO 420
800 PRINT#4,CHR$(12);
810 PRINT:PRINT"DONE SENDING FILE"
820 PRINT:PRINT"SWITCH CABLES BETWEEN PDP8 AND TEKTRONIX."
830 PRINT"AND TYPE:    @E<RETURN>"
840 PRINT"TO SAVE FILE.":PRINT
900 CLOSE1:CLOSE4
905 FL=0
910 INPUT"GET ANOTHER FILE";K$
920 IF K$="Y" THEN 130
930 END
1000 REM WAIT FOR BUFFER TO CLEAR
1010 REM ECHO FROM PDP-8E
1020 REM WAIT 5 SEC
1030 T=TI+300
1040 IF TI<T THEN 1040
1050 RETURN
```

APPENDIX B

Using the Temperature-Jump Instrument

The temperature-jump instrument has been described in some detail in Adrienne Drobnies' thesis (1979). In this appendix, I will elaborate on the instructions and descriptions not included therein. The basic block diagram for the T-jump instrument is shown in figure 1. The contents of this appendix are:

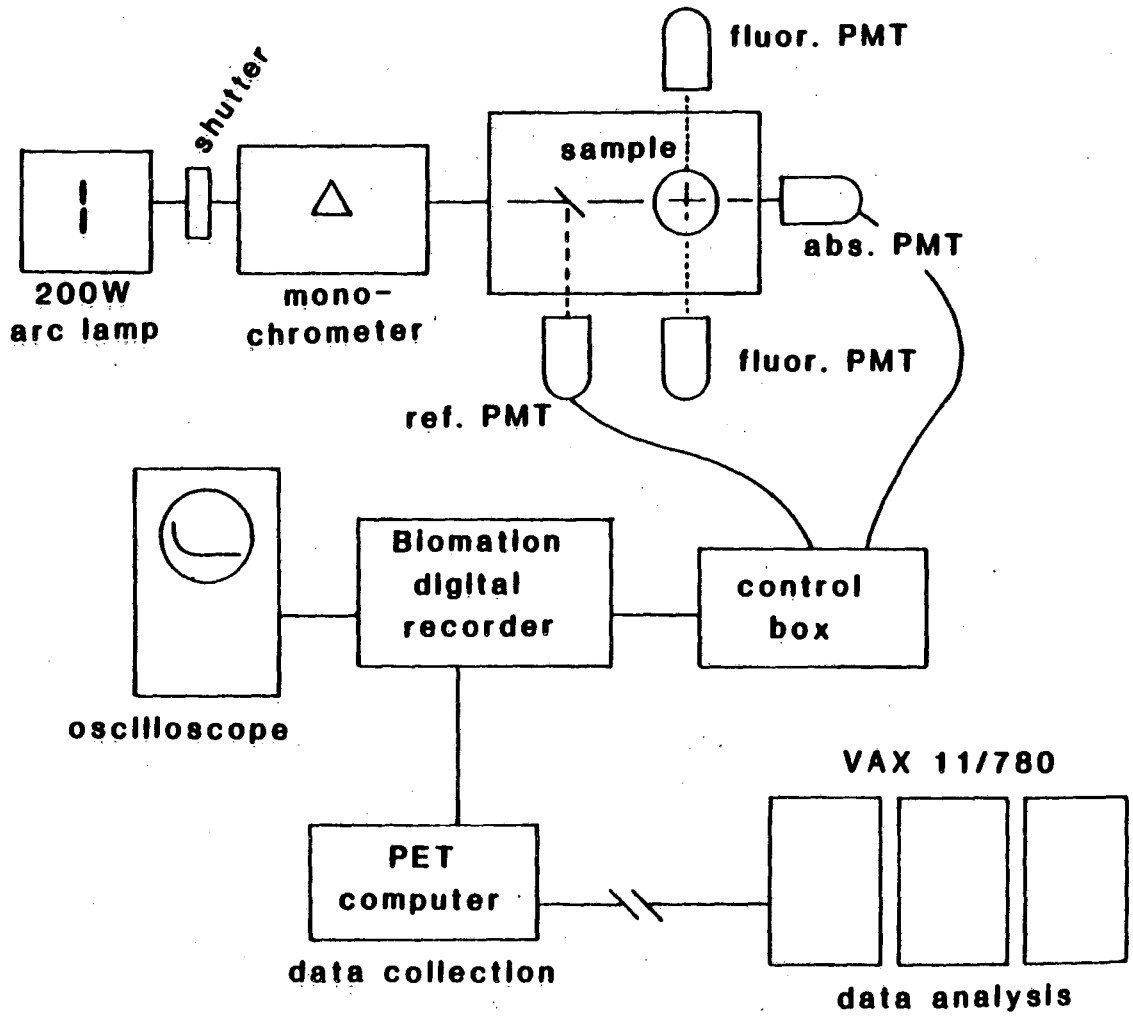
1. Sample preparation
2. Instrumental Set-Up
3. Instrument Warm-Up and Initial Settings
4. Computer Set-Up
5. Collecting Data
6. Computer Manipulations
7. Snafus (Problems and Solutions)
8. Turning Off the T-Jump
9. Computer Considerations for the VAX
10. Using the Provencher Program
11. Program Listings for the PET

1. Sample Preparation

The T-jump cell should be cleaned by soaking in approximately 25% nitric acid for 15 minutes before starting a set of experiments. This reduces the possibility of nuclease contamination. Care should be taken not to allow the exterior parts of the top or bottom electrode to come into contact with the nitric acid, as they will dissolve. After acid-washing the cell, it must be rinsed thoroughly by holding under a stream of distilled water for 10 minutes.

Before filling the cell with a sample solution, it should be clean and dry. Unscrew the top and bottom sections of the cell, and

Figure 1. A block diagram of the temperature-jump instrument.



T-JUMP INSTRUMENT

rinse all three pieces with double-distilled water, followed by 95% ethanol. The cell should then be blown dry with nitrogen. Do not let the pipette used for drying touch the cell or electrodes, as the Kel-F and the gold electrodes scratch easily. When dry, the bottom electrode is then screwed into the cell until just snug, and no more.

The sample must be degassed before it is put into the sample cell. This can be done in a number of ways, either by purging the sample with helium for approximately five minutes, or by heating the sample and shaking to remove air bubbles. The helium should first be bubbled through water to help saturate the helium with water vapor, thereby reducing evaporation of the sample. Other degassing possibilities include using a vacuum oven or sonication.

You need 0.75 ml of sample to fill the cell. Use either a plastic 1 ml disposable pipette, or a Pipetman P-1000 to fill the cell. Fill the cell slowly, inserting the pipet all the way to the bottom electrode. It is important that the cell be dry, and to avoid the formation of bubbles. You will get a mound of solution on top.

Place the top electrode in cell, and let gravity float the electrode down. It is important not to disturb the mound while doing this. Carefully screw the top down until just firm, and no more.

The cell is then placed into the T-jump. By convention, the red X directly under one cell window is placed towards the lamp. Make sure the cell is firmly seated. Screw on the metal ring which

holds the sample cell in place. If you are running at low temperature, the cell will contract, requiring you to re-tighten the ring when the temperature is reached. If you are running at high temperatures, do not tighten the ring until the temperature is reached, in order to avoid excessive strains on the cell.

The cover to the sample compartment must be replaced in such a way that the indentation on the cover lines up with the white dot on the top of the sample compartment. This completes the interlock circuit; if the interlock is broken, the high voltage source will not charge the capacitor.

Place the thermocouple through the hole in the cover, and insure that it is making a firm contact with the top electrode. You must allow a reasonable length of time for the temperature to equilibrate, usually at least one hour if the temperature deviates more than a few degrees from room temperature. Watch the digital thermometer as a guide.

2. Instrumental Set-Up

The basic connections for the components are shown in Adrienne Drobnies' thesis (1979) and the literature provided by Dia-Log. Here I will cover only the connections which are likely to need attention.

The output at the back of the control box is led to the input of the Biomation model 805. The trigger of the Biomation is connected to the BNC connector on the high-voltage capacitor housing. The Y-out on the back of the Biomation is lead to CH 1 (+) of the vertical amplifier on the Tektronix model 564 oscilloscope. The $\overline{Z}$

-out on the Biomation is led to the external trigger of the Tektronix. The 50-cable connector to the PET interface is plugged into the back of the Biomation.

When the fluorescence mode is used, the EMI high-voltage PM power supply must be connected to the fluorescence photomultiplier tubes, and the plug marked "interlock" must be plugged into the front panel of the EMI. This is an interlock which shuts the high voltage off when the sample cover is moved, preventing damage to the PM tubes due to excessive light intensity.

3. Instrument Warm-Up and Initial Settings

The following steps describe how to turn on the T-jump. The circuit breaker next to the sink should be turned on. The two plugs are labeled "HV-SUPPLY" and "T-JUMP".

To turn on the lamp: Make sure the appropriate lamp is installed, either the UV arc lamp (large lamp housing), or the tungsten lamp (small lamp housing). The H.P. 6267B DC power supply powers the lamps. The voltage control knob should be turned fully clockwise. The current control knob should be set to 8.0 to 8.2 amps for the tungsten lamp, and 8.5 amps for the Hg-Xe are lamp. Turn on the power supply. If the arc lamp is being used, the lamp must be started by pressing the "start" button on the blue box next to the lamp, labeled "LPS 255-SSP", for 1 second or less. Instructions are given on this box. Do not press this if the tungsten lamp is being used. The arc should be centered between the two electrodes. If not, shut down the power supply immediately and re-

start. Enter the time on the lamp log on the blue starter box. The arc lamp should warm up about 30 to 60 minutes.

To turn on the temperature circulating bath: Turn on the circuit breaker on the wall above the bath. Turn on the Tamson bath. If the desired temperature is below 40° C, turn on the Neslab Bath Cooler, and set the min-max switch to min. If you aren't using the Neslab cooler, remove the cooling finger from the bath. Make sure the liquid is within 1" of the top, otherwise add 95% ethanol. See Barbara Dengler for details about using the baths. Set the desired temperature by turning the top of the temperature regulator until the top of the disk is at the desired temperature. Make sure the "heating capacity" knob on the front of the Tamson is fully clockwise.

If the desired temperature is much below room temperature, turn on the nitrogen flow regulator above the temperature bath. The left-most regulator is used for this. Set to to 6 to 7 as read at the top of the ball.

The PMT power supply is on the table over the lamp power supply. Push the red rectangular plastic button labeled "NETZ", which will glow red. The upper power supply, the EMA PM25B, is used only for fluorescence, and should normally be kept off.

Choose which high voltage capacitor you wouold like, either 10, 20 or 50 nf. Change the capacitors following the instructions provided by Dia-Log. The size of the temperature jump with the 20 nf capacitor follows the equation $\Delta T(°C) = 4.3 \times 10^{-3}(V^2) + 0.02$, where the voltage V is in kilovolts. 20 KV corresponds to a jump of about

1.8°C. The calibrations of the other capacitors are not known. See Adrienne Drobnies' thesis for the calibration procedure.

The Wallis power supply to the right of the lamp power supply is used to charge the capacitor to high voltages. Turn it on by pushing the top-most black button. The red power light should come on.

The Tektronix oscilloscope is turned on by turning the "Scale Illum." knob fully clockwise; the Biomation is turned on using the power toggle switch.

The initial settings on the control box should be: Zero-Suppr. ms = DC; Risetime τ = .1 ms; Function doesn't matter; Gain = X1; Add. τ_{ref} = .1 for UV arc lamp, 0 for fluorescence, 2 for tungsten lamp; Reference = mid-scale, X1; Sample = mid-scale, X1. The dynode settings on the sample and reference PM tubes should be set to 0 dynodes; the fluorescence PM tubes, if used, should be set to 2 dynodes. The shutter should be closed, with the LED off.

The initial settings on the Biomation should be: Sensitivity = 1 volt full scale; Offset = 12 o'clock; Ground = out; DC/AC = out; Trigger = 2.00; Record Mode delayed/pretrig = in; A only/Dual TB = out; Trigger source Ext/Int = out; Couple DC/AC = in; Slope +/- = in; Trigger mode Auto = out; Norm = out; Single = in; Arm is spring loaded - press to arm Biomation. Trig level = 10 o'clock; Timebase A = 100, µsec/msec = out; Timebase B - doesn't matter; Horiz expand = X1; Horiz pos = 12 o'clock; Vert pos = 9 o'clock.

The initial settings on the Tektronix oscilloscope are: Display = non-store; Calibration = off; CRT: as required for a focused trace of correct brightness; Scale illum = 10. Time Base Time/div = 1 msec; Mode = norm; Triggering Slope = +; Coupling = DC; Source = EXT; Level - to produce stable display; Position - to center display. Vertical Amplifier: Ch. 1 (+) = DC; (-) = GND; Volts/Div = 0.1 volts; AC stab = off; Position - as required; Ch 2 (+) = GND; (-) = GND; Bandwidth = HI; Trigger = CH1; Mode = CH1.

4. Computer Set-Up

The data from the T-jump are collected by the Biomation. The data are then sent to the PET microcomputer, from where they are sent to the VAX 11/780 at the Laboratory of Chemical Biodynamics.

Turn on the CBM model 2040 floppy disk drive with the switch on the back side, lower left corner. The three LEDs should light up for a second. Turn on the "Pet Interface to Cary". Turn on the PET with the switch on back, in the lower left corner. On the teletype communication box on the table to the left of the PDP-8E, plug a jumper cord from "PET" to "VAX". Make sure the blue box labeled LEL has the green LED lit. If the red or yellow LED is lit, push the blue button. If the green LED will not light, the VAX is sick and cannot be used. Data may still be collected and saved on disk, however.

To start the program using a floppy disk (the fast way): Insert a disk with the programs into drive 0. The disk must have the programs "SET UP", "TERM", "T-JUMP" and "FILES TO VAX". Perform the

following sequence. Underlined text is printed by the PET. (If an error occurs, as indicated by the middle LED on the floppy disk lighting up, turn off the PET, wait a few seconds, turn it back on, and start over. If the error LED again lights, the disk is bad, and another must be used.)

LOAD "*",8 (the disk will make some noises)

...

READY - response by PET.

RUN

T-JUMP DATA COLLECTION SET UP

DO THIS:

1 - PUT DISK IN DRIVE 0

2 - TYPE 'TERM'

3 - TYPE 'T-JUMP'

?

You did step 1 already, so do step 2 - type "TERM". You will see ** LOADING TERM **. After several seconds, the screen will blink and show the same messages. Then type "T-JUMP". It will say:

TYPE: LOAD "T-JUMP",8

THEN: "RUN"

Do this, and you will see:

B TRANSFER FROM BIOMATION

T TALK TO VAX

F SET FILE PARAMETERS

V SEND FILE TO VAX

W WRITE FILE TO DISK OR TAPE

You are now ready.

An alternate way to start the program is to use a cassette tape. It is much slower, however the cassettes are more reliable. Put the cassette labeled "Term, T-Jump" into the cassette player. Then perform the following sequence:

LOAD

SEARCHING

FOUND TERM (after several seconds)

LOADING

READY (After a minute)

NEW

LOAD

SEARCHING

FOUND T-JUMP

LOADING

READY (after a couple of minutes)

RUN

And you are off, just as above. Rewind the cassette and return it to its case.

5. Collecting Data

In preparation to taking a measurement, you must first set the wavelength, the slit widths, the signals into the control box, the voltage of the high voltage power supply, and the Biomation.

The wavelength is set with the large micrometer knob on the front of the monochromator. The wavelength in nm is read off of the black scale, 1 mm on the micrometer = 100 nm. Thus, 500 nm cor-

responds to the knob being set to the line labeled 5. Each revolution of the knob changes the wavelength 50 nm. If the arc lamp is used, it is desirable to use a mercury line near the wavelength of interest.

Once the wavelength is selected, you must adjust the lens inside the sample box. This lens is not achromatic. Set the line marked 0 to the wavelenth chosen.

The T-jump cell presently in use has a parcular characteristic which must be taken into account. Previously, it worked quite well at the mercury line at 267 nm. Recently, however, it has shown a severe anomoly at this wavelength. A second after the temperature-jump, the absorbance increases markedly, indicated by the deflection of the control box needle to the left, usually to the extent of 3 or 4 volts. The signal slowly moves to the center as light is shined on the sample. If the shutter is closed for a while and subsequently opened, the signal is the same as it was just before the shutter was closed. This has been seen for many samples, and is therefore a characteristic of the cell. The effect is not observed at 250 or 280 nm.

The slits are set using the small micrometers located at the entrance and exit ports of the monochromator. A slits are closed when the micrometer is set to 0 at the top. The width of the slit in mm corresponds to the reading from the top in mm. One turn of the micrometer corresponds to .5 mm. Usually, the entrance slit is set larger than the exit slit, by a factor of about 2. This allows a larger bandpass, thus admitting more light, while keeping the

width of the beam shining on the sample as narrow as possible. Good starting values are 2 mm for the entrance slit and 1 mm for the exit slit. Slit width is later changed to get adequate signal-to-noise while keeping the intensity as low as possible to reduce photo-degradation of the samples.

The next step is to set the signals from the photomultiplier tubes to the control box. First, set all the dynode knobs on the PM tube housings to their minimum value. Turn the mode knob to R, and increase the number of dynodes on the reference PM tube until the signal is near 5 volts on the meter (100 μamps). Adjust the signal to 5 volts with the potientiometer labeled reference, adjusting the number of dynodes as needed. Repeat this procedure with the sample PM tube by turning the mode knob to S. Turn the mode knob to S-R and set the signal to 0 with the sample potentiometer. Finally, turn the mode knob to 10(S-R) and set the signal to 0. You are now ready to take the jump.

The stability of the arc lamp can often cause problems. The lamp stability can be observed by pressing the trigger mode Auto button. The Tektronix will display repeated traces. If this does occur, you must press the Single button, then Arm, then Auto. This is believed to be a fault in the Biomation. If still no response, perform the procedure under the Snafu section concerning the Biomation not arming. The signal flashing on the scope should be a horizontal line. Lamp instability is indicated by an oscillating signal. This is due to the arc revolving around the electrodes, and can often be reduced by adjusting the bar magnet using the adjusting

knob on top of the lamp housing. Usually the best results are obtained with the magent all the way down, adjusting the magnet by rotating the knob. A new lamp may be unstable for a number of hours. If time does not cure a new lamp, the current may be increased to up to 9.5 amps for a period of one-half to one hour. When a lamp becomes old, the only cure is to replace it with a new one. Life expectancies can vary to less than 200 hours to more than 400 hours. Often, the magnet must be adjusted before every jump for an aging lamp. Be sure to press the trigger mode Single after checking the lamp.

The voltage setting on the high voltage power supply allows voltages up to 40 kV. One turn of the potentiometer on the front panel of the HV power supply corresponds to 5 KV. Thus, if you want a 20 KV jump, set the potentiometer to 4.00. The capacitor is charged by pressing the lower black button on the panel. The voltage meter will then read the meter, and you can check the setting of the voltage. If the safety interlock is broken, namely if the cover of the sample chamber is not aligned correctly, the capacitor will not charge. If the interlock is broken while the capacitor is charged, a loud beeping will be sounded, and the voltage will drop slowly.

The final, and most confusing, step is to set up the Biomation for data collection. The two settings that need the most attention are the time/channel and the volts full scale. These can best be set by trial and error. The time base should be set so that at least one-half of the signal is essentially to baseline; the

Provencher program works best with lots of baseline. The setting of the rise time filter on the control box should be set to about the same value as the time/channel or faster. The rise time must compromise between low noise, while still being much faster than the relaxation processes you want to measure. The voltage scale should be set so the relaxation fills at least one-half of the Tektronix screen. Note - the time and voltage ranges are set at the Biomation, not the Tektronix. Leave the settings on the Tektronix at 0.1V/cm and 1 msec/cm. The Biomation is armed by pressing the button labeled Arm. The red LED will light. Make sure the Biomation will disarm by momentarily releasing and resetting the Slope +/- button. If a few repetitions does not cause the Biomation to trigger, the PET computer is acting up. See the section on the snafu's below. The Biomation can be used independently from the PET by unplugging the 50-cable connector from the back of the Biomation.

The actual temperature-jump is performed by charging the capacitor, turning the mode knob to 10(S-R), opening the shutter, setting the voltage to 0, arming the Biomation (which often disarms when the shutter is opened), and gently squeezing the rubber bulb until a quiet click is heard, and the Biomation disarms. If it has been a while since you set the reference voltage, turn the mode knob to R and quickly set the voltage to 5 volts before setting 10(S-R) to 0V. The rubber bulb must be squeezed gently, or some oscillations will show up on the trace, for some unknown reason. The relaxation will show up on the Tektronix scope.

6. Computer Manipulations

Once the data has been collected by the Biomation, it is transferred to the PET computer, from where it is sent to the VAX 11/780. The procedure for starting the program has already been discussed.

The first step is to log onto the compter. This is done by typing "T" when the menu is seen on the screen. At this point, the PET is essentially acting as a teminal to the VAX.

There are a couple of important points in this regard. Any control character is typed by typing in sequence <OFF/RVS> followed by the letter. Note this is not done simultaneously; the <OFF/RVS> key must be released prior to pushing the letter key. A reverse-field square will show up on the screen to indicate a control character. The control option can be cancelled by typing <SHIFT><OFF/RVS>. The return to the menu is done by typing <HOME>. One annoying feature of the PET is that characters sent from the VAX are missed. Usually about every fifth character is missed, making the reading of the screen a challenging exercise. This is nothing more than an inconvenience. The transfer of data to the VAX is not affected.

Logging in is done by the following sequence:

<CR>

<u>USERNAME:</u> RFP (or whatever it is.)

<u>PASSWORD:</u> PORPH (it is not echoed.)

<u>WELCOME TO VAX</u> etc.

<u>TYPE TJUMP INITIALIZE DATA COLLECTION.</u>

$ TJUMP

YOU ARE NOW READY FOR DATA COLLECTION.

IS THIS THE PET ACTING AS THE TERMINAL? Y

$ <HOME> (to get back to the menu.)

The data is transferred from the Biomation to the PET by typing "B". The screen will tell you to type "Y". Doing this starts the transfer, which will either be instantaneous, or will take about 15 seconds. I don't know the reason why it sometimes takes so long; I believe the reason is in the Biomation.

The next step is to set up the file parameters. This is done by typing "F". The program then asks for a header. If you have previously put in a header, and want to make minor changes, this can be done by typing <SHIFT><CRSR ↑>. You can then use the right and left cursor keys to type in new information, ended finally by a <CR>. The carriage return need not be done with the cursor at the end of the line.

The computer than asks for the program instructions. See the section on the Provencher program for details. Usually the sequence FTFTFFTFT___3___2___5 will work (the underlines represent two spaces). If you have already entered this line previously, just type the cursor up and carriage return.

Now you are shown the first 100 points and the last 20 points. This is done because the Biomation often sends the last point as the first point. With the Biomation used in pre-trigger mode with 2.00 on the indicator, the first 48 points should occur before the jump occurs. The point when the jump occurred is seen as

the first point which is very different. If this wrap-around problem occurs, then the first point after the jump occurs on the 49th point. Very occasionally, the number of points before the jump can be much greater than 48. This seems to be an intermittent problem in the Biomation.

The program asks you for the TO point. This is the point when the jump occurred. Using a setting of 2.00 in pre-trigger mode, it is 48 if there is no wrap-around, and 49 if there is. You then are asked for a TFIRST point. This is the first point used in the analysis. Usually the first few points are ignored, usually about five times the time/channel on the Biomation. You are then asked for TLAST, the last point to use in the analysis. Usually I use 2045, which allows some slop in the last few points in case wrap-around occurs. (The Biomation collects 2048 data points.) The program now asks for the time/channel setting on the Biomation, which must be put in exponential notation, e.g. 0.1E-3. In all of these previous settings, you may re-use an old setting by typing a carriage return. If you want to change a setting, type "Y", and the program will ask for a new value. The last step is to ask if you are using two time bases. If you are using the dual time base mode of the Biomation, type "Y" and you will be asked a couple more questions. The PET gives you one last chance to change things by echoing the settings, and asking if it is
OK. If not, you repeat the whole mess.

The next step is to enter the editor on the VAX. This is done by typing "T" to talk to the VAX. Then:

$ EDT TAPE1.DAT (TAPE1.DAT is the name of this data file.)

FILE DOES NOT EXIST. (use any name you choose.)

* I (This puts the editor in insert mode.)

<HOME> (to get back to the menu)

The actual transferring occurs when you type "V" from the menu. The PET tells you to type "Y" when ready. The file will consist of three lines of headers, and 59 lines of data. The numbers 1 to 59 will appear on the screen as the lines are sent. This only takes several seconds. If the transmission stops, as indicated by the numbers no longer appearing, see the section on Snafu's.

When the PET tells you that it completed sending the file, you must re-enter the editor, get out of insert mode, and save the file:

T (to talk to VAX)

TALKING TO VAX

<CTRL>Z (done by typing <OFF/RVS> followed by Z.)

* EX (to exit the editor.)

DISK$USERFILE1:[RFP.TJUMP]TAPE1.DAT;1 62 LINES

$

The file is now saved. You can do a quick analysis of every tenth point by typing:

$ DOIT

FILE NAME (E.G. TAPE1): TAPE1

IS THIS THE FIRST TAPE IN SET?: (Y or N, depending.)

IS THIS THE LAST TAPE IN SET?: (Y or N, depending.)

NAME FOR THIS SET OF DATA (E.G. T1T9): T1T9 (if last set.)

DONE WITH CONVERSION (after several seconds)

OUT OF DATA (after a couple of minutes.)

*** ALL DONE WITH ANALYSIS ***

$

You may look at the results by editing the file named TAPE1.SUM:

$ EDT TAPE1.SUM (see section of Provencher program.)

* 6 (results start at line 7 or so.)

* . (period will repeat the present line.)

* <CR> (to see next line)

* - (to go back one line.)

* QUIT (to get out of the editor.)

What just happened is this: The routine DOIT took the file DATA1.DAT, and if it is the first tape in this data set, copied it to file TAPES.DAT. If it was not the first, it appended TAPE1.DAT to the file TAPES.DAT. If this was the last data set, it re-named TAPES.DAT to the name you input. The Provencher program was then run on every tenth point of TAPE1, the results of which appear in TAPE1.SUM. If this was the last set, it took all of the data files of this set from TAPES.DAT and re-named it T1T9.DAT. I usually collect 5 - 10 jumps for each sample, and collect them into one large data set. By comparing the results obtained from DOIT as you go along, you can immediately assess the reproducibility of the data.

The analysis of a complete set of data, using every point, can be done by typing "ALL". You are then asked for a file name, e.g. T1T9 in the example above. The job is then submitted automatically to the batch queue, where it will take about 10 minutes for each

individual data set, longer of the computer is busy. The results will be in T1T9.ALL (summary of analysis), T1T9.OUT (the longer output), and T1T9.INP (a listing of the input data). These files may be printed by typing PRINT T1T9.ALL,T1T9.OUT,T1T9.INP.

The last option is to write the data to the PET disk or cassette tape. This is done, for example, to have a second copy of the results in case the VAX crashes, or to collect data when the VAX is not available. Simply type "W", and the program will ask whether you want to write to the disk or cassette, and will ask for a disk drive # and file name.

If you saved data on the PET disk or cassettes, they can be sent to the VAX as follows. Load TERM as you would for the T-JUMP program. Then, instead of loading T-JUMP, load FILES TO VAX, and run. The menu has options for talking to VAX, and sending the files. After logging in as was described above, talk to the VAX, enter the editor with a file name, go into insert mode, and return with <HOME>. Next, send the file by typing "S", at which time you will be asked whether you are sending from the cassette or the disk. If it is from the disk, you are asked for a file name and disk drive. The file is then sent as before. When the file is sent, you talk to VAX, leave input mode with <OFF/RVS> Z as before, followed by EX to exit the editor.

A word of caution is required here: The PET disks are well known for losing programs and data very routinely. The cassettes are more reliable, but much slower. I will generally save the data on both the disk and the cassette, relying on the cassette only if

the data on the disk is lost. If you rely on a single copy on a PET disk, you will often curse the PET at great length and emotion when your data is lost. The PET is used to it.

7. Snafus (Problems and Solutions)

There are a plethora of problems which come up because of the PET computer. Here I list several I am aware of, and the solutions.

One problem occurs when the Biomation will not trigger when performing the temperature-jump or by releasing and resetting the Slope +/- button. This is solved by typing on the PET:

```
<RUN/STOP>
READY
POKE 28723, 146
READY
POKE 28722,0
READY
GOTO 200  (the menu will appear.)
```

If the transmission stops in the middle (as seen by the numbers no longer appearing on the screen), you must turn off the power to the TNW 2000 box by turning off the circuit breaker to which it is plugged in, or by unplugging it if other things are plugged into that circuit breaker. The following sequence is then typed on the PET:

```
B*   PC   IRQ   SR   AC   XR   YR   SP
     1358 E26E ...
. X
CLOSE 4
```

OPEN 4,4

PRINT #4, "A"

GOTO 200 (menu appears)

T (to talk to the VAX.)

<OFF/RVS>Z (get out of input mode.)

* D %WH (delete buffer)

* I (re-enter input mode)

<HOME> (menu appears)

V (to send again.)

At which time you repeat the transmission. This problem sometimes never occurs in a day, and sometimes happens about every other time you send a file. It is caused by the VAX sending a control character that the PET doesn't receive.

If any problem occurs for which you appear to have no control, the surest bet is to turn off the PET, turn it back on, re-load the programs, and start fresh. Remember that the PETs are usually very unpredictable, and derive great pleasure at causing you much frustration and grief.

8. Turning Off the T-Jump

To turn off the T-jump after a day's experiments, you perform the following tasks.

Log off the VAX by typing "LO".

Turn off the PET, disk, Pet Interface to Cary 118.

Turn off the lamp power supply, and record the finishing time and number of hours on the log.

Discharge the capacitor if it is charged, and turn off the Wallis high voltage power supply. You will hear a short beep.

Turn off the Biomation and Tektronix.

Turn off the photomutiplier power supply - the red button labeled "NETZ".

Turn off the temperature bath, bath cooler, and digital thermometer.

Turn off the nitrogen flow.

Turn off the circuit breakers for the baths and thermometer.

Turn off the circuit breaker next to the sink for the HV supply and T-jump.

Remove cell, wash it, and go home.

9. Computer Considerations for the VAX

Several files are required in order to use the VAX as described above. Here I will assume the directory is DISK$USERFILE1:[RFP], although any directory will do. The sub-directory [RFP.TJUMP] should be created using CREATE/DIR [RFP.TJUMP].

The file [RFP]LOGIN.COM should contain the lines:

```
$ TJUMP:==@DISK$USERFILE1:[RFP]TJUMP
$ EDT:==EDIT/EDT/COMMAND=DISK$USERFILE1:[RFP]EDTINI.EDT
$ WRITE SYS$OUTPUT "TYPE 'TJUMP' TO INITIATE DATA COLLECTION"
```

The file [RFP]EDTINI.EDT should contain:

```
SET NONUMBERS
```

The file [RFP]TJUMP.COM should contain:

```
$ SET DEF DISK$USERFILE1:[RFP.TJUMP]
$ WRITE SYS$OUTPUT ""
$ WRITE SYS$OUTPUT "YOU ARE READY FOR TJUMP DATA COLLECTION"
$ WRITE SYS$OUTPUT "TYPE 'HELPME' FOR TO GET INSTRUCTIONS"
$ WRITE SYS$OUTPUT ""
```

```
$ HELPME:==TYPE DISK$USERFILE1:[JWN.EXPO]INSTR.DOC
$ ALL:==@DISK:USERFILE1:[RFP.TJUMP]ALL
$ DOIT:==@DISK:USERFILE1:[RFP.TJUMP]DOIT
$ INQUIRE PET "IS THIS THE PET ACTING AS TERMINAL? "
$ IF PET.NES."Y" THEN EXIT
$ SET TERM/HOSTSYNC/NOTAB/NOFORM
$ EXIT
```

The file [RFP.TJUMP]ALL.COM should contain:

```
$ INQUIRE P1 "FILE NAME "
$ OPEN/WRITE JOB JOB.COM
$ WRITE JOB "$ ASS DUMMY SYS$PRINT"
$ WRITE JOB "$ SET DEF [RFP.TJUMP]"
$ WRITE JOB "$ @ANALALL ",P1
$ CLOSE JOB
$ SUBMIT/NAME='P1' JOB
$ PURGE JOB.COM
```

The file [RFP.TJUMP]DOIT.COM should contain:

```
$ INQUIRE P1 "FILE NAME (E.G. TAPE1) "
$ INQUIRE FIRST "IS THIS THE FIRST TAPE IN SET? "
$ INQUIRE LAST "IS THIS THE LAST TAPE IN SET? "
$ IF LAST.EQS."Y" THEN INQUIRE P2 "NAME FOR THIS SET OF DATA"
$ @[RFP.TJUMP]ANAL 'P1'
$ IF FIRST.EQS."Y" THEN GOTO HERE
$ APPEND 'P1'.SUM SUM10.DAT
$ APPEND 'P1'.DAT TAPES.DAT
$ GOTO LAST_ONE
$ HERE:
$ COPY 'P1'.DAT TAPES.DAT
$ COPY 'P1'.SUM SUM10.DAT
$ LAST_ONE:
$ IF LAST.NE."Y" THEN GOTO ALL_DONE
$ RENAME TAPES.DAT 'P2'.DAT
$ RENAME SUM10.DAT 'P2'.SUM
$ ALL_DONE:
$ WRITE SYS$OUTPUT "*** ALL DONE WITH ANALYSIS ***"
$ EXIT
```

The file [RFP.TJUMP]ANAL.COM should contain:

```
$ DEASS/ALL
$ ASSIGN SYS$OUTPUT FOR006
$ ASSIGN 'P1'.DAT FOR005
$ ASSIGN 'P1'.CON FOR020
$ ASSIGN 'P1'.INP FOR010
$ RUN DISK$USERFILE1:[JWN.EXPO]CONVERT10
$ DEASS/ALL
$ ASSIGN 'P1'.CON FOR005
$ ASSIGN 'P1'.OUT FOR006
```

```
$ ASSIGN 'P1'.SUM FOR045
$ RUN DISK$USERFILE1:[JWN.EXPO]EXPO
$ DEASS/ALL
$ DELETE/NOCONFIRM 'P1'.CON;*
$ DELETE/NOCONFIRM 'P1'.INP;*
$ DELETE/NOCONFIRM 'P1'.OUT;*
```

Finally, the file [RFP.TJUMP]ANALALL.COM should contain:

```
$ ASSIGN SYS$OUTPUT FOR006
$ ASSIGN 'P1'.DAT FOR005
$ ASSIGN 'P1'.CON FOR020
$ ASSIGN 'P1'.INP FOR010
$ RUN DISK$USERFILE1:[JWN.EXPO]CONVERT
$ ASSIGN 'P1'.CON FOR005
$ ASSIGN 'P1'.OUT FOR006
$ ASSIGN 'P1'.ALL FOR045
$ RUN DISK$USERFILE1:[JWN.EXPO]EXPO
$ ASS SYS$INPUT FOR005
$ ASS SYS$OUTPUT FOR006
$ DELETE/NOCONFIRM 'P1'.CON;*
```

A problem often crops up after collecting data for a while, namely the disk quota for this particular username fills up, giving you "Disk quota exceeded" errors. This is remedied by deleting unneeded files, like output files after they have been printed. Also, data files may be deleted, but only after they have been backed-up on magnetic tape. This will allow later analysis in case you decide to use a different type of analysis, etc. This tape backup is done as follows.

After getting a new magnetic tape and mounting it on the tape drive, perform the following sequence <u>once only</u> (repeating this will destroy all data previously written to the tape):

```
$ ALLOCATE MT0:
MTA0: ALLOCATED
$ INITIALIZE/DENSITY=1600 MT0: LABEL  (LABEL is any 6-
                                       letter label used to identify the tape.)
```

Thereafter, you will allocate MTO: and perform the sequence:

$ MOUNT/DENSITY=1600 MTO: LABEL (same LABEL you set above.)

% MOUNT-I-MOUNTED, LABEL mounted on MTA0:

$ COPY T1T9.DAT MTO:* (for any data files T1T9.DAT)

...

$ DIRECTORY/PRINTER MTO: (gets a copy of files on tape)

$ DISMOUNT MTO:

$ DEALLOCATE MTO: (and remove tape from drive.)

It is wise to put a label on the tape with the following information: LABEL = LABEL (or whatever it is); DENSITY=1600; CONTENTS = T-JUMP DATA.

10. Using the Provencher Program

The data file sent to the VAX is made up as follows. The first line is a header, up to 80 characters, identifying the particular jump. The next line contains program instructions for the Provencher program, usually FTFTFFTFT 3 2 5 or something similar (see below). The next line contains the TO point, the TFIRST point, the TLAST point, the time/channel, and the number of data points (2048). The next 59 lines contain the data, 35 data points/line coded in 2-digit hexadecimal numbers with a 3-digit check sum which is the sum of the 70 1-digit hexadecimal numbers in the line. Another file can follow immediately after. The data are converted to a form suitable to the Provencher program by the program CONVER located in DISK$USERFILE1:[JWN.EXPO]CONVERT.EXE, for analyzing every point, and in DISK$USERFILE1:[JWN.EXPO]CONVERT10.EXE for analyzing

every tenth point. Examples of how this is routinely done are shown in the files ANAL.COM and ANALALL.COM listed earlier.

The programs CONVERT and CONVERT10 read the data files from unit FOR005, and write an eye-readable form of the data onto FOR010. The file FOR020 gets the file to be read by the Provencher program. If, for example, you just want to see the data for the file DATA.DAT, perform the sequence:

```
$ ASSIGN DATA.DAT FOR005
$ RUN DISK$USERFILE1:[JWN.EXPO]CONVERT
DONE WITH CONVERSION
$ DEASSIGN FOR005
$ PRINT FOR010.DAT
```

The programs CONVERT and CONVERT10 are listed at the end of this section.

The input for the Provencher program is arranged as follows. The first line of the input file contains an 80 character identification header. The second line contains 9 columns of T's or F's, selecting program options, followed by three 3-digit integers selecting the number of exponentials and the weighting of the data. The options of the first 9 columns are:

1. LAST:
   T = last data set, program will stop at end of this set.
   F = not last data set, program will look for more.
2. REGINT:
   T = data are in regular time intervals.
   F =data are not collected regularly.
3. NOBASE:
   T = the signal at infinite time is 0, i.e. there is no baseline.
   F = the baseline is not known, and the program determines it.

4. NONEG:
   T = all the exponential coefficients are positive, namely the data go from large to small values.
   F = some coefficients may be negative.
5. PRY:
   T = print out the input data.
   F = don't print out the input data.
6. PRPREL:
   T = print out information about the preliminary results in the analysis.
   F = don't, which is the value usually used.
7. PRFINL:
   T = print out information about the final iteration for the fit to the data.
   F = don't, which is also usually used.
8. PLORES:
   T = plot the resifuals for the best fit.
   F = don't.
9. REPEAT:
   T = print out a second copy of the summary of the results;
   F = don't. This is always set T to get a second copy of the results sent to file FOR045.

The three numbers following these 9 T's and F's are each three digits, and must be right-justified. The first is the number of exponential components the program searches for, usually set to at least one greater than the number you expect. It may be between 1 and 5. The second number determines the manner of weighing the data. Its values may be 1 = no weight, +2 or -2 = weight by 1/Y, +3 or -3 = weight by $1/Y^2$, and 4 = weights input in data file. The differences between +2 and -2 are small, as are the differences between +3 and -3. I don't know what the differences are. The third number is set to 5. It has no real purpose.

That takes care of the second line. The nature of the third line depends on whether the data are in regular intervals. If the data are not in regular intervals, the third line the number of points, in the first 5 columns, right-justified. The fourth card and subsequent cards contain the time values, 5 per card, in 5E15.6

format (15 columns), the exponent being right-justified if present. Following the time values are the Y-values, in 13F6.2 format, namely 13 values, using 6 columns each. (This format was changed from the original version to allow more efficient use of file space for the T-jump data, which come as 3-digit integers.) Following these are the weight, in 5E15.6 format, if the weighting function was 4.

If the data are in regular intervals, which is our case, the third card contains the number of intervals. Up to five separate intervals may be used, which allows use of the dual time base option on the Biomation. The format is I5, namely five columns, right-justified. The next line(s) contain the time for the first point in the set, the time for the last point in the set, and the number of points, in 2E15.6,I5 format, one line for each interval. The following lines contain the y data, in 13F6.2 format, as above. This is followed by the weights, if the weighting function = 4, in 5E15.6 format. An example of an input file written by program CONVERT is:

```
IDENTIFICATION FOR THIS DATA
FTFTFFTFT  3  2  5
    1
             0.         99.E-3 1990
125.  124.  122.  120.  117.  114.  112.  110.  105.  etc.
...
```

The Provencher program reads the data from file FOR005, writes the complete output to FOR006, and a one-page summary of the results to FOR045. This last file is the most useful, as the results for several consequetive data sets are listed very concisely. See the file ANALALL.COM above for an example of using the Provencher program.

The output of the program tells you the best fits assuming different numbers of exponential components, and their relative merits. For each solution, it tells you the alpha's (the exponential coefficients), the lambda's (which are $1/\tau$, where τ is the relaxation time for this component), and the estimated errors in each of these values. The standard deviations of the fits are also printed, along with lots of other information of unknown significance. See Adrienne Drobnies' thesis for an example.

The program listings for the programs CONVERT and CONVERT10 follow.

Program Convert. Program to convert temperature-jump data file into an input file for the Provencher program, analyzing every point.

```
      PROGRAM CONVER
C
C     ******************************************************************
C     * THIS PROGRAM CONVERTS T-JUMP DATA INTO THE FORM REQUIRED FOR
C     * THE PROVENCHER DISCRETE PROGRAM TO RESOLVE THE DATA INTO THE
C     * SUM OF UP TO 5 EXPONENTIALS.
C     *
C     * THE FORMAT FOR THE INPUT DATA FILE IS (FILE 5):
C     *
C     *       1.   80 CHARACTER IDENTIFICATION HEADER
C     *       2.   INSTRUCTIONS FOR PROVENCHER PROGRAM 9L1,3I3
C     *       3.   FIRST, LAST, T=0, TIME/CHANNEL, NO. OF POINTS IN
C     *                 DATA SET, NO. OF TIME BASES (1 OR 2), NO. OF POINTS
C     *                 IN FIRST TIME BASE, SECOND TIME/CHANNEL, NO. OF
C     *                 POINTS IN FIRST TIME BASE, (3I4,F8.0,I4,I4,I4,F8.0)
C     *       4    THE DATA IN HEXADECIMAL FORM, 35 TWO-DIGIT DATA POINTS
C     *                 WITH 3-DIGIT CHECKSUM PER LINE.
C     *                 CHECKSUM=SUM OF 70 1-DIG HEX NUMBERS.
C     *                 59 LINES, LAST LINE 0-FILLED.
C     *
C     * THE FORMAT FOR THE OUTPUT FILE IS (FILE 20)
C     *
C     *       1.   IDENTIFICATION HEADER
C     *       2.   INSTRUCTIONS FOR PROVENCHER PROGRAM
C     *       3.   NUMBER OF TIME BASES (1 OR 2)
C     *       4    TSTART, TEND, NO. OF POINTS IN FIRST TIME BASE
C     *       5.   TSTART, TEND, NO. OF POINTS IN SECOND TIME BASE
C     *                 (IF REQUIRED)
C     *       6.   THE DATA IN 13F6.0, 13 POINTS/RECORD
C     *            THE NEXT DATA SET FOLLOWS IMMEDIATELY IF PRESENT
C     *
C     * ANOTHER FILE ECHOS THE INPUT FILE (FILE 10).
C     *
C     *       1.   ID HEADER
C     *       2.   NO. OF POINTS, FIRST, LAST, T=0 POINTS,
C     *            TIME/CHANNEL (1 OR 2)
C     *       3    THE DATA IN 25I5, 25 POINTS/LINE
C     *
C     * THE DATA ARE READ IN AS INTEGERS IN HEXADECIMAL (Z FORMAT)
C     *       THE CONVERSION TO REAL NUMBERS IN OUTPUT IS DONE BY
C     *       INSERTING A DECIMAL POINT AS A HOLLERITH CONSTANT.
C     ******************************************************************
C     *       THIS PROGRAM WAS WRITTEN BY JEFF NELSON, OCT. 1980.
C     *       MODIFIED FOR DUAL TIME BASE OPERATION DEC. 1981.
C     ******************************************************************
C     $ VARIABLE TABLE...
C     $
C     $       HEAD1(20)   ... IDENTIFICATION HEADER, 20A4
C     $       HEAD2(20)   ... PROGRAM INSTRUCTIONS,  20A4
C     $       NDATA(2100) ... DATA IN INTEGER FORMAT
C     $       NSUM(60)    ... CHECKSUMS FOR EACH LINE
C     $       NT0         ... POINT AT WHICH T=0 (TEMP JUMP)
C     $       NFST        ... FIRST POINT TO BE USED IN ANALYSIS
```

```
C    $       NLST          ... LAST POINT TO BE USED IN ANALYSIS
C    $       TIME1         ... TIME/CHANNEL FOR FIRST TIME BASE
C    $       TIME2         ... TIME/CHANNEL FOR SECOND TIME BASE
C    $       NPOINT        ... NO. OF POINTS IN DATA SET (USU. 2048)
C    $       NTB           ... NO. OF TIME BASES (1 OR 2)
C    $       NCHAN         ... NO  OF POINTS IN FIRST TIME BASE (IF REQ'D)
C    $
C    $ FILE TABLE
C    $
C    $       FILE NO.  5 ... INPUT FILE
C    $       FILE NO.  6 ... OUTPUT FILE FOR DETECTED ERRORS IN CHECKSUMS
C    $       FILE NO. 10 ... FILE LISTING INPUT DATA
C    $       FILE NO. 20 ... OUTPUT FILE DESTINED FOR PROVENCHER PROG.
C    $
C    $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
C
      DIMENSION NDATA(2100),HEAD1(20),HEAD2(20),NSUM(60)
   10 READ(5,800,END=100)HEAD1
  800 FORMAT(20A4)
      GOTO 110
  100 STOP 'DONE WITH CONVERSION'
  110 READ(5,810)HEAD2,NFST,NLST,NT0,TIME1,NPOINT,NTB,NCHAN,TIME2
  810 FORMAT(20A4/3I4,F8.0,3I4,F8.0)
C   * NTB = NO. OF TIME BASES IS 0 OR 1 FOR 1 TB, 2 FOR 2.
C   * NCHAN, TIME2 HAVE NO MEANING IF NTB.NE.2
      IF(NTB.NE.2)NTB=1
C
C   * READ IN THE DATA
C
      READ(5,820)((NDATA(J+I-1),I=1,35),NSUM(J/35+1),
     2     J=1,NPOINT,35)
  820 FORMAT(35Z2,Z3)
C
C   * CHECK THE CHECKSUMS
C
      INDEX=0
      M2=(NPOINT+34)/35
C   * M2 IS THE NUMBER OF LINES IN THE FILE (USU. 59)
      DO 200 I=1,M2
        NTOT=0
        DO 180 J=1,35
          INDEX=INDEX+1
          NTOT=NTOT + (NDATA(INDEX).AND.'F'X) + (NDATA(INDEX)/16)
  180   CONTINUE
        IF(NTOT .NE. NSUM(I)) THEN
          WRITE(6,840)I,NTOT,NSUM(I)
  840     FORMAT(/,' **********  ERROR IN CHECKSUM **********',/,
     2             ' ********** ',I2,'TH LINE         **********',/,
     3             ' **********  NTOT=',I4,' NSUM=',I4,' **********')
          WRITE(10,840)I,NTOT,NSUM(I)
        ENDIF
  200 CONTINUE
C
C   * WRITE FILE 10 FOR LISTING TO THE LINE PRINTER
C
      WRITE(10,850)HEAD1,HEAD2,NPOINT,NT0,NFST,NLST,TIME1
  850 FORMAT(1H1,20A4/1H ,20A4/I5,' DATA POINTS'/
     2        ' T=0   POINT = ',I4/' FIRST POINT = ',I4/
     3        ' LAST  POINT = ',I4/' TIME/CHANNEL= ',E10.3)
      IF(NTB.EQ.2) WRITE(10,860)TIME2,NCHAN
  860 FORMAT(' TIME BASE 2  = ',E10.3/I5,' POINTS IN FIRST TB')
      WRITE(10,870)
  870 FORMAT(/)
      WRITE(10,880)(NDATA(I),I=1,NPOINT)
  880 FORMAT(1H ,25I5)
C
```

```
C    * WRITE FILE 20 FOR PROVENCHER PROGRAM
C
      IF (NTB  NE.  2) THEN
         NPTS=NLST-NFST+1
         TZERO=FLOAT(NFST-NT0)*TIME1
         TLAST=FLOAT(NLST-NT0)*TIME1
         WRITE(20,890)HEAD1,HEAD2,1,TZERO,TLAST,NPTS
  890    FORMAT(20A4/20A4/,I5,/ (1P2E15.7,I5) )
      ELSE
C    *    DUAL TIME BASE -
C    *       FIRST POINT OF T.B.1 IS NT0.  THIS IS 1, UNLESS THERE
C    *          THERE IS ROLLOVER FROM THE BIOMATION. WHERE NT0=2
C    *       FIRST POINT OF T.B.1 TO BE ANALYZED IS NFST,
C    *          TIME = (NFST-NT0)*TIME1.
C    *       LAST POINT OF T.B.1 IS NCHAN + NT0 (BIOMATION COLLECTS
C    *          NCHAN INTERVALS AFTER FIRST POINT IS TAKEN.)
C    *          TIME = NCHAN * TIME1.
C    *       NO.  OF POINTS IN T.B.1 IS NCHAN + NT0 - NFST + 1.
C    *       FIRST POINT OF T.B.2 IS NCHAN + NT0 + 1,
C    *          TIME = NCHAN * TIME1 + TIME2.
C    *       LAST POINT OF T.B.2 IS NLST.  TIME = NCHAN * TIME1 +
C    *          (NLST - NCHAN - NT0) * TIME2.
C    *       NO.  OF POINTS IN T.B.2 IS NLST - NCHAN - NT0.
C    *
         NPT1 = NCHAN + NT0 - NFST + 1
         NPT2 = NLST - NCHAN - NT0
         TZERO1 = FLOAT(NFST - NT0) * TIME1
         TLAST1 = FLOAT(NCHAN) * TIME1
         TZERO2 = TLAST1 + TIME2
         TLAST2 = TLAST1 + FLOAT(NLST - NCHAN - NT0) * TIME2
         WRITE(20,890)HEAD1,HEAD2,2,TZERO1,TLAST1,NPT1,
     2                TZERO2,TLAST2,NPT2
      ENDIF
      WRITE(20,900)(NDATA(I),I=NFST,NLST)
  900 FORMAT(13(I5,1H,))
      GOTO 10
      END
```

Program Convert10. Program to convert temperature-jump data file into an input file for the Provencher program, analyzing every tenth point.

```
      PROGRAM CONVER
C
C     *  *** THIS PROGRAM ANALYZES EVERY TENTH POINT ***
C
C     ******************************************************************
C     * THIS PROGRAM CONVERTS T-JUMP DATA INTO THE FORM REQUIRED FOR
C     * THE PROVENCHER DISCRETE PROGRAM TO RESOLVE THE DATA INTO THE
C     * SUM OF UP TO 5 EXPONENTIALS.
C     *
C     * THE FORMAT FOR THE INPUT DATA FILE IS (FILE 5):
C     *
C     *       1.   80 CHARACTER IDENTIFICATION HEADER
C     *       2.   INSTRUCTIONS FOR PROVENCHER PROGRAM 9L1,3I3
C     *       3.   FIRST, LAST, T=0, TIME/CHANNEL, NO. OF POINTS IN
C     *                 DATA SET, NO. OF TIME BASES (1 OR 2), NO. OF POINTS
C     *                 IN FIRST TIME BASE, SECOND TIME/CHANNEL,
C     *                  (3I4,F8.0,I4,I4,I4,F8.0).
C     *       4    THE DATA IN HEXADECIMAL FORM, 35 TWO-DIGIT DATA POINTS
C     *                 WITH 3-DIGIT CHECKSUM PER LINE.
C     *                 CHECKSUM=SUM OF 70 1-DIG HEX NUMBERS.
C     *                 59 LINES, LAST LINE 0-FILLED.
C     *
C     * THE FORMAT FOR THE OUTPUT FILE IS (FILE 20):
C     *
C     *       1.   IDENTIFICATION HEADER
C     *       2.   INSTRUCTIONS FOR PROVENCHER PROGRAM
C     *       3.   NUMBER OF TIME BASES (1 OR 2)
C     *       4.   TSTART, TEND, NO. OF POINTS IN FIRST TIME BASE
C     *       5.   TSTART, TEND, NO. OF POINTS IN SECOND TIME BASE
C     *                 (IF REQUIRED)
C     *       6.   THE DATA IN 13F6.0, 13 POINTS/RECORD.
C     *            THE NEXT DATA SET FOLLOWS IMMEDIATELY IF PRESENT
C     *
C     * ANOTHER FILE ECHOS THE INPUT FILE (FILE 10):
C     *
C     *       1.   ID HEADER
C     *       2.   NO. OF POINTS, FIRST, LAST, T=0 POINTS,
C     *            TIME/CHANNEL (1 OR 2)
C     *       3    THE DATA IN 25I5, 25 POINTS/LINE
C     *
C     * THE DATA ARE READ IN AS INTEGERS IN HEXADECIMAL (Z FORMAT)
C     *       THE CONVERSION TO REAL NUMBERS IN OUTPUT IS DONE BY
C     *       INSERTING A DECIMAL POINT AS A HOLLERITH CONSTANT.
C     *
C     *  *** THIS VERSION ANALYZES EVERY 10TH POINT. ***
C     *
C     ******************************************************************
C     *       THIS PROGRAM WAS WRITTEN BY JEFF NELSON, OCT., 1980.
C     *       MODIFIED FOR DUAL TIME BASE OPERATION DEC., 1981.
C     ******************************************************************
C     $ VARIABLE TABLE...
```

```
C    $
C    $       HEAD1(20)   ...  IDENTIFICATION HEADER, 20A4
C    $       HEAD2(20)   ...  PROGRAM INSTRUCTIONS,  20A4
C    $       NDATA(2100) ..   DATA IN INTEGER FORMAT
C    $       NSUM(60)    ...  CHECKSUMS FOR EACH LINE
C    $       NT0          ..  POINT AT WHICH T=0 (TEMP JUMP)
C    $       NFST        ...  FIRST POINT TO BE USED IN ANALYSIS
C    $       NLST        ...  LAST POINT TO BE USED IN ANALYSIS
C    $       TIME1       ...  TIME/CHANNEL FOR FIRST TIME BASE
C    $       TIME2            TIME/CHANNEL FOR SECOND TIME BASE
C    $       NPOINT       ..  NO. OF POINTS IN DATA SET (USU. 2048)
C    $       NTB          ..  NO. OF TIME BASES (1 OR 2)
C    $       NCHAN       ...  NO. OF POINTS IN FIRST TIME BASE (IF REQ'D)
C    $
C    $ FILE TABLE.
C    $
C    $       FILE NO.  5 ...  INPUT FILE
C    $       FILE NO.  6 ..   OUTPUT FILE FOR DETECTED ERRORS IN CHECKSUMS
C    $       FILE NO  10  .   FILE LISTING INPUT DATA
C    $       FILE NO. 20 ...  OUTPUT FILE DESTINED FOR PROVENCHER PROG
C    $
C    $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
C
      DIMENSION NDATA(2100),HEAD1(20),HEAD2(20),NSUM(60)
   10 READ(5,800,END=100)HEAD1
  800 FORMAT(20A4)
      GOTO 110
  100 STOP 'DONE WITH CONVERSION'
  110 READ(5,810)HEAD2,NFST,NLST,NT0,TIME1,NPOINT,NTB,NCHAN,TIME2
  810 FORMAT(20A4/3I4,F8.0,3I4,F8.0)
C     * NTB = NO. OF TIME BASES IS 0 OR 1 FOR 1 TB, 2 FOR 2.
C     * NCHAN, TIME2 HAVE NO MEANING IF NTB.NE.2
      IF(NTB.NE.2)NTB=1
C
C     * READ IN THE DATA
C
      READ(5,820)((NDATA(J+I-1),I=1,35),NSUM(J/35+1),
     2      J=1,NPOINT,35)
  820 FORMAT(35Z2,Z3)
C
C     * CHECK THE CHECKSUMS
C
      INDEX=0
      M2=(NPOINT+34)/35
C     * M2 IS THE NUMBER OF LINES IN THE FILE (USU. 59)
      DO 200 I=1,M2
        NTOT=0
        DO 180 J=1,35
          INDEX=INDEX+1
          NTOT=NTOT + (NDATA(INDEX).AND.'F'X) + (NDATA(INDEX)/16)
  180   CONTINUE
        IF(NTOT .NE. NSUM(I)) THEN
          WRITE(6,840)I,NTOT,NSUM(I)
  840     FORMAT(/,' **********  ERROR IN CHECKSUM **********',/,
     2             ' ********** ',I2,'TH LINE         **********',/,
     3             ' **********  NTOT=',I4,' NSUM=',I4,' **********')
          WRITE(10,840)I,NTOT,NSUM(I)
        ENDIF
  200 CONTINUE
C
C     * WRITE FILE 10 FOR LISTING TO THE LINE PRINTER
C
      WRITE(10,850)HEAD1,HEAD2,NPOINT,NT0,NFST,NLST,TIME1
  850 FORMAT(1H1,20A4/1H ,20A4/I5,' DATA POINTS'/
     2       ' T=0   POINT = ',I4/' FIRST POINT = ',I4/
     3       ' LAST  POINT = ',I4/' TIME/CHANNEL= ',E10.3)
```

```
      IF(NTB.EQ.2) WRITE(10,860)TIME2,NCHAN
  860 FORMAT(' TIME BASE 2 = ',E10.3/I5,' POINTS IN FIRST TB')
      WRITE(10,870)
  870 FORMAT(/)
      WRITE(10,880)(NDATA(I),I=1,NPOINT)
  880 FORMAT(1H ,25I5)
C
C   * WRITE FILE 20 FOR PROVENCHER PROGRAM
C
      IF (NTB .NE. 2) THEN
         NPTS=(NLST-NFST)/10 + 1
         TZERO=FLOAT(NFST-NT0)*TIME1
         TLAST=TZERO + 10.*FLOAT(NPTS-1)*TIME1
         WRITE(20,890)HEAD1,HEAD2,1,TZERO,TLAST,NPTS
  890    FORMAT(20A4/20A4/,I5,/ (1P2E15.7,I5) )
      ELSE
C   *    DUAL TIME BASE -
C   *       FIRST POINT OF T.B.1 IS NT0.  THIS IS 1, UNLESS THERE
C   *          THERE IS ROLLOVER FROM THE BIOMATION, WHERE NT0=2.
C   *       FIRST POINT OF T.B.1 TO BE ANALYZED IS NFST,
C   *          TIME = (NFST-NT0)*TIME1.
C   *       LAST POINT OF T.B.1 IS NCHAN + NT0 (BIOMATION COLLECTS
C   *          NCHAN INTERVALS AFTER FIRST POINT IS TAKEN.)
C   *          TIME = NCHAN * TIME1.
C   *       NO. OF POINTS IN T.B.1 IS NCHAN + NT0 - NFST + 1.
C   *       FIRST POINT OF T.B.2 IS NCHAN + NT0 + 1,
C   *          TIME = NCHAN * TIME1 + TIME2.
C   *       LAST POINT OF T.B.2 IS NLST, TIME = NCHAN * TIME1 +
C   *          (NLST - NCHAN - NT0) * TIME2.
C   *       NO. OF POINTS IN T.B.2 IS NLST - NCHAN - NT0.
C   *
         NPT1 = 1 + (NCHAN + NT0 - NFST) / 10
         TZERO1 = FLOAT(NFST - NT0) * TIME1
         TLAST1 = TZERO1 + 10. * TIME1 * FLOAT(NPT1 - 1)
C
         N2F = NFST + 10 * (NPT1 - 1) + 10
         NPT2 = 1 + (NLST - N2F) / 10
         TZERO2 = FLOAT(NCHAN) * TIME1 + FLOAT(N2F - NCHAN) * TIME2
         TLAST2 = TZERO2 + 10. * TIME2 * FLOAT(NPT2 - 1)
         WRITE(20,890)HEAD1,HEAD2,2,TZERO1,TLAST1,NPT1,
     2                TZERO2,TLAST2,NPT2
      ENDIF
      WRITE(20,900)(NDATA(I),I=NFST,NLST,10)
  900 FORMAT(13(I5,1H,))
      GOTO 10
      END
```

11. Program Listings for the PET

The following programs used for the PET are listed below. The program SET UP initializes and loads the programs from a disk. The program T-JUMP collects the data and saves it on VAX or PET disks. This is the program we have been concerned with above. TERM contains assembly-language routines to transfer the data from the Biomation, talks to the VAX, and converts the data into the file which is sent to the VAX or PET disks. It occupies memory locations 1300_{hex} to $1FFF_{hex}$, and its memory map is listed below. The assembly language program is in a folder by the T-jump. The program FILES TO VAX sends files saved on PET disk or cassettes to the VAX, and is used if data were collected while the VAX was not available.

In order to initialize a PET disk to contain the programs, you can either duplicate a disk that already has the programs (see PET 2040 disk manual), or put a new disk into drive 0,initialize it and perform the following sequence:

Load the program "SET UP" from the cassette by the PET. Then (The responsed from the PET are deleted):

```
OPEN 8,8,15
PRINT#8,"IO"
SAVE "@0:"SET UP"
```

Then, load the program TERM from the cassette labeled "TERM,T-JUMP" by inserting that cassette, and:

```
LOAD
SYS(64785)
.S "@0:TERM",08,1300,2000
```

```
.X
NEW
```

Then, load the program T-JUMP from the same cassette:

```
LOAD
SAVE "@0:T-JUMP",8
```

Finally, load the program "FILES TO VAX" from the cassette,

```
LOAD
SAVE "@0:FILES TO VAX",8
```

Which completes the process of initializing the disk. It may now be used to load the programs as described earlier, and to have data written to it.

The listings of the programs follow.

Program Set Up. Program to load the temperature-jump programs from the PET disk.

```
100 POKE 52,0:POKE 53,19
110 PRINT"3T-JUMP DATA COLLECTION SET UP"
120 PRINT
130 PRINT"DO THIS:":PRINT
140 PRINT"    1 - PUT DISK IN DRIVE 0"
150 PRINT"    2 - TYPE 'TERM'"
160 PRINT"    3 - TYPE 'T-JUMP'"
200 INPUT K$
210 IFK$="TERM"GOTO500
220 IFK$="T-JUMP"GOTO600
230 PRINT"WHAT??":GOTO200
500 CLOSE8
510 OPEN8,8,15
520 PRINT#8,"I0"
530 PRINT:PRINT"  ** LOADING TERM **"
540 LOAD"TERM",8
600 PRINT:PRINT" TYPE:  LOAD ";CHR$(34);"T-JUMP";CHR$(34);",8"
610 PRINT:PRINT" THEN:  RUN"
620 NEW
```

Program T-Jump. Program to collect data from the temperature-jump instrument.

```
100 POKE 52,0:POKE 53,19:REM SAVE MEM
110 POKE 28723,146:REM 8255 MODE
115 POKE 28722,0
120 DIM H$(3):REM FILE HEADER, ETC.
200 PRINT"3"
210 PRINT" B2 TRANSFER FROM B2IOMATION":PRINT
220 PRINT" T2 T2ALK TO VAX":PRINT
230 PRINT" F2 SET F2ILE PARAMETERS":PRINT
240 PRINT" V2 SEND FILE TO V2AX":PRINT
250 PRINT" W2 W2RITE FILE TO DISK OR TAPE":PRINT
290 GET K$:IFK$=""GOTO290
300 IFK$="B"GOTO1000
310 IFK$="T"GOTO1500
320 IFK$="F"GOTO2000
330 IFK$="V"GOTO4500
340 IFK$="W"GOTO3000
380 FNTN 200
1000 REM TRANSFER DATA FROM BIOMATION
1005 PRINT"3BIOMATION TRANSFER":PRINT
1010 PRINT"TYPE 'Y' WHEN READY"
1020 GET K$:IFK$<>"Y"GOTO1020
1025 PRINT"Y"
1030 IF(PEEK(28721)AND9)<>0 GOTO1030
1040 SYS(7232)
1040 PRINT:PRINT"TRANSFER FROM BIOMATION COMPLETE"
1060 T=TI
1070 IF(TI-T)<40 GOTO1070
1080 GOTO200
1500 REM TALK TO VAX
1510 PRINT"3TALK TO VAX. "
1520 SYS(7686)
1530 PRINT"3DONE WITH VAX"
1540 T=TI
1550 IF TI-T<40 GOTO1550
1560 GOTO 200
2000 REM GET FILE PARAMETERS-
2010 PRINT"3":PRINT" LAST HEADER ... "
2020 PRINTH$(0)
2030 INPUT H$(0)
2050 PRINT:PRINT" LAST PROG. INSTR. ... "
2060 PRINTH$(1)
2070 INPUT H$(1)
2080 PRINT:PRINT"FIRST 100 POINTS:":PRINT
2090 IN%=5119
2100 BL$="      "
2100 FNR I=1TO100
2120 IN$<IN%+1
2020 R$<STR$(PEEK(IN%))
2140 PRINTLEFT$(BL$,4-LEN(S$));S$;
2150 NEXTI
2160 IN%<6046:REM LAST 20 PTS
2160 PRINT:PRINT"LAST 20 POINTS:":PRINT
2180 FOR I=1TO20:IN%=IN%+1:S$=STR$(PEEK(IN%)):PRINTLEFT$(BL$,4-LEN(S$)),S$,
```

```
2190 NEXTI
2200 PRINT:PRINT"T0   =";T0%,"NEW T0? "
2210 GET K$:IFK$=""GOTO2210
2220 IF K$="Y"THEN INPUT" NEW=";T0%
2230 PRINT:PRINT"FIRST POINT =";TF%,"NEW FIRST? "
2240 GET K$:IFK$=""GOTO2240
2250 IF K$="Y"THEN INPUT" NEW=";TF%
2260 PRINT:PRINT"LAST POINT =";TL%,"NEW LAST? "
2270 GET K$:IFK$=""GOTO2270
2280 IFK$="Y"THENINPUT" NEW=";TL%
2290 PRINT:PRINT"TIME/CHANNEL=";TC$,"NEW T/C? "
2300 GET K$:IFK$=""GOTO2300
2310 IFK$="Y"THENINPUT" NEW=";TC$
2320 PRINT:PRINT"TWO TIME BASES?"
2330 GET K$:IFK$=""GOTO2330
2340 IFK$<>"Y"THEN NT%=1:GOTO2430
2350 NT%=2
2360 PRINT:PRINT"2ND TIME/CHANNEL=";T2$,"NEW T/C? "
2370 GET K$:IFK$=""GOTO2370
2380 IFK$="Y"THENINPUT" NEW=";T2$
2390 PRINT:PRINT"POINTS IN FIRST T.B.=";CH%,"DIFFERENT VALUE? "
2400 GET K$:IFK$=""GOTO2400
2410 IFK$="Y"THENINPUT" NEW=";CH%
2430 PRINT:PRINT:PRINTH$(0):PRINTH$(1):PRINT:PRINT"T0   ";T0%
2440 PRINT"FIRST";TF%:PRINT"LAST ";TL%
2450 PRINT:PRINT"TIME/CHANNEL  ";TC$
2460 IF NT%<2 GOTO 2490
2470 PRINT"2ND T/C        ";T2$:PRINT"POINTS IN 1ST T/C";CH%
2490 PRINT:PRINT:PRINT"ALL OK?";
2492 GETK$:IFK$=""GOTO2492
2493 IFK$="N"GOTO 2000
2500 REM SET UP H$(2)
2510 BL$="          "
2515 Z$=STR$(TF%)
2520 S$="00000"+RIGHT$(Z$,LEN(Z$)-1)
2530 H$(2)=RIGHT$(S$,4)
2540 S$="     "+STR$(TL%)
2550 H$(2)=H$(2)+RIGHT$(S$,4)
2560 S$="     "+STR$(T0%)
2570 H$(2)=H$(2)+RIGHT$(S$,4)
2580 H$(2)=H$(2)+LEFT$(BL$,8-LEN(TC$))+TC$
2590 H$(2)=H$(2)+"2048"
2600 IF NT%<>2 GOTO 2690
2610 H$(2)=H$(2)+"   2"
2620 S$="     "+STR$(CH%)
2630 H$(2)=H$(2)+RIGHT$(S$,4)
2640 H$(2)=H$(2)+RIGHT$(BL$,8-LEN(T2$))+T2$
2690 GOTO 200
3000 REM WRITE TO DISK OR TAPE
3010 PRINT"3WRITE TO T2APE OR D2ISK?"
3020 GETK$:IFK$=""GOTO3020
3030 IFK$="T"GOTO3120
3040 IFK$<>"D"GOTO3010
3050 PRINT:INPUT"DISK DRIVE";F$
3055 INPUT "FILE NAME";F2$
3060 F$=F$+":"+F2$
3070 F$=F$+",SEQ,WRITE"
3080 OPEN 10,8,2,F$
3090 GOTO 3200
3120 PRINT:INPUT"TAPE FILE NAME";F$
3130 OPEN 10,1,2,F$
3200 FORI=0TO2:PRINT#10,H$(I);CHR$(13);:NEXTI
3210 POKE7436,0:POKE7437,20
3220 FORI=1TO59
3230 SYS(7424)
3240 FORJ=0TO72:PRINT#10,CHR$(PEEK(7562+J));:NEXTJ
```

```
3250 PRINT#10,CHR$(13);
3255 PRINTI;
3260 NEXTI
3300 CLOSE10
3310 GOTO200
3500 PRINT"TRANSFER FROM DISK TO VAX"
3510 GOTO200
4000 PRINT"TRANSFER FROM DISK TO VAX"
4010 GOTO200
4500 REM SEND DATA TO VAX
4510 REM PUT HEADER INTO LOCATIONS
4520 REM $1FA0 TO $1FFF
4530 REM OR LOC 8096-8192
4540 PRINT"3TRANSFER DATA TO VAX":PRINT
4550 PRINT"TYPE 'Y' TO START"
4560 GET K$:IFK$<>"Y"GOTO4560
4570 PRINT:PRINT"SENDING TO VAX"
4600 FOR I=0 TO 2
4610 FOR J=0 TO LEN(H$(I))-1
4620 POKE 8096+J,(ASC(MID$(H$(I),J+1,1))AND127)
4630 NEXT J
4640 POKE 5052,8096AND255
4650 POKE 5053,INT(8096/256)
4660 POKE 5024,4:REM IEEE DEVICE#
4670 POKE 5025,LEN(H$(I))
4680 SYS(5028):REM SEND LINE
4690 NEXT I
4800 REM NOW SEND DATA
4810 POKE 7436,0:POKE 7437,20
4820 POKE 5052,138:POKE 5053,29
4830 POKE 5025,73
4840 FOR I=1TO59
4850 SYS(7424)
4860 SYS(5028)
4870 PRINTI;
4880 NEXT
4890 PRINT:PRINT:PRINT"DONE WITH TRANSFER"
4900 T=TI
4910 IF TI<T+40 GOTO4910
4990 GOTO 200
```

Program Term. Memory map for the assembly language routines used by the program T-Jump. Locations 1400_{hex} - $1BFF_{hex}$ are used for data storage.

```
. : 1300 48 49 FF 8D 22 E8 98 48
. : 1308 A0 00 C8 D0 03 00 00 00
. : 1310 AD 40 E8 29 40 F0 F3 AD
. : 1318 23 E8 29 F7 8D 23 E8 A0
. : 1320 00 C8 D0 03 00 00 00 AD
. : 1328 40 E8 29 01 F0 F3 AD 23
. : 1330 E8 09 08 8D 23 E8 A9 FF
. : 1338 8D 22 E8 68 A8 68 60 EA
. : 1340 24 24 24 24 24 24 24 24
. : 1348 98 48 AD 40 E8 09 02 8D
. : 1350 40 E8 A0 00 C8 D0 03 00
. : 1358 00 00 AD 40 E8 29 80 D0
. : 1360 F3 AD 40 E8 29 FD 8D 40
. : 1368 E8 AD 20 E8 49 FF 8D 98
. : 1370 13 AD 21 E8 09 08 8D 21
. : 1378 E8 A0 00 C8 D0 03 00 00
. : 1380 00 AD 40 E8 29 80 F0 F3
. : 1388 AD 21 E8 29 F7 8D 21 E8
. : 1390 A9 FF 8D 22 E8 68 A8 AD
. : 1398 98 13 60 00 EA EA EA EA
. : 13A0 04 49 EA EA AD A1 13 D0
. : 13A8 01 60 78 8D C8 13 AE A0
. : 13B0 13 A0 00 A9 11 20 50 1C
. : 13B8 20 E8 13 B9 8A 1D 29 7F
. : 13C0 20 50 1C 20 E8 13 C8 C0
. : 13C8 49 D0 F0 A9 0D 20 50 1C
. : 13D0 20 E8 13 A9 13 20 50 1C
. : 13D8 20 E8 13 20 E8 13 20 E8
. : 13E0 13 58 60 EA EA EA EA EA
. : 13E8 20 88 1C 29 7F C9 13 D0
. : 13F0 09 20 88 1C 29 7F C9 11
. : 13F8 D0 F7 60 EA EA EA EA EA

. : 1C00 A9 01 8D 32 70 A9 00 8D
. : 1C08 1B 1C A9 14 8D 1C 1C EA
. : 1C10 AD 31 70 29 02 F0 F9 AD
. : 1C18 30 70 8D 00 00 EE 1B 1C
. : 1C20 D0 0A EE 1C 1C AD 1C 1C
. : 1C28 C9 1C F0 0D A9 03 8D 32
. : 1C30 70 A9 01 8D 32 70 4C 10
. : 1C38 1C A9 00 8D 32 70 60 EA
. : 1C40 78 20 00 1C 58 60 EA EA
. : 1C48 AA AA AA AA AA AA AA AA
. : 1C50 48 AD 40 E8 29 FB 8D 40
. : 1C58 E8 8A 09 20 20 00 13 AD
. : 1C60 40 E8 09 04 8D 40 E8 68
. : 1C68 48 20 00 13 AD 40 E8 29
. : 1C70 FB 8D 40 E8 A9 3F EA 20
. : 1C78 00 13 AD 40 E8 09 04 8D
. : 1C80 40 E8 68 60 EA EA EA EA
```

```
. : 1C88 AD 40 E8 29 FB 8D 40 E8
. : 1C90 8A 09 40 20 00 13 AD 40
. : 1C98 E8 29 FD 8D 40 E8 AD 21
. : 1CA0 E8 29 F7 8D 21 E8 AD 40
. : 1CA8 E8 09 04 8D 40 E8 20 48
. : 1CB0 13 48 AD 40 E8 29 FB 8D
. : 1CB8 40 E8 AD 40 E8 09 02 8D
. : 1CC0 40 E8 AD 21 E8 09 08 8D
. : 1CC8 21 E8 A9 5F 20 00 13 AD
. : 1CD0 40 E8 09 04 8D 40 E8 68
. : 1CD8 60 AA AA AA AA AA AA AA
. : 1CE0 AA AA AA AA AA AA AA AA
. : 1CE8 AA AA AA AA AA AA AA AA
. : 1CF0 AA AA AA AA AA AA AA AA
. : 1CF8 AA AA AA AA AA AA AA AA

. : 1D00 A9 00 8D 86 1D 8D 87 1D
. : 1D08 A9 BA AA AD 23 14 A8 4A
. : 1D10 4A 4A 4A 20 60 1D 9D D0
. : 1D18 1C E8 98 20 60 1D 9D D0
. : 1D20 1C EE 0C 1D D0 0A EE 0D
. : 1D28 1D AD 0D 1D C9 1C F0 24
. : 1D30 E8 D0 D8 AD 87 1D 20 70
. : 1D38 1D 9D D0 1D E8 AD 86 1D
. : 1D40 A8 4A 4A 4A 4A 20 70 1D
. : 1D48 9D D0 1D E8 98 20 70 1D
. : 1D50 9D D0 1D 60 A9 30 E8 F0
. : 1D58 DA 9D D0 1C 4C 56 1D EA
. : 1D60 29 0F 48 18 6D 86 1D 8D
. : 1D68 86 1D 90 03 EE 87 1D 68
. : 1D70 29 0F 09 30 C9 3A 30 03
. : 1D78 18 69 07 60 EA EA EA 24
. : 1D80 56 00 00 00 00 00 09 01
. : 1D88 AA AA 30 30 30 31 30 32
. : 1D90 30 33 30 34 30 35 30 36
. : 1D98 30 37 30 38 30 39 30 41
. : 1DA0 30 42 30 43 30 44 30 45
. : 1DA8 30 46 31 30 31 31 31 32
. : 1DB0 31 33 31 34 31 35 31 36
. : 1DB8 31 37 31 38 31 39 31 41
. : 1DC0 31 42 31 43 31 44 31 45
. : 1DC8 31 46 32 30 32 31 32 32
. : 1DD0 31 30 39 AA AA AA AA AA
. : 1DD8 AA AA AA AA AA AA AA AA
. : 1DE0 AA AA AA AA AA AA AA AA
. : 1DE8 AA AA AA AA AA AA AA AA
. : 1DF0 AA AA AA AA AA AA AA AA
. : 1DF8 AA AA AA AA AA AA AA AA

. : 1E00 93 13 0D 00 EA EA EA EA
. : 1E08 58 20 E4 FF F0 12 8D 01
. : 1E10 1E C9 13 D0 01 60 20 70
. : 1E18 1E 4C 20 1E EA EA EA EA
. : 1E20 78 A2 04 A9 11 20 50 1C
. : 1E28 A0 01 20 80 1F 29 7F C9
. : 1E30 13 F0 16 C0 01 D0 07 48
. : 1E38 A9 13 20 50 1C 68 C9 00
. : 1E40 F0 21 99 A8 1F C8 4C 2A
. : 1E48 1E 20 88 1C 29 7F F0 F9
. : 1E50 C9 11 F0 07 99 A8 1F C8
. : 1E58 4C 49 1E A9 13 20 50 1C
. : 1E60 4C 2A 1E 4C 58 1F EA EA
. : 1E68 EA EA EA EA EA EA EA EA
. : 1E70 AD 01 1E 8D 00 1E C9 93
```

```
. :  1E78 D0 09 20 D2 FF A9 00 8D
. :  1E80 03 1E 60 C9 83 D0 01 60
. :  1E88 C9 12 F0 15 C9 92 D0 28
. :  1E90 AD 03 1E D0 01 60 A9 00
. :  1E98 8D 03 1E A9 14 20 D2 FF
. :  1EA0 60 8D 03 1E A9 12 20 D2
. :  1EA8 FF A9 20 20 D2 FF A9 92
. :  1EB0 20 D2 FF 60 EA EA EA EA
. :  1EB8 AE 03 1E F0 0A A2 00 8E
. :  1EC0 03 1E 29 1F 4C DE 1E A2
. :  1EC8 10 DD E7 1E D0 0A BD E6
. :  1ED0 1E 8D 00 1E 4C DE 1E EA
. :  1ED8 CA CA D0 ED 29 7F 8D 00
. :  1EE0 1E A2 04 20 50 1C 60 EA
. :  1EE8 1B 03 0A 11 FF 14 7F 94
. :  1EF0 09 1D 0C 91 0D 81 08 90
. :  1EF8 EA EA EA EA EA EA EA EA

. :  1F00 EA EA EA EA EA EA EA EA
. :  1F08 AD 00 1E 8D 02 1E A2 0A
. :  1F10 DD 47 1F D0 0A BD 46 1F
. :  1F18 8D 02 1E 20 D2 FF 60 CA
. :  1F20 CA D0 ED C9 60 90 05 29
. :  1F28 5F 4C 18 1F C9 20 B0 E8
. :  1F30 A9 12 20 D2 FF AD 00 1E
. :  1F38 09 40 20 D2 FF A9 92 20
. :  1F40 D2 FF 60 EA EA EA EA EA
. :  1F48 9D 08 11 0A 14 7F 0D 0D
. :  1F50 14 5F EA EA EA EA EA EA
. :  1F58 58 8C 60 1F A0 00 C8 C0
. :  1F60 01 F0 0E B9 A8 1F 29 7F
. :  1F68 8D 00 1E 20 08 1F 4C 5E
. :  1F70 1F 4C 08 1E EA EA EA EA
. :  1F78 EA EA EA EA EA EA EA EA
. :  1F80 98 48 A0 FB 20 88 1C F0
. :  1F88 08 8D 8F 1F 68 A8 A9 0D
. :  1F90 60 C8 D0 F0 68 A8 A9 00
. :  1F98 60 AA AA AA AA AA AA AA
. :  1FA0 AA AA AA AA AA AA AA AA
. :  1FA8 AA 35 38 0D 33 2E AA AA
. :  1FB0 AA AA AA AA AA AA AA AA
. :  1FB8 AA AA AA AA AA AA AA AA
. :  1FC0 AA AA AA AA AA AA AA AA
. :  1FC8 AA AA AA AA AA AA AA AA
. :  1FD0 AA AA AA AA AA AA AA AA
. :  1FD8 AA AA AA AA AA AA AA AA
. :  1FE0 AA AA AA AA AA AA AA AA
. :  1FE8 AA AA AA AA AA AA AA AA
. :  1FF0 AA AA AA AA AA AA AA AA
. :  1FF8 AA 4D 4F 4E 49 54 4F AA
```

Program Files to Vax. Program to send data files from the PET disk and cassette tapes to the VAX.

```
10 REM SEND T-JUMP DISK FILE TO VAX
100 POKE52,0:POKE53,19
300 PRINT"3 T2 T2ALK TO VAX":PRINT
310 PRINT" S2 S2END FILE TO VAX":PRINT
400 GET K$:IFK$=""GOTO400
410 IFK$="T"GOTO1000
420 IFK$="S"GOTO2000
430 GOTO300
1000 PRINT"3 TALK TO VAX:"
1010 SYS(7688)
1020 PRINT"3 DONE WITH VAX"
1030 T=TI
1040 IF TI<T+40 GOTO1040
1050 GOTO300
2000 REM SEND FILE TO VAX
2010 REM BUFFER STARTS AT LOC 8096
2020 REM LOC 5024 HAS IEEE DEVICE#
2030 REM LOC 5052,5053 HAS BUFFER LOC
2040 POKE 5024,4:REM IEEE DEVICE #
2050 POKE 5052,8096AND255
2060 POKE 5053,INT(8096/256)
2070 SL=8096
2075 INPUT " D2ISK OR C2ASSETTE";Q$
2080 IF Q$="C"THEN OPEN 8,1:GOTO 2500
2085 IF Q$<>"D"GOTO 2075
2090 INPUT " DRIVE #   = ";D$
2100 INPUT " FILE NAME = ";S$
2200 OPEN 8,8,14,D$+":"+S$+",SEQ,READ"
2500 FOR I=1TO62:REM 62 LINES IN FILE
2510 INPUT#8,L$
2515 PRINTI;L$
2520 FOR J=0 TO LEN(L$)-1
2530 POKE J+SL,(ASC(MID$(L$,J+1,1))AND127)
2540 NEXT J
2550 POKE 5025,LEN(L$)
2560 SYS(5028)
2580 NEXT I
2590 PRINT" DONE WITH TRANSMISSION"
2595 CLOSE 8
2600 T=TI
2610 IF TI<T+40 GOTO2610
2620 GOTO 300
```

APPENDIX C

Computer Programs to Analyze Ethidium Melting Data

The programs to analyze the ethidium melting data are located in the directory DISK$USERFILE1:[JWN.MELT] of the Biodynamics VAX 11/780. There are four of them: MELT.FOR, DYEMLT.FOR, DYEPLT.FOR and STAT.FOR. The programs are compiled and linked essentially identically:

```
$ FOR/CHECK=ALL DISK$USERFILE1:[JWN.MELT]MELT
$ LIN MELT,DISK$USERFILE1:[JWN]JWNLIB/LIB
$ FOR/CHECK=ALL DISK$USERFILE1:[JWN.MELT]DYEMLT
$ LIN DYEMLT,DISK$USERFILE1:[JWN]JWNLIB/LIB
$ FOR/CHECK=ALL DISK$USERFILE1:[JWN.MELT]DYEPLT
$ LIN DYEPLT,[JWN]JWNLIB/LIB,IGL/LIB
$ FOR/CHECK=ALL DISK$USERFILE1:[JWN.MELT]STAT
$ LIN STAT,[JWN]JWNLIB/LIB,IGL/LIB
```

However, the programs should have execuatble images under the same names, so compiling and linking should not be necessary, and you can simply run the programs. This appendix explains the use of these programs, with examples, and lists the source codes for DYEMLT and STAT. The programs are cyclic, meaning that you traverse a loop indefinitely. Each time through the loop allows you to change some parameters, while keeping others the same, allowing great flexibility on the use of the programs.

MELT.FOR subtracts the reference from the sample, and interpolates the data to regular intervals, starting at 0°C and continuing at 1°C intervals. The use of this program is essentially like MELT1 on the PDP-8E. Subtract the cell blanks, but set the

high temperature absorbance to 1., and interpolate to regular intervals of 1°C.

The program DYEMLT takes one sample from the melting curves, and determines the upper and lower baselines from the extinction coefficients input. From these baselines, the fraction double strands and fraction ethidium bound are calculated. The data must be at 260 and 283 nm, and must be interpolated to intervals of 1°C. In practice, several attempts must be made at estimating the values of the extinction coefficients for the double strands and the bound ethidium. An example of running this program is given below.

A visual check of the fit is given by running the program DYEPLT, which plots the experimental data, the baselines, and the fractions double strands and ethidium bound calculated from them. The program usually uses default plotting parameters, however the option is given to change the plot parameters to make plots for figures, or to plot the number of ethidium ions bound instead of the fraction of ethidium ions bound. The program is basically self-explainatory, and no example or listing is given here.

The program STAT performs the statistical calculations to fit the data, and plots the resulting $\ln(K_d)$ vs. 1/T graph, as well as individual experimental and calculated fractions double strands and ethidium bound. An output file can also be written, showing the calculated values for the ethidium binding constant, and the calculated and experimental data.

The details of the model are taken care of in the subroutine PARTIT in the program STAT. It is very simple to change the model

by simply defining the statistical factors g_i (see Chapter IV). Several different models can be programmed into this subroutine, and the particular model is chosen by the value of MODEL requested in the main program (see example below). Presently, there are 9 models to choose from for the oligomers rCA_5G + rCU_5G. The procedure for determining the best fit between the model and the data is explained in Chapter IV.

Examples of running the programs MELT, DYEMLT and STAT are given below. Before running these programs, it is useful to assign the plotting terminal to be used for the plots. For example, if you want to use the Tektronix terminal hooked up on the port TTG2:, simply ALLOCATE TTG2: and ASSIGN TTG2: GTERM. Replace the TTG2: by the specific port for any other plotting device.

Following the examples of running the programs are the listings for the programs DYEMLT and STAT.

An example run of program Melt.

```
$ SET DEF     DISK$USERFILE1:[JWN.MELT]
$ RUN MELT
FILENAME (.DAT EXTENSION ASSUMED IF NOT GIVEN)  R55E22.DAT

RCA5G+RCU5G+ETHID R55E22 27NOV81 2MM CELLS  2M TMTB  5 1 1.5 ETH/STR JWN
4 CUVETTES.  2 WAVELENGTHS.  260 NM.  280NM

RIGHT FILE? Y

 111 DATA SETS   MIN TEMP =  -0.1 MAX TEMP =  69.6

CORRECT W.L. 1? Y

SUBTRACT CUV #1 FROM OTHERS? Y

BLANKS TO SUBTRACT FROM CELLS 2 ... 4
BLANK =  .016
BLANK =  .016
BLANK =  .016

REL ABS TO DIVIDE INTO CELLS 2 ... 4
REL ABS = 1.
REL ABS = 1.
REL ABS = 1.

CORRECT W.L. 2? Y

SUBTRACT CUV #1 FROM OTHERS? Y

BLANKS TO SUBTRACT FROM CELLS 2 ... 4
BLANK =  .033
BLANK =  .033
BLANK =  .033

REL ABS TO DIVIDE INTO CELLS 2 ... 4
REL ABS = 1.
REL ABS = 1.
REL ABS = 1.

INTERPOLATE TO REGULAR INTERVALS? Y
STARTING TEMP (DEG C) = 0
INCREMENT = 1
FILENAME (.DAT EXTENSION ASSUMED IF NOT GIVEN): R55E22.REG
ANOTHER DATA SET? N
GOODBYE THEN
$
```

An example run of program Dyemlt.

```
: P DYEMLT

EXTINCTION COEFFICIENTS READ FROM EXTINC DAT.

2ND AND 3RD.ORDER COEFF FOR S.S. (-2.35 AND -0.0237 FOR RCA5G+RCU5G)
-2.35
-.0237

INPUT ANOTHER DATA FILE? Y
FILENAME (.DAT EXTENSION ASSUMED IF NOT GIVEN): R55E22.REG

RCA5G+RCU5G+ETHID R55E22 27NOV81 2MM CELLS .2M TMTB .5 1 1.5 ETH/STR JWN
 2 SAMPLES.      69 DATA SETS. 260 = W.L. 1,  283 = W.L. 2.

WHICH SAMPLE? 1
IDENT HEADER FOR THIS SAMPLE:
RCA5G+RCU5G+ETHIDIUM 50UM STRANDS  5 ETH/STR
PATHLENGTH OF CELL (IN CM) = .2

CHANGE THE EXTINCTION COEFF VALUES? Y

                  260 NM                                    283 NM

EXT 50 DEG  SLOPE   EXT 0 DEG   SLOPE   EXT 50 DEG  SLOPE   EXT 0 DEG   SLOPE

[   SINGLE STRAND      DOUBLE STRAND   ][   SINGLE STRAND      DOUBLE STRAND   ]

 1.400E+05 1.900E+02 1.130E+05 2.400E+02 4.170E+04-1.040E+01 4.210E+04 0.000E+00
     1         2         3         4         5         6         7         8

[   ETHIDIUM FREE       ETHIDIUM BOUND ][   ETHIDIUM FREE       ETHIDIUM BOUND ]

 1.700E+04-2.550E+01 9.100E+03-1.500E+01 5.400E+04-7.830E+01 2.200E+04 5.000E+01
     9        10        11        12        13        14        15        16

CHANGE WHICH PARAMETER? 15
OLD VALUE FOR 15 =  2.20000E+04
NEW VALUE        = 1.85E4

                  260 NM                                    283 NM

EXT 50 DEG  SLOPE   EXT 0 DEG   SLOPE   EXT 50 DEG  SLOPE   EXT 0 DEG   SLOPE

[   SINGLE STRAND      DOUBLE STRAND   ][   SINGLE STRAND      DOUBLE STRAND   ]

 1.400E+05 1.900E+02 1.130E+05 2.400E+02 4.170E+04-1.040E+01 4.210E+04 0.000E+00
     1         2         3         4         5         6         7         8

[   ETHIDIUM FREE       ETHIDIUM BOUND ][   ETHIDIUM FREE       ETHIDIUM BOUND ]

 1.700E+04-2.550E+01 9.100E+03-1.500E+01 5.400E+04-7.830E+01 1.850E+04 5.000E+01
     9        10        11        12        13        14        15        16
```

```
CHANGE WHICH PARAMETER?
MAKE FILE OF EXTINCTION COEFF DATA? Y

CALCULATE F VALUES? Y
USE ABSORBANCE AT 50 DEG FOR CONC?
CORRECT THE ABSORBANCES FOR 1 CM PATHLENGTH
ABS (260 NM) 50 DEG = 7.44
ABS (283 NM) 50 DEG = 3.45

WRITE RESULTS TO SCREEN? Y

10 LINES PRINTED, THEN OUTPUT WAITS FOR A (CR)  TYPE "R" (CR) TO END PRINTING

RCA5G+RCU5G+ETHID R55E22 27NOV81 2MM CELLS  2M TMTB  5 1 1 5 ETH/STR JWN
RCA5G+RCU5G+ETHIDIUM 50UM STRANDS  5 ETH/STR
TOT STRAND CONC =     5.00810E-05 TOT ETHIDIUM CONC =     2.52152E-05

TEMP   FRACT DBL STR  FRACT ETH BOUND   FRACT OTH

   0       0.960005       0.999801       0.999889
   1       0.958865       0.999173       0.999416
   2       0.951857       0.998481       0.998514
   3       0.945977       0.998617       0.998625
   4       0.935790       0.996667       0.996655
   5       0.928068       0.994126       0.994100
   6       0.912983       0.995596       0.995559
   7       0.901246       0.995127       0.995085
   8       0.890271       0.995176       0.995113
   9       0.869580       0.993172       0.993134
R

WRITE AN OUTPUT FILE OF RESULTS? Y
FILENAME (.DAT EXTENSION ASSUMED IF NOT GIVEN): E221

INPUT ANOTHER DATA FILE? N

CHANGE THE EXTINCTION COEFF VALUES? N

CALCULATE F VALUES? N

WRITE RESULTS TO SCREEN?
?Y

$
```

An example run of program Stat.

```
$ RUN STAT
INPUT NEW MELTING CURVES? Y
FILENAME ( DAT EXTENSION ASSUMED IF NOT GIVEN). E221
RC45G+RCU5G+ETHIDIUM 50UM STRANDS  5 ETH/STR
FILENAME ( DAT EXTENSION ASSUMED IF NOT GIVEN):
THERMO VALUES.
NO   PARAM    DELTA H        DELTA S        K(25DEG)

 1   HELIX    0.0000E+00     0.0000E+00     1.0000E+00
 2   DYE      0.0000E+00     0.0000E+00     1.0000E+00
 3   SIGMA    0.0000E+00     0.0000E+00     1.0000E+00
 4   TAU      0.0000E+00     0.0000E+00     1.0000E+00

CHANGE WHICH PARAMETER (0=DONE) : 1
NEW VALUES (1/LINE) :
-4300
-123

CHANGE WHICH PARAMETER (0=DONE) : 3
NEW VALUES (1/LINE) :
0
5.951

CHANGE WHICH PARAMETER (0=DONE) : 3
NEW VALUES (1/LINE) :
0
5.95252

CHANGE WHICH PARAMETER (0=DONE) :
WHICH MODEL NUMBER?
 1 = ALL SITES EQUAL
 2 = 1 STRONG SITE (SIGMA)
 3 = TWO STRONG SITES ON END (TAU)
 4 = STRONG SITE ON 2ND B.P.
 5 = STRONG SITE ON 3RD B.P.
 6 = COOPERATIVE BINDING (TAU)
 7 = SIGMA. TAU STRONG SITES
 8 = STRONG SITE + COOPERATIVE
 9 = TWO STRONG SITES + COOP
2

HEADER FOR THIS RUN:
DEMONSTRATION OF PROGRAM STAT
DEMONSTRATION OF PROGRAM STAT

MODEL USED WAS.  2

     DELTA H       DELTA S     K (25 DEG)

  -4.3000E+04  -1.2300E+02   3.5153E+03  HELIX
   0.0000E+00   0.0000E+00   1.0000E+00  DYE GUESS
   0.0000E+00   5.9525E+00   2.0001E+01  SIGMA
   0.0000E+00   0.0000E+00   1.0000E+00  TAU
```

```
  -1.4422E+04  -2.6496E+01   6.0506E+04  DYE. FIT.   R = 0.99952
   1.3461E+02   4.3550E-01               ERROR

          REDUCED CHI SQUARED

    MELT      F(DYE)     F(HELIX)  POINTS FIT

 OVERALL   5.052E-06   1.488E-03

       1   5.052E-06   1.488E-03      13

NAME FOR DISK FILE (RETURN = NONE)
FILENAME ( DAT EXTENSION ASSUMED IF NOT GIVEN)

MAKE PLOTS? N
MAKE FIT FROM CONSTANT PARAMETERS? N
INPUT NEW MELTING CURVES?
THERMO VALUES.
NO. PARAM   DELTA H        DELTA S        K(25DEG)

 1  HELIX  -4.3000E+04   -1.2800E+02     3.5153E+03
 2  DYE     0.0000E+00    0.0000E+00     1.0000E+00
 3  SIGMA   0.0000E+00    5.9525E+00     2.0003E+01
 4  TAU     0.0000E+00    0.0000E+00     1.0000E+00

CHANGE WHICH PARAMETER (0=DONE) :
WHICH MODEL NUMBER?
 1 = ALL SITES EQUAL
 2 = 1 STRONG SITE (SIGMA)
 3 = TWO STRONG SITES ON END (TAU)
 4 = STRONG SITE ON 2ND B P
 5 = STRONG SITE ON 3RD B P
 6 = COOPERATIVE BINDING (TAU)
 7 = SIGMA. TAU STRONG SITES
 8 = STRONG SITE + COOPERATIVE
 9 = TWO STRONG SITES + COOP
1

HEADER FOR THIS RUN


MODEL USED WAS:  1

     DELTA H      DELTA S     K (25 DEG)

  -4.3000E+04  -1.2800E+02   3.5153E+03  HELIX
   0.0000E+00   0.0000E+00   1.0000E+00  DYE GUESS
   0.0000E+00   5.9525E+00   2.0003E+01  SIGMA
   0.0000E+00   0.0000E+00   1.0000E+00  TAU

  -1.4618E+04  -2.5109E+01   1.7437E+05  DYE. FIT.   R = 0.99922
   1.7462E+02   5.6497E-01               ERROR

          REDUCED CHI SQUARED

    MELT      F(DYE)     F(HELIX)  POINTS FIT

 OVERALL   9.076E-06   3.010E-03

       1   9.076E-06   3.010E-03      13

NAME FOR DISK FILE (RETURN = NONE)
FILENAME ( DAT EXTENSION ASSUMED IF NOT GIVEN):

MAKE PLOTS?
```

Program Dyemlt. Program to convert melting curves of ethidium bromide + oligonucleotides into fractions double strands and ethidium bound.

```
      PROGRAM DYEMLT
C
C     ***********************************************************************
C     * THIS PROGRAM CALCULATES THE FRACTION OF ETHIDIUM BOUND AND THE
C     * FRACTION OF STRANDS IN DOUBLE HELICES FROM THE MELTING CURVES.
C     * THE DATA IS ON DISK FILES WRITTEN BY THE PROGRAM MELT, WHICH
C     * CORRECTS THE DATA FOR BLANKS.
C     *
C     * THE FRACTION OF ETHIDIUM BOUND AND DOUBLE STRANDS CAN BE
C     * DETERMINED BY KNOWING FOUR EXTINCTION COEFFICIENTS FOR
C     * TWO WAVELENGTHS. (260 AND 283 NM, USUALLY)
C     *
C     *       A(260) = CSS*ESS + CDS*EDS + CETH(BOUND)*EETH(BOUND) +
C     *                CETH(FREE)*EETH(FREE)
C     *       AT THE SPECIFIC TEMPERATURE.
C     *
C     *       CTOT(STRANDS) = CSS + CDS KNOWN FROM HIGH TEMP ABS
C     *       CTOT(ETH) = CETH(BOUND) + CETH(FREE)  "      "    "
C     *
C     * THE EXTINCTION COEFFICIENT DATA REQUIRED IS: (FOR BOTH W.L. S)
C     *
C     *       1.   E(SS,50 DEG C)
C     *       2.   SLOPE OF E(SS) VS  TEMP
C     *       3.   E(DS,0 DEG C)
C     *       4.   SLOPE OF E(DS) VS  TEMP
C     *       5.   E(ETH,FREE,50 DEG C)
C     *       6.   SLOPE OF E(ETH,FREE) VS  TEMP
C     *       7.   E(ETH,BOUND,0 DEG C)
C     *       8.   SLOPE OF E(ETH,BOUND) VS.  TEMP
C     *
C     * NOTE... THE EXTINCTION COEFF FOR FREE ETHIDIUM IS TREATED AS
C     *    A THIRD ORDER EXPRESSION WITH THE T**3 COEFF CONC DEPENDENT.
C     *    SEE SUBROUTINE CALCF FOR DETAILS.
C     * ALSO... THE EXTINCTION COEFF FOR THE SINGLE STRANDS OF
C     *    RCA5G + RCU5G IS FIT BY A THIRD ORDER EXPRESSION,
C     *    THE COEFFICIENTS OF WHICH IS INPUT AT THE BEGINNING
C     *    OF THE PROGRAM.  (SET TO ZERO FOR DCA5G+DCT5G.)
C     *
C     * THE DATA ARE READ FROM A FILE OF ABSORBANCES AT REGULAR
C     * TEMPERATURE INTERVALS OF 1 DEG C.
C     *
C     * THE EXTINCTION COEFFICIENT PARAMETERS MAY BE CHANGED SINGLY.
C     * OUTPUT IS TO THE SCREEN, TO ALLOW VARIATIONS OF THE EXT.
C     * COEFF'S WHICH BEST MANIPULATE THE DAT.
C     *
C     * OUTPUT MAY BE SET UP AS INPUT FOR THE STATISTICAL MODELING
C     * PROGRAM.
C     *
C     ***********************************************************************
C     * SUBROUTINES USED:
C     *
C     *    SUBROUTINE INPUT
```

```
C    *        INPUTS THE DATA GENERATED BY PROGRAM MELT.
C    *        DATA MUST BE AT REGULAR INTERVALS.  SEE SUBROUTINE FOR
C    *        FILE STRUCTURE.
C    *
C    *     SUBROUTINE EXCOEF
C    *        CHANGES ANY OF THE EXTINCTION COEFF PARAMETERS.
C    *
C    *     SUBROUTINE CALCF
C    *        CALCULATES FRACTION OF DOUBLE STRANDS, FRACTION OF DYE
C    *        BOUND, UPPER AND LOWER BASELINES FOR BOTH WAVELENGTHS
C    *
C    *     SUBROUTINE SCREEN
C    *        WRITES THE RESULTING FRACTION D.S. AND DYE BOUND TO SCREEN.
C    *
C    *     SUBROUTINE WTFILE
C    *        WRITES A FILE CONTAINING THE TEMP, FRACT D.S. AND DYE BOUND,
C    *        ABS AT EACH W.L., LOWER AND UPPER BASELINES FOR EACH W.L.
C    *
C    *     SUBROUTINE FNAME(FILENA)
C    *        GETS NAME FOR DISK FILE FOR FILE READS AND WRITES.
C    *
C    ***********************************************************************
C    * THIS PROGRAM WAS WRITTEN BY JEFF NELSON, MAY, 1981.
C    * REVISED FOR CURVATURE IN ETHID EXTINCTION COEFFICIENTS JUNE, 1982
C    ***********************************************************************
C
      COMMON /EXTINC/ E260SS,S260SS,E260DS,S260DS,E283SS,S283SS,
     2                E283DS,S283DS,E260DF,S260DF,E260DB,S260DB,
     3                E283DF,S283DF,E283DB,S283DB,A260SS,B260SS
      COMMON /DATA/ ABS(0:80,2),FRAC(0:80,2),BLINE(0:80,4),
     2              NTFST,NTLAST,HEAD1,HEAD2,CT,CD
      COMMON /OTHER/ FOTHER(0:80)
      BYTE HEAD1(80),HEAD2(80)
C
C    * READ VALUES FOR THE EXTINCTION DATA IF THE FILE EXTINC.DAT
C    * EXISTS.
C
      OPEN(UNIT=10,FILE='EXTINC.DAT',STATUS='OLD',ERR=50)
         READ(10,400)E260SS,S260SS,E260DS,S260DS
         READ(10,400)E283SS,S283SS,E283DS,S283DS
         READ(10,400)E260DF,S260DF,E260DB,S260DB
         READ(10,400)E283DF,S283DF,E283DB,S283DB
         WRITE(6,410)
      CLOSE(UNIT=10)
  400 FORMAT(4E15.6)
  410 FORMAT(/,' EXTINCTION COEFFICIENTS READ FROM EXTINC.DAT.')
C
C    * GET SECOND ORDER COEFF FOR RCA5G+RCU5G SINGLE STRANDS AT 260NM
C
   50 WRITE(6,500)
      READ(5,510)A260SS,B260SS
  500 FORMAT(/,' 2ND AND 3RD ORDER COEFF FOR S.S.',
     2          ' (-2.05 AND -0.0237 FOR RCA5G+RCU5G)')
  510 FORMAT(F15.0)
C
C    * ASK FOR INSTRUCTIONS TO INPUT DATA FILES, CHANGE EXTINCTION
C    * COEFFICIENTS, AND WRITING RESULTS TO SCREEN AND OUTPUT FILES.
C
  100 WRITE(6,600)
  600 FORMAT(/,' INPUT ANOTHER DATA FILE? ',$)
      READ(5,620)INST
  620 FORMAT(A1)
      IF (INST .EQ. 'Y') CALL INPUT
      WRITE(6,640)
  640 FORMAT(/,' CHANGE THE EXTINCTION COEFF VALUES? ',$)
      READ(5,620)INST
```

```
      IF (INST .EQ. 'Y') CALL EXCOEF
      WRITE(6,680)
  680 FORMAT(/,' CALCULATE F VALUES? ',$)
      READ(5,620)INST
      IF (INST .EQ. 'Y') CALL CALCF
      WRITE(6,700)
  700 FORMAT(/,' WRITE RESULTS TO SCREEN? ',$)
      READ(5,620)INST
      IF (INST .EQ. 'Y') CALL SCREEN
      WRITE(6,720)
  720 FORMAT(/,' WRITE AN OUTPUT FILE OF RESULTS? ',$)
      READ(5,620)INST
      IF (INST .EQ. 'Y') CALL WTFILE
      GOTO 100
      END



      SUBROUTINE INPUT
C
C     ******************************************************************
C     * THIS SUBROUTINE GETS THE NAME OF AN INPUT FILE AND PUTS THE DATA
C     * INTO THE ARRAY ABS(0:80,2)   THE NAME OF THE INPUT FILE AND THE
C     * SAMPLE WITHIN IT ARE OBTAINED FROM THE TELETYPE.
C     *
C     * INPUT FILE SPECIFICATION:
C     *
C     *       LINE 1.   'MELT NORMALIZED REGULAR' - IDENTIFIES PROCESS
C     *                 PRODUCING FILE.
C     *       LINE 2.   HEADER - 80 CHARACTERS (80A1)
C     *       LINE 3.   # OF SAMPLES, # DATA SETS, #WL, WL1, WL2
C     *                 (I1,I6,I1,2I6)
C     *       LINE 4.   BLANKS - (1,1), (1,2), (1,3), (2,1), (2,2), (2,3)
C     *                 (6F13.5)
C     *       LINE 5.   RELATIVE ABSORBANCES - (1,1), (1,2),
C     *                 (1P6E13.5)
C     *       LINE 6.   T, ABS FOR W.L. 1 AND 2, FIRST DATA SET:
C     *                 T(1,1),A(1,1),T(1,2),A(1,2),    T(2,3),A(2,3)
C     *                 6(F7.2,F8.5)
C     *
C     *                 IN LINES 4 - 6, IF THERE ARE ONLY 2 SAMPLES,
C     *                 THE VALUES CORRESPONDING TO (1,3) AND (2,3)
C     *                 ARE NOT WRITTEN.   SIMILARLY IF THERE IS ONLY
C     *                 ONE SAMPLE, THE VALUES FOR (1,2), (1,3), (2,2),
C     *                 AND (2,3) ARE NOT WRITTEN.
C     *                 (M,N) CORRESPONDS TO THE MTH W.L., NTH SAMPLE.
C     *
C     * THE DATA ARE READ INTO THE ARRAY ABS(0:80,2), WHERE THE
C     * FIRST SUBSCRIPT IS THE TEMP, THE SECOND THE W.L.
C     * (1 = 260 NM, 2 = 283 NM )
C     *
C     ******************************************************************
C
      COMMON /DATA/ ABS(0:80,2),FRAC(0:80,2),BLINE(0:80,4),
     2              NTFST,NTLAST,HEAD1,HEAD2,CT,CD
      BYTE HEAD1(80),HEAD2(80)
      CHARACTER*30 FILENA
      CHARACTER*23 CHECK
      DIMENSION A(6),T(6)
C
C     * GET THE NAME OF THE INPUT FILE.
C
  100 CALL FNAME(FILENA)
      IF (FILENA EQ. ' ') GOTO 300
      OPEN(UNIT=10, FILE=FILENA, STATUS='OLD', ERR=120,
```

```
     2      CARRIAGECONTROL='LIST')
      GOTO 150
C
C   * ERROR OPENING FILE
C
  120 WRITE(6,550)FILENA
  550 FORMAT(' *** FILE NOT FOUND: ',A,/,
     2        ' *** TRY AGAIN ')
      GOTO 100
C
C   * READ IN THE PARTICULARS, ECHO FOR CONFIRMATION
C   * FIRST LINE IN FILE MUST BE 'MELT NORMALIZED REGULAR'.
C   * OTHERWISE FILE IS NOT RIGHT TYPE.
C
  150 READ(10,590)CHECK
  590 FORMAT(A)
      IF (CHECK.NE.'MELT NORMALIZED REGULAR') THEN
         WRITE(6,595)FILENA,CHECK
  595    FORMAT(' ***** FILE NOT CORRECT TYPE *****',/,1X,A,/,1X,A)
         CLOSE(UNIT=10)
         GOTO 100
      END IF
      READ(10,600)HEAD1
  600 FORMAT(80A1)
      READ(10,620)NCUV,NDSETS,NWL,IWL1,IWL2
  620 FORMAT(I1,I6,I1,I6,I6)
      IF ( (IWL1 .NE. 260) .OR. (IWL2 .NE. 283) ) THEN
         WRITE(6,640)IWL1,IWL2
  640    FORMAT(/,' ***** WAVELENGTHS ARE NOT RIGHT - ',
     2           I4,2X,I4,' SHOULD BE 260, 283 ***** ')
         CLOSE (UNIT=10)
         GOTO 100
      ENDIF
C
C   * ECHO HEADER, # DATA SETS, W.L.'S
C
      WRITE(6,660)HEAD1,NCUV,NDSETS,IWL1,IWL2
  660 FORMAT(/,1X,80A1,/,1X,I2,' SAMPLES, ',I6,' DATA SETS,',
     2        I4,' = W.L. 1, ',I4,' = W.L. 2.')
C
C   * GET THE SAMPLE NUMBER.
C
  200 WRITE(6,680)
  680 FORMAT(/,' WHICH SAMPLE? ',$)
      READ(5,700)NSAMP
  700 FORMAT(I4)
      IF (NSAMP .EQ. 0) GOTO 300
      IF (NSAMP .GT. NCUV) GOTO 200
C
C   * GET HEADER TO IDENTIFY THIS PARTICULAR SAMPLE
C
      WRITE(6,701)
  701 FORMAT(' IDENT HEADER FOR THIS SAMPLE:')
      READ(5,702)HEAD2
  702 FORMAT(80A1)
C
C   * GET THE PATHLENGTH TO CORRECT THE ABSORBANCES.
C
      WRITE(6,705)
  705 FORMAT(' PATHLENGTH OF CELL (IN CM) = ',$)
      READ(5,710)PATH
      IF (PATH .EQ. 0.) PATH = 1.
  710 FORMAT(F10.0)
      PATH = 1./PATH
C
C   * READ IN DATA.
```

```
C    * IGNORE NEXT TWO RECORDS - THEY'RE BLANKS AND REL ABS.
C
      READ(10,720)DUM1,DUM2
  720 FORMAT(A1,/,A1)
C
C    * WE WANT TO READ COLUMNS NSAMP AND NSAMP + NCUV
C    * READ IN FIRST RECORD TO GET FIRST TEMP.
C    * CORRECT ABSORBANCES FOR PATHLENGTH
C
      NSAPCU = NSAMP + NCUV
      READ(10,760)( T(I1),A(I1),I1=1,2*NCUV )
  760 FORMAT( 6(F7.2,F8.5) )
      NTFST = INT(T(1) + 0.001)
      NTLAST = NTFST + NDSETS - 1
      ABS(NTFST,1) = A(NSAMP)*PATH
      ABS(NTFST,2) = A(NSAPCU)*PATH
C
C    * READ IN NEXT RECORDS
C
      DO N = NTFST+1, NTLAST
         READ(10,780)( A(I1),I1=1,2*NCUV )
  780    FORMAT( 6(7X,F8.5) )
         ABS(N,1) = A(NSAMP)*PATH
         ABS(N,2) = A(NSAPCU)*PATH
      END DO
C
C    * DONE READING IN FILE.  CLOSE INPUT FILE AND RETURN.
C
  300 CLOSE(UNIT=10)
      RETURN
      END



      SUBROUTINE EXCOEF
C
C    ***********************************************************************
C    * THIS SUBROUTINE PRINTS OUT THE CURRENT VALUES OF THE
C    * EXTINCTION COEFFICIENT PARAMETERS, AND ASKS IF YOU WISH
C    * TO CHANGE ANY.
C    * EACH VALUE IS NUMBERED 1 TO 16
C    ***********************************************************************
C
      COMMON /EXTINC/ E260SS,S260SS,E260DS,S260DS,E283SS,S283SS,
     2                E283DS,S283DS,E260DF,S260DF,E260DB,S260DB,
     3                E283DF,S283DF,E283DB,S283DB,A260SS,B260SS
      DIMENSION EXT(16)
      EQUIVALENCE (EXT(1),E260SS)
C
C    * WRITE OUT CURRENT VALUES WITH CODE NUMBER
C
  100 WRITE(6,600)
  600 FORMAT(/,18X,'260 NM',34X,'283 NM',/,/,
     2 1X,2( EXT 50 DEG   SLOPE    EXT 0 DEG   SLOPE    ),/,/,
     3 1X,2( [    SINGLE STRAND       DOUBLE STRAND    ]),/)
      WRITE(6,620)(EXT(I1),I1=1,8),(I1,I1=1,8)
  620 FORMAT(1X,1PSE10.3,/,1X,8(4X,I2,4X))
      WRITE(6,630)
  630 FORMAT(/,1X,2([    ETHIDIUM FREE        ETHIDIUM BOUND  ]),/)
      WRITE(6,620)(EXT(I1),I1=9,16),(I1,I1=9,16)
C
C
  200 WRITE(6,700)
  700 FORMAT(/,' CHANGE WHICH PARAMETER? ',$)
      READ(5,710,ERR=200)INST
```

```
  710 FORMAT(I3)
      IF ( (INST .GE. 1) .AND. (INST .LE. 16) ) THEN
         WRITE (6,730)INST,EXT(INST)
  730    FORMAT(' OLD VALUE FOR ',I2,' = ',1PE12.5,/,
     2          ' NEW VALUE        = ',$)
         READ(5,740,ERR=200)EXT(INST)
  740    FORMAT(E15.0)
         GOTO 100
      ELSE IF ( (INST .GE. 17) .OR. (INST .LE. -1) ) THEN
         WRITE(6,760)INST
  760    FORMAT(' *** INST = ',I3,' MUST BE BET 0 AND 16 ')
         GOTO 100
      END IF
C
C   * IF INST = 0, WE ARE DONE CHANGING VALUES
C
      WRITE(6,800)
  800 FORMAT(' MAKE FILE OF EXTINCTION COEFF DATA? ',$)
      READ(5,810)INST
  810 FORMAT(A1)
      IF (INST .EQ. 'Y') THEN
         OPEN(UNIT=10,FILE='EXTINC.DAT',STATUS='NEW',CARRIAGECONTROL=
     2        'LIST',ERR=400)
         WRITE(10,820)E260SS,S260SS,E260DS,S260DS
         WRITE(10,820)E283SS,S283SS,E283DS,S283DS
         WRITE(10,820)E260DF,S260DF,E260DB,S260DB
         WRITE(10,820)E283DF,S283DF,E283DB,S283DB
  820    FORMAT(4(1PE15.6))
         CLOSE(UNIT=10)
      END IF
      RETURN
  400 WRITE(6,850)
  850 FORMAT(' ***** ERROR OPENING FILE - NOT WRITTEN ***** ')
      RETURN
      END



      SUBROUTINE CALCF
C
C   ***********************************************************************
C   * THIS SUBROUTINE CALCULATES THE FRACTION OF THE STRANDS IN
C   * DOUBLE HELICES AND THE FRACTION OF ETHIDIUM IONS BOUND
C   * THE SET OF TWO EQUATIONS IN TWO UNKNOWNS ARE:
C   *
C   *       A11*FH + A12*FD = A13
C   *       A21*FH + A22*FD = A23
C   *
C   * WHERE FH = FRACTION OF STRANDS IN DBL HEL,
C   * FD = FRACTION OF ETHIDIUM IONS BOUND, AND
C   * A11 = CT( E260DS(T) - E260SS(T) ) ; CT = TOT CONC OF STRANDS
C   * A12 = CD( E260DB(T) - E260DF(T) ) , CD = TOT CONC OF ETHIDIUM
C   * A13 = A260(T) - CT*E260SS(T) - CD*E260DF(T) ; A260 = EXPT ABS.
C   *
C   * A21 = CT( E283DS(T) - E283SS(T) )
C   * A22 = CD( E283DB(T) - E283DF(T) )
C   * A23 = A283(T) - CT*E283SS(T) - CD*E283DF(T)
C   *
C   * ENNNXX(T) IS THE EXTINCTION COEFF DETERMINED BY THE EXT COEFF
C   * AT 0 OR 50 DEG CORRECTED BY THE SLOPE. E.G.
C   *    E260DS(T) = E260DS(0) + S260DS(T-0)
C   *
C   * ETHIDIUM AT 283 NM IS AN EXCEPTION SINCE ITS MELTING CURVE
C   *    HAS CURVATURE WHICH DEPENDS ON CONCENTRATION. HERE,
C   *    E283DF(T) = E283DF(90) + S283DF(T-90) + A3*(T-90)**3
```

```
C    *     WHERE A3 = 4.86E-6*SQRT(CD) + 8.64E-9
C    *     E283DF(90) = 5.1E4.
C    *
C    * THE TOTAL CONCENTRATIONS ARE CALCULATED FROM THE ABSORBANCES
C    * AT 50 DEG.   (ONLY SINGLE STRANDS AND FREE ETHIDIUM PRESENT)
C    *
C    *       A260 = CT*E260SS + CD*E260DF
C    *       A283 = CT*E283SS + CD*E283DF
C    *       DET = (E260SS*E283DF) - (E283SS*E260DF)
C    *       CT = (A260*E283DF) - (A283*E260DF) / DET
C    *       CD = (E260SS*A283) - (E283SS*A260) / DET
C    *
C    * THE SOLUTION IS:
C    *       DET = (A11*A22) - (A21*A12)
C    *       FH  = (A13*A22) - (A23*A12) / DET
C    *       FD  = (A11*A23) - (A21*A13) / DET
C    *
C    * THE RESULTS ARE STORED IN ARRAY F(NTEMP,1) (FH) AND
C    * FRAC(NTEMP,2) (FD).
C    *
C    ***********************************************************************
C
      COMMON /EXTINC/ E260SS,S260SS,E260DS,S260DS,E283SS,S283SS,
     2                E283DS,S283DS,E260DF,S260DF,E260DB,S260DB,
     3                E283DF,S283DF,E283DB,S283DB,A260SS,B260SS
      COMMON /DATA/ ABS(0:80,2),FRAC(0:80,2),BLINE(0:80,4),
     2              NTFST,NTLAST,HEAD1,HEAD2,CT,CD
      COMMON /OTHER/ FOTHER(0:80)
      BYTE HEAD1(80),HEAD2(80)
C
C    * DETERMINE THE TOTAL CONCENTRATION OF ETHIDIUM AND STRANDS
C    * FROM THE ABSORBANCES AT 50 DEG C AND THE EXT COEFFS.
C
      IF (50 .GT. NTLAST) THEN
   50    WRITE(6,400)NTLAST
  400    FORMAT(/ ***** TEMP DOES NOT REACH 50 DEG - MAX = ',I4,/,
     2          ' INPUT ABSORBANCES AT 50 DEG MANUALLY.',/,
     3          ' CORRECTED TO A PATHLENGTH OF 1 CM ',/,
     4          ' ABS (260 NM, 50 DEG) = ',$)
         READ(5,420,ERR=50)A260
         WRITE(6,440)
         READ(5,420,ERR=50)A283
  420    FORMAT(F10.0)
  440    FORMAT(' ABS (283 NM, 50 DEG) = ',$)
      ELSE
  100    WRITE(6,460)
  460    FORMAT(' USE ABSORBANCE AT 50 DEG FOR CONC? ',$)
         READ(5,480)INST
  480    FORMAT(A1)
         IF (INST .NE. 'Y') THEN
            WRITE(6,500)
            READ(5,510,ERR=100)A260
            WRITE(6,520)
            READ(5,510,ERR=100)A283
  500       FORMAT(' CORRECT THE ABSORBANCES FOR 1 CM PATHLENGTH.',/,
     2          ' ABS (260 NM) 50 DEG = ',$)
  510       FORMAT(F10.0)
  520       FORMAT(' ABS (283 NM) 50 DEG = ',$)
         ELSE
            A260 = ABS(50,1)
            A283 = ABS(50,2)
         ENDIF
      ENDIF
      DET = (E260SS*E283DF) - (E283SS*E260DF)
      CT = ( (A260*E283DF) - (A283*E260DF) ) / DET
      CD = ( (E260SS*A283) - (E283SS*A260) ) / DET
```

```
      IF (CD .LT. 1.E-8) CD = 0.
C
C   * CALC FH AND FD FOR EACH TEMP...
C
      DO NT = NTFST,NTLAST
C
C   *    DETERMINE EXT COEFF AT TEMP T
C
           T = FLOAT(NT)
           TZ = T - 50.
           E1SS = E260SS + S260SS * TZ + (B260SS*TZ + A260SS)*TZ*TZ
           E1DS = E260DS + S260DS * T
           E1DF = E260DF + S260DF * TZ
           E1DB = E260DB + S260DB * T
           E2SS = E283SS + S283SS * TZ
           E2DS = E283DS + S283DS * T
C           E2DF = E283DF + S283DF * TZ
           E2DB = E283DB + S283DB * T
C
C   *    ETHIDIUM FREE EXTINCTION COEFF AT 283 NM IS GIVEN BY:
C   *    E283DF(T) = E283DF(90) + S283DF*(T-90) + A3*(T-90)**3
C   *    WHERE A3 = (4.86E-6*SQRT(CD) + 8.64E-9) * E283DF(90).
C
           A3 = (4.86E-6*SQRT(CD) + 8.64E-9) * 5.1E4
           E2DF = 5.1E4 + S283DF*(T-90.) + A3*(T-90.)**3
C
C   *    IF NO ETHIDIUM IN THIS SAMPLE (DNA ONLY):
C
           IF (CD .EQ. 0.) THEN
              CT = A260 / E260SS
              FRAC(NT,2) = 0.
              FOTHER(NT) = 0.
              FH = ( ABS(NT,1) - CT*E1SS ) / ( CT*(E1DS - E1SS) )
              FRAC(NT,1) = FH
              GOTO 200
           END IF
C
C   *    CALC COEFFICIENTS FOR SIMULTANEOUS EQUATIONS
C
           A11 = CT * ( E1DS - E1SS )
           A12 = CD * ( E1DB - E1DF )
           A13 = ABS(NT,1) - CT*E1SS - CD*E1DF
           A21 = CT * ( E2DS - E2SS )
           A22 = CD * ( E2DB - E2DF )
           A23 = ABS(NT,2) - CT*E2SS - CD*E2DF
C
           DET = (A11 * A22) - (A21 * A12)
           FH  = ( (A13 * A22) - (A23 * A12) ) / DET
           FD  = ( (A11 * A23) - (A21 * A23) ) / DET
           IF (FH .GT. 99.) FH = 99.
           IF (FH .LT. -99.) FH = -99.
           IF (FD .GT. 99.) FD = 99.
           IF (FD .LT. -99.) FD = -99.
           FRAC(NT,1) = FH
           FRAC(NT,2) = FD
C
C   *    ALSO CALC F FROM MELTING CURVE AT 283 NM ONLY.
C
           FOT = ( ABS(NT,2) - CT*E2SS - CD*E2DF ) /
     2           ( CT*(E2DS - E2SS) + CD*(E2DB - E2DF) )
           FOTHER(NT) = FOT
C
C   *    SAVE UPPER AND LOWER BASELINES FOR BOTH WAVELENGTHS...
C
  200      BLINE(NT,1) = CT*E1DS + CD*E1DB
           BLINE(NT,2) = CT*E1SS + CD*E1DF
```

```
         BLINE(NT,3) = CT*E2DS + CD*E2DB
         BLINE(NT,4) = CT*E2SS + CD*E2DF
C
      END DO
C
      RETURN
      END



      SUBROUTINE SCREEN
C
C     ******************************************************************
C     * THIS SUBROUTINE WRITES THE VALUES FOR FRACTION OF STRANDS IN
C     * DOUBLE STRANDS, AND THE FRACTION OF ETHIDIUM BOUND.
C     * 10 LINES ARE PRINTED. THEN THE PROGRAM WAITS FOR A CARRIAGE
C     * RETURN TO BE TYPED BEFORE CONTINUING.
C     ******************************************************************
C
      COMMON /DATA/ ABS(0:80,2),FRAC(0:80,2),BLINE(0:80,4),
     2              NTFST,NTLAST,HEAD1,HEAD2,CT,CD
      COMMON /OTHER/ FOTHER(0:80)
      BYTE HEAD1(80),HEAD2(80)
      WRITE(6,500)HEAD1,HEAD2,CT,CD
  500 FORMAT(/,' 10 LINES PRINTED, THEN OUTPUT WAITS FOR A <CR> ',
     2          ' TYPE "R" <CR> TO END PRINTING.',//,
     2          1X,80A1,/,1X,80A1,/,' TOT STRAND CONC = ',1PE15.5,
     3          ' TOT ETHIDIUM CONC = ',E15.5,//,
     4          ' TEMP   FRACT DBL STR  FRACT ETH BOUND   FRACT OTH',/)
      NLINE = 0
      DO N = NTFST, NTLAST
  100    NLINE = NLINE + 1
         IF (NLINE .GT. 10) THEN
            READ(5,550)INST
  550       FORMAT(A1)
            IF (INST.EQ.'R') GOTO 200
            NLINE = 0
            GOTO 100
         ENDIF
         WRITE(6,600)N,FRAC(N,1),FRAC(N,2),FOTHER(N)
  600    FORMAT(1X,I4,3F15.6)
      ENDDO
  200 RETURN
      END



      SUBROUTINE WTFILE
C
C     ******************************************************************
C     * THIS SUBROUTINE WRITES A FILE CONTAINING THE FRACTION OF
C     * DOUBLE STRANDS AND DYE BOUND, AS WELL AS THE ABSORBANCE
C     * DATA AND LOWER AND UPPER BASELINES FOR BOTH WAVELENGTHS.
C     *
C     * FILE SPECIFICATIONS:
C     *
C     *    LINE 1 - 'DYE MELTING CURVE' TO IDENTIFY THE FILE AS COMING
C     *             FROM THIS PROGRAM.
C     *    LINE 2 - ID HEADER FOR INPUT FILE
C     *    LINE 3 - ID HEADER FOR THIS SAMPLE.
C     *    LINE 4 - NTFST, NTLAST, 2I5,' FIRST AND LAST TEMP'
C     *    LINE 5 - TOTAL STRAND CONC. 1PE15.5,' TOTAL STRAND CONC'
C     *    LINE 6 - TOTAL DYE CONC. 1PE15.5,' TOTAL DYE CONC'
C     *    LINE 7 - COLUMN HEADERS
```

```
C    *    LINE 3 - TEMP, FH, FD, ABS(260), ABS(283),
C    *             LOW BL, HIGH BL (260),  LOW BL, HIGH BL (283)
C    *             I5,8F9.5
C    *
C    *
C    * THE FILE NAME IS GOTTEN FROM THE TERMINAL.
C    ***********************************************************************
C
      COMMON /EXTINC/ E260SS,S260SS,E260DS,S260DS,E283SS,S283SS,
     2                E283DS,S283DS,E260DF,S260DF,E260DB,S260DB,
     3                E283DF,S283DF,E283DB,S283DB,A260SS,B260SS
      COMMON /DATA/ ABS(0:80,2),FRAC(0:80,2),BLINE(0:80,4),
     2              NTFST,NTLAST,HEAD1,HEAD2,CT,CD
      COMMON /OTHER/ FOTHER(0:80)
      BYTE HEAD1(80),HEAD2(80)
      CHARACTER*10 FILENA
C
C    * GET NAME FOR FILE.
C
  100 CALL FNAME(FILENA)
      IF (FILENA .EQ. ' ') GOTO 300
      OPEN(UNIT=10, STATUS='NEW', FILE=FILENA, CARRIAGECONTROL='LIST',
     2     ERR=150)
      GOTO 200
  150 WRITE(6,500) FILENA
  500 FORMAT(' ***** ERROR OPENING FILE ***** ',A,/,' TRY AGAIN.')
      GOTO 100
C
C    * WRITE THE HEADERS.
C
  200 WRITE(10,600)
  600 FORMAT('DYE MELTING CURVE')
      WRITE(10,610)HEAD1,HEAD2
  610 FORMAT(80A1)
      WRITE(10,620)NTFST,NTLAST
  620 FORMAT(2I5,' FIRST AND LAST TEMP')
      WRITE(10,640)CT,CD
  640 FORMAT(1PE15.6,' TOTAL STRAND CONC',/,E15.6,' TOTAL DYE CONC')
      WRITE(10,660)
  660 FORMAT(' TEMP  F HELIX    F DYE  ABS 260  ABS 283  LOW 260',
     2          HI 260  LOW 283   HI 283')
      DO I = NTFST, NTLAST
         WRITE(10,680)I,FRAC(I,1),FRAC(I,2),ABS(I,1),ABS(I,2),
     2                BLINE(I,1),BLINE(I,2),BLINE(I,3),BLINE(I,4)
  680 FORMAT(I5,8F9.4)
      END DO
      CLOSE(UNIT=10)
  300 RETURN
      END
```

```
      SUBROUTINE FNAME(FILE)
C
C    ******************************************************************
C    * THIS SUBROUTINE GETS THE NAME FOR A FILE, ADDING
C    * THE EXTENSION .DAT IF NONE IS GIVEN. THE FILE
C    * NAME IS RETURNED IN VARIABLE FILE (CHARACTER VAR).
C    * IF NO NAME IS GIVEN (ALL BLANKS), FILE RETURNS ' '.
C    ******************************************************************
C
         CHARACTER*30 FILE
         WRITE(6,1840)
 1840    FORMAT(' FILENAME (.DAT EXTENSION ASSUMED IF NOT GIVEN): ',$)
         READ(5,1850,ERR=900)FILE
 1850    FORMAT(A)
C
C    *   FIND LAST NON-BLANK CHARACTER, AND ADD .DAT EXT. IF NOT INCL.
C
         IN = LEN(FILE)
         DO WHILE (FILE(IN:IN) .EQ. ' ')
            IN = IN - 1
            IF (IN .EQ. 0) GOTO 200
         END DO
         LASTCH = INDEX(FILE,'.')
         IF (LASTCH .EQ. 0) FILE = FILE(1:IN)//'.DAT'
      RETURN
  200    FILE = ' '
      RETURN
  900    WRITE(6,6000)
 6000    FORMAT(/,' ********************************',/,
     2           ' ********************************',/,
     3           ' ***** ERROR IN INPUT VALUE *****',/,
     4           ' ********************************',/,
     5           ' ********************************')
      RETURN
      END
```

Program Stat. Program to calculate experimental ethidium melting curves using a statistical model for ethidium binding to oligonucleotides. The input files are written by program Dyemlt.

```
      PROGRAM STAT
C
C     ***********************************************************************
C     * THIS PROGRAM CALCULATES THE BINDING PARAMETERS OF A DYE (ETHIDIUM)
C     * INTO OLIGONUCLEOTIDES, USING A STATISTICAL MODEL OF DYE BINDING.
C     *
C     * THE MODEL IS DEFINED IN SUBROUTINE PARTIT, WHEREIN DETAILS ON
C     * THE MODEL MAY BE FOUND.
C     *
C     * NOTES:
C     *    ONLY NON-SELFCOMPLEMENTARY OLIGOMERS MAY BE FIT BY THIS
C     *       PROGRAM.  CHANGES IN SUBROUTINE PARTIT MUST BE MADE
C     *       TO RUN SELF-COMPLEMENTARY OLIGOMERS.
C     *    DATA MUST BE IN REGULAR INTERVALS AT INTEGRAL DEG. C, BETWEEN
C     *       0 AND 80 DEG C, INCLUSIVE
C     *    UP TO 12 MELTING CURVES MAY BE ANALYZED SIMULTANEOUSLY.
C     *    THERMODYNAMIC PARAMETERS FOR DOUBLE STRAND FORMATION, AND
C     *       MODEL PARAMETERS SIGMA AND TAU, ARE INPUT AND HELD CONSTANT.
C     *       THE PARAMETERS OF DELTA H AND DELTA S FOR DYE BINDING ARE
C     *       THE BEST FIT FOR THESE INPUT PARAMETERS AND THE MODEL
C     *
C     * DESCRIPTION OF PARAMETERS:
C     *
C     * /BLKI/
C     *    NMELTS     - NUMBER OF MELTING CURVES INPUT (.LE. 12).
C     *    NTFST(I)   - LOWEST TEMP OF MELTING CURVE I.
C     *    NTLAST(I)  - HIGHEST TEMP. OF I.
C     *    NTFIT1(I)  - FIRST POINT USED IN FITTING FOR BINDING CONSTANTS
C     *    NTFIT2(I)  - LAST POINT ...
C     *    CSTRAN(I)  - CONC OF NUCLEIC ACID STRANDS.
C     *    CDYE(I)    - CONC OF DYE (ETHIDIUM).
C     *    MODEL      - NUMBER DESIGNATING WHICH MODEL TO USE IN
C     *                 SUBROUTINE PARTIT.  (SEE SUBROUTINE FOR DETAILS).
C     *
C     * /BLKII/
C     *    FHEXP(I,J)  - EXPERIMENTAL FRAC. DOUBLE STRANDS FOR MELTING
C     *                  CURVE I, TEMPERATURE J (DEG C).  J BET. 0 AND 80
C     *    FDEXP(I,J)  - EXPERIMENTAL FRAC. DYE BOUND
C     *    FHCALC(I,J) - CALCULATED FRAC. DOUBLE STRANDS.
C     *    FDCALC(I,J) - CALCULATED FRAC. DYE BOUND
C     *    AKEQM(I,J)  - DYE BINDING EQUILIBRIUM CONSTANT WHICH FITS
C     *                  FDEXP(I,J).
C     *    CHIFH(I)    - REDUCED CHI SQUARED FOR FIT OF FRAC. D.S. OF
C     *                  MELTING CURVE I.
C     *    CHIFD(I)    - REDUCED CHI SQUARED FOR FIT TO FRAC. DYE BOUND
C     *    CHIFHT      - TOTAL REDUCED CHI SQUARED FOR FRAC. D.S. FOR
C     *                  ALL THE MELTING CURVES.
C     *    CHIFDT      - TOTAL REDUCED CHI SQUARED FOR FRAC. DYE BOUND.
C     *
C     * /BLKIII/
C     *    DHHEL  - DELTA H FOR DOUBLE STRAND FORMATION.
C     *    DSHEL  - DELTA S FOR DOUBLE STRAND FORMATION.
C     *    DHDYE  - GUESS FOR DYE BINDING DELTA H.
C     *    DSDYE  - GUESS FOR DYE BINDING DELTA S.
```

```
C     *     DHSIG   - DELTA H FOR SIGMA PARAMETER IN MODEL (SEE PARTIT).
C     *     DSSIG   - DELTA S ...
C     *     DHTAU   - DELTA H FOR TAU PARAMETER IN MODEL (SEE PARTIT).
C     *     DSTAU   - DELTA S ...
C     *
C     * /BLKIV/
C     *     TINV(300) - 1/T (DEG K-1) FOR ALL THE AKEQM'S.
C     *     ALNK(300) - LN(AKEQM) FIT TO EXPT'L FRAC. DYE BOUND.
C     *     NREG(I)   - HOW MANY POINTS FROM MELTING CURVE I ARE
C     *                 IN TINV AND ALNK.
C     *     DHBEST    - DELTA H FOR DYE BINDING FIT TO ALNK VS. TINV
C     *     DSBEST    - DELTA S FOR DYE BINDING FIT.
C     *     ERRDH     - ERROR IN FITTING DHBEST.
C     *     ERRDS     - ERROR IN FITTING DSBEST
C     *     R         - LINEAR CORRELATION COEFFICIENT
C     *
C     * MISC
C     *     IDENT(I,80) - 80 CHARACTER ID HEADER FOR MELTING CURVE I
C     *     HEADER(80)  - 80 CHARACTER ID HEADER FOR FIT TO DHBEST AND
C     *                   DSBEST
C     *
C     * SUBROUTINES REQUIRED:
C     *
C     *     SUBROUTINE PARTIT(MODEL,MODE,KH,KD,CT,CD,SIGMA,TAU,
C     *                       FH,FD,P,Q,ITER)
C     *        SOLVES FOR EITHER KD (MODE = 2) OR FD (MODE = 1)
C     *        USING THE MODEL SPECIFIED IN MODEL.  THE PARTITION
C     *        FUNCTION IS RETURNED IN Q.  THE FRACTIONAL POPULATIONS
C     *        OF DOUBLE STRANDS WITH J DYES BOUND (J BET. 0 AND 10.
C     *        INCLUSIVE) ARE RETURNED IN P(J).  ITER RETURNS AS THE NUMBER
C     *        OF ITERATIONS REQUIRED.  ITER = 0 IF NO FIT IS ACHIEVED
C     *
C     *     SUBROUTINE INPUT(HEADER,IDENT)
C     *        INPUTS NEW MELTING CURVES, VALUES FOR DOUBLE STRAND
C     *        FORMATION, SIGMA AND TAU, AND THE CHOICE OF MODELS.
C     *
C     *     SUBROUTINE FIT
C     *        FITS AKEQM, THE EQUILIBRIUM CONSTANT FOR DYE BINDING, TO
C     *        EACH VALUE OF THE EXPERIMENTAL FRAC. DYE BOUND.
C     *
C     *     SUBROUTINE BESTK
C     *        CALCULATES THE BEST FIT OF DELTA H AND S FOR DYE BINDING,
C     *        BY DOING A LINEAR LEAST SQUARES FIT TO LN(AKEQM) VS. 1/T
C     *
C     *     SUBROUTINE THEORY
C     *        CALCULATES VALUES FOR FRAC. DOUBLE STRAND AND FRAC. DYE
C     *        BOUND USING BEST PARAMETERS DETERMINED BY SUBROUTINE BESTK
C     *
C     *     SUBROUTINE PRNT(HEADER,IDENT)
C     *        PRINTS RESULTS TO THE SCREEN, AND OPTIONALLY TO A DISK FILE.
C     *
C     *     SUBROUTINE PLOT(HEADER,IDENT)
C     *        PLOTS LN(AKEQM) VS. TINV FOR ALL THE MELTING CURVES
C     *        TOGETHER, AND PLOTS SEPARATELY EXPERIMENTAL AND CALCULATED
C     *        FRACTION DOUBLE STRAND AND DYE BOUND.
C     *
C     *     SUBROUTINE FNAME(FILE)
C     *        GETS THE NAME OF A DISK FILE FOR INPUT OR OUTPUT.
C     *
C     *     SEVERAL PLOTTING SUBROUTINES ARE USED FOR PLOTTING.
C     *
C     ***********************************************************************
C     * THIS PROGRAM WAS WRITTEN BY JEFF NELSON, JUNE, 1982.
C     ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
```

```
     2                CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
      BYTE HEADER(80),IDENT(12,80)
C
  100 CALL INPUT(HEADER,IDENT)
      CALL FIT
      CALL BESTK
      CALL THEORY
      CALL PRNT(HEADER,IDENT)
      CALL PLOT(HEADER,IDENT)
      WRITE(6,1000)
 1000 FORMAT(' MAKE FIT FROM CONSTANT PARAMETERS? ',$)
      READ(5,1020)INST
 1020 FORMAT(A1)
      IF (INST .EQ. 'Y') THEN
         WRITE(6,1040)
 1040    FORMAT(' VALUES FOR DH, DS FOR DYE BINDING, 1/LINE ')
         READ(5,1060)DHZ,DSZ
 1060    FORMAT(F15.0)
         CALL FIT
         CALL BESTK
         DHBEST = DHZ
         DSBEST = DSZ
         ERRDH = 0.
         ERRDS = 0.
         R = 0.
         CALL THEORY
         CALL PRNT
         CALL PLOT
      END IF
      GOTO 100
      END



      SUBROUTINE INPUT(HEADER,IDENT)
C
C    ***********************************************************************
C    * THIS SUBROUTINE INPUTS THE DYE MELTING CURVES, FROM FILES
C    * SPECIFIED FROM THE TERMINAL.
C    * UP TO 12 MAY BE INPUT
C    *
C    * THE FILES WERE CREATED BY PROGRAM DYEMLT, WHICH DETERMINES THE
C    * FRACTION DOUBLE STRANDS AND FRACTION DYE BOUND FROM MELTING
C    * CURVES.
C    *
C    * FILE STRUCTURE:
C    *
C    *    LINE 1 - "DYE MELTING CURVE"
C    *    LINE 2 - ID HEADER FROM MELTING CURVE FILE.
C    *    LINE 3 - ID HEADER FOR THIS PARTICULAR SAMPLE.
C    *    LINE 4 - NTFST,NTLAST,2I5," FIRST AND LAST TEMP"
C    *    LINE 5 - TOTAL STRANC CONC., 1PE15.5," TOTAL STRAND CONC"
C    *    LINE 6 - TOTAL DYE CONC., 1PE15.5," TOTAL DYE CONC"
C    *    LINE 7 - COLUMN HEADERS
C    *    LINE 8 - TEMP, FH, FD, ABS(260), ABS(283),
C    *             LOW BL, HIGH BL (260), LOW BL, HIGH BL (283),
C    *             I5,8F9.5
C    *
C    * ENTHALPY AND ENTROPY DATA ARE ALSO INPUT, AS WELL AS THE MODEL
```

```
C     * TO BE USED.
C     *
C     * UPON RETURN FROM A LOOP FROM A PROGRAM, NEW MELTING CURVE
C     * DATA MAY BE INPUT, NEW ENTHALPIES AND ENTROPIES, AND THE
C     * PROGRAM GOES AGAIN.
C     *
C     ***********************************************************************
C     * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JUNE, 1982.
C     ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     2               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
      BYTE HEADER(80),IDENT(12,80)
      CHARACTER*20 FILE
      CHARACTER*17 CHECK
C
   50 WRITE(6,1000)
 1000 FORMAT(' INPUT NEW MELTING CURVES? ',$)
      READ(5,1010)INST
 1010 FORMAT(A1)
      IF (INST .EQ. 'Y') THEN
         NZZ = 0
  100    CALL FNAME(FILE)
         IF (FILE .EQ. ' ') GOTO 200
         OPEN(UNIT=10,FILE=FILE,STATUS='OLD',ERR=120)
         READ(10,1020)CHECK
 1020    FORMAT(A)
         IF (CHECK .NE. 'DYE MELTING CURVE') THEN
            WRITE(6,1030)FILE,CHECK
 1030       FORMAT(' ***** FILE NOT RIGHT TYPE *****',/,1X,A,/,1X,A)
            CLOSE(UNIT=10)
            GOTO 100
         END IF
         NZZ = NZZ + 1
         READ(10,1010)NDUM
         READ(10,1040)(IDENT(NZZ,I),I=1,80)
         WRITE(6,1045)(IDENT(NZZ,I),I=1,80)
         READ(10,1050)NTFST(NZZ),NTLAST(NZZ)
         READ(10,1060)CSTRAN(NZZ)
         READ(10,1060)CDYE(NZZ)
         READ(10,1010)NDUM
         READ(10,1070)(FHEXP(NZZ,I),FDEXP(NZZ,I),
     2                 I=NTFST(NZZ),NTLAST(NZZ))
         CLOSE(UNIT=10)
 1040    FORMAT(80A1)
 1045    FORMAT(1X,80A1)
 1050    FORMAT(2I5)
 1060    FORMAT(E15.0)
 1070    FORMAT(5X,2F9.0)
         IF (NZZ .EQ. 12) GOTO 200
         GOTO 100
  120    WRITE(6,1100)FILE
 1100    FORMAT(' ***** FILE OPEN ERROR ***** ',A)
         GOTO 100
      END IF
  200 IF (NZZ .EQ. 0) THEN
         WRITE(6,1150)
 1150    FORMAT(' *** OLD MELTING CURVES USED AGAIN ***')
      ELSE
         NMELTS = NZZ
```

```
      END IF
C
C   * GET THE THERMODYNAMIC VALUES...
C
      EQHEL = EXP((DSHEL - DHHEL/298.15)*0.5033)
      EQDYE = EXP((DSDYE - DHDYE/298.15)*0.5033)
      EQSIG = EXP((DSSIG - DHSIG/298.15)*0.5033)
      EQTAU = EXP((DSTAU - DHTAU/298.15)*0.5033)
      WRITE(6,1200)
 1200 FORMAT(' THERMO VALUES: '/
     2        ' NO.  PARAM    DELTA H        DELTA S        K(25DEG) '/)
      WRITE(6,1210)1,'HELIX',DHHEL,DSHEL,EQHEL
      WRITE(6,1210)2,'DYE  ',DHDYE,DSDYE,EQDYE
      WRITE(6,1210)3,'SIGMA',DHSIG,DSSIG,EQSIG
      WRITE(6,1210)4,'TAU  ',DHTAU,DSTAU,EQTAU
 1210 FORMAT(I3,2X,A5,2X,3(1PE11.4,3X))
  250 WRITE(6,1230)
 1230 FORMAT(/,' CHANGE WHICH PARAMETER (0=DONE) : ',$)
      READ(5,1240,ERR=280)INST
 1240 FORMAT(I5)
      IF (INST .EQ. 1) THEN
         WRITE(6,1300)
         READ(5,1310,ERR=280)DHHEL,DSHEL
 1300    FORMAT(' NEW VALUES (1/LINE) : ')
 1310    FORMAT(E15.0)
      ELSE IF (INST .EQ. 2) THEN
         WRITE(6,1300)
         READ(5,1310,ERR=280)DHDYE,DSDYE
      ELSE IF (INST .EQ. 3) THEN
         WRITE(6,1300)
         READ(5,1310,ERR=280)DHSIG,DSSIG
      ELSE IF (INST .EQ. 4) THEN
         WRITE(6,1300)
         READ(5,1310,ERR=280)DHTAU,DSTAU
      ELSE
         GOTO 300
      END IF
      GOTO 250
  280 WRITE(6,1350)
 1350 FORMAT(' *** ERROR IN INPUT VALUE ***')
      GOTO 50
  300 CONTINUE
C
C   * GET MODEL NUMBER FOR SUBROUTINE PARTIT
C
      WRITE(6,1400)
 1400 FORMAT(' WHICH MODEL NUMBER? '/,
     2        '  1 = ALL SITES EQUAL'/,    2 = 1 STRONG SITE (SIGMA)'/,
     3        '  3 = TWO STRONG SITES ON END (TAU)'/,
     4        '  4 = STRONG SITE ON 2ND B.P.'/,
     5        '  5 = STRONG SITE ON 3RD B.P.'/,
     6        '  6 = COOPERATIVE BINDING (TAU)'/,
     7        '  7 = SIGMA/ TAU STRONG SITES'/,
     8        '  8 = STRONG SITE + COOPERATIVE'/,
     9        '  9 = TWO STRONG SITES + COOP')
      READ(5,1420,ERR=280)MODEL
 1420 FORMAT(I5)
      IF (MODEL .EQ. 0) GOTO 50
C
C   * GET HEADER FOR THIS RUN, AND RETURN.
C
      WRITE(6,1500)
 1500 FORMAT(/,' HEADER FOR THIS RUN: ')
      READ(5,1510)HEADER
 1510 FORMAT(80A1)
      RETURN
```

```
      END


      SUBROUTINE FIT
C
C     ******************************************************************
C     * THIS SUBROUTINE DETERMINES THE VALUE FOR KD, THE EQUILIBRIUM
C     * CONSTANT FOR BINDING A DYE (ETHIDIUM), FOR EACH EXPERIMENTAL VALUE
C     * OF FD, FRACTION OF DYE BOUND.  RESULTS ARE STORED IN THE ARRAY
C     * AKEQM.
C     *
C     * THE FITTING IS DONE FOR POINTS WITHIN A RANGE OF FD, AS
C     * SPECIFIED IN THE PARAMETERS FDMIN AND FDMAX.
C     *
C     * THE FITTING ROUTINE IN PARTIT IS A NEWTON'S METHOD.  THE INITIAL
C     * GUESS FOR KD IS FROM DHDYE AND DSDYE IF THE FIRST TEMP FOR THAT
C     * MELTING CURVE.  OTHERWISE IT IS THE VALUE DETERMINED FOR THE
C     * PREVIOUS TEMPERATURE.
C     *
C     * SUBROUTINES USED:
C     *
C     *    SUBROUTINE PARTIT(MODEL,MODE,KH,KD,CT,CD,SIGMA,TAU,
C     *       FH,FD,P,Q,ITER)
C     *       CALCULATES KD TO FIT THE EXPERIMENTAL FD, BY SETTING
C     *       MODE = 2.  MODEL SPECIFIES WHICH MODEL PARTIT SHOULD USE.
C     *
C     ******************************************************************
C     * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JULY, 1982.
C     ******************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINY(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
      PARAMETER (FDMIN = 0.2, FDMAX = 0.9)
      DIMENSION P(0:10)
      REAL KH,KD
C
      DO I=1,NMELTS
         J = 0
         DO WHILE (FDEXP(I,J) .GT. FDMAX)
            J = J + 1
         END DO
         NTFIT1(I) = J
         KD = 0.
         DO WHILE (FDEXP(I,J) .GT. FDMIN)
            TEMP = FLOAT(J) + 273.15
            KH = EXP((DSHEL - DHHEL/TEMP)*0.5033)
C    *       USE DSDYE AND DHDYE FOR FIRST GUESS; OTHERWISE, PREVIOUS KD
C             IF (KD .EQ. 0.) KD = EXP((DSDYE - DHDYE/TEMP)*0.5033)
C    *       USE KD = 1.E5 AS INITIAL GUESS
            IF (KD .EQ. 0.) KD = 1.E5
            SIGMA = EXP((DSSIG - DHSIG/TEMP)*0.5033)
            TAU = EXP((DSTAU - DHTAU/TEMP)*0.5033)
            FD = FDEXP(I,J)
            CT = CSTRAN(I)
            CD = CDYE(I)
  100       CALL PARTIT(MODEL,2,KH,KD,CT,CD,SIGMA,TAU,
     2                  FH,FD,P,Q,ITER)
C    *       ITER = 0 OR -1 IS ERROR CONDITION
```

```
            IF (ITER .EQ. -1) THEN
               WRITE(6,1200)MODEL
 1200          FORMAT(' *** NO MODEL NUMBER',I3,' IN PARTIT. SET TO 1')
               MODEL = 1
               GOTO 100
            ELSE IF (ITER .EQ. 0) THEN
               WRITE(6,1210)I,J,FD
 1210          FORMAT(' *** NO CONVERGENCE,',I3,'TH MELT, T = ',I4,
     2                ', FD = ',F9.5)
               AKEQM(I,J) = 0.
            ELSE
               AKEQM(I,J) = KD
            END IF
            NTFIT2(I) = J
            J = J + 1
         END DO
      END DO
      RETURN
      END


      SUBROUTINE BESTK
C
C     ***********************************************************************
C     * THIS SUBROUTINE DETERMINES THE BEST VALUE FOR THE EQUILIBRIUM
C     * CONSTANT BY DOING LINEAR REGRESSION ON LN(KD) VS. 1/T.
C     *
C     * INPUT VARIABLES:
C     *
C     *    AKEQM(I,J) - KD FOR MELT I, TEMP J.
C     *    NTFIT1(I)  - THE FIRST TEMP FIT FOR MELT I.
C     *    NTFIT2(I)  - THE LAST TEMP FIT FOR MELT I.
C     *
C     * OUTPUT VARIABLES:
C     *
C     *    ALNK   - ARRAY OF LN(KD)'S
C     *    TINV   - ARRAY OF 1/T (KELVIN)'S
C     *    DHBEST - DELTA H DETERMINED HEREIN.
C     *    DSBEST - DELTA S
C     *    DHSIG  - ERROR IN DHBEST
C     *    DSSIG  - ERROR IN DSBEST
C     *    R      - THE CORRELATION COEFF
C     ***********************************************************************
C     * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JULY, 1982.
C     ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
C
C     * CREATE ARRAYS TINV AND ALNK.
C
      IND = 0
      DO I = 1,NMELTS
         NREG(I) = NTFIT2(I) - NTFIT1(I) + 1
         DO J = NTFIT1(I),NTFIT2(I)
C     *      AKEQM = 0. SIGNALS NO FIT.
            IF (AKEQM(I,J) .EQ. 0.) THEN
               NREG(I) = NREG(I) - 1
               GOTO 100
```

```
            END IF
            IND = IND + 1
            TINV(IND) = 1. / (FLOAT(J) + 273.15)
            ALNK(IND) = ALOG( AKEQM(I,J) )
  100     END DO
      END DO
C
C   * DO LINEAR REGRESSION ANALYSIS.
C
      MODE = 0
      CALL LINREG(TINV,ALNK,DUM,IND,MODE,A,SIGMAA,B,SIGMAB,R)
      DHBEST = -B*1.987
      ERRDH = 1.987*SIGMAB
      DSBEST = 1.987*A
      ERRDS = 1.987*SIGMAA
C
      RETURN
      END



      SUBROUTINE THEORY
C
C   ***********************************************************************
C   * THIS SUBROUTINE CALCULATES THE VALUES FOR THE FRACTION DYE BOUND,
C   * AND THE FRACTION DOUBLE STRANDS, USING THE VALUES FOR DHBEST
C   * AND DSBEST DETERMINED FROM LN(KD) VS. 1/T.
C   *.
C   * SUBROUTINE PARTIT DOES THE CALCULATIONS.
C   ***********************************************************************
C   * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JULY, 1982.
C   ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,
     2               VARNCE,R
      REAL KH,KD
      DIMENSION P(0:10)
C
C   * INITIALIZE ARRAYS WITH -1.
C
      DO I = 1,NMELTS
         DO J = 0,80
            FDCALC(I,J) = -1.
            FHCALC(I,J) = -1.
         END DO
      END DO
C
C   * CALCULATE FOR EACH EXPERIMENTAL POINT
C
      DO I=1,NMELTS
         CT = CSTRAN(I)
         CD = CDYE(I)
         FD = .5
         DO JT = NTFST(I),NTLAST(I)
            TEMP = FLOAT(JT) + 273.15
            KH = EXP((DSHEL - DHHEL/TEMP)*0.5033)
            KD = EXP((DSBEST - DHBEST/TEMP)*0.5033)
            SIGMA = EXP((DSSIG - DHSIG/TEMP)*0.5033)
```

```
            TAU = EXP((DSTAU - DHTAU/TEMP)*0.5033)
            CALL PARTIT(MODEL,1,KH,KD,CT,CD,SIGMA,TAU,
     2                  FH,FD,P,Q,ITER)
            IF (ITER .EQ. 0) THEN
               WRITE(6,800)I,JT,FD
  800          FORMAT(' *** NO CONVERGENCE FOR FD, MELT(,I3,
     2                ',TEMP(,I3,'), FD =(,F9.5)
               FD = 0.
               FH = 0.
            END IF
            FDCALC(I,JT) = FD
            FHCALC(I,JT) = FH
         END DO
      END DO
C
C    * CALCULATE CHI SQUARED FOR THIS FIT
C
      CHIFDT = 0.
      CHIFHT = 0.
      NTOT = 0
      DO I=1,NMELTS
         CHIFD(I) = 0
         CHIFH(I) = 0.
         DO J=NTFIT1(I),NTFIT2(I)
            DIFF = FDEXP(I,J) - FDCALC(I,J)
            CHIFD(I) = CHIFD(I) + DIFF*DIFF
            DIFF = FHEXP(I,J) - FHCALC(I,J)
            CHIFH(I) = CHIFH(I) + DIFF*DIFF
         END DO
         CHIFDT = CHIFDT + CHIFD(I)
         CHIFHT = CHIFHT + CHIFH(I)
         NT = NTFIT2(I) - NTFIT1(I) + 1
         CHIFD(I) = CHIFD(I) / FLOAT(NT-2)
         CHIFH(I) = CHIFH(I) / FLOAT(NT-2)
         NTOT = NTOT + NT
      END DO
      CHIFDT = CHIFDT / FLOAT(NTOT - 2)
      CHIFHT = CHIFHT / FLOAT(NTOT - 2)
      RETURN
      END



      SUBROUTINE PRNT(HEADER,IDENT)
C
C    ***********************************************************************
C    * THIS SUBROUTINE PRINTS THE RESULTS TO THE SCREEN OR TO DISK FILES
C    * FOR LATER PRINTING.
C    ***********************************************************************
C    * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JULY, 1982.
C    ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
      BYTE HEADER(80),IDENT(12,80)
      CHARACTER*30 FILE
      LOGICAL SCREEN
C
C    * PRINT OUT SUMMARY OF FIT, CHI SQUARED TO SCREEN
```

```
C
      SCREEN = .TRUE.
      IWRITE = 6
      GOTO 200
C
C   * DETERMINE IF DISK FILE IS DESIRED.
C
  100 WRITE(6,1000)
 1000 FORMAT(/,' NAME FOR DISK FILE (RETURN = NONE)')
      CALL FNAME(FILE)
      IF (FILE .EQ. ' ') GOTO 200
      OPEN(UNIT=10, FILE=FILE, STATUS='NEW', ERR=120)
      IWRITE = 10
      GOTO 200
C
  120 WRITE(6,1080)FILE
 1080 FORMAT(' ***** ERROR OPENING FILE ***** ',A)
      GOTO 100
  200 CONTINUE
C
C   ******************************************************************
C   * WRITE OVERALL CHARACTERISTICS
C   ******************************************************************
C
      WRITE(IWRITE,2000)HEADER
 2000 FORMAT(1H1,80A1)
      WRITE(IWRITE,2020)MODEL
 2020 FORMAT(/,' MODEL USED WAS ',I2)
      EH = EXP((DSHEL - DHHEL/298.15)*0.5033)
      ED = EXP((DSDYE - DHDYE/298.15)*0.5033)
      ES = EXP((DSSIG - DHSIG/298.15)*0.5033)
      ET = EXP((DSTAU - DHTAU/298.15)*0.5033)
C
      WRITE(IWRITE,2050)DHHEL,DSHEL,EH,DHDYE,DSDYE,ED,
     2                  DHSIG,DSSIG,ES,DHTAU,DSTAU,ET
 2050 FORMAT(/,'        DELTA H       DELTA S     K (25 DEG)',/,
     2          1PE14.4,E13.4,E13.4,'  HELIX',/,
     3            E14.4,E13.4,E13.4,'  DYE GUESS',/,
     4            E14.4,E13.4,E13.4,'  SIGMA',/,
     5            E14.4,E13.4,E13.4,'  TAU ')
C
      EB = EXP((DSBEST - DHBEST/298.15)*0.5033)
      WRITE(IWRITE,2100)DHBEST,DSBEST,EB,R,ERRDH,ERRDS
 2100 FORMAT(/,1PE14.4,E13.4,E13.4,'  DYE, FIT,  R =',0PF8.5,/,
     2          1PE14.4,E13.4,13X,'  ERROR')
C
      WRITE(IWRITE,2150)
 2150 FORMAT(/,12X,'REDUCED CHI SQUARED',//,
     2        5X,'MELT      F(DYE)     F(HELIX)  POINTS FIT',/)
      WRITE(IWRITE,2200)CHIFDT,CHIFHT
 2200 FORMAT('   OVERALL',1PE12.3,E12.3,/)
      WRITE(IWRITE,2250)(I,CHIFD(I),CHIFH(I),NREG(I),I=1,NMELTS)
 2250 FORMAT(I9,1PE12.3,E12.3,I7)
      IF (SCREEN) THEN
         SCREEN = .FALSE.
         GOTO 100
      END IF
C
C   ******************************************************************
C   * WRITE PARTICULARS FOR EACH FILE.
C   ******************************************************************
C
      DO I=1,NMELTS
         WRITE(IWRITE,3000)(IDENT(I,J),J=1,80),CSTRAN(I),CDYE(I),NREG(I)
 3000    FORMAT(1H1,80A1,//,1PE13.5,' STRAND CONC',/,E13.5,' DYE CONC',
     2          /,I13,' POINTS FITTED')
```

```
         WRITE(IWRITE,3020)
 3020    FORMAT(/,' TEMP      EXPT FD CALC         EXPT FH CALC')
         NSTART = NTFIT1(I) - 15
         IF (NSTART .LT. 0) NSTART = 0
         NSTOP = NTFIT2(I) + 15
         IF (NSTOP .GT. 80) NSTOP = 80
         DO J = NSTART,NSTOP
            WRITE(IWRITE,3100)J,FDEXP(I,J),FDCALC(I,J),
     2                        FHEXP(I,J),FHCALC(I,J)
 3100       FORMAT(I5,F9.4,F8.4,F12.4,F8.4)
         END DO
      END DO
C
C    * DONE WRITING
C
      IF (IWRITE .EQ. 10) CLOSE (UNIT=10)
  300 RETURN
      END



      SUBROUTINE PLOT(HEADER,IDENT)
C
C    ***********************************************************************
C    * THIS SUBROUTINE MAKES PLOTS OF THE LN(K) VS. 1/T, AND THE
C    * EXPERIMENTAL AND CALCULATED FRACTION DYE BOUND AND DOUBLE HELIX.
C    ***********************************************************************
C
      COMMON /BLKI/ NMELTS,NTFST(12),NTLAST(12),NTFIT1(12),NTFIT2(12),
     2              CSTRAN(12),CDYE(12),MODEL
      COMMON /BLKII/ FHEXP(12,0:80),FDEXP(12,0:80),FHCALC(12,0:80),
     2               FDCALC(12,0:80),AKEQM(12,0:80),CHIFH(12),CHIFD(12),
     3               CHIFHT,CHIFDT
      COMMON /BLKIII/ DHHEL,DSHEL,DHDYE,DSDYE,DHSIG,DSSIG,DHTAU,DSTAU
      COMMON /BLKIV/ TINV(400),ALNK(400),NREG(12),DHBEST,DSBEST,ERRDH,
     2               ERRDS,R
      BYTE HEADER(80),IDENT(12,80)
      LOGICAL FIRST,CHANGE
      DIMENSION X(81),Y(81),YFIT(81)
      DATA FIRST,CHANGE/.TRUE.,.FALSE./
      DATA GXMIN,GXMAX,GYMIN,GYMAX/20.,115.,20.,95./
      DATA GXMIN2,GXMAX2,GYMIN2,GYMAX2/20.,35.,15.,75./
      DATA LABSIZ,NTISIZ,TICLEN,TICLAB/1.1,1.1,.2,./
C
C    * DETERMINE IF ANY PLOTS ARE DESIRED
C
      WRITE(6,1000)
 1000 FORMAT(/,' MAKE PLOTS? ',$)
      READ(5,1010)INST
 1010 FORMAT(A1)
      IF (INST .NE. 'Y') GOTO 300
      IF (FIRST) THEN
   50    WRITE(6,1020)
 1020    FORMAT(' PLOTTER OPTIONS (E.G. 4014,2) ',$)
         READ(5,1040,ERR=50)N1,N2
 1040    FORMAT(2I5)
         WRITE(6,1060)
 1060    FORMAT(' MAKE PLOT PARAMETERS CHANGABLE? ',$)
         READ(5,1010)INST2
         IF (INST2 .EQ. 'Y') CHANGE = .TRUE.
         CALL GRSTRT(N1,N2)
         CALL TXAM
         FIRST = .FALSE.
      END IF
C
```

```
C    * MAKE PLOT OF LN(K) VS. 1/T.
C
      WRITE(6,1100)
 1100 FORMAT(' PLOT LN(K) VS. 1/T? ',$)
      READ(5,1010)INST
      IF (INST .EQ. 'Y') THEN
         IF (CHANGE) THEN
            WRITE(6,1120)
 1120       FORMAT(' USE DEFAULT VALUES? ',$)
            READ(5,1010)INST2
            IF (INST2 .NE. 'Y') THEN
               WRITE(6,1130)
 1130          FORMAT(' CHANGE THE VALUES? ',$)
               READ(5,1010)INST3
               IF (INST3 .EQ. 'Y') THEN
  200             WRITE(6,1140)GXMIN,GXMAX,GYMIN,GYMAX
 1140             FORMAT(' OLD VALUES FOR GXMIN, GXMAX, GYMIN, GYMAX ',
     2                  /,1PE12.4,/,E12.4,/,E12.4,/,E12.4,/,
     3                  ' NEW VALUES (1/LINE) ',/)
                  READ(5,1150,ERR=200)GXMIN,GXMAX,GYMIN,GYMAX
 1150             FORMAT(F15.0)
                  WRITE(6,1160)XMIN,XMAX,YMIN,YMAX
 1160             FORMAT(/,' OLD VALUES FOR XMIN, XMAX, YMIN, YMAX ',/,
     2                    1PE12.4,/,E12.4,/,E12.4,/,E12.4,/,
     3                    ' NEW VALUES (1/LINE) ',/)
                  READ(5,1150,ERR=200)XMIN,XMAX,YMIN,YMAX
               END IF
               GOTO 300
            END IF
         END IF
C    *   DETERMINE RANGE OF DATA
         NTOT = 0
         DO I = 1, NMELTS
            NTOT = NTOT + NREG(I)
         END DO
         CALL MINMAX(NTOT,XSM,XBIG,TINV)
         CALL MINMAX(NTOT,YSM,YBIG,ALNK)
         XMIN = (XSM - 0.1*(XBIG - XSM)) * 1.E3
         XMAX = (XBIG + 0.1*(XBIG - XSM)) * 1.E3
         YMIN = YSM - 0.1*(YBIG - YSM)
         YMAX = YBIG + 0.1*(YBIG - YSM)
C    *   SET UP AXIS LABELS.
  300    CALL AXIS(XMIN,XMAX,XLABST,XLABIN)
         CALL AXIS(YMIN,YMAX,YLABST,YLABIN)
         XTICIN = 0.5*XLABIN
         XTICST = XLABST
         IF ( (XTICST - XTICIN) .GE. XMIN ) XTICST = XTICST - XTICIN
         YTICIN = 0.5*YLABIN
         YTICST = YLABST
         IF ( (YTICST - YTICIN) .GE. YMIN ) YTICST = YTICST - YTICIN
C    *   MAKE AXES.
         IF (CHANGE) THEN
            WRITE(6,1190)
 1190       FORMAT(' MAKE AXES? ',$)
            READ(5,1195)INST3
 1195       FORMAT(A1)
            IF (INST3 .NE. 'Y') GOTO 400
         END IF
         CALL GRID(GXMIN,GXMAX,GYMIN,GYMAX,XMIN,XMAX,YMIN,YMAX,
     2             XTICST,XTICIN,YTICST,YTICIN,TICLEN)
         CALL LABEL(GXMIN,GXMAX,GYMIN,GYMAX,XMIN,XMAX,YMIN,YMAX,
     2              LABSIZ,TICLAB,XLABST,XLABIN,4,2,
     3              YLABST,YLABIN,6,2,
     4              0,DUM,0,DUM,0,DUM,0)
C
C    *   GRAPH EACH MELTING CURVE.
```

```
C
  400     NCHAR = 0
          IND = 0
          DO I = 1, NMELTS
             NCHAR = NCHAR + 1
             IF (NCHAR .GE. 8) NCHAR = 1
             DO J = 1, NREG(I)
                IND = IND + 1
                X(J) = TINV(IND)*1.E3
                Y(J) = ALNK(IND)
             END DO
             CALL POINT(GXMIN,GXMAX,GYMIN,GYMAX,XMIN,XMAX,YMIN,YMAX,
     2                  NREG(I),NCHAR,1,X,Y)
          END DO
C
C    * PLOT LEAST SQUARES LINE.
C
          CALL LEAST(GXMIN,GXMAX,GYMIN,GYMAX,XMIN,XMAX,YMIN,YMAX,
     2               DSBEST/1.987, -DHBEST*1.E-3/1.987)
C
C    * FINISH PLOT.
C
          CALL HOME
          CALL GRSEND
        END IF
C
C    * DETERMINE IF MELTING CURVES ARE TO BE PLOTTED.
C
        XMIN2 = 0.
        XMAX2 = 60.
        YMIN2 = -0.1
        YMAX2 = 1.1
        DO I = 1, NMELTS
          WRITE(6,1200)I
 1200     FORMAT(' PLOT MELTING CURVE #',I2,' ("R" = RETURN) ?',$)
          READ(5,1210)INST
 1210     FORMAT(A1)
          IF (INST .EQ. 'R') GOTO 800
          IF (INST .NE. 'Y') GOTO 500
C
C    *    PLOT FRACTION DYE BOUND.
C
          WRITE(6,1250)
 1250     FORMAT(' PLOT FRAC. DYE BOUND?',$)
          READ(5,1210)INST
          IF (INST .EQ. 'Y') THEN
             NCHAR = 1
             NSIZ = 1
             IF (CHANGE) THEN
                WRITE(6,1260)
 1260           FORMAT(' SYMBOL #: ',$)
                READ(5,1270)NCHAR
 1270           FORMAT(I5)
                WRITE(6,1280)
 1280           FORMAT(' SYMBOL SIZE: ',$)
                READ(5,1270)NSIZ
             END IF
             IND = 0
             DO J = NTFST(I),NTLAST(I)
                IND = IND + 1
                X(IND) = FLOAT(J)
                Y(IND) = FDEXP(I,J)
                YFIT(IND) = FDCALC(I,J)
             END DO
             CALL PLOTIT(GXMIN2,GXMAX2,GYMIN2,GYMAX2,
     1                   XMIN2,XMAX2,YMIN2,YMAX2,
```

```
     2                    XMIN2,10.,0.,  1,1 ,
     3                    1,2.,XMIN2,20.,2,-1,
     4                    0.,  5,4,1,
     5                    0,'TEMP',0,'DYE',0,DUM,1,
     6                    IND,NCHAR,NSIZ,X,Y,
     7                     .FALSE.,DUM1,DUM2)
            CALL POINT(GXMIN2,GXMAX2,GYMIN2,GYMAX2,
     1                  XMIN2,XMAX2,YMIN2,YMAX2,
     2                  IND,-1,1,X,YFIT)
            CALL HOME
            CALL GRSEND
         END IF
C
C    *   PLOT FRACTION DOUBLE STRANDS
C
         WRITE(6,1350)
 1350    FORMAT(' PLOT FRAC. DOUBLE STRAND? ',$)
         READ(5,1210)INST
         IF (INST .EQ. 'Y') THEN
            NCHAR = 3
            NSIZ = 1
            IF (CHANGE) THEN
               WRITE(6,1360)
 1360          FORMAT(' SYMBOL #: ',$)
               READ(5,1370)NCHAR
 1370          FORMAT(I5)
               WRITE(6,1380)
 1380          FORMAT(' SYMBOL SIZE ',$)
               READ(5,1370)NSIZ
            END IF
            IND = 0
            DO J = NTFST(I),NTLAST(I)
               IND = IND + 1
               X(IND) = FLOAT(J)
               Y(IND) = FHEXP(I,J)
               YFIT(IND) = FHCALC(I,J)
            END DO
            CALL PLOTIT(GXMIN2,GXMAX2,GYMIN2,GYMAX2,
     1                  XMIN2,XMAX2,YMIN2,YMAX2,
     2                  XMIN2,10.,0.,  1,1 ,
     3                  1,2.,XMIN2,20.,2,-1,
     4                  0.,  5,6,1,
     5                  0,'TEMP',0,'HELIX',0,DUM,1,
     6                  IND,NCHAR,NSIZ,X,Y,
     7                   .FALSE.,DUM1,DUM2)
            CALL POINT(GXMIN2,GXMAX2,GYMIN2,GYMAX2,
     1                  XMIN2,XMAX2,YMIN2,YMAX2,
     2                  IND,-1,1,X,YFIT)
            CALL HOME
            CALL GRSEND
         END IF
C
  500 END DO
C
  800 RETURN
      END



      SUBROUTINE PARTIT(MODEL,MODE,KH,KD,CT,CD,SIGMA,TAU,
     2            FH,FD,P,Q,ITER)
C
C    ***********************************************************************
C    * THIS SUBROUTINE CALCULATES THE PARTITION FUNCTION, THE FRACTION
C    * DOUBLE STRANDS AND DYE BOUND, AND POPULATION DISTRIBUTIONS FOR
```

```
C     * ALL THE SPECIES.  UP TO 10 DYE MOLECULES PER DOUBLE STRAND MAY
C     * BE SPECIFIED BY THE MODEL.  USUALLY MUCH FEWER ARE USED
C     *
C     * DESCRIPTION OF PARAMETERS:
C     *
C     *    MODEL  - WHICH MODEL TO USE TO GET COEFFICIENTS FOR PARTITION
C     *             FUNCTION.  (MODEL = 1,2,3, ETC.)
C     *    MODE   - 1 TO SOLVE FOR FH, FD GIVEN KH, KD, CT, CD.
C     *             2 TO SOLVE FOR KD GIVEN KH, CT, CD, FD.
C     *    KH     - EQUILIBRIUM CONST. FOR DOUBLE STRAND FORMATION.
C     *    KD     - EQUILIBRIUM CONST. FOR DYE BINDING.  THIS IS SOLVED
C     *             TO FIT THE FRACTION DYE BOUND (FD) IF MODE = 2.
C     *    CT     - TOTAL STRAND CONCENTRATION.
C     *    CD     - TOTAL DYE CONCENTRATION.
C     *    SIGMA  - MODEL PARAMETER, USUALLY THE ENHANCEMENT OF THE STRONG
C     *             BINDING SITE TO THE OTHERS.
C     *    TAU    - ANOTHER MODEL PARAMETER, INSERTED FOR FLEXIBILITY.
C     *    FH     - FRACTION DOUBLE STRANDS, CALCULATED HERE.
C     *    FD     - FRACTION DYE BOUND, FITTED IF MODE = 1, USED TO
C     *             FIT KD TO THE EXPERIMENTAL FD IF MODE = 2.
C     *    P(I)   - ARRAY OF POPULATIONS OF DOUBLE STRANDS WITH I DYE
C     *             MOLECULES BOUND.
C     *    Q      - PARTITION FUNCTION
C     *    ITER   - NUMBER OF ITERATIONS REQUIRED FOR FIT.  ITER = 0
C     *             IF THE FIT DID NOT CONVERGE WITHIN A CERTAIN NUMBER
C     *             OF ITERATIONS.  ITER = -1 IF THERE IS NO MODEL
C     *             CORRESPONDING TO THE VALUE OF MODEL.
C     *              ITER = -2 IF MODE IS NOT 1 OR 2.
C     *
C     * DESCRIPTION OF THE MODEL:
C     *
C     *    THE STATISTICAL WEIGHT OF THE SPECIES ARE:
C     *
C     *       SINGLE STRANDS: S.W. = 1 (REFERENCE)
C     *       DOUBLE STRANDS: S.W.(0) = KH
C     *       DOUBLE STRANDS WITH I MOLECULES BOUND: G(I)*KH*S**I
C     *          WHERE G(I) IS A DEGENERACY FACTOR DESCRIBING THE
C     *          POSSIBLE WAYS TO BIND I MOLECULES.  G(I) IS INFLUENCED
C     *          BY THE MODEL, AND PARAMETERS SIGMA AND TAU.
C     *          S = KD*(1.-FD)*CD
C     *       UP TO N DYE MOLCULES MAY BE BOUND, DETERMINED BY THE MODEL
C     *
C     *    ONCE THE G(I) ARE SPECIFIED, THE CALCULATIONS ARE:
C     *
C     *       A VALUE FOR FD IS ASSUMED.
C     *
C     *       S = CD*(1.-FD)*KD
C     *
C     *            N
C     *       Q = SUM (KH * G(I) * S**I)
C     *           I=0
C     *
C     *               KH * G(I) * S**I
C     *       P(I) = ---------------- = FRACTION OF D.S. SPECIES WITH I
C     *                    Q            MOLECULES BOUND.
C     *
C     *            (1.+2.*Q*CT - SQRT(1. + 4.*Q*CT))
C     *       FH = ---------------------------------
C     *                          2.*Q*CT
C     *
C     *                         FH*CT     N
C     *       FCTN(FD) = FD - ----- * SUM (I * P(I))
C     *                         CD      I=1
C     *
C     *    IF THE FD WAS CORRECT, FCTN = 0.  IF NOT, A NEWTON'S METHOD IS
C     *       USED TO CALCULATE A NEW FD, AND ITERATED UNTIL FD CONVERGES
```

```
C     *     THE SAME PROCEDURE IS USED TO FIT KD, EXCEPT THE NEWTON'S
C     *        METHOD VARIES KD TO MAKE FCTN(FD) = 0 FOR THE EXPT'L FD
C     *
C     * SUBROUTINES USED:
C     *
C     *     FUNCTION FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
C     *        CALCULATES FH,Q,P(I), AND FCTN AS ABOVE.   (N IS THE NUMBER
C     *        OF DYE MOLECULES ALLOWED BY THE MODEL.)
C     *
C     * THE MODEL SPECIFICS ARE DESCRIBED IN THE SECTIONS SETTING THE
C     *     VALUES FOR THE G(I)'S AND N, THE TOTAL NUMBER OF DYES THAT
C     *     CAN BIND
C     *
C     ***********************************************************************
C     * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JUNE, 1982
C     ***********************************************************************
C
      PARAMETER (FINC = .001, FCRIT=0.00001, ITMAX=100, AKINC=1.E-4,
     2           AKCRIT = 1.E-5)
      REAL KH,KD,KD2,KDNEW
      DIMENSION P(0:10)
      DIMENSION G(0:10)
C
C     * INITIALIZE G AND P ARRAYS
C
      DO I=0,10
         P(I) = 0
         G(I) = 0
      END DO
C
C     * SET UP MODEL PARAMETERS G(I)...
C
      IF (MODEL .EQ. 1) THEN
C     *     BINDING ETHIDIUM TO RCA5G + RCU5G OR DCA5G + DCT5G
C     *     ALL SITES EQUAL MODEL.
         G(0) = 1.
         G(1) = 6.
         G(2) = 10.
         G(3) = 4.
         N = 3
      ELSE IF (MODEL .EQ. 2) THEN
C     *     PYR-PUR SITE STRONGER THAN REST BY SIGMA.
         G(0) = 1.
         G(1) = SIGMA + 5.
         G(2) = 4.*SIGMA + 6.
         G(3) = 3.*SIGMA + 1.
         N = 3
      ELSE IF (MODEL .EQ. 3) THEN
C     *     TWO END SITES STRONGER THAN REST BY TAU.
         G(0) = 1.
         G(1) = 2.*TAU + 4.
         G(2) = TAU*TAU + 6.*TAU + 3.
         G(3) = 2.*TAU*TAU + 2.*TAU
         N = 3
      ELSE IF (MODEL .EQ. 4) THEN
C     *     SECOND SITE STRONGER BY SIGMA.   (ARBITRARY MODEL)
         G(0) = 1.
         G(1) = SIGMA + 5.
         G(2) = 3.*SIGMA + 7.
         G(3) = SIGMA + 3.
         N = 3
      ELSE IF (MODEL .EQ. 5) THEN
C     *     THIRD SITE STRONGER BY SIGMA.   (ARBITRARY MODEL)
         G(0) = 1.
         G(1) = SIGMA + 5.
         G(2) = 3.*SIGMA + 7.
```

```
         G(3) = 2.*SIGMA + 2.
         N = 3
      ELSE IF (MODEL .EQ. 6) THEN
C     *  COOPERATIVITY PARAMETER TAU FOR BINDING TO ADJACENT SITES.
         G(0) = 1.
         G(1) = 6.
         G(2) = 4.*TAU + 6.
         G(3) = 2.*TAU*TAU + 2.*TAU
         N = 3
      ELSE IF (MODEL .EQ. 7) THEN
C     *  STRONG SITE ON LEFT END BY SIGMA
C     *  STRONG SITE ON RIGHT BY TAU
         G(0) = 1.
         G(1) = SIGMA + TAU + 4.
         G(2) = SIGMA*TAU + 3.*SIGMA + 3.*TAU + 3.
         G(3) = 2.*SIGMA*TAU + SIGMA + TAU
         N = 3
      ELSE IF (MODEL .EQ. 8) THEN
C     *  STRONG SITE (SIGMA) + COOPERATIVITY (TAU)
         G(0) = 1
         G(1) = SIGMA + 5.
         G(2) = SIGMA*(TAU + 3.) + 3.*TAU + 3.
         G(3) = SIGMA*(TAU*TAU + 2.*TAU) + TAU*TAU
         N = 3
      ELSE IF (MODEL .EQ. 9) THEN
C     *  TWO STRONG SITES (SIGMA) + COOPERATIVITY (TAU)
         G(0) = 1.
         G(1) = 2.*SIGMA + 4.
         G(2) = SIGMA*SIGMA + 2.*TAU*SIGMA + 2.*TAU + 4.*SIGMA + 1.
         G(3) = 2.*SIGMA*TAU*( SIGMA + TAU )
         N = 3
      ELSE
         ITER = -1
         RETURN
      END IF
C
C     * CALCULATE FD FOR THIS KH,KD,CT,CD
C
      IF (MODE .EQ. 1) THEN
         ITER = 1
C     *  MAKE SURE FIRST GUESS IS BETWEEN 0.1 AND 0.9
         IF ( (FD .GE. 0.9) .OR. (FD .LE. 0.1) ) FD = 0.5
         DO WHILE (ITER .LE. 20)
            FDIFF = FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
            IF (FD .GT. 0.5) THEN
               FD2 = FD - FINC
               FDIFF2 = FCTN(KH,KD,CT,CD,FH,FD2,P,Q,G,N)
               DERIV = (FDIFF - FDIFF2) / FINC
               FDNEW = FD - FDIFF/DERIV
            ELSE
               FD2 = FD + FINC
               FDIFF2 = FCTN(KH,KD,CT,CD,FH,FD2,P,Q,G,N)
               DERIV = (FDIFF2 - FDIFF) / FINC
               FDNEW = FD - FDIFF/DERIV
            ENDIF
            IF (ABS(FD - FDNEW) .LT. FCRIT) THEN
               FDIFF = FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
               RETURN
            ENDIF
            IF (FDNEW .GT. 1.) FDNEW = 1.0
            IF (FDNEW .LT. 0.) FDNEW = 0.
            FD = FDNEW
            ITER = ITER + 1
         END DO
C
C     *  IF THERE IS NO CONVERGENCE AFTER 20 TRIES, DO A LINEAR SEARCH
```

```
C     *     FCTN(FD) DEC  AS FD DEC.  FCTN(1) = 1. FCTN(0) = -2. OR SO.
C     *     SET FD = 1.. DECREASE BY 0.1 UNTIL FCTN  LT  0   THEN DO
C     *     LINEAR INTERPOLATION TO FIND NEW APPROXIMATE FD. UNTIL IT
C     *     CHANGES LESS THAN FCRIT
C     *     ALSO. AFTER FINDING A NEW APPROXIMATE FD. DECREASE THE INTERVAL
C     *     BETWEEN FD1 AND FD2. KEEPING FCTN(FD1) .GT. 0  AND
C     *     FCTN(FD2) .LE. 0.
C
            FD2 = 1.
            CFD2 = 1
            DO WHILE (CFD2 .GE. 0 )
               FD1 = FD2
               CFD1 = CFD2
               FD2 = FD2 - 0.1
               CFD2 = FCTN(KH,KD,CT,CD,FH,FD2,P,Q,G,N)
            END DO
            DO WHILE (ITER  LE.  ITMAX)
               FDNEW = FD2 + (0  - CFD2) * (FD1 - FD2) / (CFD1 - CFD2)
               IF ( (ABS(FDNEW - FD2) .LT.  FCRIT)  .OR.
     2              (ABS(FDNEW - FD1) .LT.  FCRIT) ) THEN
                  FD = FDNEW
                  FDIFF = FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
                  RETURN
               END IF
               FDIFF = FCTN(KH,KD,CT,CD,FH,FDNEW,P,Q,G,N)
               IF (FDIFF  LE.  0.) THEN
                  FDCHG = FDNEW - FD2
                  FD2 = FDNEW
                  CFD2 = FDIFF
                  IF ( (FD2 + FDCHG) .LT.  FD1 ) THEN
                     FD1 = FD2 + FDCHG
                     CFD1 = FCTN(KH,KD,CT,CD,FH,FD1,P,Q,G,N)
                     DO WHILE (CFD1  LT  0 )
                        FD2 = FD1
                        CFD2 = CFD1
                        FD1 = FD1 + FDCHG
                        CFD1 = FCTN(KH,KD,CT,CD,FH,FD1,P,Q,G,N)
                     END DO
                  END IF
               ELSE
                  FDCHG = FDNEW - FD1
                  FD1 = FDNEW
                  CFD1 = FDIFF
                  IF ( (FD1 + FDCHG)  GT  FD2 ) THEN
                     FD2 = FD1 + FDCHG
                     CFD2 = FCTN(KH,KD,CT,CD,FH,FD2,P,Q,G,N)
                     DO WHILE (CFD2 .GT.  0.)
                        FD1 = FD2
                        CFD1 = CFD2
                        FD2 = FD2 + FDCHG
                        CFD2 = FCTN(KH,KD,CT,CD,FH,FD2,P,Q,G,N)
                     END DO
                  END IF
               END IF
               ITER = ITER + 1
            END DO
C
C     *     NO CONVERGENCE AFTER ITER TRIES   ITER = 0 SIGNALS THIS
C
            ITER = 0
            RETURN
         ELSE IF (MODE  EQ.  2) THEN
            ITER = 1
C     *     FIT KD TO THE EXPERIMENTAL FD.
            IF (KD .LE. 0 ) KD = 1.E4
            DO WHILE (ITER  LE.  ITMAX)
```

```
            FDIFF = FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
            KD2 = KD + KD * AKINC
            FDIFF2 = FCTN(KH,KD2,CT,CD,FH,FD,P,Q,G,N)
            DERIV = (FDIFF2 - FDIFF) / (KD * AKINC)
            KDNEW = KD - FDIFF/DERIV
            IF (ABS(KDNEW-KD)/KD .LT. AKCRIT) THEN
               KD = KDNEW
               FDIFF = FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
               RETURN
            END IF
            IF (KDNEW .LE. 0.) KDNEW = KD*0.1
            KD = KDNEW
  100       ITER = ITER + 1
         END DO
C    *    NO CONVERGENCE AFTER ITER TRIES   ITER = 0 SIGNALS THIS
         ITER = 0
         RETURN
      ELSE
         ITER = -2
      END IF
      RETURN
      END



      FUNCTION FCTN(KH,KD,CT,CD,FH,FD,P,Q,G,N)
C
C    ***********************************************************************
C    * THIS FUNCTION CALCULATES FCTN = FD - FH*CT/CD*SUM(I*P(I)).
C    * SEE SUBROUTINE PARTIT FOR DETAILS.
C    *
C    * FH, P AND Q ARE DETERMINED IN THIS ROUTINE.
C    ***********************************************************************
C    * THIS SUBROUTINE WAS WRITTEN BY JEFF NELSON, JUNE, 1982.
C    ***********************************************************************
C
      DIMENSION P(0:10), G(0:10)
      REAL KH,KD
      S = KD * (1.-FD) * CD
      Q = 0.
      ZZ = 0.
      SZ = 1.
      DO I=0,N
         QZ = KH * G(I) * SZ
         P(I) = QZ
         Q = Q + QZ
         SZ = SZ * S
      END DO
      DO I=0,N
         P(I) = P(I) / Q
         ZZ = ZZ + FLOAT(I)*P(I)
      END DO
      FH = (1.+2.*Q*CT - SQRT(1.+4.*Q*CT)) / (2.*Q*CT)
      FCTN = FD - FH*CT*ZZ/CD
      RETURN
      END
```

REFERENCES

Adhya, S. & Gottesmann, M. (1978) Ann. Rev. Biochem. 47, 967-996.

Albergo, D. D., Marky, L. A., Breslauer, K. J. & Turner, D. H. (1981) Biochemistry 20, 1409-1413.

Appleby, D. W. & Kallenbach, N. R. (1973) Biopolymers 12, 2093-2120.

Applequist, J. & Damle, V. (1963) Biopolymers 12, 2093-2120.

Applequist, J. & Damle, V. (1965) J. Amer. Chem. Soc. 87, 1450-1458.

Azbel, M. Ya. (1973) Phys. Rev. Lett. 31, 589-592.

Azbel, M. Ya. (1975) J. Chem. Phys. 62, 3635-3641.

Azbel, M. Ya. (1979) Proc. Nat. Acad. Sci. 76, 101-105.

Bernasconi, C. F. (1976) Relaxation Kinetics, Academic Press, Inc., New York.

Bevington, P. R. (1969) Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York, chapter 6.

Bloomfield, V. A., Crothers, D. M. & Tinoco, I., Jr. (1974) Physical Chemistry of Nucleic Acids, Harper & Row, New York.

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

Borer, P. N., Kan, L. S. & Ts'o, P. O. P. (1975) Biochemistry 14, 4847-4863.

Breslauer, K. J. & Bina-Stein, M. (1977) Biophys. Chem. 7, 211-216.

Breslauer, K. J., Sturtevant, J. M. & Tinoco, I., Jr. (1975) J. Mol. Biol. 99, 549-565.

Bresloff, J. & Crothers, D. M. (1975) J. Mol. Biol. 95, 103-123.

Cohen, G. & Eisenberg, H. (1968) Biopolymers 6, 1077-1100.

Craig, M. E., Crothers, D. M. & Doty, P. (1971) J. Mol. Biol. 62, 383-401.

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

Dahl, K. S. (1981) Ph.D. Thesis, University of California, Berkeley, California.

Dahl, K. S., Pardi, A. & Tinoco, I., Jr. (1982) Biochemistry 21, 2730-2737.

Davanloo, P,. & Crothers, D. M. (1976) Biochemistry 15, 5299-5305.

Dewey, T. G. & Turner, D. H. (1979) Biochemistry 18, 5757-5762.

Douthart, R. J., Burnett, J. P., Weasley, F. W., & Frank, B. H. (1973) Biochemistry 12, 214-220.

Drobnies, A. E. (1979) Ph.D. Thesis, University of California, Berkeley, California.

Felsenfeld, G. & Miles, H. T. (1967) Ann. Rev. Biochem. 36, 407-448.

Fink, T. R. & Crothers, D. M. (1972a) Biopolymers 11, 127-136.

Fink, T. R. & Crothers, D. M. (1972b) J. Mol. Biol. 66, 1-12.

Fixman, M. & Freire, J. J. (1977) Biopolymers 16, 2693-2704.

Freier, S M. Albergo, D. D. & Turner, D. H. (1982) Biopolymers, submitted.

Fritz, H. J., Belagaje, R., Brown, E. L., Fritz, R. H., Jones, R. A., Lees, R. G. & Khorana, H. G. (1978) Biochemistry 17, 1257-1267.

Gilbert, W. (1976) in RNA Polymerase, Losick, R. & Chamberlain, M. J., Eds., Cold Spring Harbor Laboratory, New York, pp. 195-205.

Goeddel, D. V., Yansura, D. G. & Caruthers, M. H. (1977) Biochemistry 16, 1765-1772.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. f. M., van Boom, J. H. & Altona, C. (1979) Nature 281, 235-236.

Handbook of Biochemistry and Molecular Biology, Nucleic Acids, 3rd ed., Vol. I (1975) CRC Press, Cleveland, Ohio, p. 589.

Helfgott, D. C. & Kallenback, N. R. (1979) Nucleic Acids Res. 7, 1011-1017.

Hoggett, J. G. & Maass, G. (1971) Ber. Bunsenges. Phys. Chem. 75, 45-54.

Huang, C. & Hearst, J. E. (1980) Anal. Biochem. 103, 127-139.

Jain, S. C., Tsai, C.-C. & Sobell, H. M. (1977) J. Mol. Biol. 114, 317-331.

Kallenbach, N. R. (1968) J. Mol. Biol. 37, 445-466.

Kastrup, R. V., Young, M. A. & Krugh, T. R. (1978) Biochemistry 17, 4855-4865.

Khorana, H. G. (1968) Pure Appl. Chem. 17, 349-381.

Kreishman, G. P, & Chan, S. I. (1971) J. Mol. Biol. 61, 45-58.

Krugh, T. R., Laing, J. W. & Young, M. A. (1976) Biochemistry 15, 1224-1228.

Krugh, T. R. & Reinhardt, C. G. (1975) J. Mol. Biol. 97, 133-162.

Krugh, T. R., Wittlin, F. N. & Cramer, S. P. (1975) Biopolymers 14, 197-210.

Lee, C.-H. & Tincoc, I., Jr. (1978) Nature 274, 609-610.

Levine, M. D. (1974) Ph.D. thesis, University of California, Berkeley.

LePecq, J.-B. & Paoletti, C. (1967) J. Mol. Biol. 27, 87-106.

Lyubchenko, Y. L., Vologodskii, A. V. & Frank-Kamenetskii, M. D. (1978) Nature 271, 28-31.

McCann, J., Choi, E., Yamasaki, E. & Ames, B. N. (1975) Proc. Nat. Acad. Sci. 72, 5135-5139.

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

Marmur, J. & Doty, P. (1962) J. Mol. Biol. 5, 109-118.

Martin, F. H. & Tinoco, I., Jr. (1980) Nucleic Acids Res. 8, 2295-2299.

Martin, F. H., Uhlenbeck, O. C. & Doty, P. (1971) J. Mol Biol. 57, 201-215.

Nelson, J. W., Martin, F. H. & Tinoco, I., Jr. (1981) Biopolymers 20, 2509-2531.

Nelson, J. W. & Tinoco, I., Jr. (1982) Biochemistry, in press.

Pardi, A. (1980) Ph.D. Thesis, University of California, Berkeley, California.

Pardi, A., Martin, F. H. & Tinoco, I., Jr. (1981) Biochemistry 20, 3986-3996.

Pardi, A. & Tinoco, I., Jr. (1982) Biochemistry 21, 0000.

Pardi, A., Morden, K. M., Patel, D. J. & Tinoco, I., Jr. (1982) Biochemistry, in press.

Patel, D. J. & Canuel, L. L. (1976) Proc. Nat. Acad. Sci. 73, 3343-3347.

Patel, D. J. & Canuel, L. L. (1979) Eur. J. Biochem. 96, 267-276.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K. & Breslauer, Kenneth J. (1982a) Biochemistry 21, 445-451.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982b) Biocmemistry 21, 428-436.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982c) Biochemistry 21, 437-444.

Podder, S. K. (1971) Eur. J. Biochem. 22, 467-477.

Poland, D. (1974) Biopolymers 13, 1859-1871.

Pörschke, D. (1971) Biopolymers 10, 1989-2013.

Pörschke, D. & Eigen, M. (1971) J. Mol. Biol. 62, 361-381.

Pörschke, D., Uhlenbeck, O. C. & Martin, F. H. (1973) Biopolymers 12, 1313-1335.

Pörschke, D. (1974) Biophys. Chem. 2, 83-96.

Provencher, S. W. (1976a) Biophys. J. 16, 27-41.

Provencher, S. W. (1976b) J. Chem. Phys. 64, 2772-2777.

Ravetch, J., Gralla, J. & Crothers, D. M. (1974) Nucleic Acids Res. 1, 109-127.

Reinhardt, C. G. and Krugh, T. R. (1977) Biochemistry 16, 2890-2895.

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

Rigler, R., Rabl, C. & Jovin, T. M. (1974) Rev. Sci. Instrum. 45, 580-588.

Romaniuk, P. J., Hughes, D. W., Grégoire, R. J., Bell, R. A. & Neilson, T. (1979) Biochemistry 18, 5109-5116.

Schaller, H. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 401-408.

Sims, J., Koths, K. & Dressler, D. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 349-365.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77-84.

Thomas, G. & Roques, B. (1972) FEBS Letters 26, 169-175.

Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) Nature [New Biol.] 246, 40-41.

Tsai, C.-C., Jain, S. C. & Sobell, H. M. (1977) J. Mol. Biol. 114, 301-315.

Uhlenbeck, O. C., Martin, F. H. & Doty, P. (1971) J. Mol. Biol. 57, 217-229.

Van de Sande, J. H., Caruthers, M. H., Kumar, A. & Khorana, H. G. (1976) J. Biol. Chem. 251, 571-586.

Vizard, D. L., White, R. A. & Ansevin, A. T. (1978) Nature 275, 250-251.

Vologodskii, A. V. & Frank-Kamenetskii, M. D. (1978) Nucleic Acids Res. 5, 2547-2556.

Wakelin, L. P. G. & Waring, M. J. (1974) Mol. Pharmacol. 9, 544-561.

Waring, M. J. (1965) J. Mol. Biol. 13, 269-282.

Waring, M. J. (1966) Biochim. Biophys. Acta 114, 234-244.

Warshaw, M. M. (1965) Ph.D. thesis, University of California, Berkeley, California.

Wells, L. R. D., Blakesley, R. W., Hardies, S. C., Horn, G. T., Larson, J. E., Selsing, E., Burd, J. F., Chan, H. W., Dodgson, J. B., Jensen, K. F., Nes, I. F. & Wartell, R. M. (1977) CRC Crit. Rev. Biochem. 4, 305-340.

Wells, R. D., Larson, J. R., Grant, R. C., Shortle, B. R. & Cantor, C. R. (1970) J. Mol. Biol. 54, 465-497.

Wiesehahn, G., Cech, T. R. & Hearst, J. E. (1976) Biopolymers 15, 1591-1613.

Young, M. A. & Krugh, T. R. (1975) Biochemistry 14, 4841-4847.

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720