KINBERG-CALHOUN, JUDY

PHYSICAL MAPPING OF A VIRAL GENOME BY SITE SPECIFIC ENHANCEMENT AND STUDIES OF THE EFFECT OF CIRCULAR STRUCTURE AND PERMUTATION ON THE RATE OF DNA RENATURATION

City University of New York PH.D. 1980

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs ✓
2. Colored illustrations _____
3. Photographs with dark background ✓
4. Illustrations are poor copy _____
5. Print shows through as there is text on both sides of page _____
6. Indistinct, broken or small print on several pages _____ throughout _____
7. Tightly bound copy with print lost in spine _____
8. Computer printout pages with indistinct print _____
9. Page(s) _____ lacking when material received, and not available from school or author _____
10. Page(s) _____ seem to be missing in numbering only as text follows _____
11. Poor carbon copy _____
12. Not original copy, several pages with blurred type _____
13. Appendix pages are poor copy _____
14. Original copy with light type _____
15. Curling and wrinkled pages _____
16. Other _____

University Microfilms International

300 N. ZEEB RD. ANN ARBOR MI 48106 (313) 761-4700

PHYSICAL MAPPING OF A VIRAL GENOME BY SITE SPECIFIC ENHANCEMENT AND STUDIES OF THE EFFECT OF CIRCULAR STRUCTURE AND PERMUTATION ON THE RATE OF DNA RENATURATION

By

Judy Kinberg-Calhoun

A dissertation submitted to the Graduate Faculty in Biomedical Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1980

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

6/27/80
date

Chairman of Examining Committee

6/30/80
date

Executive Officer

Supervisory Committee

The City University of New York

ABSTRACT

PHYSICAL MAPPING OF A VIRAL GENOME
BY SITE SPECIFIC ENHANCEMENT
AND
STUDIES OF THE EFFECT OF CIRCULAR
STRUCTURE AND PERMUTATION ON THE
RATE OF DNA RENATURATION

by

Judy Kinberg-Calhoun

Advisor: James G. Wetmur, Ph.D.

The studies reported here exploit and explore certain of the physical properties of DNA. One set of studies is aimed at producing fine structure physical genomic maps and takes advantage of the relative instability of mismatched base pairs. The feasibility of the method is tested using a deletion mutant of bacteriophage T7, and initial experiments are performed using T7 point mutants. A separate set of studies examines the effect of excluded volume on the rate at which DNA strands renature. DNA molecules of varying shape or permutation are used to

explore this phenomenon.

For the feasibility studies of site specific enhancement mapping, heteroduplex DNA molecules were formed using T7 deletion mutant C63 and the wild type bacteriophage. These heteroduplexes contain a deletion loop of approximately 600 bases whose location is known. First, the heteroduplexes were incubated in 0.03 M chloroacetaldehyde in a solvent consisting of 2.4 M tetraethylammonium chloride (Et_4NCl), at pH 4.7. Chloroacetaldehyde reacts with and covalently modifies deoxycytosine and deoxyadenosine residues in DNA. These reactions occur much more readily with single stranded DNA than double stranded DNA. Chloroacetaldehyde will also react at or near defects in the DNA structure such as the ends, single strand breaks, and mismatch sites in DNA heteroduplexes. In 2.4 M Et_4NCl, DNA melting is independent of base composition. At pH 4.7 in this solvent, chloroacetaldehyde reacts with cytosine and adenosine residues at the same rate. Second, the modified substrates are incubated with endonuclease S1. This enzyme specifically cleaves single stranded regions of DNA. In a heteroduplex molecule of T7 containing a deletion loop of 600 bases, reaction with chloroacetaldehyde results in denaturation from the site of the deletion loop.

Endonuclease S1 is better able to recognize and cleave the DNA strand opposite the deletion loop when the heteroduplex has been pretreated with chloroacetaldehyde than in the absence of such pretreatment.

Site specific enhancement mapping was also tested by forming heteroduplex DNA molecules using the DNA from an amber mutant of bacteriophage T7 and from its spontaneously occurring revertant. The conditions of chloroacetaldehyde or S1 incubation were varied in an attempt to find conditions in which heteroduplexes containing a single base mismatch could be cleaved by endonuclease S1. We were unable to find conditions which allowed detection of any such cleavage.

The possibility of an effect of excluded volume on DNA renaturation rates was tested by examining the rate at which circularly permuted linear molecules renature with one another and by comparing the rates at which a linear DNA molecule renatures with its linear complement and its circular complement. The results of trials in which molecules ranging from 9 to 42% out-of-phase were allowed to renature before being examined in the electron microscope are consistent with there being little or no detectable effect of circular permutation on DNA renatur-

ation rates. When a single stranded linear DNA molecule was presented with two topologically distinct but otherwise identical complements, renaturation was approximately three times faster with the linear complement than with the circular one. This result is consistent with an effect of excluded volume on the rate of DNA renaturation.

Finally, an excluded volume theory has been derived which takes into account the probability of random overlap of segments of two DNA chains which are interacting to form a nucleation site. Such random overlaps decrease the probability that the nucleation configuration is allowed. The experimental results agree qualitatively and semi-quantitatively with the excluded volume theory.

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my advisor, Dr. James G. Wetmur, for his support and guidance throughout this project. He was never too busy to sit down and talk about the experiments, and our discussions always proved enlightening.

Special thanks are due to Dr. Kilbourne. His willingness to accept me as a student and his assemblage of a fine faculty have given me two of the most important things in my life -- one of which has been the opportunity to earn this degree.

DEDICATION

This thesis is dedicated to my best friend, my husband David, for his support, encouragement, and love. He always manages to keep his wits about him , while all around him I am losing mine.

This thesis is also dedicated to my wonderful parents. Their love and belief in me is a constant source of strength.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

(LIST OF FIGURES CONTINUED)

(LIST OF FIGURES CONTINUTED)

CHAPTER 1

GENERAL INTRODUCTION

The work to be presented in the following chapters consists of a group of studies designed to explore two aspects of the behavior of nucleic acids in solution. In Chapter 2 an approach is described that has been used to develop a physical technique to generate fine structure maps of a bacteriophage genome. The study is based on the thermodynamic effect of base mismatches on nucleic acid stability. In Chapter 3 the adaptation of this method for use with the RNA containing viruses is discussed. Chapters 4 and 5 contain descriptions of two effects of DNA topology on the kinetics of renaturation. In Chapter 4 the effect of circular permutation on DNA renaturation rates is discussed. Chapter 5 contains a consideration of the effect of excluded volume on DNA renaturation rates. In this introductory chapter I intend to survey each of the fields relevant to the studies to be discussed in order to provide a background for each of my projects.

The DNA in all experiments is from one of three sources: bacteriophage T7, bacteriophage ØX174, or the animal virus SV40. Phage T7 is a virulent bacteriophage that infects certain strains of Escherichia coli (E. coli) B, C, and

K12 (Studier, 1979). It was isolated independently by Demerec and Fano (1945) and by Delbrück (1946). Its polyhedral head contains a single molecule of double stranded DNA which codes for 30-35 proteins.

Bacteriophage øX174 is also a virulent phage that infects certain strains of E. coli. Infection is initiated by the injection of the single stranded circular viral DNA into the host cell. During virus replication the single stranded DNA circle is converted to a double strand form. Both kinds of DNA have been isolated. The entire genome of øX174 has been sequenced (Sanger et al., 1977, 1978).

SV40 is an animal virus belonging to the papovavirus group. It has been important experimentally as a model of eukaryotic gene organization and gene expression, as well as for its oncogenic potential. Monkey cells which are permissive for SV40 infection allow a lytic mode of viral replication. In non-permissive cells such as those from mice and rats, in which no progeny virus is produced and all cells survive infection, or in semipermissive cells such as those from humans and hamsters, in which few progeny viruses are produced and many cells survive infection, part or all of the SV40 genome can become integrated into the host cell genome and lead to transformation of the cell (Kelly and Nathans, 1977). Its double stranded circular DNA molecule

is associated with histones derived from the host cell. The entire genome has been sequenced (Reddy et al., 1978). Some of the pertinent physical characteristics of T7, ØX174, and SV40 are summarized in Table 1.

A. Mapping bacteriophage or viral genomes.

The ability to correlate a specific biological function with a unique segment in a chromosome is the definition of genome mapping. The literature on mapping is vast. Although the principles governing mapping were initially established with eukaryotic organisms, beginning with the work of T.H Morgan (1911), my work and this discussion is restricted to the mapping of bacteriophages. I will confine this section to a presentation of the two main approaches used in generating genomic maps and to a discussion of the strengths and weaknesses of each. This overview is not intended to be exhaustive. The two approaches that have been used are: genetic mapping and more recently, physical mapping.

In genetic mapping the relative order and linkage of genetic markers is deduced by observing phenotypic alterations in recombinants resulting from genetic exchange. Genetic mapping is also capable of producing fine-structure maps within genes. This is its outstanding strength and one example should suffice to illustrate the method. Benzer, as early as 1961, made a detailed genetic map of the

TABLE 1

PHYSICAL CHARACTERISTICS OF THE DNA OF ESCHERICHIA COLI PHAGES T7 AND øX174 AND ANIMAL VIRUS SV40

Characteristic	T7 (reference)	øX174 (reference)	SV40 (reference)
Number of bases or base pairs	40,000 (1)	5386 (2)	5226 (3)
Number of bases in terminal redundancy	260 (4) 70 (5,6) 150-160 (7) 159 (Dunn, cited in (7))	none	none
Topology	Linear, nonpermuted, fully double stranded (4)	Single stranded and double stranded circle (8)	Double stranded circle (9)
Percent GC	48 (10)	42 (11)	41 (12)

References for Table 1

(1) McDonell et al., 1977.
(2) Sanger et al., 1978.
(3) Reddy et al., 1978.
(4) Ritchie et al., 1967.
(5) Ludwig and Summers, 1975.
(6) Dreiseikelmann and Wackernagel, 1978.
(7) Rosenberg et al., 1979.
(8) Sinsheimer et al., 1962.
(9) Vinograd and Lebowitz, 1966.
(10) Lunan and Sinsheimer, 1956.
(11) Hayashi et al., 1963.
(12) Danna and Nathans, 1972.

bacteriophage T4 r_{11} region, which consists of two cistrons that control phage growth in E. coli strain K. He made use of point and deletion mutants. The deletion mutants were first ordered on the genome by making two factor crosses and looking for wild type recombinants. Benzer reasoned that if two deletions overlap one another on the genome, wild type recombinant progeny would never be found; if the deletions did not overlap then wild type recombinants would be produced. In this way he was able to locate a series of overlapping deletion mutants. To examine the r_{11} region in even greater detail, he crossed point mutants first with mutants having relatively large regions deleted, in order to assign the point mutant to a particular segment. Then he crossed the point mutant with mutants containing progressively smaller deletions of the appropriate region. The failure to find wild type recombinants indicated overlap between the point mutant and the deletion mutant and served to localize the point mutant to a region the size of a small deletion. Finally, using two factor crosses between pairs of point mutants assigned to the same small region, he determined their relative order. This powerful approach produced a fine structure map of the A and B cistrons of the r_{11} region of T4. Fine structure genetic maps may also be constructed using 3 factor crosses not only to

order genetic markers but also to calculate map distances between them.

The dependence on recombination inherent in genetic mapping is responsible for both the strength and the weakness of the technique. Because it depends on recombination it is ineffective for viruses which do not recombine, or those which do so only rarely. Moreover, the very ability to recombine carries with it a unique set of problems: the occurrence of multiple cross-over events and of interference. Multiple cross-over events can lead to inaccurate map distances. Genetic mapping becomes ineffective for distantly linked markers because multiple crossover events between them become frequent. Another complication arises because assumptions about the randomness of crossing over are not met. Non randomness was first observed among eukaryotes where cross-over rarely occurs near the centromere. Pairs of loci in this region show little recombination and if compared by recombination frequency appear to be closer than they actually are. One factor affecting non randomness of crossing over is the presence of recombinational hot spots in the phage genome. Just what it is that confers hot spot status remains to be defined, but it may be that the DNA sequence in such a region has a high affinity

for the breaking and joining enzymes involved in recombination. Vegetative recombination in phage lambda is under the control of two systems: the Rec system of the host and the Red system of the phage. McMilin et al. (1974) looked for recombinational hotspots by performing crosses between lambda phages that had been deleted for the lambda recombination genes. They found about a fifteenfold increase in recombination at a site 10% in from the right end of lambda. In contrast to lambda, øX174 relies on the host recombination system for most of its recombinant formation (Benbrow et al., 1974). Benbrow et al. (1974) found a hot spot for recombination in gene A of øX174. The frequency of recombination within gene A is increased by about one to two orders of magnitude in Rec^+ cells compared to Rec^- cells. The existence of the hot spot could be explained if there is a gap in this gene near the origin of replication. Another aspect of randomness is whether or not there is interference. In other words, does the occurrence of one cross-over event make it either more or less probable that a second cross-over will take place somewhere else. The observed number of double cross-overs divided by the expected number of double cross-overs is known as the ratio of coincidence, s. When exchanges have no influence on one another s should equal 1. For bacteriophage s is

typically greater than 1, indicating that the number of double exchanges is greater than expected (negative interference). Pairs of loci exhibiting negative interference seem to be closer than they actually are. To a great extent negative interference may be due to the manner in which recombinational intermediates are processed. Because of the phenomenon of interference, map distances obtained by genetic mapping are not readily converted into physical distances. Because the frequency of recombination and patterns of interference differ among bacteriophage or even in different regions of one genome, map distances are not comparable or additive between bacteriphage or within one genome.

The second class of techniques, those utilizing physical methods, avoids some of the problems of genetic mapping, but has its own limitations. Physical mapping does not depend on recombination and therefore is potentially useful for mapping the RNA viruses, and it is free from the effects of interference. An early example of the method is the heteroduplex mapping of Westmoreland et al. (1968) and Davis and Davidson (1968). In order to demonstrate that a certain b2 marker in bacteriophage lambda actually represented a deletion they denatured and renatured a mixture of DNAs from a wild type lambda phage and one carrying the b2 marker. Fifty percent of the molecules will renature with

their original complementary strand and are of no interest. However, fifty percent of the renatured molecules will be heteroduplexes, containing one strand from each parent. Westmoreland et al. (1968) examined the DNA in the electron microscope using the formamide spreading technique, and saw that the completely double stranded heteroduplex molecules were interrupted at a certain point by a single stranded loop of DNA. This loop of DNA from the wild type phage corresponded to the material that had been deleted in the lambda b2 phage. When the loop was measured, it was found to represent about 17% of the total chromosome length. This supported evidence from three factor crosses that suggested that lambda b2 had about 17% of the chromosome deleted. Furthermore, the distance of this loop from the end of the DNA could be obtained in numbers that represented real physical distance. Data of this type, because it measures real physical distance, is comparable among all organisms.

Denaturation mapping, another physical technique which also makes use of the electron microscope, takes advantage of the fact that Adenine-Thymine (A·T) nucleotide pairs are less stable that Guanine·Cytosine (G·C) pairs. Inman (1966) heated lambda DNA in the presence of formaldehyde just enough to melt the A·T pairs. He found when he looked at the DNA in the electron microscope that there were sites of denaturation that were characteristic for phage lambda The distance

of these denatured regions from the ends of the duplex and from each other could be measured in a manner analogous to that for heteroduplex mapping and a characteristic pattern determined. Again, these represent real distances along the chromosome, just as in heteroduplex mapping.

Still a third approach to physical mapping was taken by Lai and Nathans (1974). They used restriction enzyme fragments of wild type SV40 to correct various SV40 mutational defects. The restriction fragments were independently ordered on the SV40 genome. By determining which fragments could correct each of five temperature sensitive mutations carried by five different SV40 isolates the site of each mutation was localized to its own particular region of the genome. The technique is called mapping by marker rescue and has been widely used.

The one shortcoming shared by these physical mapping techniques is the degree of resolution attainable. Resolution is usually limited to regions of the DNA the size of restriction fragments as in marker rescue experiments, or the size of a loop of DNA visible in the electron microscope as in heteroduplex mapping. The aim of my work was to develop a method of mapping that would be able to generate fine structure genome maps comparable in resolution to those achieved by genetic mapping, but which did not depend on recombination. In other words, a physical

mapping technique that combined the best features of physical and genetic mapping. The technique involves a chemical modification of heteroduplex DNA molecules and is discussed in detail in Chapters 2 and 3.

B. The excluded volume effect in DNA renaturation.

DNA renaturation consists of formation of double stranded helical DNA molecules from complementary single strands of DNA. It is a second order reaction with respect to the concentration of complementary DNA strands. The renaturation rate depends on the complexity (N) and length (L) of the DNA molecules as well as on the temperature, ionic strength, pH, and viscosity of their environment. Complexity refers to the total number of bases in non-repeating sequence. One of the most important applications of DNA renaturation kinetics has been to determine the complexity of nucleic acids.

The mechanism of DNA renaturation has been reviewed by Wetmur (1976). The renaturation process consists of a slow step followed by a rapid reaction that produces the final duplex molecule. The slow, rate limiting step is a nucleation event - the formation of two to three base pairs in register. Following nucleation the rest of the bases pair very rapidly along the entire length of the strand in what has been termed a zippering reaction. At a constant

DNA nucleotide concentration, the greater the DNA complexity, the lower the concentration of a particular nucleation site. The second order rate constant, k_2, should therefore be inversely proportional to N. Because the rapid zippering of the molecule encompasses the entire length of the DNA, k_2 should also be directly proportional to L_s, the length of the shorter DNA strand in the reaction. Based on these assumptions, Wetmur and Davidson (1968) predicted that

$$k_2 = k_N \frac{L_s}{N} \tag{1}$$

where k_N is the nucleation rate constant, which depends on environmental factors. When DNA that has been hydrolyzed, sheared, or sonicated to create fragments of different length is renatured, the second order renaturation rate constant is found to obey the relationship:

$$k_2 = k'_N \frac{L_s^{0.5}}{N} \tag{2}$$

where k'_N is a length independent nucleation rate constant. The dependence of the renaturation rate constant on the square root of the length of the DNA was an unexpected finding. In order to explain it, Wetmur and Davidson (1968) postulated that the excluded volume of the DNA polymers affected their rate of renaturation. Physical excluded volume refers to the fact that no two elements either on

the same or different molecules may occupy the same space at the same time. The existence of a steric hinderance effect could limit access to some parts of molecules and thereby render those portions unavailable for reaction. Flory (1949) described the property of thermodynamic excluded volume and presented an estimate of its influence on chain configurations in solution. In all solvents in which a polymer is soluble, except for the ideal solvent (at the theta temperature), the thermodynamic excluded volume will be the same order of magnitude as the physical excluded volume. In terms of the reacting single strands of DNA under consideration in this work, it means that only those elements of the molecule which are on or close to the surface of the random DNA coil at any one time are available to participate in a nucleation event. This could explain why the renaturation rate constant is proportional not to the full length of a DNA strand but to a somewhat shorter length, specifically the square root of the length of the DNA random coil. No one has been able to assign a numerical value to the magnitude of the effect of excluded volume on polymer interactions, and the present work will not attempt to do so either. A number of workers have developed various theoretical models of the kinetic excluded volume effect (Wetmur, 1971; Morawetz et al., 1973; Cho and

Morawetz, 1973; Horie and Mita, 1978) with varying degrees of success. In spite of all this there has never been any satisfactory physical evidence for the effect of excluded volume on DNA renaturation rates. The work described in Chapter 5 provides that evidence. In Chapter 4 experiments are presented which demonstrate that circular permutation of DNA does not affect the rate of renaturation.

CHAPTER 2

PHYSICAL MAPPING: SITE SPECIFIC ENHANCEMENT WITH DELETION MUTANTS

INTRODUCTION

Recent progress has been made in adapting physical mapping techniques for use with segments of DNA too small to be seen as loops in the electron microscope (less than 300 bases). Shenk et al., (1975a,b) used the single strand specific endonuclease S1 from Aspergillus oryzae to locate a deletion of about 190 base pairs and an insertion of about 50 base pairs in heteroduplexes of SV40 DNA molecules first made linear by cleavage at their single Eco R1 site. Densitometer tracings of these samples after agarose gel electrophoresis clearly showed two predominant fragments with the expected mobilities. A mixture of each parental homoduplex denatured and renatured by itself did not yield detectable quantities of such fragments after S1 digestion. However, when the same approach was used to cleave an SV40 heteroduplex molecule containing a much smaller mismatch, presumably consisting of a single base pair, the noise level due to non-specific cleavage by S1 was very close to the level of correct cleavage products.

Several laboratories have addressed the question of

the smallest single strand DNA region the S1 nuclease will attack. Wiegand et al. (1975) showed that as few as 12 bases could serve as a substrate as judged by the inability of phage lambda DNA to cyclize after S1 digestion. They also demonstrated, as did Shishido and Ando (1975), that S1 can cleave DNA opposite a nick. Both groups used double stranded DNA from phage T5 which contains 4 major naturally occurring single strand nicks (Johnston et al., 1977). The pattern obtained in neutral sucrose gradients of T5 DNA before and after incubation with S1 were consistant with what would be expected if S1 cleaved opposite some of the nicks. A nick renders DNA more susceptible to S1 cleavage than does a base pair mismatch. The evidence for this is that S1 cuts at nicks even when they occur in a G·C-rich region (Nichols and Donelson, 1977) while no significant cutting above background seemed to occur in the single base pair mismatch SV40 heteroduplexes used by Shenk et al. (1975a,b), with one exception. In the one heteroduplex where greater cleavage did appear to take place, it is quite possible that the mismatch occurred in a relatively A·T-rich region on the DNA.

Site specific enhancement mapping, the approach we have developed for mapping a viral genome discussed here and in Chapter 3, should possess the ability to map alter-

ations in DNA as small as a single base pair regardless of whether or not they occur in an A·T-rich region. Its strategy is to first modify the region containing the aberant base pair so that it is made easily susceptible to endonucleolytic cleavage. Once this cleavage has been achieved, restriction enzyme digestion and analysis by agarose gel electrophoresis would permit an accurate determination of the site of the original base pair mismatch.

The agent of the DNA modification is chloroacetyldehyde. Barrio et al. (1972) showed that chloroacetaldehyde reacts with both adenosine and cytidine at low pH and hence possesses the capacity to modify all base pairs. This modification occurs much more readily with unpaired than with paired bases. In chloroacetaldehyde modification a five membered ring is covalently added to deoxyadenosine and deoxycytidine (Figure 1) in such a way that the product of the reaction can no longer base pair. Unlike the interaction of formaldehyde with DNA (Luskashin et al., 1976; von Hippel and Wong, 1971; McGhee and von Hippel, 1977a,b), the reaction with chloroacetaldehyde is irreversible.

In order to assure that all bases are equally susceptible to modification by chloroacetaldehyde, conditions were sought that eliminated the difference in melting temperature (Tm) between A·T and G·C pairs. Melchior and

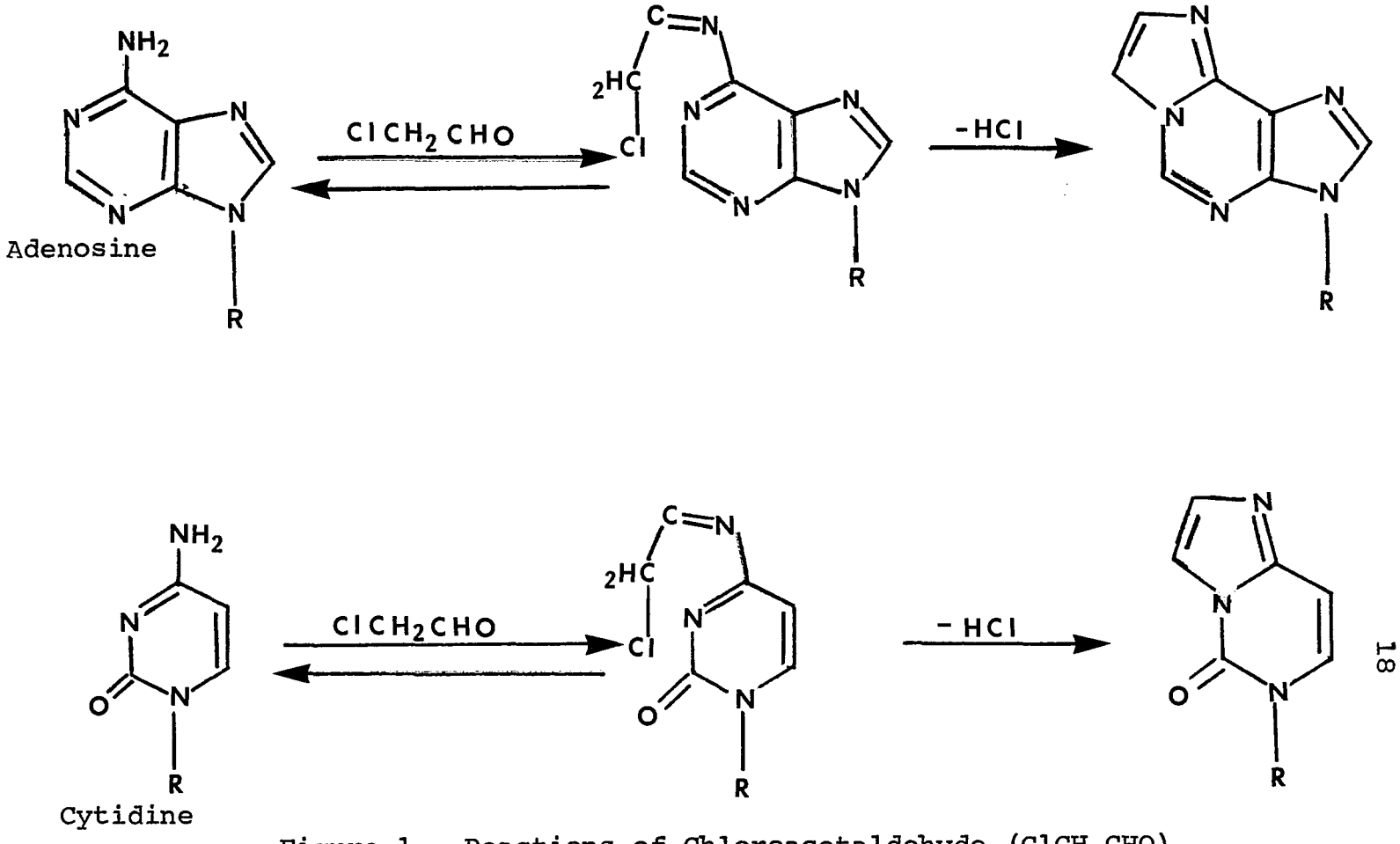


Figure 1. Reactions of Chloroacetaldehyde ($ClCH_2CHO$)

von Hippel (1973) developed an aqueous solvent system that reduced the differential stability of dA·dT and dG·dC pairs without affecting the structure and stability of the native DNA double helix at room temperature. They found that in a solution of 2.42 M tetraethylammonium chloride (Et_4NCl), DNAs ranging in content from 42% to 72% G·C had superimposible melting profiles with a very narrow transition width. The interpretation of this observation is that in the presence of 2.42 M Et_4NCl, A·T base pairs are as thermally stable as G·C pairs. In this solvent the effect of neighboring base pairs on mismatch stabilities would be eliminated.

Ruyechan (1976) showed that at pH 4.74 in 2.4 M Et_4NCl, the rate of the reaction of each of the two DNA bases with chloroacetaldehyde is equivalent. The energy of activation for the reaction is 23.4 kcal/mol for adenosine and 21.2 kcal/mol for cytidine (Ruyechan and Wetmur, personal communication). Because every base pair contains either cytosine or adenine, chloroacetaldehyde in 2.4 Et_4NCl and over a wide temperature range is a base composition independent reagent for probing for defects in native DNA structure. We will show that DNA heteroduplex molecules containing a mismatched region can be chemically modified by chloroacetaldehyde in the presence of 2.4 M Et_4NCl and that such molecules may be more efficiently cleaved by endonuclease

S1 than molecules which have not been so modified. This method has the potential for ultimately rendering single base pair mismatches, described in Chapter 3, as sensitive to S1 cleavage as deletion or insertion loops. The location of a mismatch could then be determined by restriction enzyme analysis.

MATERIALS AND METHODS

Bacterial and phage strains. E. coli B (sus^-) is used as the permissive host for T7 deletion mutant C63 and T7 revertants of amber mutants. It was obtained from C. Wu.

Bacteriophage T7 deletion mutant C63 and the wild type T7 were a generous gift from F.W. Studier. The wild type phage was obtained by Studier in 1961 from Sinsheimer who had obtained it from Delbrück. Bacteriophage T7 deletion mutant C63 is missing a 604 base pair segment extending from 6.1 map units to 7.61 map units at the left end of the genome (Studier et al., 1979b).

T7 4R was isolated as a spontaneously occurring revertant of T7 amber mutant 4, isolated in this laboratory by JKC.

Buffers, media, enzymes.

Buffers

phage buffer: 0.15 M NaCl, 0.001 M $MgCl_2$, 0.01 M Tris-HCl (pH 7.7)

DNA buffer:	0.15 M NaCl, 0.001 M Tris-HCl (pH 7.8), 2 x 10^{-3} M Ethylene-diamine-tetraacetate (EDTA)
S1 digestion buffer:	0.03 M Sodium Acetate pH 4.6, 0.05 M NaCl; 1 x 10^{-3} $ZnCl_2$ added prior to digestion
Mbo I digestion buffer:	0.075 M NaCl, 0.01 M Tris-HCl (pH 7.4), 0.01 M $MgCl_2$; 0.001 M dithiothreitol, 100 µg Bovine Serum Albumen (BSA)/ml added prior to digestion.
Hpa I digestion buffer:	0.02 M Tris-HCl (pH 7.4), 0.01 M $MgCl_2$, 0.001 M dithiothreitol, 0.006 M KCl
Electrophoresis buffer:	0.04 M Tris-Acetate (pH 7.9), 0.005 M Sodium Acetate, 0.001 M EDTA, 0.5 µg Ethidium Bromide/ml

Media

M9 minimal medium:	1g/l NH_4CL, 3g/l KH_2PO_4, 6g/l Na_2HPO_4, 4g/l glucose, 1 ml/l 1 M $MgSO_4$; 2 µg/ml sterile thymidine added after autoclaving the medium
LB rich medium:	10g/l Bactotyrptone, 5g/l Yeast extract, 10g/l NaCl, 0.1% glucose,

pH adjusted to 7.0 with 1N NaOH

B2 low phosphate medium: 4.2 g/l bis (2-Hydroxyethyl) imino-tris (Hydroxy-methyl) methane, 1 g/l NH_4Cl, 2.5 g/l NaCl, 1.5 g/l KCl , 1 ml/l 1 M $MgSO_4$, 1 ml/l 0.001 M $FeCl_3$, concentrated HCl added to adjust pH to 7.0. After autoclaving, 4 g/l sterile glucose and 2.5 ml/l 0.064 M sterile phosphate buffer are added.

Enzymes

Endonuclease S1: 5000 units/ml purchased from Sigma. One unit is defined as the amount that causes 1 µg of DNA to become perchloric acid soluble per minute at pH 4.6 and 37°C.

Restriction enzyme Mbo I: 1800 units/ml purchased from New England Biolabs. One unit is defined as the amount required to digest 1 µg of lambda DNA in 15 minutes at 37°C in a total reaction mixture of 0.05 ml

Restriction enzyme Hpa I: 800 units/ml purchased from Bethesda Research Labs. One unit is defined as the amount of enzyme required to completely digest 1 μg of lambda DNA or equivalent in one hour at 37°C in a total reaction mixture of 0.05 ml.

Preparation of bacteriophage stock. E. coli B was grown at 30°C in a water bath shaker in LB medium. At 2 x 10^8 cells/ml T7 phage were added at a multiplicity of infection (MOI) of 2-4. The flask was shaken at 30°C until lysis occurred, usually within 2-4 hours. When lysis appeared complete, NaCl was added to a concentration of 1 M, and the lysate was centrifuged for 10 minutes at 7000 rpm to remove cell debris. The supernatant was centrifuged for three hours at 17,000 rpm to pellet the phage. The pellet was resuspended in phage buffer and purified in a CsCl buoyant density gradient (final density = 1.5 g/ml), by centrifugation at 40,000 rpm for 16 hours.

^{32}P labeling of T7 in vivo. E. coli B was labeled according to the procedure of McDonell et al. (1977). Cells were grown in a water bath shaker at 30°C in low phosphate

B2 medium. Exponentially growing cells were diluted 4-fold into fresh B2 medium and $^{32}PO_4$ (ICN Pharmaceuticals) was added to give a final concentration of 10-25 μCi/ml. When the cells reached a concentration of 3-5 x 10^8/ml, 5 x 10^7 purified T7 phage particles/ml were added and the culture was shaken until lysis was complete. Purification was the same as for phage stock.

Heteroduplex formation (wt/C63). Aliquots from wild type (wt) and C63 phage stocks were mixed together in ratios ranging from 1:1 to 1:25 depending on the experiment, in 0.01 M EDTA, 0.3% Sodium Dodecyl Sulfate (SDS). To this, 0.1 volume of 1 M NaOH was added to release and denature the DNA. After incubation for 5 minutes at 37°C, the solution was neutralized by the addition of 0.1 volume 2 M Tris-HCl and adjusted to 0.4 M NaCl. Renaturation was for 5 minutes at 70°C.

The DNA was phenol extracted by the procedure of Mandell and Hershey (1960) followed by 5 chloroform extractions. Phenol extraction was necessary in order to remove an exonuclease which was not inactivated during the DNA extraction. This empirically developed procedure permitted the recovery of intact DNA. The extracted DNA was dialyzed into DNA buffer.

Determination of DNA integrity after alkali treatment. A Beckman Model E Analytical Ultracentrifuge equipped with ultraviolet optics was used to determine the integrity of T7 DNA during and after alkali treatment. 70 μl of a solution containing 1.5 μg of T7 wt DNA was loaded into the well of a banding forming type III charcoal-filled Epon centerpiece of a 30 mm cell. Sedimentation was in alkaline cesium chloride (3 M CsCl, 0.1 M NaOH, pH 13). Photographic negatives taken during centrifugation were scanned on a Canalco Model J Densitomoter.

Preparation of Et_4NCl. Et_4NCl was purified by treating concentrated aqueous solutions of Et_4NCl with Norit A, Carbon Decolorizing, Neutral (Fisher) followed by filtration. Filtration was repeated until no more Norit A was deposited on the filter paper. The molarity of the resulting Et_4NCl solution was determined using the formula (Chang et al., 1974):

$$M_{(Et_4NCl)} = \frac{\eta Et_4NCl \text{ sample} - \eta H_2O}{0.033} \qquad (3)$$

Stock solutions were made between 3.0 M and 3.4 M and were stored at room temperature.

Chemical modification with Chloroacetaldehyde. Heteroduplex DNA (A_{260}=1) was mixed with Et_4NCl (final concentration

2.4 M), 0.1 M acetate buffer pH 4.7 and chloroacetaldehyde (final concentration = 0.03 M) Incubation, unless otherwise noted, was for 10 minutes at 55°C. The reaction was stopped by quenching in ice. The DNA was dialyzed overnight into S1 buffer without zinc.

The melting curve for the reaction of DNA with chloroacetaldehyde in the presence of 2.4 M Et_4NCl was obtained using a Gilford Model 2400 spectrophotometer with a Tamson constant temperature circulator, and attached automatic temperature programmer from NesLab Instruments, Inc. The cuvette temperature, the reference line maintained by an automatic reference compensator, and the increase in A_{260} as the sample was heated were recorded.

Endonuclease S1 Digestion. $ZnCl_2$ was added to the DNA solution to a final concentration of 2×10^{-3} M. S1 enzyme from 0 units/μg DNA to 10 units/μg DNA was added and incubation was for 7 minutes at 37°C. The reaction was stopped by the addition of 0.05 volume 1 M Tris base, 0.1 M EDTA.

Restriction enzyme digestion. Restriction endonuclease Mbo I buffer or restriction enzyme Hpa I buffer was added to the DNA solution and 0.5 units/μg DNA of restriction enzyme Mbo I or 2 units/μg DNA of Hpa I were

added. Incubation was for two hours at 37°C.

Agarose Gel Electrophoresis. DNA samples were adjusted to 10% Ficoll and 5µl of 0.125% bromophenol blue solution was added to help follow the loading of the DNA into the gel slots. Vertical slab gels (Hoeffer) were used for electrophoresis. The gel concentration was 1.2% agarose (Sigma or SeaKem) with 0.5µg Ethidium Bromide/ml, and electrophoresis was for 14 to 16 hours at 40 volts.

Autoradiography. After electrophoresis the gels were dried in a Biorad gel dryer and exposed at 4°C using X-Omat R Film, XR-1 (Kodak) X-Ray film and Rarex B aluminized-polyester X-Ray intensifying screens (GAF Corporation).

Nuclear Magnetic Resonance (NMR). The nuclear magnetic resonance spectrum was kindly obtained by H. Wyssbrod on the consortium-operated Varian HR-220 field-sweep spectrometer located at the Rockefeller University. Tetrasilyl Phosphate (TSP) was used as an internal reference.

RESULTS AND DISCUSSION

Characterization of the DNA. The DNA from phage grown as described in the Materials and Methods section and banded in a CsCl buoyant density gradient was subjected to digestion with restriction enzyme Hpa I or Mbo I (Dpn II). Dpn II and

and Mbo I are isoschizomers (McDonell et al., 1977). The restriction pattern of T7 DNA with each of these enzymes is known (McDonell et al., 1977), and was used to verify that our samples were T7 DNA. Figure 2 compares (A) the Mbo I restriction pattern of T7 wt and C63 DNA and (B) the Hpa I restriction pattern of T7 wt and C63 DNA. Figure 3 shows restriction digestion maps of T7 DNA for Mbo I and Hpa I as determined by McDonell et al., (1977). Table 2 gives the number of base pairs in the various restriction fragments as determined by McDonell et al.,(1977). Mutant C63 is missing a 604 base pair region extending from 6.1 map units to 7.61 map units on the T7 DNA (Studier et al., 1979b). Reference to Figure 3 will show that cleavage of this mutant DNA by restriction enzyme Mbo I affects only fragment B, which will migrate slightly ahead of wt fragment B during agarose gel electrophoresis. Cleavage of C63 DNA with restriction enzyme Hpa I results in the loss of fragments F and H since the deletion removes the Hpa I recognition site that lies between them. Loss of both fragments F and H may be detected on an agarose gel. Loss of fragments F and H is accompanied by the appearance of a new band

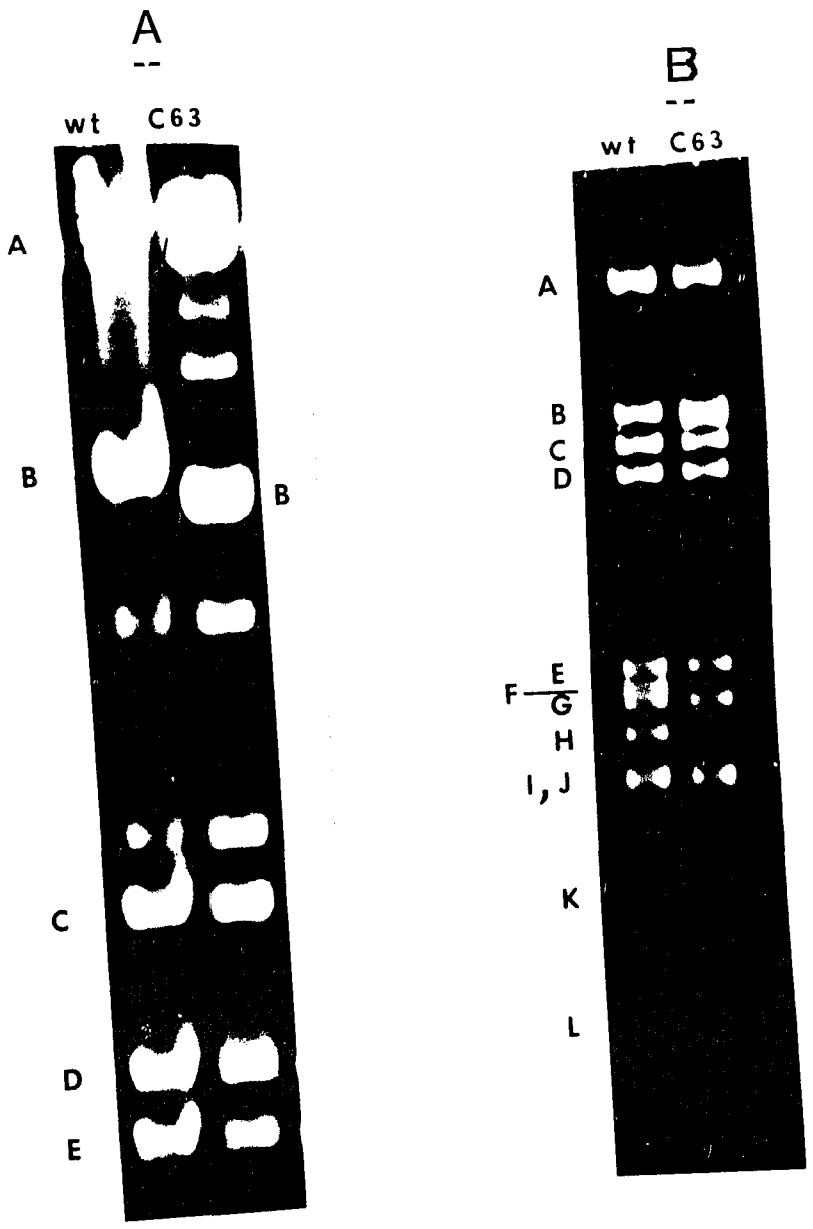


Figure 2. Agarose gel electrophoresis of bacteriophage T7 wt and C63 DNA digested by two different restriction enzymes.

A. Restriction enzyme Mbo I. Unlabeled bands are partial digestion products.

B. Restriction enzyme Hpa I.

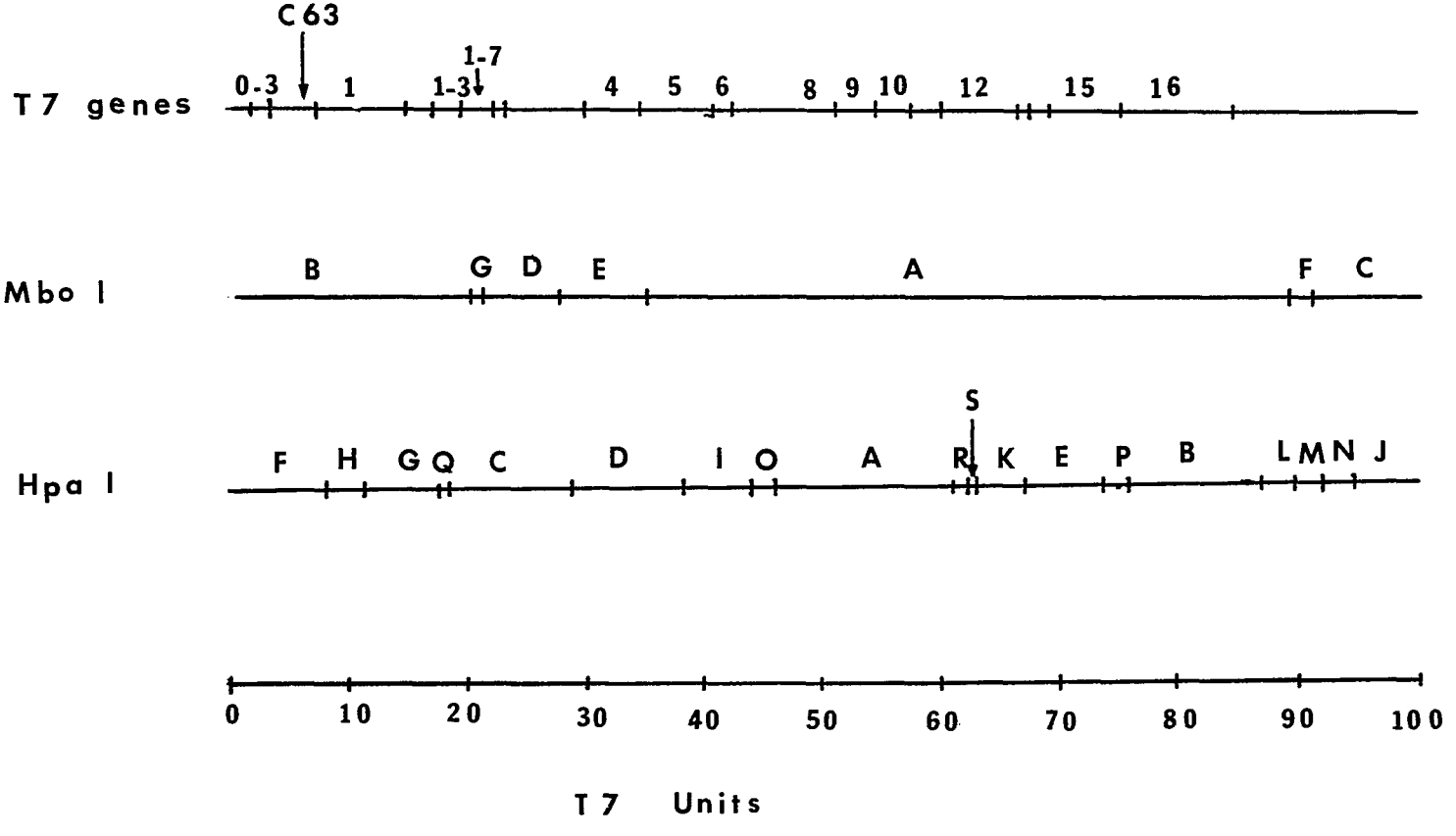


Figure 3. Mbo I and Hpa I Restriction Maps of Bacteriophage T7. Adapted from McDonell et al., 1977.

TABLE 2

APPROXIMATE NUMBER OF BASE PAIRS OF T7 DNA RESTRICTION FRAGMENTS

Fragment	Hpa I Base Pairs	Fragment	Mbo I (Dpn II) Base Pairs
A	5956	A	21,428
B	4368	B	8268
C	4072	C	3820
D	3820	D	3116
E	2640	E	2840
F	2500	F	408
G	2464	G	120
H	2312	Total	40,000
I	2144		
J	2116		
K	1756		
L	1384		
M	996		
N	892		
O	840		
P	604		
Q	440		
R	412		
S	284		
Total	40,000		

Adapted from McDonell *et al.*, 1977.

consisting of F joined to H. This fragment has a mobility close to that of band B and is detected as a widening of band B. The agarose gel patterns in Figure 2 are consistent with the known T7 restriction patterns for wt and C63 DNAs.

The method used for heteroduplex formation is alkaline denaturation (pH 13) of a mixture of the two sources of DNA, followed by renaturation. Alkaline denaturation will not affect intact DNA and can in fact reveal any pre-existing DNA depurination sites. Thermal denaturation, in contrast, is a more harsh treatment which could depurinate some sites in the DNA, rendering the molecule susceptible to breakage in subsequent handling. In order to detect any damage in the DNA, a sample of a phage preparation was subjected to velocity sedimentation in alkaline CsCl in the Model E Analytical Ultracentrifuge. Figure 4 is a densitometer tracing of T7 wt DNA in alkaline CsCl. The alkaline sedimentation coefficient determined from this data was $S_{20, w} = 39.3s$, which compares favorably with the published value of 37.2s (Studier, 1965) for T7 DNA. DNA molecules containing

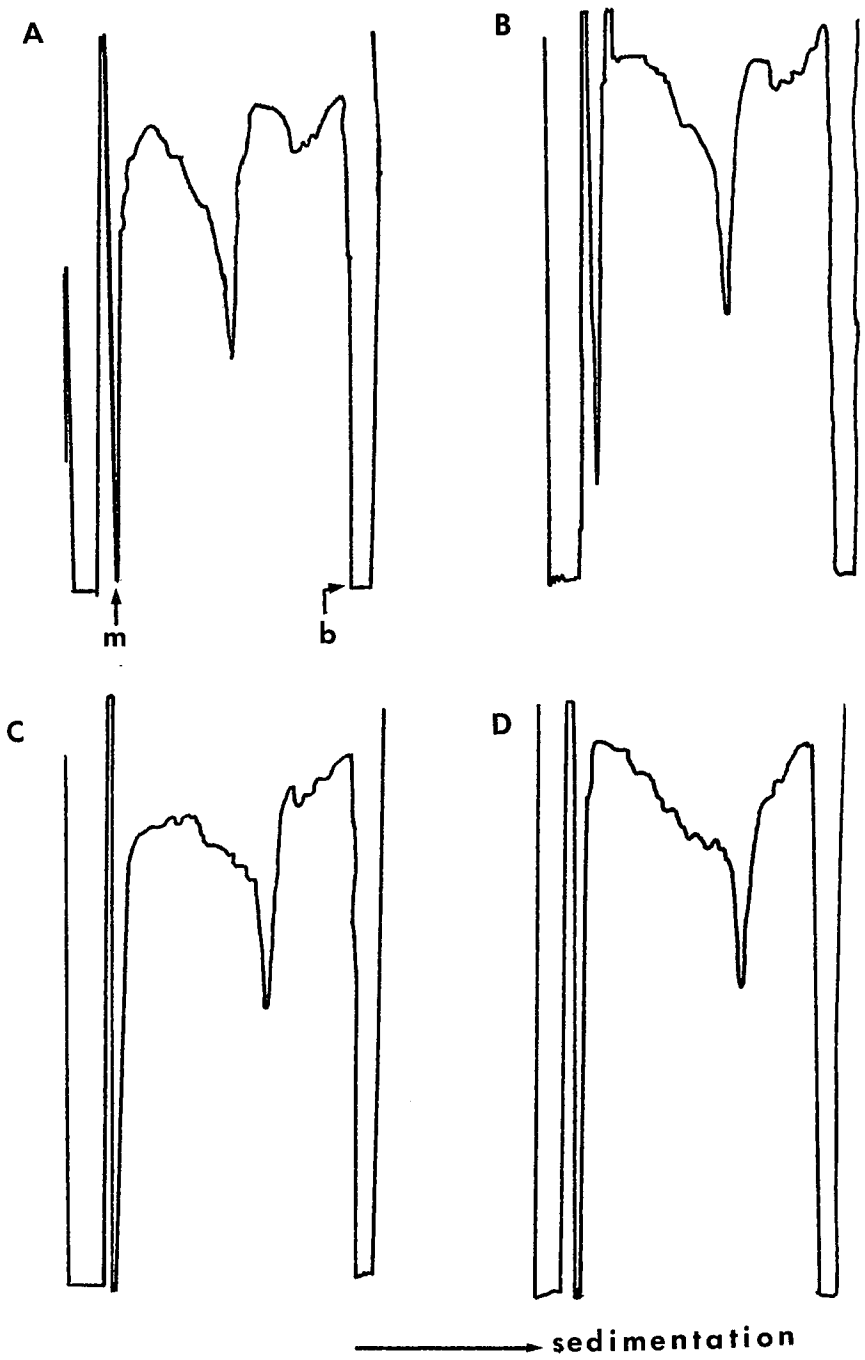


Figure 4. Densitometer tracings of an analytical velocity sedimentation of T7 wt DNA in alkaline CsCl. m = meniscus, b = bottom.

depurination sites or single strand breaks will interfere with the detection of site specific modification of DNA. The bases around a break may be even more susceptible to chloroacetaldehyde modification and to S1 cleavage than mismatched bases. The mismatched bases in heteroduplexes are at the site of the mutation in presumably one location. The single strand breaks or depurination sites are accidental and random, and will make the detection of specific mismatches more difficult. It is important therefore, to work with undamaged DNA. As shown in Figure 4, most of the DNA appears to sediment as a homogeneous species. A diffuse band would have indicated the existence of a population of different length molecules. The small amount of damaged DNA in the sample shown in Figure 4 appears as a shoulder on the main DNA band. This sample is representative of the DNA used in the experiments described in this chapter.

Characterization of the reagents. Chloroacetaldehyde could be subject to polymerization or oxidation to chloroacetic acid. Either reaction leads to the loss of the aldehyde group and prevents the reaction with DNA from occuring. Any loss of the aldehyde group may be detected by proton NMR. Figure 5 is an NMR spectrum of the starting material. Numerical data related to this spectrum is

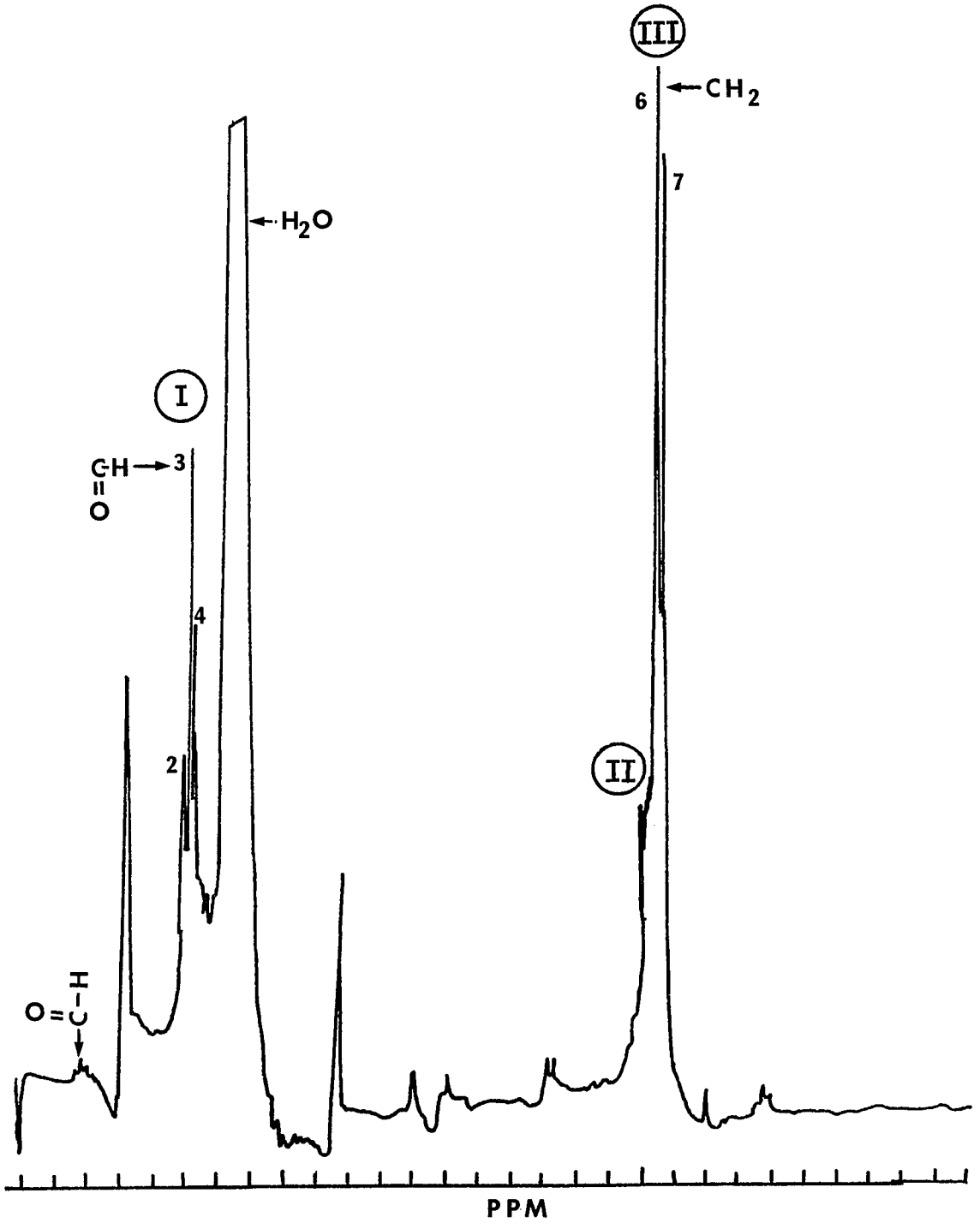


Figure 5. Proton NMR of chloroacetaldehyde.

given in Table 3.

TABLE 3

ANALYSIS OF NMR OF CHLOROACETALDEHYDE

CHEMICAL SHIFTS

Group	Height	δ (PPM)
$ClCH_2$	711	3.174
CHO	204	4.817
	445	4.797
	271	4.777
reference: TSP	---	9.183

INTEGRATIONS

Group	Area	δ (PPM) from	to
CHO	1.0 (set)	4.747	4.847
$ClCH_2$	2.5	3.074	3.204
$ClCH_2$ shoulder	1.3	3.204	3.530

The integrated value of the aldehyde peak (I) was set equal to 1.0. The integrated value for the methylene group (III) was determined only under the peaks labeled 6 and 7 and was found to be 2.5. This value is consistent with the structure of chloroacetaldehyde. The only other non-solvent resonances observed were a shoulder with small peak (II) on the

methylene resonance which may be due to chloroacetic acid, and two very small high field patterns which could be due to a contamination by ethanol or to unreacted acid aldehyde. Neither of these contaminants is present in sufficient quantity to affect the reaction of chloroacetaldehyde with DNA.

Optically pure Et_4NCl was prepared as described in Materials and Methods. 2.4 M Et_4NCl is a base composition independent solvent.

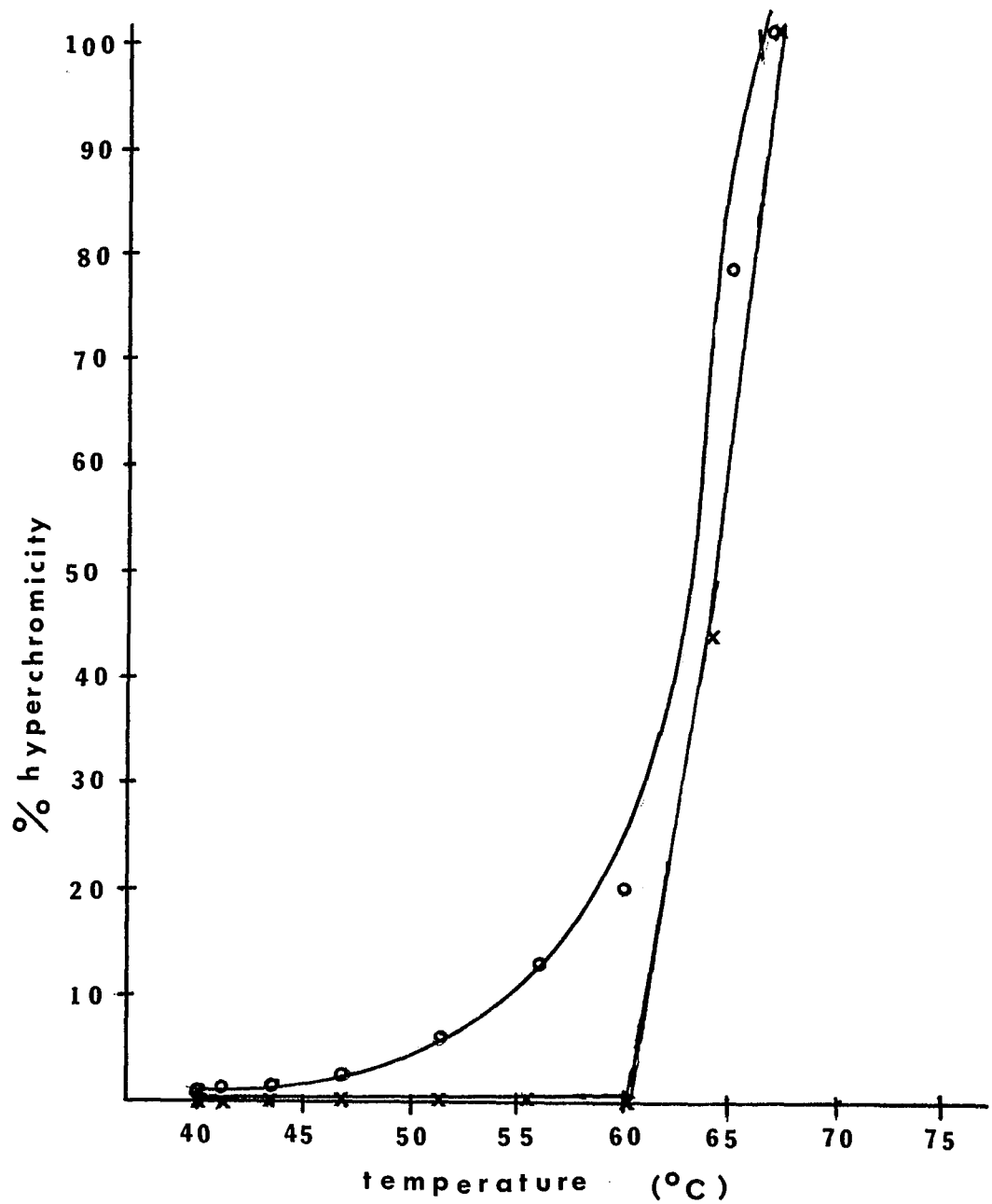
Reaction of DNA with chloroacetaldehyde in 2.4 M Et_4NCl. Figure 6 shows a melting curve of Calf Thymus DNA in 2.4 M Et_4NCl at pH 4.7. The melting temperature of 63°C is characteristic of all DNAs in this solvent (Melchior and Von Hippel, 1973). The very narrow transition width for this DNA of heterogeneous base composition confirms the fact that in this solvent A·T and G·C pairs are equal in thermal stability. Figure 6 also shows a melting curve for Calf Thymus DNA in 2.4 M Et_4NCl pH 4.7 in the presence of 0.05 M chloroacetaldehyde. This pH was chosen because chloroacetaldehyde reacts with A and C bases at the same rate in this solvent (Ruyechan, 1976). The melting curves were carried out at high DNA concentration in 1mm path length cells where the absorbance of either solution without DNA does not significantly increase during a similar

melting experiment. The results indicate that reaction of chloroacetaldehyde with DNA occurs most readily if the DNA is denatured. Below 63°C a small amount of reaction (most likely at the ends) is seen, consistent with breathing of the DNA. The melting temperature of the DNA is not seriously affected by the presence of the chloroacetaldehyde.

The bases at the ends of DNA molecules and those at either end of a defect within the DNA molecule are more susceptible to thermal perturbation than base pairs within the body of the DNA molecule (Johnston et al., 1977). Gently heating the DNA will tend to unpair these end bases first. The results in Figure 6 indicate that chloroacetaldehyde reaction at defects in DNA in 2.4 M Et_4NCl at pH 4.7 may be achievable without seriously affecting the integrity of the rest of the molecule. Temperatures at which the premelting absorbance increase occurs in Figure 6 are good candidates for the temperatures at which selective and base composition independent modification of specific heteroduplex mismatches may occur.

S1 cleavage of a deletion loop. Chloroacetaldehyde can be used to enlarge a mismatch region in conjunction with Et_4NCl. One goal of this work is to detect site specific

Figure 6. Calf thymus DNA in 2.4 M Et_4NCl, pH 4.7. In the absence of chloroacetaldehyde (x) and in 0.05 M chloroacetaldehyde (o). The temperature was increased 0.8 ^{o}C per minute.


100
90
80
70
60
50
40
30
20
10
% hyperchromicity
40
45
50
55
60
65
70
75
temperature (°C)

modifications by obtaining endonuclease S1 cleavage at the site of modification followed by determination of the molecular weight of the fragments produced. A 2.4 M Et_4NCl solution of DNA containing mismatches cannot simply be heated to open up the mismatch and cut with S1. Et_4NCl is a protein denaturant and even concentrations as low as 0.36 M will inactivate endonuclease S1 (Ruyechan, 1976). Therefore, Et_4NCl must be removed by dialysis before S1 can be used. To determine the ability of endonuclease S1 to make a double stranded cut in the region of a deletion loop, and to show that chloroacetaldehyde modification of DNA does not interfere with this reaction, the following experiment was performed, as outlined in Figure 7. A heteroduplex of T7 DNA (^{32}PC63/wt) was incubated with or without chloroacetaldehyde followed by incubation with or without S1. All samples were digested with restriction enzyme Mbo I and subjected to agarose gel electrophoresis. An autoradiogram made after electrophoresis is shown in Figure 8. The amount of DNA needed to detect the proportion of the population of heteroduplexes that were cleaved by S1 results in overloading of the high molecular weight fragments on the agarose gel. For this reason, in this print and this print only, the film was dodged in order to be able to display the high molecular weight fragments

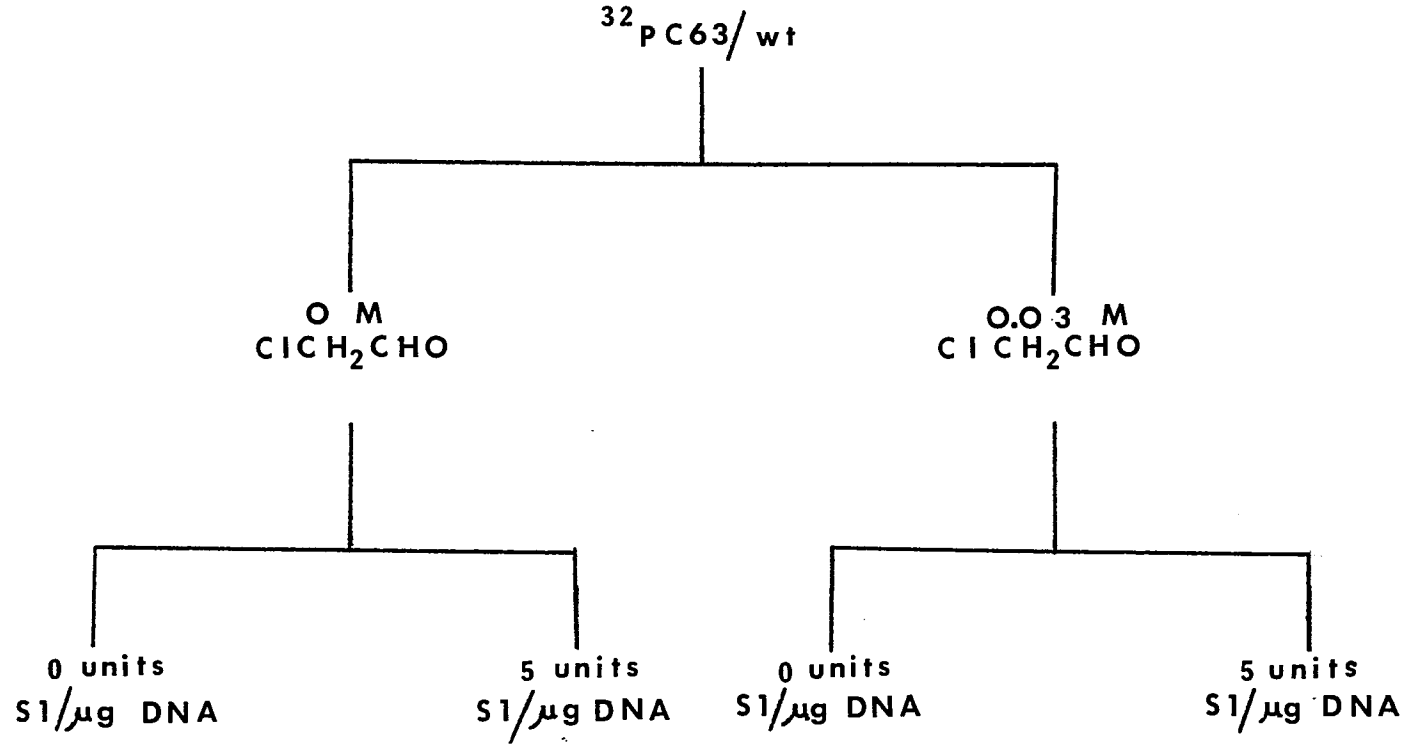
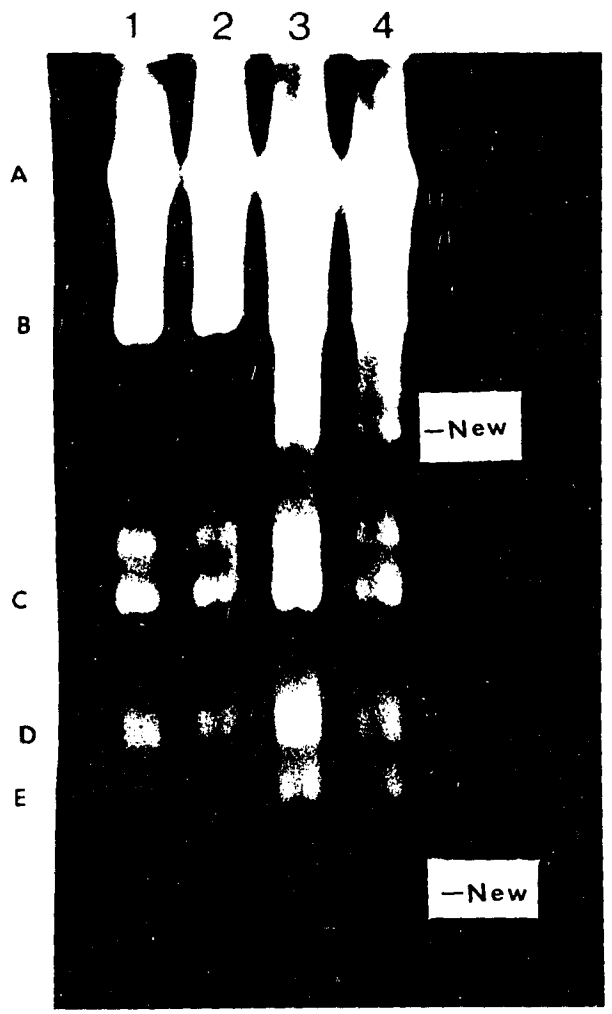


Figure 7. Protocol for determining the ability of endonuclease S1 to cleave a chloroacetaldehyde-modified deletion loop.

as well as the low molecular weight ones. That is, different regions of the print were exposed to different amounts of light. The intensities of the spots in the print do not, therefore, reflect intensities in the negative in a simple way. Only the seven usual Mbo I bands are seen when unmodified (Lane 1) or modified (Lane 2) heteroduplex DNA was cut with restriction enzyme in the absence of prior treatment with S1. This demonstrates that in the absence of S1 no cleavage occurs at the deletion loop. This is in contrast to the case where unmodified (Lane 3) or modified (Lane 4) heteroduplex DNA is incubated with 5 units S1/μg DNA prior to restriction enzyme digestion. After agarose gel electrophoresis two new bands are found. The distance migrated by each fragment resulting from Mbo I digestion of heteroduplex DNA (C63/wt) is plotted as a function of the logarithm of its molecular weight in Figure 9a. The size of the larger new fragment was estimated by graphical interpolation and the size of the smaller new fragment by extrapolation, in each case using data obtained by using the five other Mbo I fragments as internal "standards". Figure 9b is a schematic representation of the generation of the two new fragments. These results show that the new bands have the molecular weight predicted for the products of cleavage at the site of the deletion loop.

Figure 8. Autoradiogram of the cleavage of a deletion loop by endonuclease S1. ^{32}PC63/wt heteroduplex DNA. Specific activity of starting heteroduplex was 2×10^4 CPM/μg DNA. No chloroacetaldehyde, no S1 (Lane 1); 0.03 M chloroacetaldehyde, no S1 (Lane 2); no chloroacetaldehyde, 5 units S1/μgDNA (Lane 3); 0.03 M chloroacetaldehyde, 5 units S1/μg DNA (Lane 4). 4×10^4 CPM per Lane. 1.2 % agarose gel, electrophoresed 17 hours at 40 volts. Autoradiogram exposed for 8 hours at 4°C.


1
2
3
4
A
B
C
D
E
—New
—New

Any small differences between predicted and observed fragment sizes may be due to non-linearity of the graph in Figure 9a, or to the fact that after S1 cleaves the DNA it may nibble on the ends of the double stranded molecule (Shenk et al., 1975b). Clearly though, these fragments correspond in size to DNA from Mbo I fragment B with a deletion loop in the position described for mutant C63 (see Figure 9b).

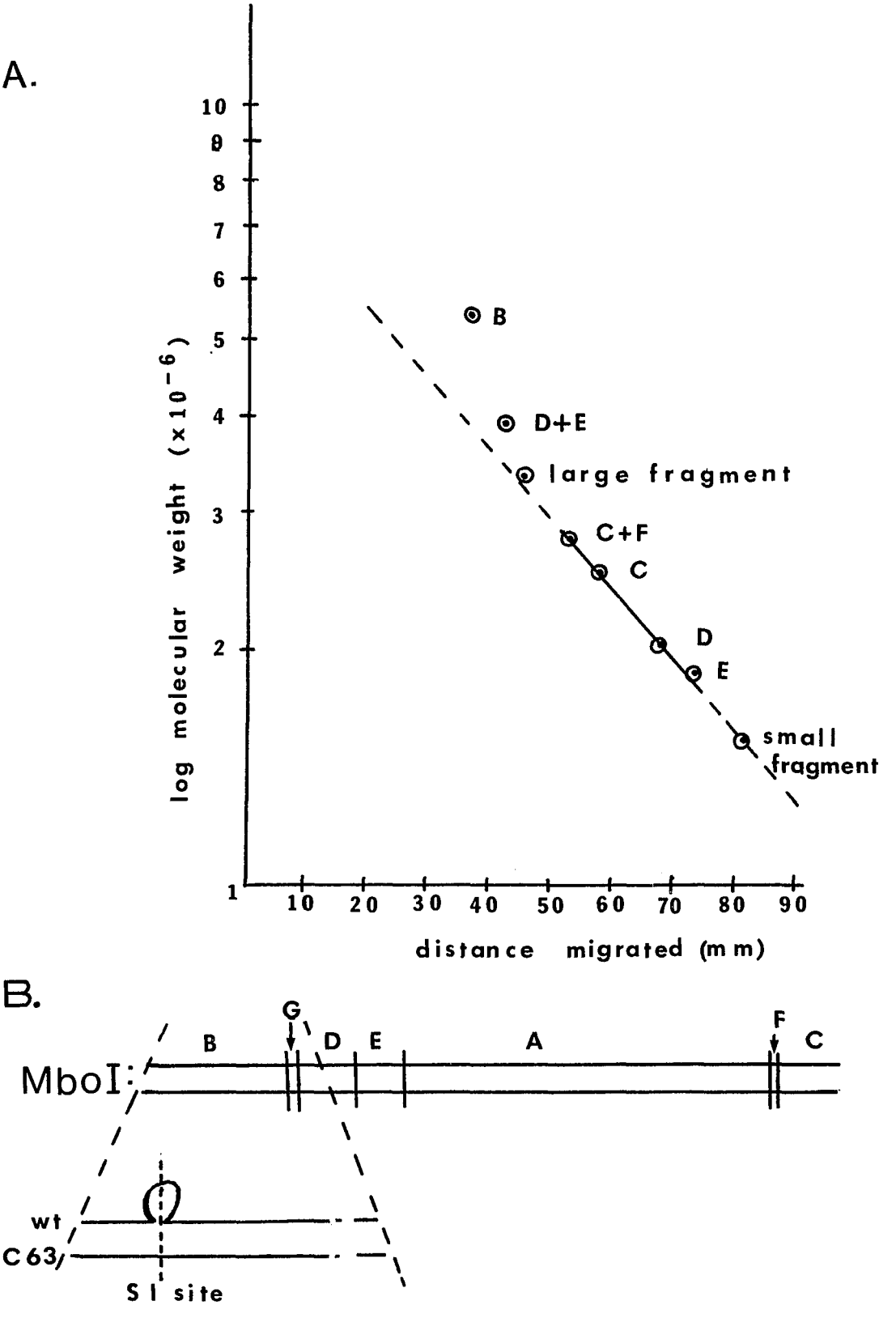
These results confirm those of Shenk et al (1975a,b) and extend them to a DNA molecule ten times longer than SV40. A deletion loop of this size is an adequate substrate for S1 cleavage even without modification. No enhancement is apparant in the chloroacetaldehyde treated DNA because 5 units of S1 is excessive and will cleave the DNA at this site of the single stranded loop even without modification of the DNA.

Detection of chloroacetaldehyde site specific enhancement by endonuclease S1. As discussed in the Introduction, Shenk et al (1975a,b) established conditions for cleaving insertion or deletion loops in SV40 with S1. They routinely used 135 units S1/μg DNA, where one unit was defined as the amount of S1 that releases 1.0 nmoles of nucleotides per minute at 37°C. As a first approximation in our studies,

Figure 9. Relative mobilities of T7 C63/wt Mbo I fragments after endonuclease S1 digestion.

A. Distance migrated of Mbo I generated restriction fragments of ^{32}PC63/wt T7 DNA following cleavage by endonuclease S1. Fragment A is off-scale on this graph. Calculations based on measurements made on the autoradiograms in Figure 8.

B. Schematic representation of the generation of two new C63/wt DNA fragments after cleavage by endonuclease S1.


A.
log molecular weight (x10⁻⁶)
10
9
8
7
6
5
4
3
2
1
B
D+E
large fragment
C+F
C
D
E
small fragment
10 20 30 40 50 60 70 80 90
distance migrated (mm)
B.
G
F
B
D
E
A
C
MboI:
wt
C63
S1 site

we decided to use the same amount of S1 as Shenk et al. (1975a) did. This necessitated converting Shenk's S1 units into our (Sigma) S1 units in order to have a basis for comparison. This was done as follows:

1 nmole DNA = 0.34 μg DNA
1 Shenk unit S1 = 0.34 μg DNA released
3 Shenk units = 1 Sigma unit, therefore

135 Shenk units = 45 Sigma units

However, we found that in our hands, and combined with chloroacetaldehyde modification of the DNA, this was far in excess of what was needed to cleave a deletion loop in a DNA heteroduplex molecule ten times the size of SV40. We were able to get cleavage using as little as 1 unit S1/ μg modified DNA.

An experimental protocol for the detection of chloroacetaldehyde induced site specific enhancement is shown in Figure 10. Heteroduplex ^{32}PC63/4R and homoduplex ^{32}PC63/C63 were incubated in the presence or absence of chloroacetaldehyde and subjected to various levels of endonuclease S1 digestion. All samples were then digested with restriction enzyme Mbo I and electrophoresed on a 1.2% agarose gel. Figure 11 is an autoradiogram of the agarose gel after electrophoresis. The Figure shows that as the concentration of S1 is increased from 1 unit of S1 per microgram of DNA up to 4 units of S1 per microgram of DNA both the chloroacetaldehyde modified

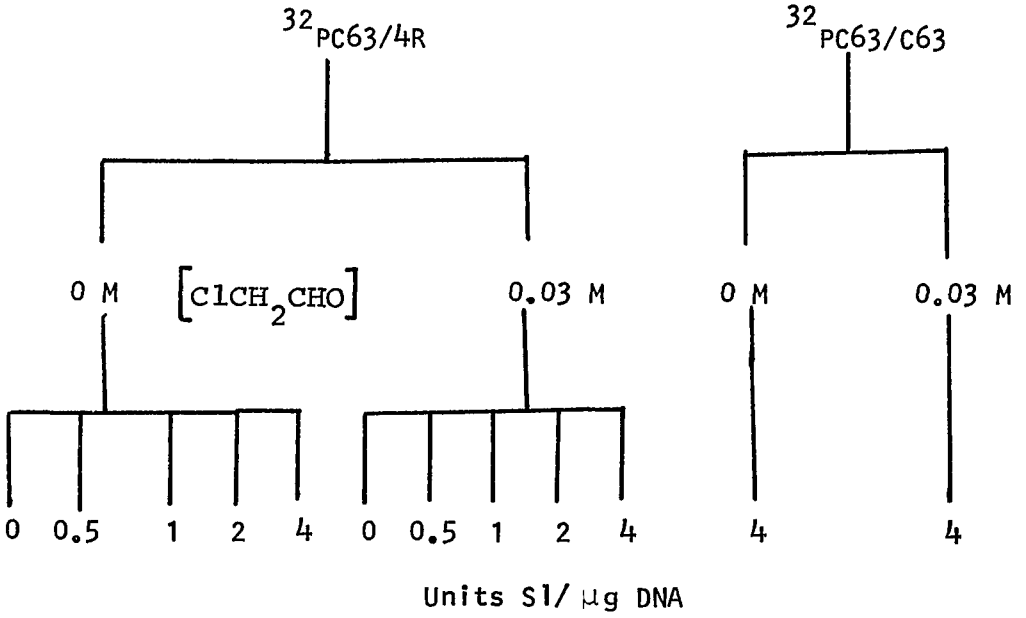


Figure 10. Protocol for site specific enhancement of a deletion loop. All samples are subsequently digested with restriction endonuclease Mbo I and subjected to agarose gel electrophoresis.

and unmodified heteroduplex molecules become more susceptible to cleavage as evidenced by the emergence of the two new bands of DNA after restriction enzyme digestion. The smaller of the new bands is not detected until at least two units of S1/μg DNA are used (Lane 10). Its smaller size requires that a proportionately greater amount of it be generated compared to the larger new band before it can be seen in an autoradiogram. Two other facts emerge from this autoradiogram. First, as the S1 concentration is increased, more random cleavage of the DNA occurs (compare, for instance, Lanes 7, 9 and 11), and secondly, that modification by chloroacetaldehyde results in greater background after S1 digestion (compare Lanes 9 and 10). Because of the large background effect as more S1 is used, the larger new band tends to get buried. For this reason only the areas under the peaks of the two smallest Mbo I fragments (D and E) and the area under the new small band have been calculated. Densitometer tracings of these three bands are shown for the autoradiogram (Figure 11) in Figure 12. These results were converted from transmittence to density, and the background under the peaks was subtracted by assuming true minima occurred between peaks C and D and between peak E and the new band. The values obtained from these area measurements are presented in

Figure 11. Autoradiogram of site specific enhancement of a deletion loop. ^{32}PC63/C63 homoduplex and ^{32}PC63/4R heteroduplex DNA. Specific activity of starting duplexes: 2.3 x 10^4CPM/μg DNA.

^{32}PC63/C63: No chloroacetaldehyde, 4 units S1/μg DNA (Lane 1); 0.03 M chloroacetaldehyde, 4 units S1/μg DNA (Lane 2).

^{32}PC63/4R: No chloroacetaldehyde, no S1 (Lane 3); 0.03 M chloroacetaldehyde, no S1 (Lane 4); no chloroacetaldehyde, 0.5 units S1/μg DNA (Lane 5); 0.03 M chloroacetaldehyde, 0.5 units S1/μg DNA (Lane 6); no chloroacetaldehyde, 1 unit S1/μg DNA (Lane 7); 0.03 M chloroacetaldehyde 1 unit S1/μg DNA (Lane 8); no chloroacetaldehyde, 2 units S1/μg DNA (Lane 9); 0.03 M chloroacetaldehyde, 2 units S1/μg DNA (lane 10); no chloroacetaldehyde, 4 units S1/μg DNA (Lane 11); 0.03 M chloroacetaldehyde, 4 units S1/μg DNA (Lane 12).

5 x 10^4 CPM/Lane. 1.2% agarose gel, electrophoresis 16½ hours at 40 volts. Autoradiogram exposed for 7½ hours at 4°C.

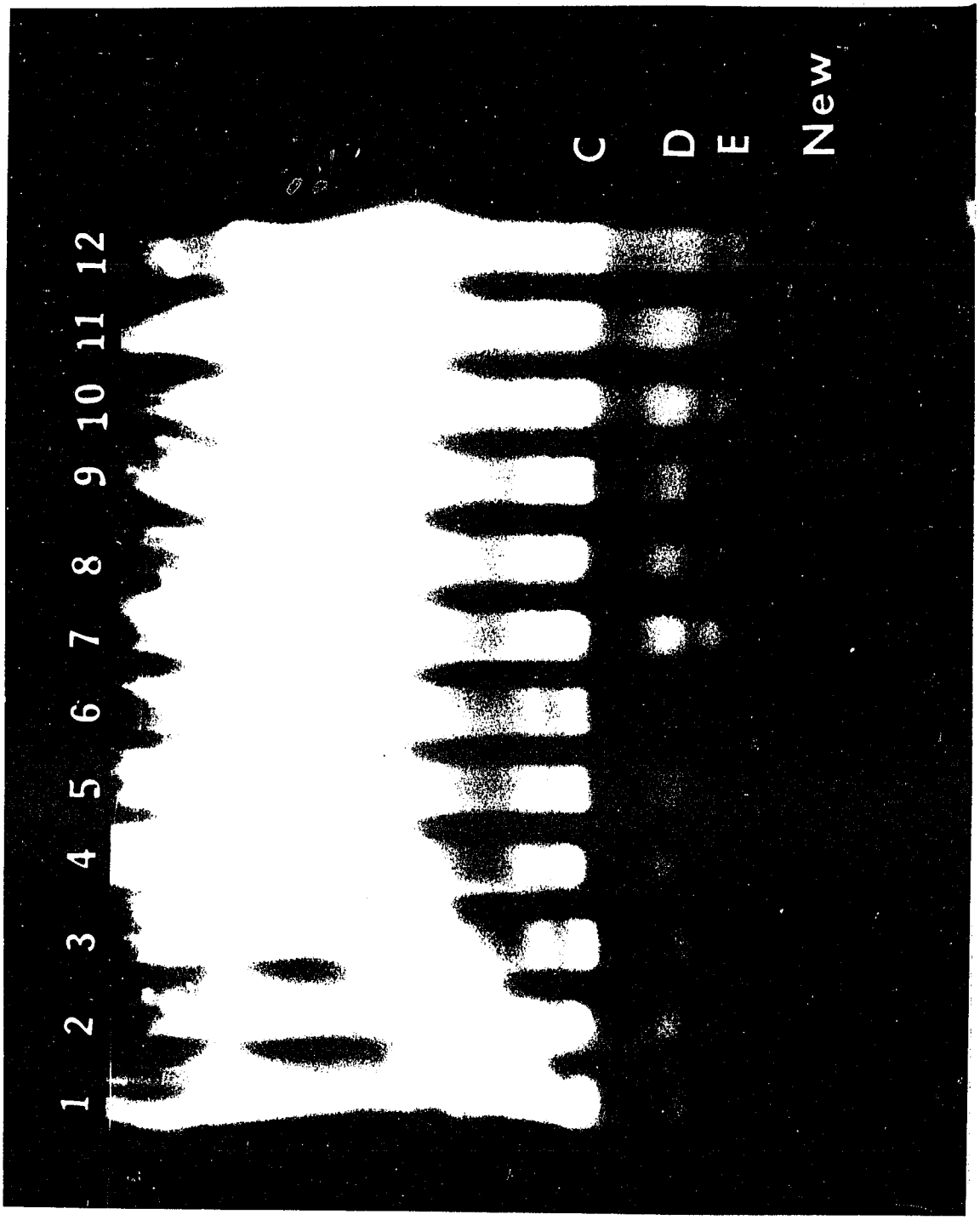
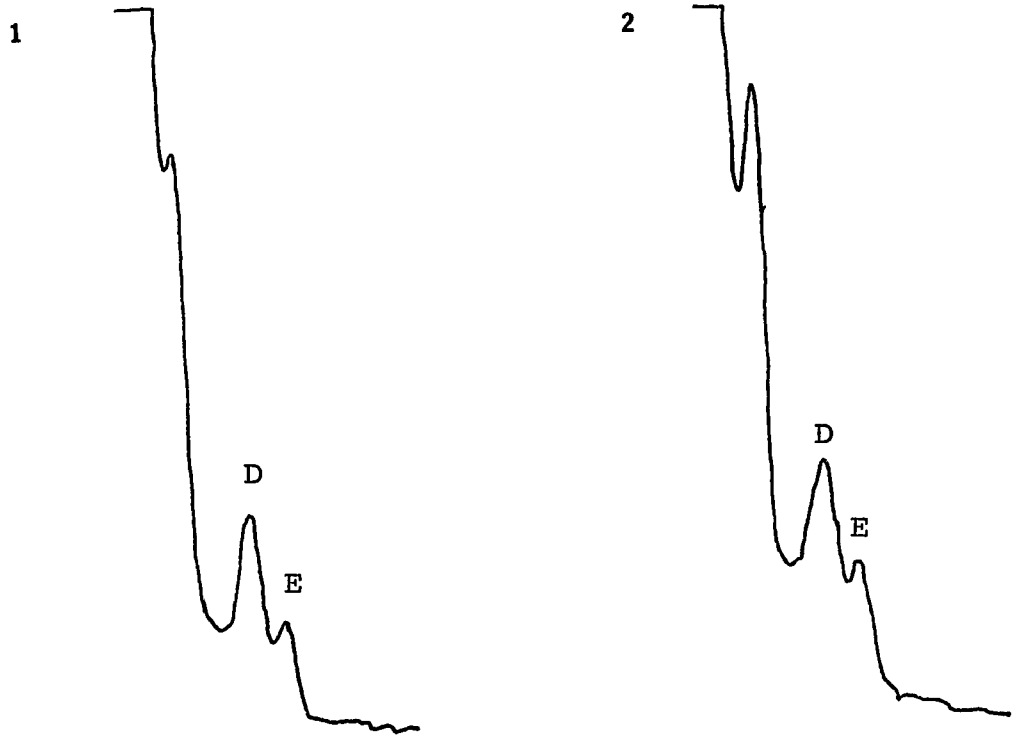

1 2 3 4 5 6 7 8 9 10 11 12
C
D
E
New

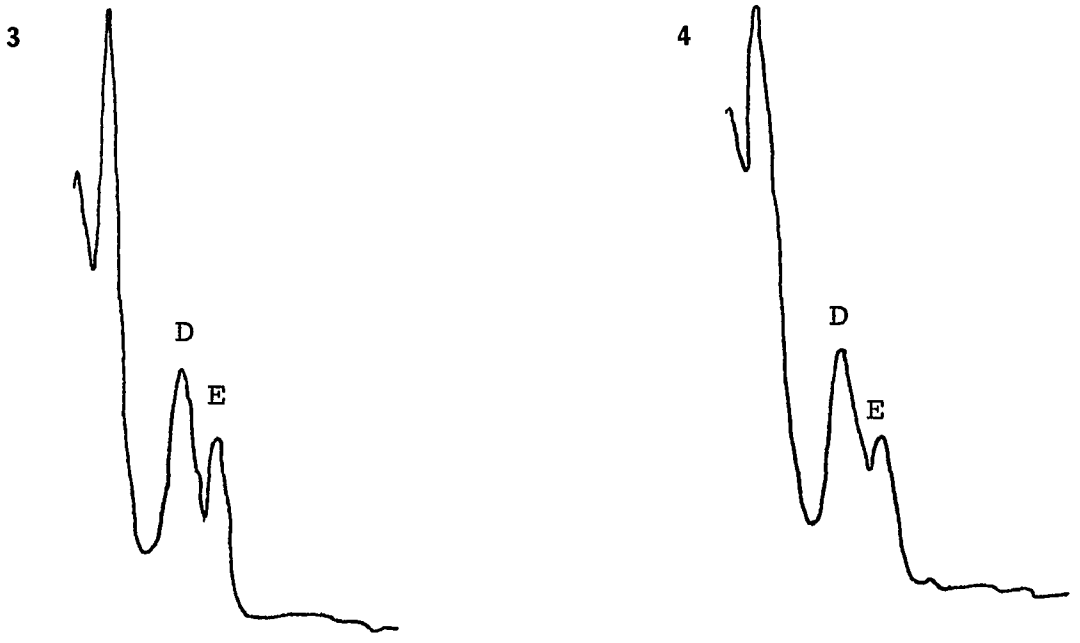
Figure 12. Densitometer tracings of site specific enhancement of a deletion loop. The number of the tracing refers to the lane on the gel shown in Figure 11. Lanes 1 and 2 are controls using the homoduplex ^{32}PC63/C63 with and without chloroacetaldehyde. A description of the contents of each lane accompanies the tracing.



No $ClCH_2CHO$ 0.03 M $ClCH_2CHO$

T7 ^{32}PC63/C63 DNA

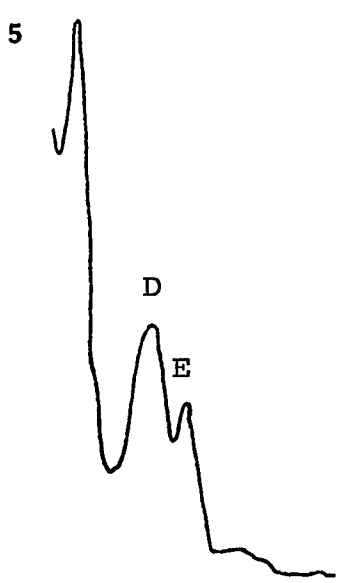
4 units S1/μg DNA



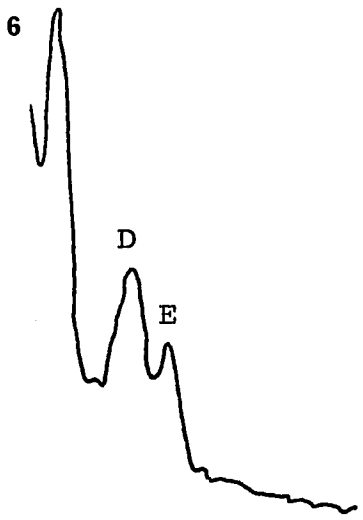
No $ClCH_2CHO$ 0.03 M $ClCH_2CHO$

T7 ^{32}PC63/4R DNA

0 units S1/μg DNA



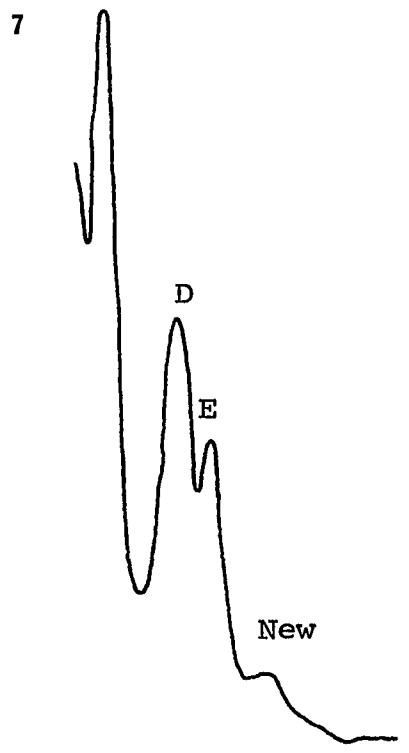
No $ClCH_2CHO$



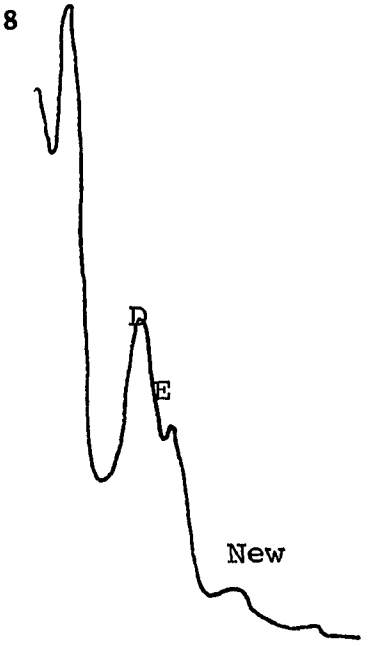
0.03 M $ClCH_2CHO$

T7 ^{32}PC63/4R DNA

0.5 units S1/μg DNA



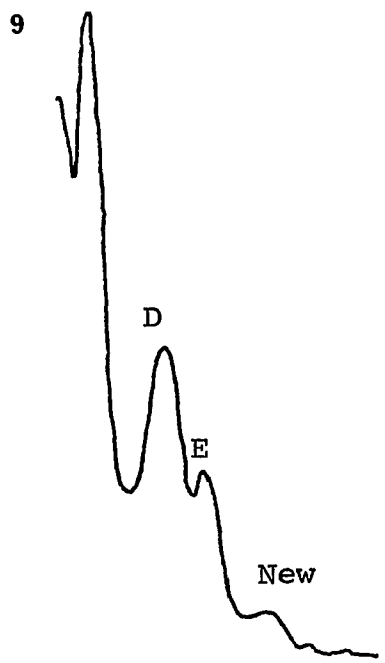
No $ClCH_2CHO$



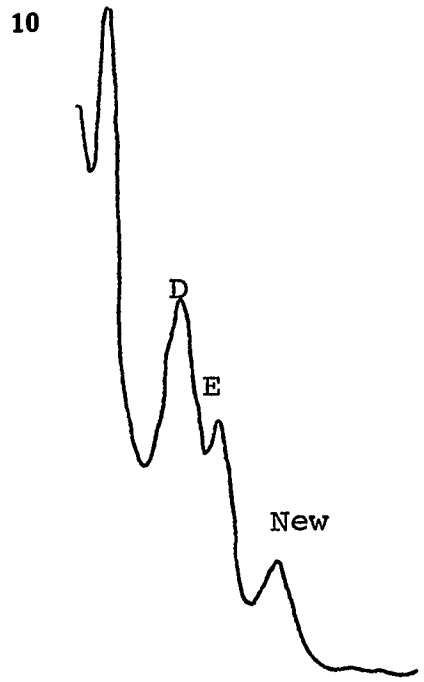
0.03 M $ClCH_2CHO$

T7 ^{32}PC63/4R DNA

1 unit S1/μg DNA



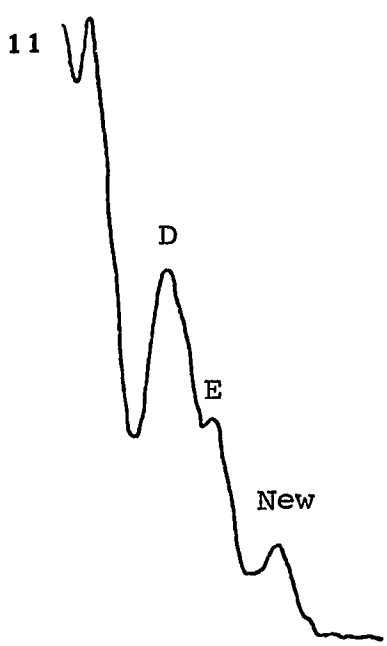
No $ClCH_2CHO$



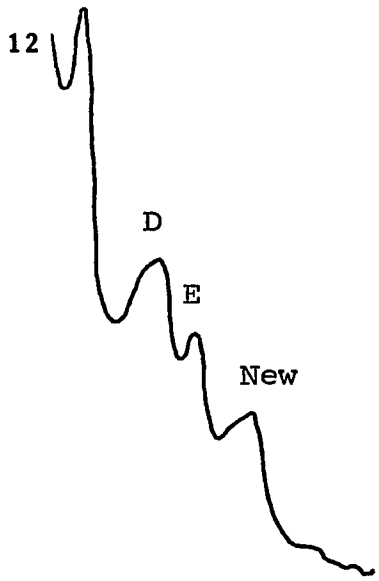
0.03 M $ClCH_2CHO$

T7 ^{32}PC63/4R DNA

2 units S1/μg DNA



No $ClCH_2CHO$



0.03 M $ClCH_2CHO$

T7 ^{32}PC63/4R DNA

4 units S1/μg DNA

Table 4. This table is representative of the results obtained in similar experiments. It clearly shows that treatment of a heteroduplex with chloroacetaldehyde leads to enhanced site specific cleavage by endonuclease S1 when there is insufficient S1 to completly cleave the heteroduplex. In order to confirm that these peak densities were in the linear range of the film, the tracings of radioactive standards exposed on the same film were converted from transmittance to density values and compared to the data peaks. Additional discussion of site specific enhancement appears in Chapter 3.

TABLE 4

AREAS UNDER THE MBO I-GENERATED PEAKS OF ^{32}PC63/4R HETERODUPLEX DNA TREATED WITH VARIOUS CONCENTRATIONS OF ENDONUCLEASE S1

(Data from Figure 12)

‡ $ClCH_2CHO$:	-	+	-	+	-	+	-	+	-	+
Units S1/μg DNA :	0	0	0.5	0.5	1	1	2	2	4	4
Band areas										
D + E	225	270	240	201	432	258	250	299	397	180
New	<5	<5	<5	<5	<5	23	12	32	29	55
Ratio										
New/D+E	<0.03	<0.03	<0.03	<0.03	<0.03	0.089	0.048	0.107	0.073	0.305
Enhancement										
$ClCH_2CHO$ +/-	No		No		Yes		Yes		Yes	

‡ - = 0 M $ClCH_2CHO$, + = 0.03 M $ClCH_2CHO$

CHAPTER 3

PHYSICAL MAPPING: SITE SPECIFIC ENHANCEMENT WITH POINT MUTANTS

INTRODUCTION

As indicated in Chapters 1 and 2, the power of site specific enhancement mapping lies in its ability to generate fine structure physical maps. Experiments aimed at fine structure mapping of DNA are described in this chapter. An additional benefit deriving from the physical basis of the technique might be, with appropriate modification, its adaptability for use with the RNA containing viruses. Variables involved in optimizing the technique for DNA as well as for RNA genomes are considered in the Discussion.

The strategy for mapping point mutants using the site specific enhancement technique is essentially the same as that employed for deletion mutants. Amber mutants and their spontaneously occurring revertants are used to form heteroduplexes. The formation of heteroduplex DNA between amber mutants and their revertants assures, as much as is possible, that the heteroduplex molecule will be base paired at all loci except at the site of the amber mutation itself.

MATERIALS AND METHODS

E. coli O11', a thy^-, Su_{11} derivative of *E. coli* B was constructed in Brenner's laboratory and obtained from F.W. Studier. It is used as the permissive strain for T7 amber mutants.

Amber mutants in genes 4. 8, 12, and 16 are located as shown in Figure 3, Chapter 2. They were a generous gift from F.W. Studier.

T7 revertants 4R, 8R, 12R, and 16R were isolated in this laboratory as spontaneously occurring revertants.

E. coli O11' was grown at 30°C in a water bath shaker in LB medium (Chapter 2). At 2×10^8 cells/ml T7 amber mutant 4 was added at an MOI of 2 - 4. The flask was shaken at 30°C until lysis occurred.

All other materials and methods are described in Chapter 2.

RESULTS

The conditions of chloroacetaldehyde or S1 incubation were varied in an attempt to find conditions in which heteroduplex molecules containing a single base pair mismatch could be cleaved by endonuclease S1. In all experiments DNA is extracted from the phage by alkali denaturation.

All heteroduplexes are extracted with phenol, using the method of Mandell and Hershey (1960) followed by 5 chloroform extractions. In the experiments reported below, DNA was added directly to a solution of Et_4NCl and chloroacetaldehyde.

The amber mutation used in all these experiments is in gene 4. This gene falls entirely within Band E of an Mbo I digest (Chapter 2, Figure 3). Although we don't know precisely where in gene 4 the mutation is located, S1 cleavage anywhere in the gene will produce two bands from fragment Mbo I-E. Both new bands will migrate faster than E, the fastest migrating band on the gel. These bands would contain approximately 475 bases and 2360 bases each if cleavage occurred at the rightmost or leftmost end of gene 4, or 1400 bases if cleavage were around the center of gene 4, and some number of bases intermediate to these if cleavage occurred anywhere else on the gene. With the appropriate electrophoresis conditions all fragments of these sizes are capable of being detected in an agarose gel.

A heteroduplex was formed from a 1:20 mixture of ^{32}P T7-4R and cold T7-4 DNAs. Figures 13 and 14 show a photograph of the gel after electrophoresis and of the autoradiogram respectively. Incubation conditions with chloroacetaldehyde and endonuclease S1 are given in the caption

to Figure 13.

The DNA samples were treated with chloroacetaldehyde under conditions which would modify about 2 to 3 residues if a mismatch defect is thermodynamically equivalent to an end of a DNA molecule (see Discussion). The modified mismatch defect should form a substrate that can be cut efficiently by S1. The concentration of S1 was varied from 0 units/μg DNA to 120 units/μg DNA and incubation was for seven minutes at 37°C. Shenk et al. (1975a) used 45 units S1/μg DNA (Shenk's units have been converted to our (Sigma) units, see Discussion) for 20 minutes at 25°C to cut a deletion loop 32 bases long. Although it is difficult to compare these conditions since the S1 incubation temperatures differ, our range of concentrations should include conditions that are comparable to those of Shenk et al. (1975a). Both homoduplex and heteroduplex molecules in the absence of chloroacetaldehyde are badly damaged by 120 units S1/μg DNA (Lanes 1 and 9) which could be explained if the T7 DNA preparations contained single strand breaks. In two experiments where degradation was indicated, the heteroduplex DNA was again alkali denatured, renatured, cut with Mbo I, and subjected to agarose gel electrophoresis. No indication of degradation was observed. Lanes 3 and 4 were not treated with S1 and as expected did not

show any new bands. Lanes 6, 8 and 10 were incubated with chloroacetaldehyde and then treated with 4 units/μg, 40 units/μg, and 120 units/μg of S1 respectively (as were 5, 7 and 9, which were incubated without chloroacetaldehyde). The DNA in Lanes 6, 8 or 10 had the potential to contain new bands. No new bands appear on the autoradiogram of this gel (Figure 14). A gel that has run this far should still contain a 2400 to 1400 base pair fragment if it were present in the preparation. The experimental conditions (various incubation times and temperatures) were modified according to the rationale set forth in the Discussion. The results in Figures 13 and 14 were the best we were able to achieve. Table 5 summarizes experimental conditions that were tried.

DISCUSSION

A. Physical Mapping of Point Mutations in DNA Viruses Using Chloroacetaldehyde

The experiments in Chapter 2 demonstrated that the region around a deletion loop in a heteroduplex molecule can be chemically modified by chloroacetaldehyde, making it a better substrate for S1 cleavage than unmodified duplexes. Following S1 cleavage, the site of the original DNA defect can be located by analyzing agarose gel

Figure 13. Agarose gel electrophoresis of heteroduplex ^{32}P4R/4.

Homoduplex ^{32}P4R/4R: no chloroacetaldehyde, 120 units S1/μg DNA (Lane 1); 0.03 M chloroacetaldehyde, 120 units S1/μg DNA (Lane 2).

Hteroduplex ^{32}P4R/4: no chloroacetaldehyde, no S1 (Lane 3); 0.03 M chloroacetaldehyde, no S1 (Lane 4); no chloroacetaldehyde, 4 units S1/μg DNA (Lane 5), 0.03 M chloroacetaldehyde, 4 units S1/μg DNA (Lane 6); no chloroacetaldehyde, 40 units S1/μg DNA (Lane 7); 0.03 M chloroacetaldehyde, 40 units S1/μg DNA (Lane 8); no chloroacetaldehyde, 120 units S1/μg DNA Lane 9); 0.03 M chloroacetaldehyde, 120 units S1/μg DNA (Lane 10).

1.2% agarose, 14 hours, 40 volts.

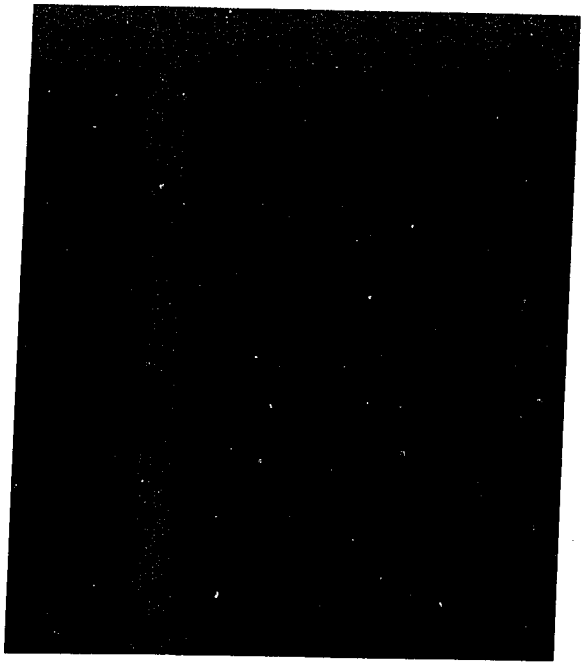


Figure 14. Autoradiogram of agarose gel electrophoresis of heteroduplex ^{32}P4R/4. 5 x 10^4 CPM/ Lane. Contents of each lane and duration of electrophoresis same as for Figure 13. This is an autoradiogram of the gel shown in Figure 13. Autoradiogram exposed 24 hours.

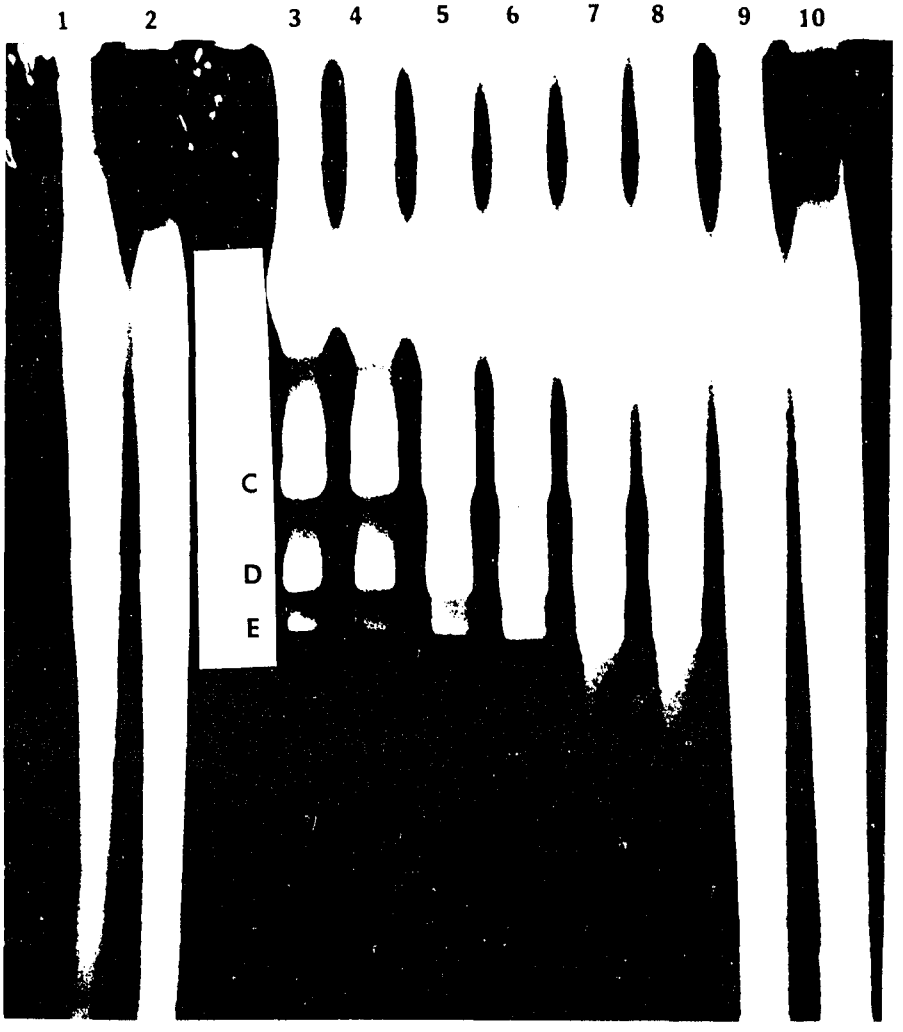

1
2
3
4
5
6
7
8
9
10
C
D
E

TABLE 5

EXPERIMENTAL CONDITIONS USED TO MAP POINT MUTATIONS BY SITE SPECIFIC ENHANCEMENT

Experiment	Chloroacetaldehyde incubation Concentration (M)	Temp. (^{o}C)	Time (min.)	ǂUnits S1 µg/DNA
1	0	20	480	50
	"	37	"	"
	0.02	20	"	"
	"	37	"	"
2	0	55	10	0
	"	"	"	4
	"	"	"	40
	"	"	"	120
	0.03	"	"	0
	"	"	"	4
	"	"	"	40
	"	"	"	120
3	0	55	0	40
	0.03	"	10	"
	"	"	15	"
	"	"	20	"
4	0.03	55	0	0
	"	"	"	5
	"	"	10	0
	"	"	"	5
	"	"	20	0
	"	"	"	5

ǂ All S1 incubations were for 7 minutes at 37^{o}C.

electrophoretic patterns of restriction enzyme fragments. We attempted to find appropriate conditions to apply this technique to a DNA heteroduplex with a defect involving only a single base pair. Heteroduplex formation and restriction enzyme digestion and analysis are exactly the same in both cases. The critical steps for which conditions must be determined are: modification by chloroacetaldehyde, and cleavage with S1.

Three variables must be considered with respect to chloroacetaldehyde modification of a molecule containing a single base pair mismatch: the temperature, the chloroacetaldehyde concentration, and the reaction time. Salt concentration (2.4 M Et_4NCl) and pH (4.7) are fixed by the requirement for maintaining base composition independent reaction conditions.

Two temperature (T)-dependent variables may be expected to affect the rate of reaction of chloroacetaldehyde with native DNA near a defect in a molecule, such as a mismatch or end. One variable is the rate constant, k, for reactions of chloroacetaldehyde with denatured DNA. k will behave according to

$$k = Ae^{-E^{\ddagger} / RT} \qquad (4)$$

where A is the Arrhenius (preexponential) factor, R is the gas constant, and $E^{\ddagger}$ is the activation energy. As

noted in Chapter 2, k values have been measured at various temperatures for reactions with adenosine and cytidine residues. $E^{\ddagger}$ is approximately 22 kcal/mole. The second variable is the equilibrium constant for base-pair formation, s, which will behave according to

$$s = s_o \, e^{-\Delta H / RT} \tag{5}$$

where ΔH is the enthalpy of melting of DNA(-6.5 kcal/mole in 2.4 M Et_4NCl (Klump, 1977)) and s_o is determined so that s = 1 at the melting temperature. Using an end of a molecule as a model for a defect, the average number of base pairs irreversibly unpaired by reaction with chloroacetaldehyde, $\langle x \rangle$, is given by

$$\langle x \rangle = \frac{skCt}{(s-1)^2} \tag{6}$$

where C is the chloroacetaldehyde concentration and t is the time. A derivation of equation 6, adapted from Ruyechan (1976) is given in Appendix A. Chloroacetaldehyde is always present in excess. The extent of its reaction with DNA therefore is proportional to the product of the chloroacetaldehyde concentration and reaction time. These are not independent variables. Our normal reaction conditions (see Figure 8, Chapter 2 and Figures 13 and 14, Chapter 3) involve the use of 0.03 M chloroacetaldehyde at 55°C for 10 minutes. These reaction conditions are expected to lead

to $\langle x \rangle \approx 1$ for ends of molecules. If a defect such as a deletion loop were identical to an end, then an average of one base pair would be opened on either side of the deletion loop. Reaction for longer periods of time did not improve the site specific enhancement with heteroduplexes containing a deletion loop and served only to increase the background of randomly cleaved molecules. These results, reported in Chapter 2, imply that the use of a molecular end as a model for a deletion loop defect leads to satisfactory agreement between experiment and theory.

A single base-pair mismatch might be expected to be more difficult to detect than a deletion loop (Shenk et al., 1975a). The reaction of chloroacetaldehyde with DNA bases near a mismatch defect (signal) might be expected to depend upon temperature in a different manner than reaction of chloroacetaldehyde with DNA bases transiently open due to "breathing" (noise). If so, the optimization of the reaction with respect to signal to noise ratio might be achieved by either raising or lowering the temperature and compensating for the effect of temperature by altering the reaction time. Equation 6 was used to calculate a range of reaction times for temperatures other than 55°C. Site specific enhancement was attempted at temperatures from 55°C (Tm - 8°C) to 20°C (Tm - 43°C). When these temperatures

and various times were used, comparable levels of random degradation (noise) were observed following endonuclease S1 digestion. These results indicate that equation 6 may describe the temperature and time dependence for chloroacetaldehyde reactions with random DNA bases in duplex DNA as well as reaction with DNA bases near a defect. If so, the signal to noise ratio could not be improved by altering the reaction conditions. The limits for detecting a defect by site-specific modification would then be the size of the genome and the nature of the defect. Any thermodynamic differences between the DNA destabilization introduced by deletion loops versus mismatches involving transitions or transversions cannot be affected in a predictable manner by the alteration of the chloroacetaldehyde reaction conditions. The effect of the size of the genome could be altered by working with isolated restriction fragments. If further work on chloroacetaldehyde-induced site specific modification were to be attempted, however, consideration should be given to the use of a bacteriophage system with a lower complexity (N) than bacteriophage T7.

With respect to the second step, cleavage with endonuclease S1, Wiegand et al. (1975), Shishido and Ando (1975), and Shenk et al. (1975a) reported that S1 can cleave the DNA opposite a nick, suggesting that the minimum

substrate size may be only one nucleotide in length. However, a nicked molecule is a poor substrate. Cleavage is inefficient and requires large amounts of S1, or long incubation times. Shenk et al. (1975a,b) used S1 to cleave heteroduplex DNA containing a single base pair mismatch. Even in the best example given there was a very high signal to noise ratio after cleavage. Dodgson and Wells (1977a,b) systematically examined the question of size by synthesizing molecules containing 1 to 6 mispaired bases in a G·C duplex. Their work indicated that a single stranded region of DNA consisting of 1 or 2 bases lying within a duplex molecule is a very poor substrate for endonucleolytic cleavage by S1. Sensitivity to S1 increased as the size of the single stranded region was increased to three or more bases. It is very difficult to extract any but the most general sort of guidelines from these and other experiments because incubation conditions are not comparable, nor are units of S1 defined with any consistancy among laboratories. We have chosen to try to modify our single base mismatch heteroduplex so that the single stranded region is approximately 2 to 3 bases long. We were limited in that additional chloroacetaldehyde reaction resulted in the creation of additional S1 substrates.

B. Other Single Strand Specific DNA Modifying Agents

Although there is no evidence at the present time to implicate chloroacetaldehyde in the failure to attain the goal of site specific cleavage of a heteroduplex containing one unpaired base, it is prudent to consider the possible use of alternative modifying agents. For a modifying agent to be suitable for the purposes of this study, it must have the following attributes: (a) be specific for, or preferentially bind to, unpaired rather than paired bases under conditions in which denaturation of secondary structure is minimal; (b) form a stable complex with the base or bases; (c) prevent the re-pairing of bases; (d) leave the DNA free of single strand breaks after reaction. Several suitable candidates exist. The first is the aliphatic dialdehyde, glyoxal (CHOCHO). It adds to and dissociates from adenosine and cytidine rapidly but reacts more slowly and stably with guanosine (Hsu et al., 1973). Under appropriate conditions of glyoxal concentration, pH, and temperature, it binds preferentially to guanine residues in DNA or RNA and is stable from 20 to 50 hours at 20°C (Hsu et al., 1973; McMaster and Carmichael, 1977). This binding occurs when bases are unpaired and is low or absent in double stranded DNA (Nakaya et al., 1968). Binding introduces an additional

ring onto guanine residues and sterically hinders the formation of G·C pairs (see Figure 15a). Although it affects only G·C pairs, unlike chloroacetaldehyde which affects both A·T and G·C, it may still be useful for site-specific enhancement. The reaction is efficient at pH 7 - 8, somewhat gentler conditions for DNA than pH 4.7 of the chloroacetaldehyde reaction. Reaction with glyoxal does not introduce single strand breaks into DNA as assayed by agarose gel electrophoresis (McMaster and Carmichael, 1977).

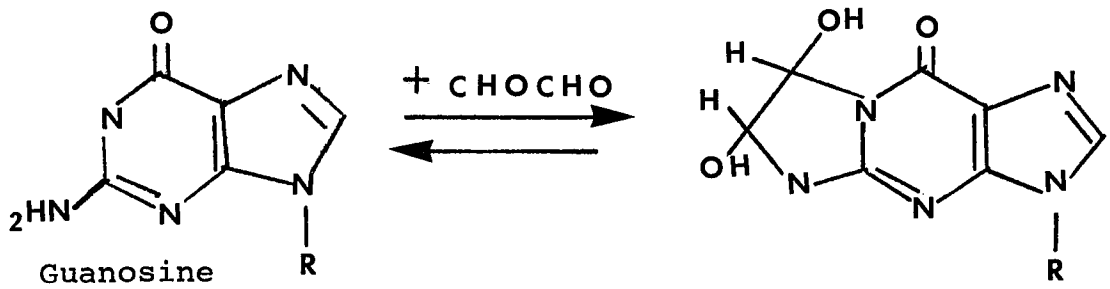
Another candidate for modifying agent is the carbodiimide, N-cyclohexyl-N'-β-(4 methylmorpholinium) ethylcarbodiimide. It reacts with thymine, guanine, and uracil under mild conditions (pH 8) (Lebowitz et al., 1976). The reaction with thymine is illustrated in Figure 15b. It reacts quickly with single stranded DNA and very slowly with native DNA (Metz and Brown, 1969b). Once bound to the base, re-pairing is hindered. It has been used with both DNA and RNA and does not introduce single strand breaks (Lebowitz et al., 1976) When reacted with tRNA only the anticodon loop became modified (Metz and Brown, 1969b). Wurst et al. (1978) showed the same sort of specificity using S1 as the probe of secondary structure. Carbodiimide also satisfies the criteria for a modifying agent and could be used instead of chloroacetaldehyde in our system.

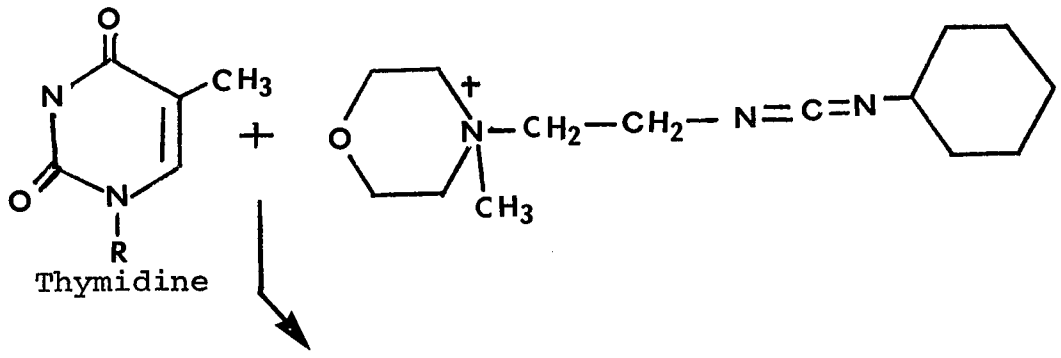
Figure 15. Modification of DNA bases by glyoxal and carbodiimide.

A. Glyoxal modification of guanosine.

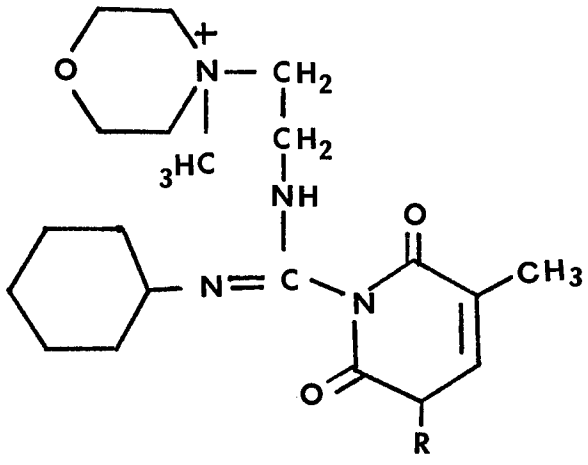
B. Carbodiimide modification of thymidine.

A





B



C. Genetic Mapping of RNA Viruses

Mapping RNA viral genomes has progressed at a slower pace than that of their DNA counterparts. One reason is that many techniques so successfully applied to DNA viruses, such as marker rescue, are not suitable for RNA. Cells appear to lack the capability to promote recombination of RNA and therefore, with a few exceptions, to be discussed below, mapping by recombination analysis is not possible with these viruses.

Baltimore (1971) classified animal viruses into 6 groups based on their mode of transmission of genetic information. His scheme will be used as the basis for the organization of this section of the Discussion. However, the points to be made are applicable to all RNA containing viruses, not just those infecting animals. The RNA viruses are contained in classes III, IV, V, and VI of Baltimore's scheme.

Class III ($^{\pm}$RNA $\rightarrow$ mRNA). Representative members: reovirus, orbivirus, wound tumor virus, and rice dwarf virus.

All known members of this class have segmented genomes, each of which seems to be monocistronic (Both et al., 1975). Members of Class III, and certain members of Class V, which also have segmented genomes, have higher

recombination frequencies than viruses with uninterrupted genomes. This apparent recombination is due to random reassortment of the RNA pieces at maturation.

The most thoroughly studied member of this group is probably reovirus. McCrae and Joklik (1978) identified the polypeptides encoded by each of the 10 double stranded segments of reovirus by isolating the segments, denaturing them, and using them individually in an *in vitro* wheat germ translating system. Peptide maps (generated by staphylococcal V8 protease), of *in vitro* translation products were compared to authentic reovirus proteins for positive identification.

Another approach to mapping was undertaken as the result of a report by Ramig *et al.* (1977) that the double stranded RNA segments and the polypeptides of the three reovirus serotypes could be distinguished by their electrophoretic mobilities on polyacrylamide gels. Sharpe *et al.* (1978) analyzed the polyacrylamide gel patterns of recombinants between serotypes to construct a map of the reovirus genome in a manner analogous to that used for influenza virus (see Class V). This method has allowed ts mutants to be mapped onto specific segments of the reovirus genome.

Ito and Joklik (1972) found that some hybrid reovirus RNA species composed of the plus strand of a ts mutant and

the wt minus strand migrated significantly more slowly during polyacrylamide gel electrophoresis than either the corresponding homologous species or the reverse hybrids (ts minus strand, wt plus strand). This finding suggested that retarded electrophoresis was caused by the presence of a mutation. The presence of an unpaired base (in the hybrid RNA) in certain critical locations could lead to the formation of loops which would slow the electrophoretic migration rate (Ito and Joklik, 1972). Scheurch and Joklik (1973) examined hybrids of reovirus wt and 35 ts mutants and concluded that anomalous electrophoretic migration of hybrid RNA molecules could serve as a physical marker for the presence of mutations. However, not all hybrid molecules exhibit altered electrophoretic mobility.

All 3 methods were used to assign mutations or polypeptides to the appropriate RNA segments. Once a method is available for detecting the location of the loops postulated to exist in hybrid RNA molecules in the third technique, fine structure mapping may be possible for those hybrids which possess altered electrophoretic mobilities.

Class IV ($^{+}$RNA $\longrightarrow$ $^{-}$RNA $\longrightarrow$ mRNA). Representative members: picornaviruses (enteroviruses such as polio virus, coxsackie virus, echo virus, rhinovirus), and probably arbovirus and togavirus.

The single stranded RNA genome of members of this class has the same polarity as the mRNA, and the RNA is infectious.

The mapping of poliovirus will serve as one example of a picornavirus within this group. The techniques applicable to the mapping of arboviruses and togaviruses are similar to those used for mapping Class V unsegmented genomes, such as VSV (see below). Poliovirus capsid proteins may be synthesized from a single precursor molecule which is later cleaved into smaller ones (Summers and Maizel, 1968; Jacobson and Baltimore, 1968). This posttranslational cleavage provides one basis for determining the position of regions of viral RNA coding for the capsid proteins. Taber (1971) and Rekosh (1972) used the drug pactamycin to prevent initiation of protein synthesis while permitting elongation to continue. The further a "cistron" is from the initiation site, the more frequently it is translated by ribosomes during the 10-12 minutes after addition of drug before protein synthesis completely ceases. By comparing the fraction of total ^{35}S-methionine incorporated into polypeptides following addition of the drug to the fraction of radioactivity found in the absence of the drug, a measure of the relative distances of "cistrons" from the origin of initiation was made. Mapping

by posttranslational cleavage is not a fine structure mapping technique.

Contrary to the opening statement of Section C, several reports of recombination between poliovirus mutants appeared in the 1960's (Hirst, 1962; Ledinko, 1963; and Cooper, 1968). Cooper (1968) produced a genetic map of poliovirus based on recombinational analysis of mixed infections using three factors crosses with different ts mutants. Double mutants were isolated in 10-30x excess over what was expected from spontaneous mutation as judged from self crosses. This apparent RNA-RNA recombination is unprecedented and is difficult to explain. Recombinants could be used to generate fine structure maps. In spite of this advantage, recombinational mapping of Class IV viruses has not progressed very much since the work of Cooper (1968)

Class V ($\bar{}$RNA $\longrightarrow$ mRNA.) Representative members: rhabdovirus, paramyxovirus, orthomyxovirus.

Members of this class contain single stranded RNA whose polarity is opposite to that of mRNA. Rhabdoviruses such as vesicular stomatitis virus (VSV), rabiesvirus, and paramyxoviruses such as sendai and newcastle disease viruses (NDV) have a single long RNA genome. Influenza

virus, bunyavirus, and arenavirus all have segmented genomes.

VSV mRNA is thought to be initiated at a single site on the genome and cleaved into the five known monocistronic mRNAs (Rhodes and Banerjee, 1976). Abraham and Banerjee (1976) used ultraviolet (UV) irradiation of VSV RNA to determine the order of the VSV genes. UV irradiation produces pyrimidine dimers which block RNA transcription. The UV dose required to inhibit mRNA transcription is proportional to the target size of the mRNA and therefore to the size of the gene. However, when the experiment was performed it was discovered that the target sizes of the genes were not proportional to their physical sizes. In fact, after analysis it appeared that the transcription of each gene was dependent on the prior expression of all 3'-adjacent genes. This confirmed the idea of a single initiation site for mRNA transcription and allowed the order of the genes in the VSV genome to be determined. This technique does not generate fine structure maps.

The eight single strands of influenza RNA (Palese and Schulman, 1976a) can undergo reassortment (Sugiura and Kilbourne, 1966) during mixed infection. Palese and Schulman (1976b) used differences in the electrophoretic patterns of the RNA segments of different strains of

influenza A virus to identify genome segments coding for hemagglutinin and neuraminidase. They extended this method to derive a complete map of the influenza genome correlating protein products to the genome segment from which they were derived (Ritchey et al., 1976).

Class VI ($^{+}$RNA $\longrightarrow$ $^{-}$DNA $\longrightarrow$ $^{\pm}$DNA $\longrightarrow$ mRNA).

Representative members: Retroviruses are grouped by virtue of morphological differences into A-type particles which are immature B-type particles; B-type particles such as murine mammary tumor virus; and C-type particles which include all known sarcoma and leukemia viruses.

All members of this group contain an RNA-dependent DNA polymerase (reverse transcriptase). During viral replication the genome exists as a double stranded DNA intermediate. The viruses contain two single strand RNA genomes with poly A at the 3' end.

A multi-step procedure has been used to map the src, env, pol, and gag genes of rous sarcoma virus (RSV). Random fragments of poly A containing ^{32}P-labeled genomic RNA are grouped according to size after purification on an oligo dT column. The RNA segments are identified by their RNase-T1-resistant oligonucleotides. The position of each segment relative to the 3' (poly A) end was determined in this way (Coffin and Billeter, 1976). Genetic

function was determined by comparing nondefective RSV with deletion mutants of RSV to identify which oligonucleotides are missing and which genetic function is missing.

Gag and pol were mapped by isolating recombinant viruses with different parental-type oligonucleotide maps. By correlating recombinant markers with recombinant oligonucleotides the genes could be ordered relative to the 3' end (Wang et al., 1976). This kind of analysis, because it involves recombination, may be extended to fine structure mapping. It is likely that recombination occurs during the time the genome exists as a double stranded DNA intermediate.

Genomes of RNA viruses which do not recombine have been mapped by taking advantage of unique properties possessed by the various classes of viruses. Fine structure analysis is only possible through detection of changes in ordered oligonucleotide fingerprints or in the sequence of the genome itself. RNA mapping could benefit from an adaptation of the site specific enhancement technique.

D. Adaptation of Site Specific Enhancement Mapping for Use with RNA Containing Viruses

As outlined in Chapter 2, the process of site specific enhancement mapping involves:

1. forming heteroduplex DNA containing a single base pair mismatch

2. chemical modification of the mismatch region with chloroacetaldehyde in order to convert the heteroduplex from a poor to a good substrate for endonuclease S1. Modification is carried out in a base composition independent solvent

3. cleavage of single stranded nucleic acid with endonuclease S1 at the specific site of the mismatch.

4. cleavage with a sequence specific (restriction enzyme) endonuclease

I will address each of these steps in turn with respect to RNA.

1. Heteroduplex formation. This might at first glance seem difficult to accomplish with single stranded RNA viruses. However, it has been shown for several members of the paramyxovirus group, at least, that full-length double stranded RNA can be isolated (Kolakofsky et al., 1974). The mechanism is presumably that a certain proportion of virions contain the "wrong" ($^{+}$RNA) strand due to

incorrect packaging. Another instance in which heteroduplex formation is possible is with the group of viruses containing double stranded RNA. Members of this group are found infecting vertebrates (reovirus), plants (wound tumor virus), bacteria (ø6 of Pseudomonas), and fungi (Penicillium chrysogenum mycophage). All have segmented genomes containing from 2-15 segments (Wood, 1973). With other RNA virus groups, double stranded RNA could be made by using intracellular RNA intermediates present during viral replication.

2. Chemical modification. The requirements are (a) a solvent in which the base pair stability of A·U equals that of G·C and (b) an agent specific for unpaired purines or pyrimidines, that binds irreversibly and prevents bases from re-pairing.

Melchior and Von Hippel (1973) studied solvent systems containing different small alkyl ammonium ions known to preferentially bind A·T-rich regions in DNA. Ruyechan (1976) conducted a series of studies of the behavior of double stranded RNA in solutions containing various amounts of Et_4NCl in an attempt to find a base composition independent solution for RNA melting. He used RNAs from NDV and Penicillium chrysogenum mycophage. He found that at concentrations between 3.24 M and 3.33 M Et_4NCl, base

composition effects are eliminated for RNA. This is about 1.35 times the concentration needed to achieve the same effect with DNA. The RNA melted at 29°C in 3.33 M Et_4NCl. The renaturation rate maximum at 20- 25°C below Tm is below room temperature for RNA-RNA when Et_4NCl is raised to 3.24 M. These low temperatures are incompatible with site specific modifications using chloroacetaldehyde. In another set of studies, Wetmur (unpublished) has tested tetramethylammonium chloride (Me_4NCl) for use as an RNA base composition independent reagent. He has shown that the difference in base pair stability in RNA-RNA disappears in 4 M Me_4NCl and that under these conditions melting occurs at 92.7°C. Me_4NCl solutions appear to be good candidates for a chemical modification solvent for RNA-RNA heteroduplexes.

Chloroacetaldehyde, the chemical agent used in the DNA mapping studies (Chapters 2 and 3) was initially studied by Barrio et al. (1972) using RNA. At 37°C the chloroacetaldehyde modification reaction has a pH optimum of 4.5 for adenosine and 3.5 for cytidine in aqueous solutions. At 52°C the rates of reaction of poly rA and poly rC with chloroacetaldehyde were equivalent at a pH near 4.7 (Ruyechan, 1976). It seems quite likely that conditions can be found in which chloroacetaldehyde is a base independent reagent for modifying RNA.

3. Cleavage of single stranded RNA. Chemical modification by chloroacetaldehyde results in a region of single strandedness in the heteroduplex molecule. In DNA-DNA hybrids endonuclease S1 was used to cleave this region. Gonda et al. (1978) reported cleaving DNA-RNA hybrids using S1. A number of studies exist showing that S1 can be used to preferentially cut single stranded RNA (Vournakis et al., 1976, Flashner and Vournakis, 1977, Harada and Dahlberg, 1975, Wurst et al., 1978, and Wrede et al., 1979).

4. Cleavage with sequence specific enzymes. At present there are no known counterparts of restriction enzymes that act on double stranded RNA. However, this mapping system may not have to depend on sequence specific enzymes for its success. The double stranded RNA viruses exist in a naturally "cleaved" state in that they all have segmented genomes (Wood, 1973). Although the paramyxovirus genome is a single polynucleotide molecule, its molecular weight of around 7×10^6 daltons for most members (Davis et al., 1973) makes it small enough so that the two pieces of RNA created after S1 digestion should be able to enter a gel and be detected after electrophoresis without additional cleavage.

CHAPTER 4

DNA RENATURATION RATES: THE EFFECT OF CIRCULAR PERMUTATION

INTRODUCTION

DNA renaturation (reassociation) is a second order reaction characterized by a rate constant, k_2. This rate constant depends upon the length of the reacting single strands (L) and the complexity of the DNA (N) as well as upon various environmental factors such as temperature, viscosity, and ionic strength. For a detailed discussion of how these variables are related to the mechanism of DNA renaturation, see Wetmur (1976). The DNA renaturation reaction may be separated into a rate determining nucleation event, involving the formation of the first few base pairs, and a faster propagation reaction involving base pair formation to the end of the molecule. The greater the DNA complexity, the lower the concentration of a particular nucleation site at constant DNA nucleotide concentration. All other variables being constant, k_2 should be inversely proportional to N.

Because propagation proceeds to the maximum extent possible, k_2 should be directly proportional to L_s, the length of the shorter of two reacting complementary DNA

strands. The only exception should be a decrease in base pair formation of 33 percent when molecules of equal length but undetermined origin are allowed to react (Wetmur and Davidson, 1968). Thus one expects

$$k_2 = k_N \frac{L_s}{N} \quad (1)$$

where k_N is the nucleation rate constant which should depend only on environmental factors. Wetmur and Davidson (1968) and Wetmur (1971) observed, however, that

$$k_2 = k'_N \left(\frac{L_s}{N}\right)^{0.5} \quad (2)$$

where k'_N is a length-independent nucleation rate constant. The unexpected nature of the dependence of k_2 on L was proposed to be a manifestation of an excluded volume effect. Although no experimental evidence has been found which contradicts the excluded volume hypothesis, neither does there exist any evidence requiring an excluded volume effect on DNA renaturation. In this and the subsequent chapter we compare DNA reassociation reactions which differ only in the topology of the reacting strands. We also present a quantitative excluded volume theory consistent with both the observations relating k_2 to strand topology and previous observations relating k_2 to strand lengths. Taken together, these experimental and theoretical results greatly strengthen the argument for an excluded volume

effect on DNA renaturation reactions.

The work to be described in this chapter examines the effect of circular permutation of DNA molecules on their rate of renaturation. It was undertaken in order to help resolve a disagreement in the literature concerning the influence of excluded volume on the renaturation of circularly permuted DNA. In 1970 Lee et al. investigated some of the physical characteristics of the DNA of coliphage 15 of E. coli. In addition to its other properties Lee et al. (1970) found that the DNA was circularly permuted and terminally redundant. Electron microscopy of denatured and renatured coliphage 15 DNA revealed that 40% of the renatured molecules were double stranded circles with single stranded branches. The methods used for detecting circular permutation were essentially those first used by MacHattie et al. (1967). Circular molecules of renatured coliphage 15 DNA arose as the result of renaturation of two strands which began at different points in the DNA sequence. The separation of the branches is a reflection of the distance between the starting points in the renatured duplex. The shorter the interbranch distance, the closer together the two starting points, i.e., the more in-register the two renaturing linear molecules were initially. Measurements of the longer distance between the two branches of 150

circular molecules resulted in a distribution skewed in favor of the renaturation of closely in-register molecules. This could indicate either that the starting population of the DNA was itself biased in its collection of circular permutations or that starting with a perfectly random set of circularly permuted molecules, some factor favored renaturation of molecules that were closely in-register. Lee et al. (1970) favored this later hypothesis. They reasoned that the skewed distribution was due to the phenomenon of excluded volume. Excluded volume implies that points near the center of a random coil are less available to participate in the initial nucleation event in renaturation than are points close to the outside. Random walk theory shows that it is the topological ends of the random coil rather than the topological center that are more likely, on the average, to be found on the outside of the coil, and therefore be available for reacting with complementary DNA strands. With respect to the set of circularly permuted coliphage 15 DNA molecules, for two complementary strands with widely spaced beginning points, the end of one molecule will be complementary to the center of the other and vice versa (see Figure 16a). On the other hand, for strands with close beginning points the topological ends are homologous (see Figure 16b). According to the elementary

A.

B.

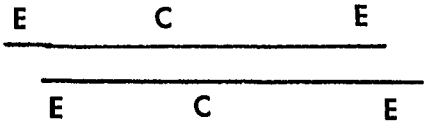


Figure 16. Schematic representation of DNA duplexes with different beginning sequences.

A. Molecules with widely spaced starting points.

B. Molecules with closely spaced starting points.

E = End, C = Center

excluded volume argument presented above, because of their accessibility and ease of interpenetration these later two strands will nucleate more easily and therefore renature more rapidly than the former pair. It is this later population of molecules that is most heavily represented in the skewed distribution of coliphage 15 DNA renatured molecules. Two questions that arise in response to this argument are (1) is there independent verification that the starting population of coliphage 15 molecules really is randomly distributed with respect to permutation and (2) is it true that renaturation of in-register molecules proceeds at a faster rate than out-of-register molecules. I will reserve comment on these questions for the Results and Discussion section.

Tye et al. (1974a) used the DNA from Salmonella phage P22 for their work. This DNA, like that of coliphage 15, is circularly permuted and has terminally redundant ends. They observed that in the majority of wt P22 chromosomes only 20% of the DNA ever appeared in an end position. The extent of permutation, in other words, was quite limited. They present evidence that DNA packaging is not totally random, but rather is initiated at a specific site on the precursor concatameric DNA and proceeds unidirectionally and sequentially. One possible explanation for such a

packaging mechanism would involve sequential encapsulation from the end of a concatemer replicated from a rolling circle (Gilbert and Dressler, 1968) with a unique origin of replication. Given a small terminal repetition (2% of the genome, or about 800 bases for P22) and a concatemer (an end-to-end polymer of the phage genome) only a few headfuls long (no more than 10 for phage P22), then sequential cutting from a unique site will result in limited circular permutation. Streisinger et al (1967) and Tye et al. (1974b) showed that a deletion in the phage P22 genome is compensated for by a lengthening of the region of terminal repetition by an amount equal to the length of the deletion. Deletion mutants offer a way to obtain a more representative collection of circularly permuted DNA. Phage P22bp1 with a net deletion of approximately 5% (Tye et al., 1974a) has 60 to 70% of its genome positioned at the ends in a representative collection of molecules Tye et al. (1974a) used P22 wt, and deletion mutants P22bp1 and P22bp5 (14% deletion) to perform the same type of denaturation-renaturation experiment as Lee et al. (1970) did. Histograms of the circular molecules were made by dividing the longer distance around the circle between the two branch points by the circumference of the molecule. They found a skewed distribution for wt P22 in which molecules

with close branch points were overrepresented, just as Lee et al. (1970) found for coliphage 15. However, Tye et al. (1974a) could now unambiguously attribute this nonrandom distribution to the nonrandom distribution of circular permutations in the renaturing population. In the case of the two deletion mutants known to constitute a more random collection of circularly permuted molecules, the histograms showed a uniform (non skewed) distribution of branch distances. Tye et al. (1974a) then recalled the excluded volume argument of Lee et al. (1970), i.e., that because of the effect of excluded volume in a population of circularly permuted molecules, those strands with beginning points near each other will preferentially nucleate and renature, giving rise to a skewed distribution. Since the random set of circularly permuted molecules of Tye et al. (1974a) (P22bp1 and bp5) did not give rise to such a skewed population, they concluded that excluded volume did not affect DNA renaturation rates. Here, then, is the point of contention.

The experiments to be discussed in this chapter were designed to examine the question of whether the extent of staggering of beginning points in renaturing DNA molecules affects their rate of renaturation and whether this is a valid measure of the effect of excluded volume on DNA

renaturation rates.

MATERIALS AND METHODS

Bacteriophage and virus. Bacteriophage øx174 RF I DNA and ^{3}H-labelled SV40 DNA component I were from Bethesda Research Labs.

Restriction endonucleases and buffers. Restriction endonucleases Ava I and II were purchased from New England Biolabs. Eco Rl, Hpa II, and Bam Hl were from Bethesda Research Labs.

øx174 RF I DNA was digested with either Ava I or Ava II; SV40 DNA was digested with either Eco Rl, Hpa II or Bam Hl. Appropriate DNAs were digested using 1.5 units of restriction enzyme per microgram of DNA at 37°C for one hour. The buffer used for each restriction enzyme is as follows:

Ava I, Ava II - 60 mM NaCl, 6 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 6 mM 2-mercaptoethanol, 100 µg BSA/ml

Eco Rl - 50 mM NaCl, 100 mM Tris-HCl (pH 7.2), 5 mM $MgCl_2$, 2 mM 2-mercaptoethanol

Hpa II - 20 mM Tris-HCl (pH 7.4), 7 mM $MgCl_2$, 1 mM dithiothreitol

Bam H1 - 20 mM phosphate (K^+) (pH 7.0), 100 mM NaCl, 7 mM $MgCl_2$, 2 mM 2-mercaptoethanol

Digestion was terminated by adding EDTA and NaCl to final concentrations of 0.02 M and 0.2 M respectively. The DNA was extracted with phenol using the procedure of Mandell and Hershey (1960) followed by 5 chloroform extractions. Dialysis was overnight at 4^oC in DNA buffer (Chapter 2).

Nucleic acid concentrations. Double-stranded DNA concentrations were determined spectrophotometrically using a Beckman Model 25 recording spectrophotometer. Relative concentrations of various circular permutations of SV40 DNA produced by restriction endonuclease digestion were determined by scintillation counting of aliquots of the sample in a Beckman LS 9000.

Single strand breaks. The SV40 starting DNA was determined to be primarily covalently closed circular DNA by gel electrophoresis. After linearization of øx174 RF and SV40 DNA with a restriction endonuclease these DNAs were denatured and renatured as described below, mounted for electron microscopy using the aqueous Kleinschmidt technique described by Davis, Simon and Davidson (1971), and observed in an AEI Model EM 801 electron microscope. Photographs were taken at a magnification of 6300x. The

photographic negatives were displayed on a blackboard with a lantern-slide projector and the lengths of the DNA molecules were traced with a map measuring device (Keufel and Esser). Means and standard deviations of the lengths were determined using standard statistical analysis. Only enzymes leaving DNA in good condition were used in subsequent experiments.

Denaturation-Renaturation. Either two aliquots of SV40 DNA made linear by cleavage with different single site restriction enzymes were mixed in a ratio of 1:1, or two aliquots of øx174 DNA made linear by cleavage with different single site restriction enzymes were mixed in a ratio of 1:1, or control samples of each restricted linear DNA by itself were treated as follows: EDTA was added to a final concentration of 0.01 M, and the solution was warmed at 37°C for 5 minutes. One-tenth volume of 1 M NaOH was added to denature the DNA. After 5 minutes at 37°C, 0.1 volume of 2 M Tris-HCl was added to neutralize the solution. Conditions for renaturation are given for each experiment in Table 7. Control samples consisted of each øx174 RF sample denatured and renatured by itself.

THEORETICAL

General aspects of the theory The derivation of

excluded volume effects on DNA renaturation rates (below) involves the use of the distribution (probability) functions for end-to-end distances of polymers (Tanford, 1961). The random flight distribution function, W(h), for end-to-end distance h is graphed for a linear polymer in Figure 17. It can be seen that end-to-end distances approaching zero, i.e., a circular molecule, as well as end-to-end distances approaching the contour length of the molecule, i.e., a rigid rod, are improbable. The equation which is graphed in Figure 17 is,

$$W(h)\,dh = 4\pi\left(\frac{3}{2\pi\,\langle h^2\rangle}\right)^{3/2}\left(e^{-3h^2/2\,\langle h\rangle}\right)(h^2 dh) \qquad (7)$$

where $\langle h^2\rangle$ is the mean square end-to-end distance. For molecules consisting of σ segments, each of length l_e, the contour length of the molecule, L, is σl_e. Let $\langle h_{mn}^2\rangle$ be the mean square distance between any two points m and n on the chain. It can be shown (Tanford, 1961) that

$$\langle h^2\rangle = L l_e = \sigma l_e^2 \qquad (8a)$$

or, more generally

$$\langle h_{mn}^2\rangle = L_{mn} l_e \qquad (8b)$$

A covalently closed circular molecule may be described (Sharp and Bloomfield, 1968) by the same distribution function W(h), where however,

$$\langle h_{mn}^2\rangle = L_{mn}\left(\frac{L - L_{mn}}{L}\right) l_e \,. \qquad (9)$$

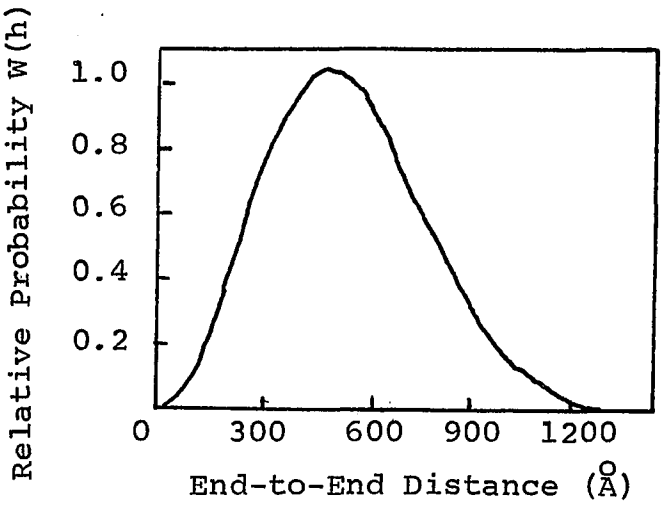


Figure 17. The distribution function W(h) for end-to-end distances of a polymer. Relative probability W(h) is in arbitrary units. The calculation is for a DNA chain of 1000 bases in a denaturing solvent. The maximum in the plot occurs at

$$h_{max} = \sqrt{\frac{2}{3} L l_e}$$

L = number of bases x base spacing. Assume 1000 bases, 5Å per base. So L = 5 x 10^{-5}cm

l_e = number of segments. About 8 x 10^{-7} cm (Wetmur,1971).

h_{max} is about 500Å.

Adapted from Tanford, 1961.

Consider two polymer chains (1 and 2) which meet at an origin O (see Figure 18a). The probability of overlap of two segments L_m and L_n bases distant from O on the respective chains is given by

$$\theta_{m,n} = V^* 4\pi \int_0^\infty \frac{W_m(h) \cdot W_n(h) \cdot h^2 dh}{(4\pi h^2)^2} \qquad (10a)$$

Solving (10a) we find

$$\theta_{m,n} = V^* \left(\frac{3}{2\pi l_e}\right)^{3/2} \left(\frac{1}{L_m+L_n}\right)^{3/2} = V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \left(\frac{1}{\sigma_m+\sigma_n}\right)^{3/2} \qquad (10b)$$

If equation 10b is true in general it must also be true for the special case where L_m and L_n add up to a single polymer chain ($L_m + L_n = L$). Equation 10b may then be simplified to:

$$\theta \text{ ends} = V^* \left(\frac{3}{2\pi L l_e}\right)^{3/2} = V^* j \qquad (11)$$

where V* is the excluded volume of an end segment and j is the Jacobson-Stockmayer factor. Wang and Davidson (1966b) experimentally verified the Jacobson-Stockmayer factor for the interaction of the ends of a lambda DNA molecule. The Jacobson-Stockmayer factor must be taken into account when doing DNA ligation for cloning. We are aware that intramolecular excluded volume will affect the chain statistics, but for simplicity we continue without taking this factor into account. In fact, the magnitude of intramolecular excluded volume for single stranded DNA under renaturation conditions is unknown.

Interactions between two linear molecules. Now consider four chains of a,b,c, and d segments respectively, meeting at O (see Figure 18b). This point of intersection is going to be the nucleation site for DNA renaturation. Let two chains be connected to produce $\sigma_m = a+b$. Let the other two chains be connected to produce $\sigma_n = c+d$. The total probability of new overlaps resulting from this requirement for intersection of the two polymers at O may be determined by integrating equation 10b over all of the possible overlaps.

$$\theta = V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \left[\left(\int_o^a d\sigma_m + \int_o^b d\sigma_m \right) \left(\int_o^c \sigma_n + \int_o^d \sigma_n \right) \left(\frac{1}{\sigma m + \sigma n} \right)^{3/2} \right] \qquad (12)$$

Evaluation of equation (12) gives

$$\theta = 4V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \left[2 \left(\sqrt{a} + \sqrt{b} + \sqrt{c} + \sqrt{d}\right) - \left(\sqrt{a+c} + \sqrt{a+d} + \sqrt{b+c} + \sqrt{b+d}\right) \right] \qquad (13)$$

Equation (13) will be used for all θ calculations involving linear DNA chains.

Effect of length on DNA renaturation. Let us consider the evaluation of the overlap function (equation 13) for two DNA chains of identical number of segments (σ). As extreme cases we will consider the nucleation event for renaturation (the origin) to occur at the end of both molecules ($a = c = \sigma$, $b = d = 0$) and the nucleation event for renaturation to occur at the middle of both molecules ($a = b = c = d = \frac{\sigma}{2}$). The results of overlap calculations

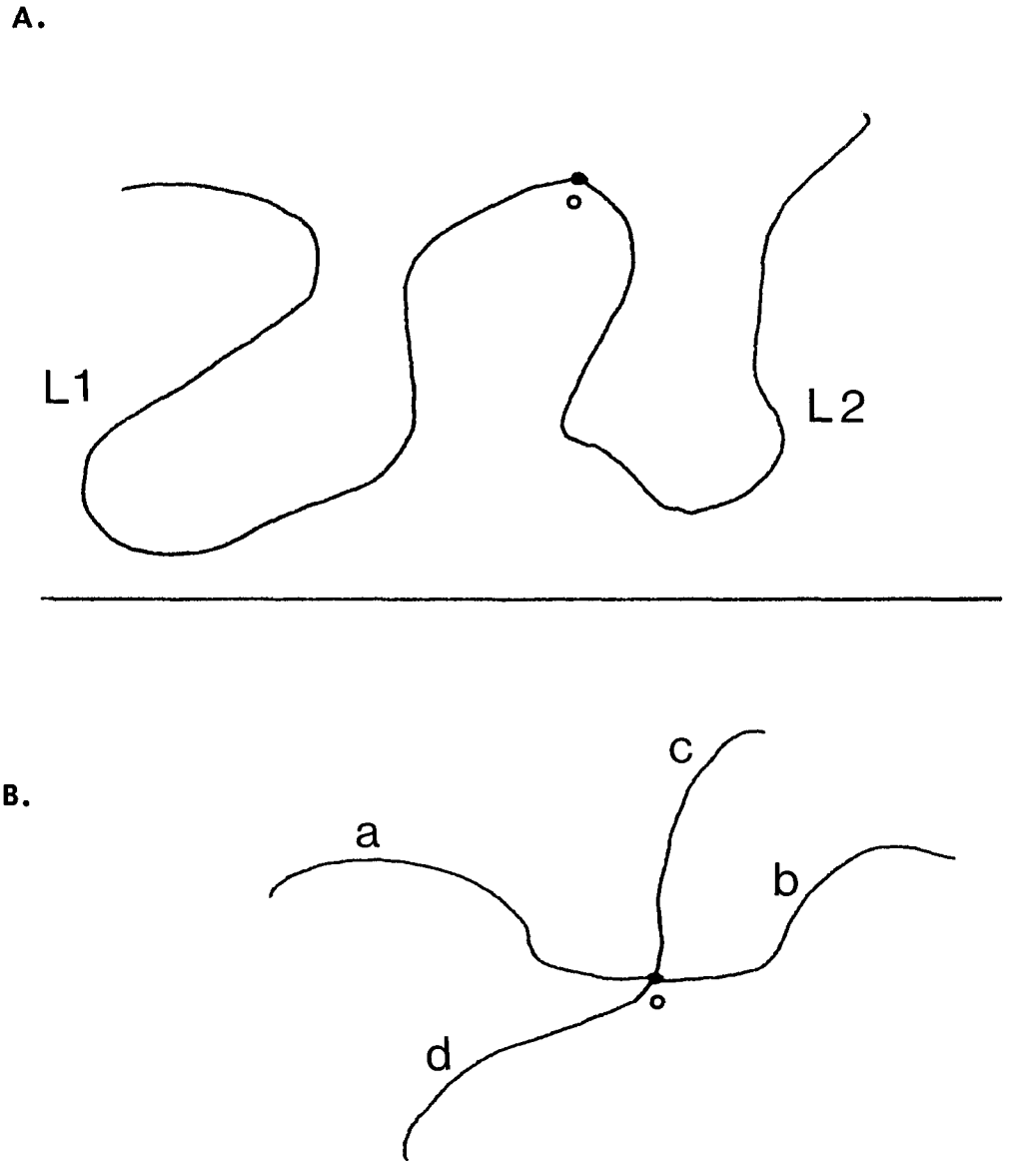


Figure 18. Schematic representation of polymer chains meeting at origin 0.

A. Two polymer chains

B. Four polymer chains

using equation 13 are:

$$\theta \text{ center} = \left[4V^*\left(\frac{3}{2\pi l_e^2}\right)^{3/2}\right] 1.6\sqrt{\sigma} \qquad (14a)$$

$$\theta \text{ ends} = \left[4V^*\left(\frac{3}{2\pi l_e^2}\right)^{3/2}\right] 0.6\sqrt{\sigma} \qquad (14b)$$

The total number of overlaps possible in solution cannot be increased by connecting the segments of a polymer together to form chains. Therefore, let the probability of finding any conformation with a particular origin configuration be proportional to $\frac{1}{1+\theta}$. The probability of reaction at the origin is then given by $\frac{\theta_1}{(\theta+1)}$ where θ_1 is $4V^*\left(\frac{3}{2\pi l_e^2}\right)^{3/2}$. Referring back to equations (1) and (2), we find that

$$k_N = \frac{k_N' \theta_1}{\sqrt{s}\ (\theta + 1)} \qquad (15)$$

where s is the number of bases in a segment. For the renaturation reactions with nucleation events at the centers or ends of the molecules we find that

$$k_N \approx \frac{k'_N}{\sqrt{L}} \cdot \frac{1}{1.6} \qquad (16a)$$

for center reactions, and

$$k_N \approx \frac{k'_N}{\sqrt{L}} \cdot \frac{1}{0.58} \quad \text{for end reactions.} \qquad (16b)$$

The predicted dependence of k_N on $\sqrt{L}$ is precisely that seen experimentally by Wetmur and Davidson (1968) and scales the same as previous excluded volume calculations

by Wetmur (1971) for DNA renaturation kinetics. The difference between reaction rates for nucleation at the end or center of DNA molecules has not been observed, but will be discussed below.

Reactions between short and long DNA strands. Evaluation of equation (13) with long and short (L_s) strands (see Appendix B) gives, for the same cases of reaction at the end and center,

$$k_N \approx \frac{k'_N}{\sqrt{L_s}} \cdot \frac{1}{2.8} \qquad (17a)$$

for center reactions and

$$k_N \approx \frac{k'_N}{\sqrt{L_s}} \cdot \frac{1}{1} \quad \text{for end reactions.} \qquad (17b)$$

This dependence on L_s is exactly as seen experimentally by Wetmur (1971) and similar to that seen experimentally by Hinnebusch, Clark, and Klotz (1978). This result will be considered again in the discussion below.

Circular permutations. One of the experiments described below deals with reactions between linear single stranded DNAs which are circularly permuted. Integration of equation (13) over all possible nucleation sites and division by the number of nucleation sites gives the average value of θ, $<\theta>$. $<\theta>$ depends on L and P where P is the number of bases by which the reacting strands are out of phase. We find (see Appendix C for derivation of

equation 18)

$$\langle\theta\rangle = \frac{C}{L}\left\{8L^{3/2}-(2L-P)^{3/2}+P^{3/2}+(L-P)^{3/2}-(L+P)^{3/2}-\frac{3}{2}(L-P)^{3/2}-\frac{3}{2}(L+P)^{1/2}(L-P)-\frac{3}{2}P(2L-P)^{1/2}-\frac{3}{2}P^{3/2}\right\} \quad (18)$$

where C is a constant. Evaluation of $\langle\theta\rangle$ for P = 0 and $P = \frac{L}{2}$ are almost identical (see Appendix C). This small difference would be difficult to determine experimentally. This means that we predict that DNA renaturation rates between circularly permuted molecules will not depend on the extent of the circular permutation (phase). Results presented in this chapter are in agreement with this prediction.

Circular molecules. We have reevaluated θ for the case where one molecule is linear and one molecule is circular. The circular molecule is described by the same distribution function W(h) but with altered $\langle h^2\rangle$ (see equation 9). The result of this calculation (see Appendix D) for a linear molecule interacting with a circular molecule is

$$\theta_{Circ} = 4V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \cdot \sqrt{\theta}\left\{\pi-\sin^{-1}\left(\frac{1}{\sqrt{1+\frac{4a}{\sigma}}}\right)-\sin^{-1}\left(\frac{1}{\sqrt{1+\frac{4b}{\sigma}}}\right)\right\} \quad (19)$$

where the linear strand has σ segments, the same as the circular strand, and $\sigma = a + b$. Again, considering end and center nucleations for the linear strands we find

$$k_N \approx \frac{k'_N}{\sqrt{L}} \cdot \frac{1}{1.91} \quad \text{for center reactions, and} \tag{20a}$$

$$k_N \approx \frac{k'_N}{\sqrt{L}} \cdot \frac{1}{1.11} \quad \text{for end reactions.} \tag{20b}$$

Thus the rate of renaturation of circular or linear molecules with linear molecules may be compared by looking at equations (16a,b) and (20a,b). For experiment to agree with theory, circular-linear reactions must be slower than linear-linear reactions. We shall see in Chapter 5 that this is indeed the case.

The theory above deals with various effects of strand length and strand topology on DNA renaturation rates. Applicability of this theory to other polymer systems will be considered in the discussion.

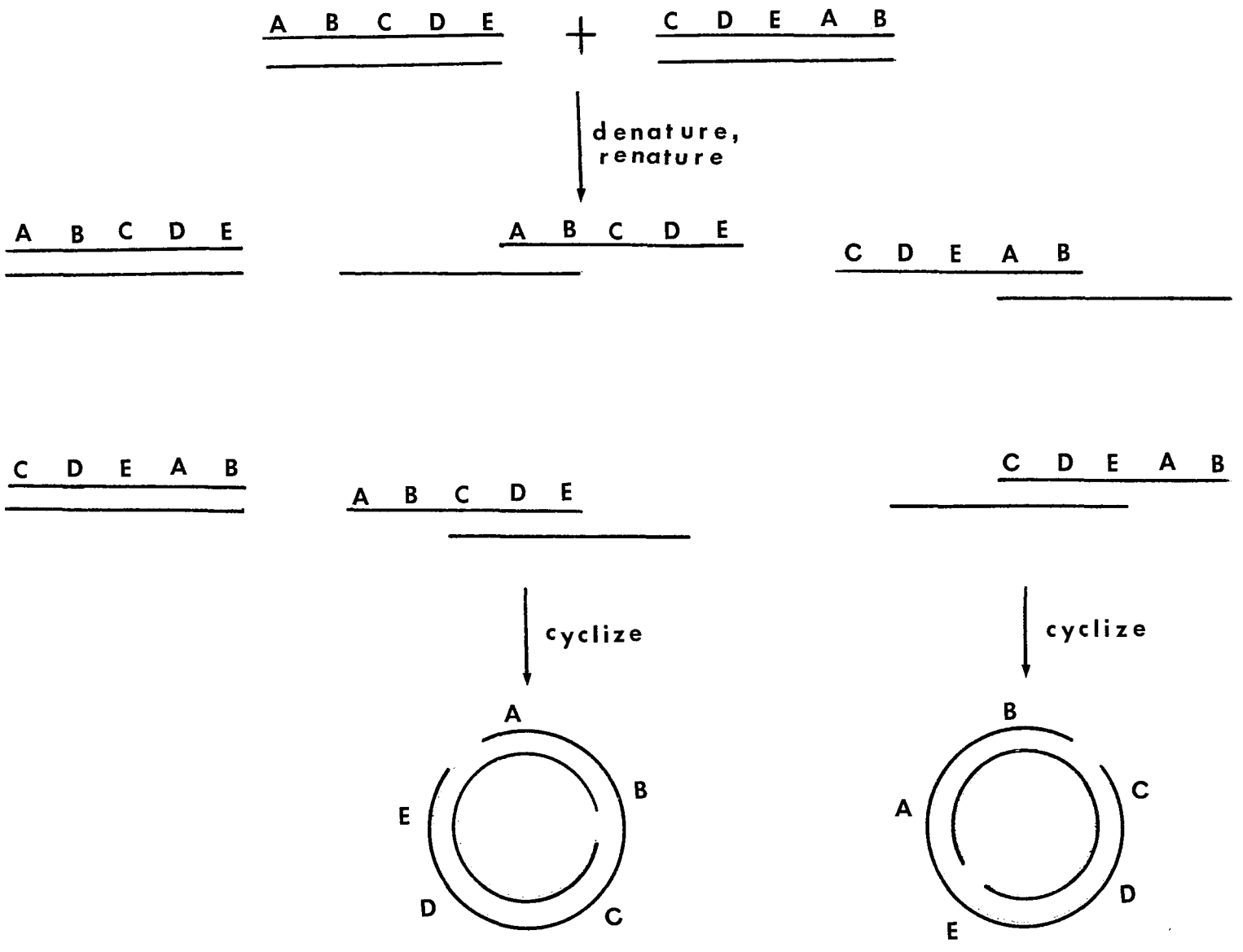
RESULTS AND DISCUSSION

There are a number of restriction enzymes which have only a single site on SV40 or øX174 DNA. Because the entire DNA sequence has been determined for øX174 and SV40 (Sanger et al., 1978, Reddy et al., 1978), and the recognition sequence is available for each of the enzymes used here (Roberts, 1980), the exact location of cleavage by each of these enzymes on øX174 or SV40 is known. A number of these enzymes which act at a single site have been used to prepare circularly permuted DNA molecules.

The products of denaturing and renaturing a mixture of circularly permuted DNA molecules are illustrated in Figure 19. Renaturation between complementary DNA strands which are in-phase results in the formation of fully duplex linear molecules. When two complementary, but out-of-phase, DNA strands renature with one another, the initial product is a partially duplex molecule with complementary single stranded tails; the final product is a double stranded circle with two staggered nicks. The experiments to be described here used dilute DNA solutions to promote cyclization of partially duplex molecules before either of the tails can react with a third strand or with another partially duplex molecule. The requirements for such conditions may be calculated from the results of Wang and Davidson (1966a,b).

Restriction maps. Maps of øX174 and SV40 DNA showing the sites of the relevant restriction enzymes are presented in Figure 20a and b respectively. In øX174 the 5' terminus of the recognition sequence for Ava I is at nucleotide 162; for Ava II it is at nucleotide 5042. The zero position is the Pst I cleavage site, and the total number of nucleotides in the DNA is 5386 (Fuchs et al., 1978). In SV40 the 5' terminus of the recognition sequence for Eco R1 is at nucleotide 1700, for Hpa II it is at 264, and for Bam H1 it is at 2451. The zero position is near the origin of

Figure 19. Protocol for testing the effect of circular permutation on DNA renaturation rates. The products of the renaturation reaction are shown. The double stranded linear starting molecules were generated from the same pool of øx174 RF I or SV 40 molecules, but were cleaved with different single site restriction enzymes.


A B C D E
C D E A B
denature, renature
A B C D E
A B C D E
C D E A B
C D E A B
A B C D E
C D E A B
cyclize
cyclize
A
B
E
C
D
B
A
C
D
E

DNA replication and the total number of nucleotide pairs in the DNA is 5226 (Fuchs et al., 1978). The Eco R1-cut and the Hpa II-cut SV40 DNAs are 27 percent out of phase with one another; the Hpa II-cut and the Bam HI-cut SV40 DNAs are 42 percent out of phase with each other; the Ava I and Ava II-cut øX174 DNAs are 9 percent out of phase with one another. These three combinations span most of the possible range of phases which would be found in a sample with random circular permutation.

Validity of the electron microscopic assay. The electron microscope is used to quantify the products of various denaturation-renaturation mixtures. In order to be certain that electron microscopy is free from bias and accurately reflects the number of linear and circular molecules in a solution, an equimolar mixture of Eco RI cut ^{3}H-labeled SV40 linear DNA and open circular ^{3}H-labeled SV40 DNA prepared by DNase I nicking was made. The 1:1 mixture was assured by scintillation counting of aliquots of each DNA. To prevent operator bias during sample examination in the electron microscope, adjustments were made at one grid square before translation to an adjacent square for photographing before observing the field. Table 7, line 1 confirms that neither the linear nor the circular DNA is preferentially

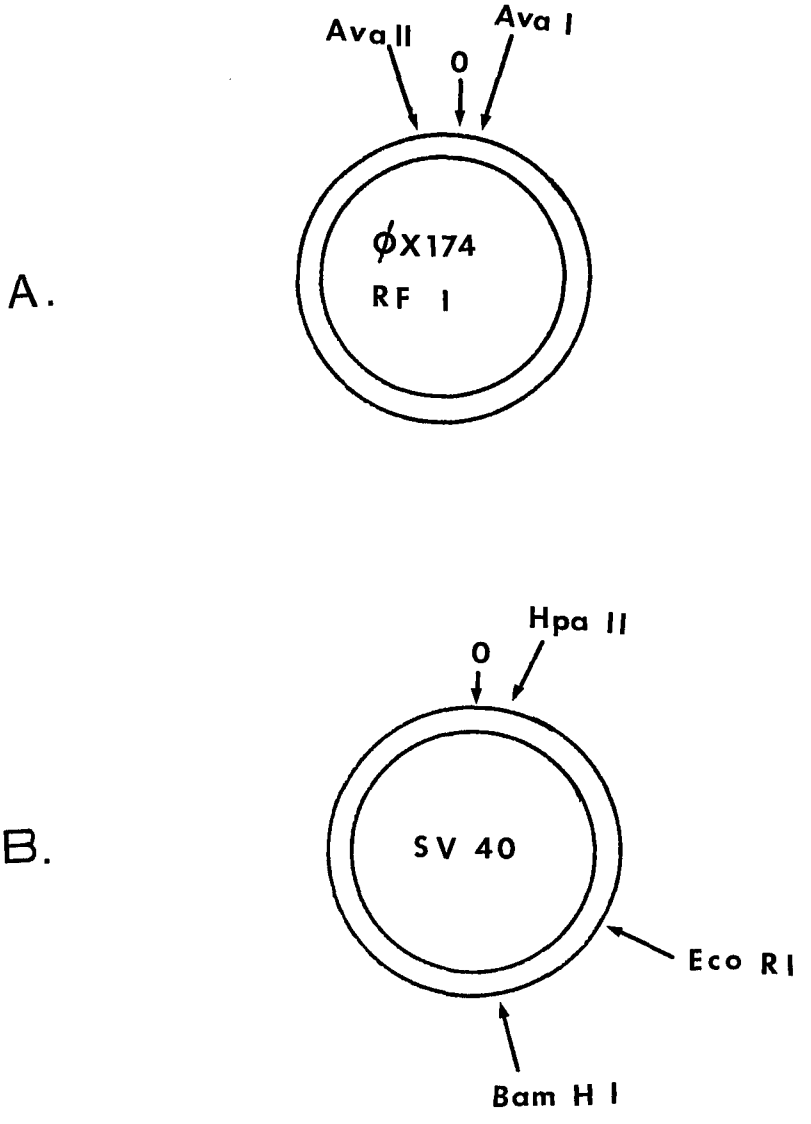


Figure 20. øx174 and SV 40 selected restriction enzyme maps.

A. øx174 RF I DNA.

B. SV 40 DNA

detected using the electron microscopy procedure described here.

Integrity of DNA following restriction enzyme digestion. Samples of linear øX174 RF DNA with different beginning sequences are generated by cleaving one aliquot of øX174 RF I DNA with restriction enzyme Ava I and another aliquot with restriction enzyme Ava II. Samples of ^{3}H-labeled SV40 linear DNA are generated in exactly the same way using restriction enzymes Eco Rl, Hpa II, or Bam Hl. Histograms of the DNAs cut with the restriction enzyme were determined both before and after a denaturation-renaturation cycle. The mean and standard deviation values for øX174 DNA cut with three different restriction enzymes is given in Table 6.

The length distribution after digestion with Pst I but before denaturation and renaturation is composed of shorter-than-full-length DNA. This indicates that some DNA degradation accompanied digestion with Pst I. The further decrease in average DNA length following denaturation-renaturation suggests that single strand breaks had been introduced into the molecule. Pst I was not used any further in these experiments.

It is of interest to note here that restriction enzyme

TABLE 6

MEAN LENGTH AND STANDARD DEVIATION OF ØX174 DNA CLEAVED BY DIFFERENT RESTRICTION ENZYMES

Restriction Digestion	Number of Molecules	Mean Length (μ)	Standard Deviation (μ)
Ava I, not D/R‡	118	1.47	0.24
Ava I, D/R	110	1.30	0.36
Ava II, not D/R	90	1.51	0.19
Ava II, D/R	106	1.20	0.33
Pst I, not D/R	46	1.05	0.22
Pst I, D/R	100	0.64	0.31

‡D/R = denatured, renatured as described in Materials & Methods section.

Sst II was also tried on øX174 DNA. According to Fuchs et al. (1978), there should be a single Sst II site in øX174 at nucleotide 2862. In our hands the Sst II did not cut our preparation of øX174 DNA, although it was active on lambda DNA (data not shown). Restriction enzyme Sst II was not used any further in these experiments.

Effect of circular permutation on DNA renaturation rates. To determine the effect of circular permutation on DNA renaturation rates equimolar amounts of circularly permuted øX174 linear DNAs are mixed together and denatured, or two of the circularly permuted SV40 DNA samples are mixed together and denatured. Renaturation conditions for each trial are given in Table 7 along with the fraction of the molecules that were observed in the electron microscope to be circles.

Mixtures of SV40 DNA molecules that are either 42% (Table 7, line 2) or 27% (Table 7, line 5) out of register produced essentially the same number of linear and circular molecules when renatured in 1 M NaCl and sufficient time (about 5 half times) to assure renaturation and almost complete cyclization. øX174 molecules 9% out of register (Table 7, line 10) produced 44% circles. The decrease in the fraction of øX174 circular molecules compared to SV40 may reflect the somewhat less favored cyclization of a

molecule that has single strand tails at each end amounting to only 9% of the total length of the molecule, and the fact that more stringent renaturation conditions were used. Two minutes for renaturation in 0.4 M NaCl is only 1 half time for øX174. In light of the qualifications, these three data points are in good agreement with one another. In another trial with SV40 DNA, conditions for renaturation were chosen that might be expected to enhance any excluded volume effect (Table 7, lines 3 and 6). Lower salt (0.15 M) has the effect of permitting the negatively charged DNA strands to repel each other more, and the lower temperature of 39°C slows down renaturation. The lower rate of renaturation than at Tm - 25°C may itself reflect an excluded volume phenomenon. Even with these renaturation conditions the relative populations of linear and circular molecules are unaffected by circular permutation of 27% or 42%. Finally, the results of renaturing SV40 DNA in 50% formamide, 0.2 M Tris-HCl, 0.01 M EDTA pH 8 (Table 7, lines 4 and 7, are consistent with little or no effect of 27% or 42% circular permutation on DNA renaturation rates. Taken together, the SV40 and øX174 data indicate that circular permutation due to phase shifts ranging from 9% to 42% (out of a possible 50%) does not affect the rate of DNA renaturation. Furthermore, neither salt concentration,

TABLE 7

RENATURATION OF CIRCULARLY PERMUTED DNA

DNA Source		Renaturation Conditions of 2 μg DNA/ml			
		Salt	T (°C)	t (min.)	f_c
SV40 Eco RI, DNase I	(1)	Not denatured / renatured			0.490 ± 0.020
SV40 Hpa II, Bam HI	(2)	1 M	65	5	0.500 ± 0.030
	(3)	0.15 M	39	60	0.440 ± 0.030
	(4)	50% $HCONH_2$	25	60	0.460 ± 0.030
SV40 Eco RI, Hpa II	(5)	1 M	65	5	0.490 ± 0.020
	(6)	0.15 M	39	60	0.480 ± 0.020
	(7)	50% $HCONH_2$	25	60	0.520 ± 0.050
øX174 Ava I	(8)	0.4 M	70	2	0.013 ± 0.007
Ava II	(9)	0.4 M	70	2	0.010 ± 0.006
Ava I, Ava II	(10)	0.4 M	70	2	0.440 ± 0.025

Measure N molecules. f_c^m = fraction found circular

$$f_c = f_c^m \pm \sqrt{f_c^m (1-f_c^m)/N}$$

temperature, nor use of organic solvents alters this conclusion.

Figure 21 combines the results of Lee et al. (1970), Tye et al. (1974a), and this work on one graph for purposes of comparison. Our approach to examining the effect of circular permutation of DNA renaturation rates was to use DNA samples that have been made linear with the use of restriction enzymes. This is in contrast to previous work which relied on obtaining a random collection of circularly permuted molecules by extracting DNA from phage known to contain a linear chromosome with a circular genetic map. The advantage of our strategy is that we are able to clearly define our starting material in terms of beginning points, and we are able to control the proportions of each DNA component in the renaturation mixture. Examination of the summary graph (Figure 21) indicates that our results are consistent with those of Tye et al. (1974a). It seems likely, based on what is now known about the packaging of the DNA of phage with circularly permuted genomes, that the explanation for the findings of Lee et al. (1970) is that the wild type population of coliphage 15 DNA was not a truly random collection of circularly permuted molecules. In addition, our results and those of Tye et al. (1974a) suggest that the effect of excluded volume on DNA

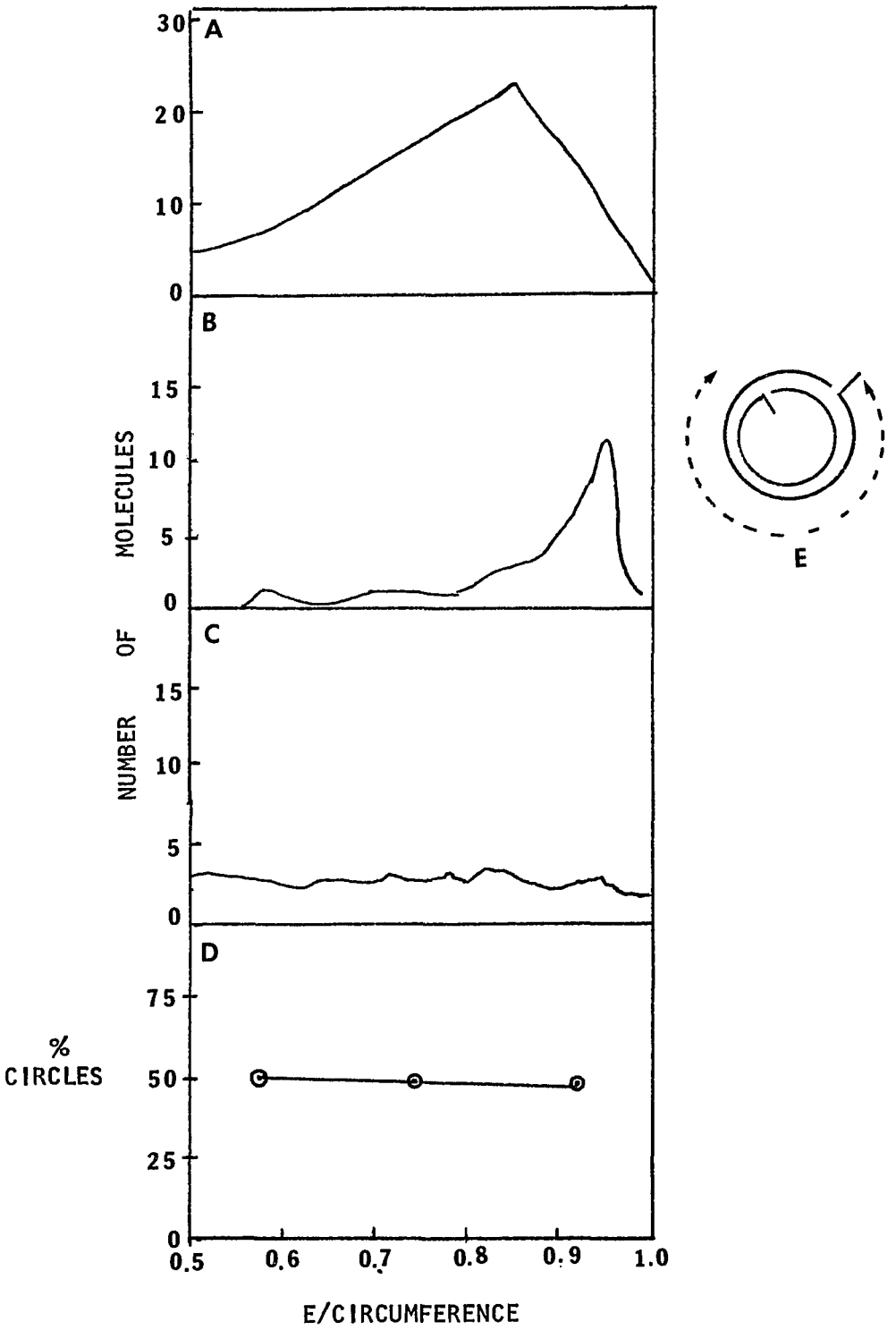
Figure 21. Composite graph of the effect of circular permutation on DNA renaturation.

A. Coliphage 15 DNA. Taken from Lee et al., 1970.

B. P22 wt DNA. Taken from Tye et al., 1974a.

C. P22 bp5 DNA. Taken from Tye et al. 1974a.

D. øx174 DNA. This work.


A
B
C
D
30
20
10
0
15
10
5
0
15
10
5
0
75
50
25
0
MOLECULES
NUMBER OF
%
CIRCLES
0.5
0.6
0.7
0.8
0.9
1.0
E/CIRCUMFERENCE
E

renaturation rates cannot be detected by comparing the rate at which different pairs of out-of-phase DNA molecules renature with one another. These results are in complete accord with the excluded volume theory resulting in equation 18 which indicates that the average overlap of two reacting DNA strands is essentially independent of the percent out-of-phase, 100P/L. A method that does measure an excluded volume effect on DNA renaturation rates is presented in Chapter 5.

CHAPTER 5

DNA RENATURATION RATES: THE EFFECT OF EXCLUDED VOLUME

INTRODUCTION

In the previous chapter, an excluded volume theory was proposed which made certain predictions concerning the effects of DNA strand topology on DNA renaturation rates. One of these predictions, the absence of an observable effect of circular permutation on DNA renaturation rates, was tested and found to be correct. A second prediction dealt with the effect of single stranded circular DNA structure on renaturation kinetics. In this case, a positive result is expected if the excluded volume argument is valid. The experiments described below, designed to test this hypothesis, provide evidence to support the excluded volume theory. Finally, certain predictions are made concerning the effects of single strand DNA lengths on DNA renaturation rates. These predictions are compared with previously reported experimental results in the Discussion. All of our experimental results as well as those in the literature have been found to be in agreement with the excluded volume theory presented in the previous chapter.

MATERIALS AND METHODS

Nucleic acids and restriction endonucleases. Bacteriophage øX174 RF I DNA was purchased from Bethesda Research Labs, and viral DNA was purchased from Miles Laboratories. Restriction endonucleases Ava I and II were purchased from New England Biolabs and used as described in Chapter 4.

Nucleic acid concentrations. A Beckman Model 25 double beam spectrophotometer was used to determine the absorbance at 260 nanometers (A_{260}) of various preparations of DNA.

A Gilford Model 2400 spectrophotometer with a Tamson constant temperature circulator, and attached automatic temperature programmer from NesLab Instruments, Inc. was used in the study to determine the correction factor for absorbance of single stranded DNA at A_{260}. The cuvette temperature, the reference line maintained by an automatic reference compensator, and the increase in A_{260} as the sample was heated were recorded. The procedure was as follows: The A_{260} at room temperature of an aliquot of øX174 viral DNA (which is single stranded) was recorded. The temperature was raised and the concomitant increase in absorbance was followed. When non-specific base interactions were fully melted out the absorbance stabilized

and its value was recorded. This value for fully denatured DNA was corrected back to native DNA by assuming the hyperchromic shift had been 36%. This is a relatively constant value for the melting of double stranded DNA (Wetmur, 1976). The viral DNA was in effect being treated as if it were one strand of a double stranded molecule. The ratio of the actual A_{260} of native øX174 viral DNA to the corrected native A_{260} provided the correction factor needed to determine the concentration of the single stranded viral DNA spectrophotometrically. This correction factor is salt dependent. In the salt concentration used for these experiments the correction factor was determined to be 1.325; that is, for single stranded viral DNA, an A_{260} of 1.325 is equal to 50 µg DNA/ml.

Single strand breaks. A Beckman Model E Analytical Ultracentrifuge equipped with ultraviolet optics was used to determine the percent of nicked single strand DNA in the øX174 viral DNA preparation. The sedimentation coefficients in alkali of øX174 single stranded circular and linear DNAs are 15s and 13s respectively (Pagano and Hutchison, 1971). Thirty µl of a solution containing 1.5 µg of øX174 viral DNA was loaded into the well of a band forming type III charcoal-filled Epon centerpiece of a 30 mm cell. Sedimentation was in alkaline cesium

chloride (3 M CsCl, 0.1 M NaOH, pH 13). Photographic negatives taken during centrifugation were scanned on a Canalco Model J Microdensitometer. The amount of breakage of the viral DNA was determined from the relative areas of the circular and linear single stranded peaks.

Denaturation and renaturation. Experiments to examine the rate of renaturation of a single stranded linear molecule with either its linear or its circular complement were performed as follows: øX174 RF I was digested with 1.5 units of Ava I per microgram of DNA at 37°C for 60 minutes. This enzyme cleaves RF I once, generating a linear duplex. Digestion was stopped by adding EDTA to a final concentration of 0.02 M and NaCl to a final concentration of 0.2 M. The DNA was extracted with phenol following the procedure of Mandell and Hershey (1960) in order to remove protein. Following overnight dialysis in the cold against DNA buffer (Chapter 2), an aliquot of this DNA was then mounted on parlodion covered grids for electron microscopy (AEI Model EM 801) using the Kleinschmidt technique as modified by Davis, Simon, and Davidson (1971) to determine the extent of restriction enzyme digestion. The remaining DNA was divided into two aliquots. øX174 viral DNA was added to one aliquot of linear double stranded øX174 DNA in the ratio of 1:2. Both aliquots were denatured with

alkali, final concentration 0.1 M with 0.01 M EDTA at 37°C for five minutes, and then neutralized with 0.1 volume of 2 M Tris-HCl, final pH 7.5. The DNA, with a plus strand concentration of about 2 μg/ml, was renatured at 70°C for 2 or 10 minutes. Aliquots of each sample were prepared for examination in the electron microscope. The number of circular and linear molecules was determined by inspection. Molecular lengths were calculated by first tracing photographs of the molecules.

RESULTS

The DNA of bacteriophage øX174 is available in two forms, as a single stranded covalently closed circle isolated from the virus and as a double stranded circle isolated from infected cells. The circular single stranded viral DNA has been completely sequenced (Sanger et al., 1977, 1978); it contains 5386 bases and has a molecular weight of 1.7×10^6 daltons. After infection this viral DNA is converted into a double stranded covalently closed circular form called RF I. One strand of RF I is identical to the viral DNA; the other strand is complementary to it. Both forms of øX174 DNA were used in these studies. In order to study the effect of circular versus linear DNA structures on the rate of renaturation of DNA, øX174

RF I DNA was cut with Ava I to produce linear molecules. Two parts of these molecules were mixed with one part of øX174 viral DNA, denatured and renatured. The reaction and expected products are outlined in Figure 22. With the two to one ratio of Ava I cut øX174 RF DNA to viral DNA the negative DNA strand from the restricted øX174 RF DNA would be as likely to encounter the positive linear complementary strand as to encounter the positive circular complementary strand of viral øX174 DNA. These reactions will lead to formation of linear and circular duplex molecules respectively. An aliquot of the same preparation of Ava I-cut øX174 RF I DNA was denatured and renatured in order to determine the extent of single strand breaks in the molecule.

Characterization of the nucleic acids. The electron microscopy of native DNA products was carried out using the same random sampling procedure described in Chapter 4; the extent of digestion of øX174 RF I DNA by Ava I was determined by electron microscopy and is reported in Table 8. This result was unaffected by denaturation and renaturation for slightly greater than one half-time. Digestion was from 97 to 98 percent complete. Analysis of histograms of the denatured and renatured molecules indicated few single-strand breaks in the restricted

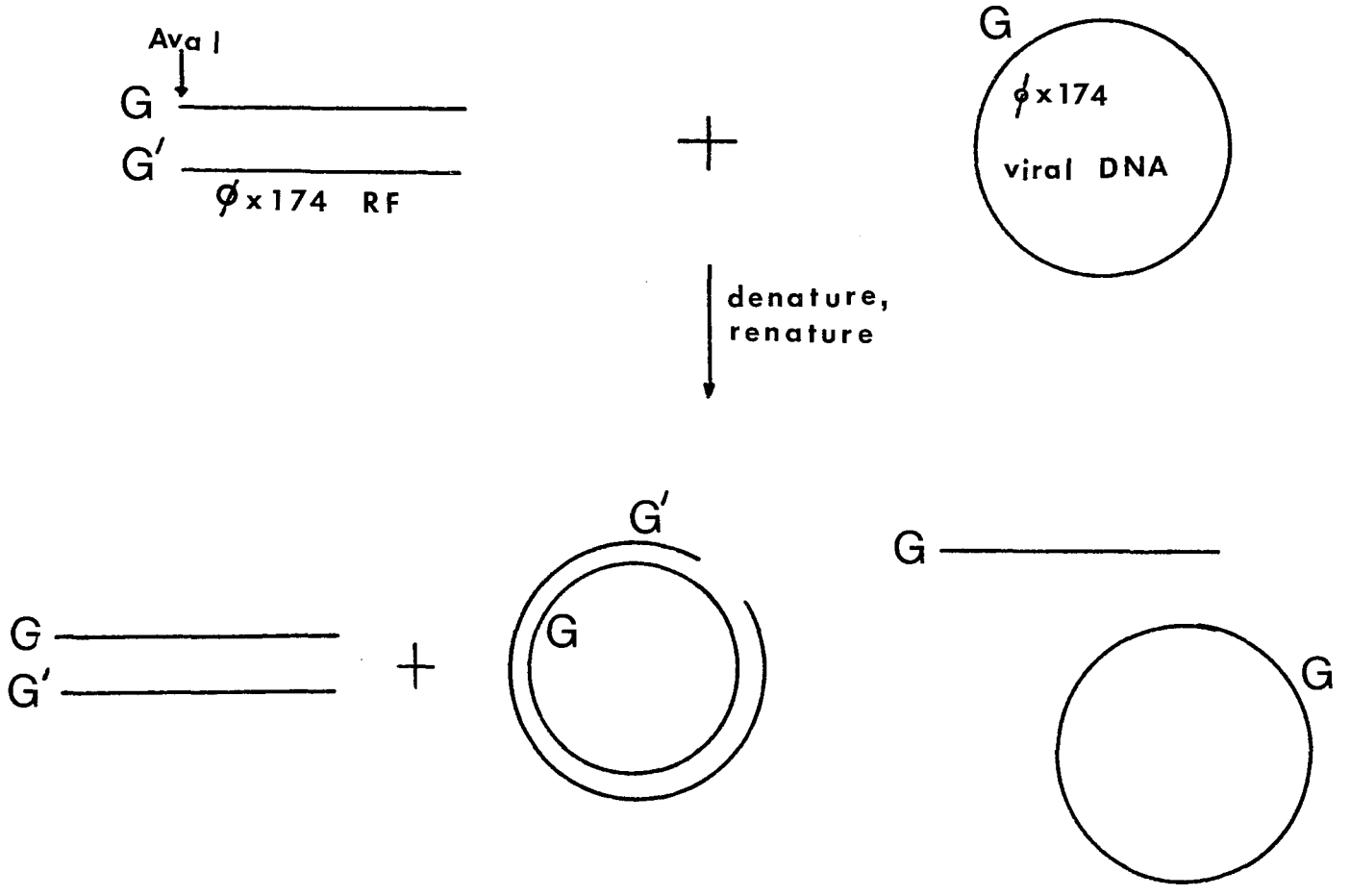


Figure 22. Protocol to examine the rate of renaturation of a linear molecule with its linear and its circular complements.

DNA. Histograms of Ava I-cut øX174 RF I DNA are analyzed in Table 6, Chapter 4. The histograms were essentially the same before and after denaturation and renaturation.

The øX174 viral DNA used in these experiments was determined by analytical band velocity sedimentation to be 81 percent intact. The remaining 19 percent of the DNA, containing one break at a random location, would be expected to renature with the rate of a linear DNA with complementary in-phase linear DNA but to cyclize rapidly to a circular form. These kinetics are the same as those of the circularly permuted DNAs studied in Chapter 4.

The double stranded linear RF I and the single stranded circular viral DNA are added to a reaction mixture in the ratio of 2:1. If this ratio is not achieved, then the distribution of double stranded linear and circular molecules obtained after renaturation could be attributed to the presence of unequal concentrations of reacting species in the mixture rather than to a difference in their abilities to renature with each other. DNA concentrations were measured spectrophotometrically using a correction factor which was determined for single stranded DNA hypochromicity by equating the absorbance of native and denatured single stranded DNA at temperatures above Tm.

Renaturation of linear and circular DNA. The results of two separate experiments with linear and circular DNA are given in Table 8. In one trial 123 out of 354 molecules were circles; in another 103 out of 275 molecules were circles. The technique used to mount the DNA on the grids does not spread single stranded DNA, so all the molecules seen are duplexes. The renaturation reactions were done in 0.4 M NaCl at 70°C for the indicated times. Incubation for slightly greater than one half time (2 minutes) resulted in the same fraction of circular molecules as incubation for more than five half times (10 minutes). The results clearly indicated a preference for reaction between linear molecules. The relative renaturation rate constants k_L/k_C, for renaturation of a linear molecule with complementary linear or circular molecules, are also given in Table 8. These rate constants are calculated by taking into account the fraction of RF I DNA uncut by Ava I, the fraction of viral DNA which is broken to yield linear molecules and the amounts of each strand remaining at all stages of the renaturation reaction. The derivation of equation 21 in Table 8 is given in Appendix E. This calculation indicates that reactions between linear molecules proceeds approximately three times faster than reactions between linear and circular molecules. This experimental

TABLE 8

LINEAR AND CIRCULAR MOLECULES COUNTED AFTER DENATURATION AND RENATURATION OF A MIXTURE OF øX174 RF I AND VIRAL DNA

DNA SOURCE	RENATURED	TIME (min.)	NUMBER OF MOLECULES Linear	Circular	f_c	k_L/k_c
øX174 RF I						
Ava I	No	---	484	13	0.026 ± 0.007	---
Ava I	Yes	2	410	9	0.021 ± 0.007	---
Ava I, øX174 v. DNA	Yes	2	231	123	0.350 ± 0.030	3.8
øX174 RF I						
Ava I	No	---	407	8	0.019 ± 0.007	---
Ava I	Yes	2	398	8	0.020 ± 0.007	---
Ava I, øX174 v. DNA	Yes	2	172	103	0.370 ± 0.030	3.2
Ava I, øX174 v. DNA	Yes	10	161	99	0.380 ± 0.030	2.9

f = fraction broken øX174 v. DNA f_c: corrected for uncut DNA

$$k_L/k_c = \frac{\log f_c}{\log \left\{ \frac{(1 - f_c) - ff_c}{1-f} \right\}} \qquad (21)$$

result agrees qualitatively as well as semi-quantitatively with the excluded volume theory. Analyses of equations (11a,b) and(16a,b) indicate that linear-linear reactions might be expected to be up to two-fold faster than linear-circular reactions. This result showing an effect of DNA strand topology on DNA renaturation rates is another indication, along with the form of the effect of DNA strand length on renaturation rates, that an excluded volume effect is involved in the mechanism of DNA renaturation.

DISCUSSION

An excluded volume theory for DNA renaturation kinetics is proposed in Chapter 4. In order to be complete, such a theory needs to account for both the previously published data concerning the effects of strand lengths on renaturation rates and the new data in this and Chapter 4 concerning the effects of DNA strand topology on renaturation rates.

The theory predicts that the rate of renaturation of DNA will depend on the square root of the length of the reacting single strands ($L^{\frac{1}{2}}$). This prediction is clearly in agreement with the experimental results of Wetmur and Davidson (1968) and subsequent investigators

(for a review, see Wetmur, 1976). The theory in Chapter 4 is not unique in making the $L^{\frac{1}{2}}$ prediction. Other theories based on hard sphere or random coil models of DNA (Wetmur and Davidson, 1968; Wetmur, 1971) have also led to the same conclusion.

The theory in Chapter 4 also predicts that renaturation will depend on the square root of the length of the shorter of two reacting DNA single strands ($L_s^{\frac{1}{2}}$). This prediction is clearly in agreement with the experimental results of Wetmur (1971) and subsequent investigators (for a review, see Wetmur, 1976: Hinnebusch, et al., 1978). In this case, the simple hard sphere model of Wetmur and Davidson (1968) failed to agree with experiment and the more complex random coil model -based theory of Wetmur (1971) agreed over a limited range of ratios of long to short reacting DNA strands. One reason that the theory in Chapter 4 might be expected to be more correct than the theory of Wetmur (1971) is that Wetmur's theory involves interactions at the limit of the distribution functions (W(h)) where these functions are expected to be the least reliable as models of real DNA configurations. The theory in Chapter 4 uses distribution function data

(W(h)) from the entire range of the functions. Whatever the reason, the theory in Chapter 4 is in better agreement with experiments relating DNA renaturation rates to strand lengths than any previously published theory.

Morawetz et al. (1973) published an excluded volume theory for the interactions of polymers in solution. Two versions of the theory involved a hard sphere model and a random coil model for the polymers. Cho and Morawetz (1973) then made two types of polymers to test the theory. They prepared acrylamide containing small amounts of either N-acrylyl-o-acetyl-3, 5-dinitrotyrosine methyl ester (to be hydrolyzed) or N-acrylyl-4-aminomethyl pyridine (to act as a catalyst). Three different lengths of catalyst-containing polyer were prepared. When mixed with the ester containing polymer, all three catalytic polymers were about equally effective in inducing the hydrolysis of the ester. Furthermore, the hydrolysis rates with the polymers were not substantially different than were found with monomers of ester and catalyst. Because this result did not agree with their theory, Cho and Morawetz concluded that excluded volume did not play

a part in this type of polymer-polymer reaction. The lack of length dependence in the experiments of Cho and Morawetz could easily be attributed to the fact that the length of the ester-containing polymer was in every case equal to or shorter than the length of the catalytic polymer and that the length of the ester-containing polymer (L_S) was invariant. However, the relatively high rates of reaction of the polymer residues are best explained if there were little or no excluded volume in this system. This would occur in the model in Chapter 4 if V* is small enough so that θ is less than 1 (equation 13) leading to no length dependence for k_N (equation 15). DNA molecules are polyelectrolytes. It would be interesting to reinvestigate a system of polymers similar to those studied by Cho and Morawetz under conditions where the lengths of both polymers as well as their charge densities could be varied in an attempt to bridge the gap between the experimental results with acrylamide polymers in solvents close to the ideal solvent and DNA renaturation for which no ideal solvent is known.

The theory in Chapter 4 predicts that circular permutation of DNA will have no effect on DNA renaturation rates. This prediction, albeit negative, is

borne out by the data in Chapter 4 as well as the data of Tye et al. (1974a) and has been adequately discussed in Chapter 4.

Finally, the theory makes a new and testable prediction, namely, that circular single stranded DNA will renature slower than linear single stranded DNA. The results in this chapter are in agreement with this prediction.

Any DNA renaturation mechanism must involve a preequilibrium plus a rate determining step or a diffusion controlled rate-determining step. Excluded volume is the only way in which length may be involved in the preequilibrium step. No reasonable model exists for a length dependence at the level of the rate determining step following a preequilibrium. It is hard to imagine the remainder of a freely jointed molecule affecting the formation of the second (or third) base pair during a nucleation event and there is theoretical argument to the contrary (Wetmur and Davidson, 1968). This leaves models involving diffusion control as the only alternatives to an excluded volume model. A DNA renaturation mechanism which involves translational diffusion as a rate limiting step might be expected to lead to a rate constant inversely

proportional to solvent viscosity and proportional to the square root of the length of the shorter of two reacting complementary DNA strands and independent of circular permutation. All of these predictions agree with experiment. Wetmur and Davidson (1968) dismissed this mechanism because the absolute rate of renaturation was too slow to be consistent with translational diffusion and especially because the temperature dependence of DNA renaturation is totally inconsistent with a translational diffusion mechanism. The temperature profile for DNA renaturation implies a nucleation step with a preequilibrium followed by the rate determining step. Segmental diffusion control could only make the expected absolute rate of renaturation increase and be less in agreement with experiment and could not resolve the inconsistency between a diffusion-controlled mechanism and the observed temperature profile. A final prediction of a translational diffusion model would be more rapid renaturation rates for more compact circular DNA, a result opposite to that observed in the experiments described in this chapter. The existence of a unified excluded volume theory which explains all the known data concerning DNA strand length and topology effects on DNA renaturation rates does not

prove that the theory is correct. However, until contradictory data are obtained or another equally satisfactory theory is derived, an excluded volume effect must be accepted as contributing to the mechanism of DNA renaturation.

APPENDIX A

DERIVATION OF EQUATION 6

Opening from the end of a DNA molecule without chemistry. The number of bases opened at one end of a DNA molecule whose base pairs all have equivalent thermal stabilities may be expressed as follows: Let

$$p^x = 1/s^x \tag{A1}$$

be the probability that a base pair x - 1 bases from the molecular ends is in an open or denatured conformation. s is defined in equation 5. The average number of bases open at the end, <x> is given by

$$\langle x \rangle = \sum_{x=0}^{\infty} x p^x \Big/ \sum_{x=0}^{\infty} p^x = \frac{p}{1-p} = \frac{1}{s-1} \tag{A2}$$

This result was first obtained by Zimm (1960).

Opening from the end of a DNA molecule with chemistry. A system in which bases are modified as they become unpaired near the ends of the DNA molecule calls for a more complex treatment, which includes the time dependence of end opening as a result of the chemical reaction occurring. Let Ci be the concentration of an open end terminated by a modified base at position i, the most internal modification. k is defined in equation 4. C is the chloroacetaldehyde concentration. Reaction of chloroacetaldehyde with a DNA

end can be described by the following set of differential equations:

$$-\frac{dC_o}{dt} = kC\ C_o \sum_{x=1}^{\infty} p^x = \frac{kCp}{1-p}\ C_o \qquad \text{(A3a)}$$

$$\frac{dCi}{dt} = \sum_{j=1}^{i} kCp^{i-j+1} C_{j-1} - \frac{kCp}{1-p}\ Ci \qquad \text{(A3b)}$$

$$i \neq 0$$

The solutions to equations A3a and b are

$$C_o = C_o^o e^{-\frac{kCpt}{1-p}} \qquad \text{(A4a)}$$

$$Ci = C_o^{\,o}\ e^{-\frac{kCpt}{1-p}}\ p^i \sum_{x=1}^{i} \frac{(kCt)^x\ (i-1)!}{x!\ (x-1)!\ (i-x)!} \qquad \text{(A4b)}$$

$C_o^{\,o}$ is the initial concentration of ends. The average number of bases open at the end, <x>, is given by

$$\langle x \rangle = \sum_{x=0}^{\infty} xCi / \sum_{x=0}^{\infty} Ci = \sum_{x=0}^{\infty} Ci/C_o^{\,o} \qquad \text{(A5)}$$

Substitution of equations A4 a and b into A5 yields

$$\langle x \rangle = pCkt/(1-p)^2 = \frac{skCt}{(s-1)^2} \qquad \text{(A6)}$$

which is identical to equation 6.

APPENDIX B

DERIVATION OF EQUATIONS 17 a AND b

Equation 15 may be rewritten

$$k_N = \frac{k_{N'}}{\sqrt{s}} \frac{\theta_1}{\theta+1} \approx \frac{k_{N'}}{\sqrt{s}} \frac{\theta_1}{\theta} \quad \text{where } s = \frac{L}{\sigma}$$

and according to equation 13,

$$\frac{\theta}{\theta_1} = 2(\sqrt{a} + \sqrt{b} + \sqrt{c} + \sqrt{d}) - (\sqrt{a+c} + \sqrt{a+d} + \sqrt{b+c} + \sqrt{b+d}).$$

17a: $a = b$, $c = d$, $a >> c$ $(c = \sigma_s/2)$.

$$\frac{\theta}{\theta_1} = 2(2\sqrt{a} + 2\sqrt{c}) - 4\sqrt{a+c} = 4\sqrt{a} + 4\sqrt{c}$$

$$- 4\sqrt{a}\left(1 + \frac{c}{2a}\ldots\right) = 4\sqrt{c}\left(1 - \frac{1}{2}\sqrt{\frac{c}{a}}\ldots\right)$$

$$\frac{\theta}{\theta_1} = 4\sqrt{c} = 4\sqrt{\sigma_s}/2 = 2\sqrt{2}\ \sqrt{\sigma_s}$$

or

$$k_N = \frac{k_{N'}}{\sqrt{L_s}} \frac{1}{2\sqrt{2}} = \frac{k_{N'}}{\sqrt{L_s}} \frac{1}{2.8} \qquad (17a)$$

17b: $a = \sigma_L$, $c = \sigma_s$, $b = d = 0$.

$$\frac{\theta}{\theta_1} = 2\sqrt{\sigma_L} + 2\sqrt{\sigma_s} - \sqrt{\sigma_L + \sigma_s} - \sqrt{\sigma_L} - \sqrt{\sigma_s}$$

$$\frac{\theta}{\theta_1} = \sqrt{\sigma_L} + \sqrt{\sigma_s} - \sqrt{\sigma_L}\left(1 + \frac{\sigma_s}{2\sigma_L}\right) = \sqrt{\sigma_s}\left(1 - \tfrac{1}{2}\sqrt{\frac{\sigma_s}{\sigma_L}}\right)$$

$$\frac{\theta}{\theta_1} = \sqrt{\sigma_s}$$

or

$$k_N = \frac{k_{N'}}{\sqrt{L_s}} \tag{17b}$$

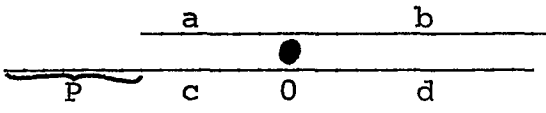
APPENDIX C

DERIVATION OF EQUATION 18 AND CONCLUSION

$$\text{Let } I = \frac{\theta\sqrt{s}}{\theta_1} \cong 2(\sqrt{a} + \sqrt{b} + \sqrt{c} + \sqrt{d}) - (\sqrt{a+c} + \sqrt{a+d} + \sqrt{b+c} + \sqrt{b+d})$$

where a, b, c, d have units of bases.

a. First integration:



$a = x;\ b = L - x$

$c = x + P;\ d = L - x - P$

range: $a = 0$ to $a = L - P$

$$I_1 = \int_0^{L-P} [2\ (\sqrt{x} + \sqrt{L - x} + \sqrt{x+P} + \sqrt{L-x-P})$$

$$- (\sqrt{2x+P} + \sqrt{L-P} + \sqrt{L+P} + \sqrt{2L-2x-P})]dx$$

Now $\int \sqrt{a+bx}\ dx = \frac{2}{3b}(a+bx)^{3/2}$ so

$$I_1 = \left[\frac{2}{3}\ 2\ (x^{3/2} - (L-x)^{3/2} + (x+P)^{3/2} - (L-x-P)^{3/2})\right]_0^{L-P}$$

$$- \frac{2}{3}\left[\left(\frac{2x+P}{2}\right)^{3/2} + \frac{3}{2}\left(\sqrt{L-P} + \sqrt{L+P}\right)x - \left(\frac{2L-2x-P}{2}\right)^{3/2}\right]_0^{L-P}$$

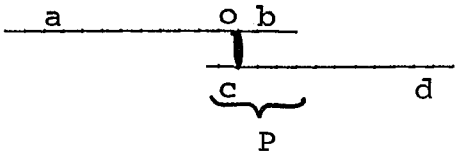
$$I_1 = \frac{2}{3}\left\{2\left((L-P)^{3/2} - 0 - P^{3/2} + L^{3/2} + L^{3/2} - P^{3/2} - 0 + (L-P)^{3/2}\right)\right\}$$

$$- \frac{2}{3}\left\{\left(\frac{(2L-P)}{2}\right)^{3/2} - \frac{P}{2}^{3/2} + \frac{3}{2}\left(\sqrt{L-P} + \sqrt{L+P}\right)(L-P) - 0 - \frac{P}{2}^{3/2} + \left(\frac{2L-P}{2}\right)^{3/2}\right\}$$

$$I_1 = \frac{2}{3}\left\{4(L-P)^{3/2} + 4L^{3/2} - 4P^{3/2} - (2L-P)^{3/2} + P^{3/2}\right.$$

$$\left. - \frac{3}{2}(L-P)\left((\sqrt{L+P} + \sqrt{L-P})\right)\right\}$$

b. Second integration



$a = x, \quad b = L-x$

$c = x-L+P, \quad d = 2L-x-P$

range: $a = L-P$ to $a = L$

$$I_2 = \frac{2}{3}\left[2(x^{3/2} - (L-x)^{3/2} + (x-L+P)^{3/2} - (2L-x-P)^{3/2})\right.$$

$$\left. - \left(\left(\frac{2x-L+P}{2}\right)^{3/2} + \frac{3}{2}\left((2L-P)^{\frac{1}{2}} + P^{\frac{1}{2}}\right)x - \left(\frac{3L-2x-P}{2}\right)^{3/2}\right)\right]$$

$$I_2 = \left\{\frac{2}{3}\right. \quad 2(L^{3/2} - (L-P)^{3/2} + P^{3/2} + 0 + 0 + P^{3/2} - (L-P)^{3/2} + L^{3/2})$$

$$- \left[\left(\frac{L+P}{2}\right)^{3/2} - \left(\frac{L-P}{2}\right)^{3/2} + \frac{3}{2}P\left(\sqrt{2L-P} + \sqrt{P}\right) - \left(\frac{L-P}{2}\right)^{3/2} + \left(\frac{L+P}{2}\right)^{3/2}\right]$$

$$I_2 = \left\{ \frac{2}{3} \; 4L^{3/2} - 4(L-P)^{3/2} + 4P^{3/2} + (L-P)^{3/2} \right.$$

$$\left. - (L+P)^{3/2} - \frac{3P}{2} \left(\sqrt{2L-P} + \sqrt{P}\right) \right\}$$

c. Combining the integrals to include all permutations

$$I_1+I_2 = \frac{2}{3} \left\{ 8L^{3/2} - (2L-P)^{3/2} + P^{3/2} + (L-P)^{3/2} \right.$$

$$\left. - (L+p)^{3/2} - \frac{3}{2} \left(P^{3/2} + P\sqrt{2L-P} + (L-P)^{3/2} + (L-P)\sqrt{L+P}\right) \right\}$$

$\langle\theta\rangle = \frac{c}{L} (I_1+I_2)$ where c is a constant. Then

$$\langle\theta\rangle = \frac{c}{L} \left\{ 8L^{3/2} - (2L-P)^{3/2} + P^{3/2} + (L-P)^{3/2} - (L+P)^{3/2} \right.$$

$$\left. - \frac{3}{2} \left((L-P)^{3/2} + (L-P)(L+P)^{\frac{1}{2}} + P(2L-P)^{\frac{1}{2}} + P^{3/2} \right) \right\}$$

or

$$\langle\theta\rangle = \frac{c}{L} \left\{ 8L^{3/2} - (2L-P)^{3/2} - \frac{P}{2}^{3/2} - \left(\frac{L-P}{2}\right)^{3/2} \right.$$

$$\left. - (L+P)^{3/2} - \frac{3}{2} (L-P)(L+P)^{\frac{1}{2}} - \frac{3}{2} P(2L-P)^{\frac{1}{2}} \right\} \qquad (18)$$

Examples

i) Let P = 0 (in phase)

$$\langle\theta\rangle = c\sqrt{L} \left\{ 8 - 2^{3/2} - 0 + \tfrac{1}{2} - 1 - \frac{3}{2} - 0 \right\}$$

$$\theta = c\sqrt{L} \; (5-2\sqrt{2}) \; = \; 2.2c\sqrt{L}$$

ii) Let P = L/2 (180^{o} out of phase)

$$\langle\theta\rangle = c\sqrt{L}\left\{8 - \left(\frac{3}{2}\right)^{3/2} - \frac{\left(\frac{1}{2}\right)^{3/2}}{2} - \frac{\left(\frac{1}{2}\right)^{3/2}}{2} - \left(\frac{3}{2}\right)^{3/2}\right.$$

$$\left. - \left(\frac{3}{2}\right)\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)^{\frac{1}{2}} - \frac{3}{2}\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)^{\frac{1}{2}}\right\}$$

$$\langle\theta\rangle = c\sqrt{L}\left\{8 - 9\frac{\sqrt{6}}{4} - \frac{\sqrt{2}}{4}\right\} = 2.15c\sqrt{L}$$

APPENDIX D

DERIVATION OF EQUATIONS 19 AND 20a,b.

$$\theta = V^*\left(\frac{3}{2\pi l_e^2}\right)^{3/2} I_1 \left\{ \left(\int_o^a d\sigma_i + \int_o^b d\sigma_i\right) \left(\frac{1}{\sigma_1 + \sigma_j^*}\right)^{3/2} \right\}$$

where $\sigma_j^* = \sigma_j$ for linears

and $\sigma_j^* = \sigma_j\ ((\sigma - \sigma_j)/\sigma)$ for circles,

and where $I_1 = \int_o^c d\sigma_j + \int_o^d d\sigma_j$ for linears

and $I_1 = 2\int_o^{\sigma/2} d\sigma_j$ for circles.

In general:

$$\theta = 2V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} I_1 \left\{ \frac{2}{\sqrt{\sigma_j^*}} - \frac{1}{\sqrt{\sigma_j^* + a}} - \frac{1}{\sqrt{\sigma_j^* + b}} \right\}$$

For the linear case, as found before,

$$\theta_{Lin} = 4V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \left\{ 2(\sqrt{a} + \sqrt{b} + \sqrt{c} + \sqrt{d}) - (\sqrt{a+c} + \sqrt{a+d} + \sqrt{b+c} + \sqrt{b+d}) \right\}$$

Now for the special case of a linear-circular reaction

$$\theta_{Circ} = 4V^* \left(\frac{3}{2\pi l_e^2}\right)^{3/2} \int_0^{\sigma/2} d\sigma_j \left\{ \frac{2}{\sqrt{\sigma_j^*}} - \frac{1}{\sqrt{\sigma_j^* + a}} - \frac{1}{\sqrt{\sigma_j^* + b}} \right\}$$

$$\text{Now } \int \frac{dx}{\chi} = \frac{1}{\sqrt{-c}} \text{ Sin}^{-1} \left(\frac{-2cx - b}{\sqrt{b^2 - 4ac}} \right)$$

$$\text{with } \chi = a + bx + cx^2, c < 0$$

$$\theta_{Circ} = 4V^* \left(\frac{3}{2\pi l_e^2} \right)^{3/2} \sqrt{\sigma} \left[2 \text{ Sin}^{-1} \left(\frac{2x-1}{\sigma} \right) - \text{Sin}^{-1} \left(\frac{\left\{ \frac{2}{\sigma} x - 1 \right\}}{\sqrt{1+4a/\sigma}} \right) - \text{Sin}^{-1} \left(\frac{\left\{ \frac{2}{\sigma} x - 1 \right\}}{\sqrt{1+4b/\sigma}} \right) \right]_0^{\sigma/2}$$

$$\theta_{Circ} = 4V^* \left(\frac{3}{2\pi l_e^2} \right)^{3/2} \sqrt{\sigma} \left\{ \pi - \text{Sin}^{-1} \left(\frac{1}{\sqrt{1+4a/\sigma}} \right) - \text{Sin}^{-1} \left(\frac{1}{\sqrt{1+4b/\sigma}} \right) \right\} \tag{19}$$

Case 1: $a = b = \sigma/2$

$$\frac{\theta}{\theta_1} = 1.91 \sqrt{\sigma}; \qquad k_N \approx k'_N \frac{1}{\sqrt{L}} \frac{1}{1.91} \tag{20a}$$

Case 2: $a = \sigma, b = 0$

$$\frac{\theta}{\theta_1} = 1.11 \sqrt{\sigma} \qquad k_N \approx k'_N \frac{1}{\sqrt{L}} \frac{1}{1.11} \tag{20b}$$

APPENDIX E

DERIVATION OF EQUATION 21

Circles

C_o = initial concentration = $C_I^o + C_B^o$

C_I = intact viral DNA

C_B = broken viral DNA

$C_B^o = f\,C_o$; $C_I^o = (1 - f)\,C_o$

Linears

C_o = initial concentration = $C_L^o = C_{L'}^o$

C_L = linear (+)

$C_{L'}$ = linear (-)

Rate Equations

$$\frac{-dC_I}{dt} = k_c\, C_I\, C_{L'}$$

$$\frac{-dC_B}{dt} = k_L\, C_B\, C_{L'}$$

$$\frac{-dC_L}{dt} = k_L\, C_L\, C_{L'}$$

Solution

$$\frac{d\ln C_I}{k_c} = \frac{d\ln C_B}{k_L} = \frac{d\ln C_L}{k_L}$$

$$\frac{\ln C_I/(1-f)C_o}{k_c} = \frac{\ln C_B/fC_o}{k_L} = \frac{\ln C_L/C_o}{k_L}$$

$$\frac{C_I}{(1-f)C_o} = \left(\frac{C_L}{C_o}\right)^{k_c/k_L} \; ; \; \frac{C_I}{C_o} = (1-f)\left(\frac{C_L}{C_o}\right)^{k_c/k_L}$$

$$\frac{C_I}{C_o} + \frac{C_B}{C_o} = (1-f)\left(\frac{C_L}{C_o}\right)^{k_c/k_L} + f\left(\frac{C_L}{C_o}\right)$$

Now $$\frac{C_I}{C_o} + \frac{C_B}{C_o} = 1-f_c \; ; \; \frac{C_L}{C_o} = f_c$$

So $$(1-f_c) = (1-f)\,(f_c)^{k_c/k_L} + ff_c$$

$$\frac{(1-f_c) - ff_c}{1-f} = (f_c)^{k_c/k_L}$$

$$\log\left[\frac{(1-f_c) - ff_c}{1-f}\right] = \frac{k_c}{k_L}\log f_c$$

$$\frac{k_L}{k_c} = \frac{\log f_c}{\log\left\{\dfrac{(1-f_c) - ff_c}{1-f}\right\}} \qquad (21)$$

REFERENCES

Abraham, G., A.K. Banerjee (1976) Proc. Natl. Acad. Sci. 73: 1504-1508.

Baltimore, D. (1971) Bact. Rev. 35: 235-241.

Barrio, J.R., J.A. Secrist III, N.J. Leonard (1972) Biochem. Biophys. Res. Commun. 46: 597-604.

Beemon, K., P. Duesberg, P. Vogt (1974) Proc. Natl. Acad. Sci. 71: 4254-4258.

Benbow, R.M., A.J. Zuccarelli, G.C. Davis, R.L. Sinsheimer (1974) J. Virol. 13: 898-907.

Benbow, R.M., A.J. Zuccarelli, R.L. Sinsheimer (1975) Proc. Natl. Acad. Sci. 72: 235-239.

Benzer, S. (1961) Proc. Natl. Acad. Sci. 47: 403-415.

Both, G.W., S. Lavi, A.J. Shatkin (1975) Cell 4: 173-180.

Britten, R.J., D.E Kohne (1968) Science 161: 529-540.

Burton, A., R.L. Sinsheimer (1963) Science 142: 961.

Chang, C.-T., T.C. Hain, J.R. Hutton, J.G. Wetmur (1974) Biopolymers 13: 1847-1858.

Cho, J.-R., H. Morawetz (1973) Macromolec. 6: 628-631.

Coffin, J.M , M.A. Billeter (1976) J. Molec. Biol. 100: 293-318.

Cooper, P.D (1968) Virol. 35: 584-596.

Cooper, P.D , S. Steiner-Pryor, P.D. Scotti, D. Delong (1974) J. Gen. Virol. 23: 41-49.

Cross, R.K., B N. Fields (1972) Virol. 50: 799-809.

Danna, K J., D. Nathans (1972) Proc. Natl. Acad. Sci. 69: 3097-3100.

Davis, B.D., R. Dulbecco, H.N. Eisen, H.S. Ginsberg, W.B. Wood, Jr. (1973) Microbiology p. 1336, Harper and Row, Maryland.

Davis, R.W., M. Simon, N. Davidson (1971) Methods in Enzymology (Grossman, L. & K. Moldane, eds.) 21D: 413-428, Academic Press, New York.

Davis, R.W., N. Davidson (1968) Proc. Natl. Acad. Sci. 50: 243-250.

Delbrück, M. (1946) Biol. Rev. Cambridge Phil. Soc. 21: 30-40.

Demerec, M., U. Fano (1945) Genetics 30: 119-136.

Dodgson, J.B , R.D. Wells (1977a) Biochem. 16: 2367-2374.

Dodgson, J.B., R.D. Wells (1977b) Biochem. 16: 2374-2379.

Dreiseikelmann, B., W. Wackernagel (1978) Mol. Gen. Genet. 159: 321-328.

East, J.L., D.A. Kingsbury (1971) J. Virol. 8: 161-173.

Eigner, J. P. Doty (1965) J. Molec. Biol. 12: 549-580.

Fenner, F., B.R. McAuslan, C.A. Mims, J. Sambrook, D.O. White (1974) The Biology of Animal Viruses, Second Edition, Academic Press, New York, London.

Fields, B.N., W.K. Joklik (1969) Virol. 37: 335-342.

Flashner, M., J.N. Vournakis (1977) Nucl Acid Res. 4: 2307-2319.

Flory, P J. (1949) J. Chem. Phys. 17: 303-310.

Freifelder, D. (1970) J. Molec. Biol. 54: 567-577.

Freifelder, D , A.K. Kleinschmidt, R.L Sinsheimer (1964) Science 146: 254-255.

Fuchs, C., E.C. Rosenvold, A. Honigman, W. Szybalski (1978) Gene 4: 1-23

Gilbert, W., D. Dressler (1968) Cold Spring Harbor Symp. Quant. Biol. 33: 474-484.

Gold, M., M. Gefter, R. Hausmann, J. Hurwitz (1966) J. Genl. Physiol. 49 (6 part 2): 5-28.

Gonda, T.J , R H. Symons (1978) Virol 88: 361-370.

Harada, F., J.E. Dahlberg (1975) Nucl. Acid Res. 2: 865-871.

Hayashi, M., M.N. Hayashi, S. Spiegelman (1963) Proc. Natl. Acad. Sci. 50: 664-672.

Hayes, W. (1968) The Genetics of Bacteria and Their Viruses, John Wiley & Sons, New York.

Helling, R B., H M. Goodman, H.W. Boyer (1974) J. Virol. 14: 1235-1244.

Hinnebusch, A.G., V.E. Clark, L.C. Klotz (1978) Biochem. 17: 1521-1529.

Hirst, G.K. (1962) Cold Spring Harbor Symp. Quant. Biol. 27: 303-308.

Horie, K., I. Mita (1978) Macromolec. 11: 1175-1179.

Hsieh, T., J.C Wang (1976) Biochem. 15: 5776-5782.

Hsu, M., H. Kung, N. Davidson (1973) Cold Spring Harbor Symp. 38: 943-950.

Ikegami, N., P.J. Gomatos (1968) Virol. 36: 447-458.

Inman, R.B. (1966) J. Molec. Biol. 18: 464-476.

Ito, Y., W.K. Joklik (1972) Virol. 50: 202-208.

Jacobson, M.F., D. Baltimore (1968) Proc. Natl. Acad. Sci. 61: 77-84.

Johnston, J.V., B.P. Nichols, J.E. Donelson (1977) J. Virol. 22: 510-519.

Kelly, T.J., D. Nathans (1977) Adv. in Virus Res. 21: 85-173.

Klump, H. (1977) Biochim Biophys. Acta 475: 605-610.

Kolakofsky, D., E. Boy de la Tour, A. Bruschi (1974) J. Virol. 14: 33-39

Lai, C.J., D. Nathans (1974) Virol 60: 466-475.

Lai, C.J., D. Nathans (1975) Cold Spring Harbor Symp. Quant. Biol. 39: 53-60.

Lang, D. (1970) J. Molec. Biol. 54: 557-565.

Lebowitz, J., C.G. Garon, M.C.Y. Chen, N.P. Salzman (1976) J. Virol. 18: 205-210.

Ledinko, N. (1963) Virol. 20: 107-119.

Lee, C.S., R.W. Davis, N. Davidson (1970) J. Molec. Biol. 48: 1-22.

Ludwig, R.A , W.C. Summers (1975) Virol. 68: 360-373.

Lunan, K.D., R.L. Sinsheimer (1956) Virol. 2: 455-462.

Luskashin, A.V., A.V. Vologodskii, M.D. Frank-Kamenetskii, Y.L. Lyubchenko (1976) J. Molec. Biol. 108: 665-682.

MacHattie, L.A , D.A. Ritchie, C.A. Thomas, Jr., C.C. Richardson (1967) J. Molec. Biol. 23: 355-363

Mandell, J.D., A.D. Hershey (1960) Anal. Biochem. 1: 66-77.

Marmur, J., P. Doty (1961) J. Molec Biol. 3: 585-594.

McCorquodale, D.J. (1975) CRC Crit. Rev. Micro. 4: 101-159.

McCrae, H., W K. Joklik (1978) Virol. 89: 578-593.

McDonell M.W., M.N. Simon, F.W. Studier (1977) J. Molec. Biol. 110: 119-146.

McGhee J.D., P.H. von Hippel (1977a) Biochem. 16: 3267-3276.

McGhee, J D., P.H. von Hippel (1977b) Biochem. 16: 3276-3293.

McMaster, G.K., G.C. Carmichael (1977) Proc. Natl. Acad. Sci. 74: 4835-4838.

McMilin, K.D., M.M. Stahl, F.W Stahl (1974) Genetics 77: 409-423.

Melchior, W.B., P.H. von Hippel (1973) Proc. Natl. Acad. Sci. 70: 298-302.

Metz, D.H., G.L. Brown (1969a) Biochem. 8: 2312-2328.

Metz, D.H., G.L. Brown (1969b) Biochem. 8: 2329-2341.

Morawetz, H., J.-R. Cho, P.J Gans (1973) Macromolec. 6: 624-627.

Morgan, T.H. (1911) J. Exper. Zool. 11: 365-411.

Nakaya, K., O. Takenaka, H. Horinishi, K. Shibata (1968) Biochim. Biophys. Acta 161: 23-31.

Nichols, B.P., J.E. Donelson (1977) J. Virol. 22: 520-526.

Nonoyama, M., A.F. Graham (1970a) J. Virol. 6: 693-694.

Nonoyama, M., Y. Watanabe, A.F. Graham (1970b) J. Virol. 6: 226-236.

Orosz, J.M. (1975) DNA Iodination, Provirus Isolation, Excluded Volume and Electrostatic Effects of DNA Renaturation Rates. Ph.D thesis, U. of Illinois-Urbana.

Pagano, J S , C.A. Hutchison III (1971) Methods in Virology V., pp. 79-123, Academic Press, New York, London.

Palese, P., J.L. Schulman (1976a) J. Virol 17: 876-884.

Palese, P., J.L Schulman (1976b) Proc. Natl. Acad. Sci. 73: 2142-2146.

Palese, P., M.B. Ritchey, J.L. Schulman (1977) Virol. 76: 114-121.

Ramig, R.F., R.K. Cross, B.N. Fields (1977) J. Virol. 22: 726-733.

Reddy, V.B., B. Thimmappaya, R. Dhar, K.N. Subramanian, B.S. Sain, J. Pan, P.K. Ghosh, M.L. Cellma, S.M. Weissman (1978) Science 200: 494-502.

Rekosh, D. (1972) J. Virol. 9: 479-487.

Rhodes, O.P., A.K. Banerjee (1976) J. Virol. 17: 33-42.

Ritchey, M.B , P. Palese, J.L. Schulman (1976) J. Virol. 20: 307-313.

Ritchie, D.A., C A. Thomas, Jr., L.A. MacHattie, P.C. Wensink (1967) J. Molec. Biol. 23: 365-376.

Roberts, R.J (1980) Nucl. Acids Res. 8: r63-r79.

Rosenberg, A.H., Simon, M.N., Studier, F.W. (1979) J. Molec. Biol. 135: 907-915.

Ruyechan, W.T. (1976) Selected Interactions of Nucleic Acids with Proteins and Small Molecules. Ph.D. thesis U. of Illinois.

Sanger, F , G.M Air, B.G. Barrell N.L. Brown, A.R. Coulson, J.C Fiddes, C.A Hutchison III, P.M. Slocombe, M. Smith (1977) Nature 265: 687-695.

Sanger, F., A.R. Coulson, T. Friedman, G.M. Air, B.G. Barrell, N.L. Brown, J.C Fiddes, C.A. Hutchison III, P.M. Slocombe, M. Smith (1978) J. Molec. Biol. 125: 225-246.

Scheurch, A.R , W.K. Joklik (1973) Virol. 56: 218-229.

Sharp. P.A., V.A. Bloomfield (1968) J. Chem. Phys. 49: 4564-4566.

Sharpe, A. R.F. Ramig, T.A. Mustoe, B N. Fields (1978) Virol. 84: 63-74.

Shenk, T.E., C. Rhodes, P. Rigby, P. Berg (1975a) Proc. Natl. Acad. Sci. 72: 989-993.

Shenk, T E., C. Rhodes, P. Rigby, P. Berg (1975b) Cold Spring Harbor Symp. Quant. Biol. 39: 61-67.

Shishido, K., T. Ando (1975) Biochim. Biophys. Acta 390: 125-132.

Siebenlist, U. (1979) Nucl. Acids Res. 6: 1895-1907.

Sinsheimer, R.L. (1959) J. Molec. Biol. 1: 43-53.

Sinsheimer, R.L., B. Starman, C. Nagler, S. Guthrie (1962) J. Molec. Biol. 4: 142-160.

Streisinger, G., J. Emrich, M.M Stahl (1967) Proc. Natl. Acad. Sci. 57: 292-295.

Studier, F.W. (1965) J. Molec. Biol. 11: 373-390.

Studier, F.W. (1972) Science 176: 367-376.

Studier, F.W. (1979a) Virol. 95: 70-84.

Studier, F.W., A.H. Rosenberg, M.N Simon, J.J. Dunn (1979b) J. Molec. Biol. 135: 917-937.

Sugiura, A., E.D. Kilbourne (1966) Virol. 29: 84-91.

Summers, D.F., J.V. Maizel (1968) Proc. Natl. Acad. Sci. 59: 966-971.

Taber, R., D. Rekosh, D. Baltimore (1971) J. Virol. 8: 395-410.

Tai, H.T., C.A. Smith, P.A. Sharp, J. Vinograd (1972) J. Virol. 9: 317-325.

Tanford, C. (1961) Physical Chemistry of Macromolecules. John Wiley & Sons, Inc., New York.

Tye, B.K., J.A. Huberman, D. Botstein (1974a) J. Molec. Biol. 85: 501-532.

Tye, B.K., R.K. Chan, D. Botstein (1974b) J. Molec. Biol. 85: 485-500.

Vinograd, J , J. Lebowitz (1966) J. Gen. Physiol. 49: 103-125.

von Hippel P.H., K -Y. Wong (1971) J. Molec. Biol. 61: 587-613.

Vournakis, J.N , M.S. Flashner, M. Katopas. G A. Kitos, N.C. Vamvakopoulous, M.S. Sell. R.M. Wurst (1976) Prog. Nucl. Acids Res. Molec. Biol. 19: 233-252.

Wang, L.-H., P.H. Duesberg (1974) J. Virol. 14: 1515-1529.

Wang, L.-H. P.H. Duesberg, K. Beemon, P.K Vogt (1975) J. Virol 16: 1051-1070.

Wang, L.-H., D. Galehouse, P. Mellon, P. Duesberg. W.S. Mason, P.K. Vogt (1976) Proc. Natl. Acad Sci. 73: 3952-3956.

Wang, J.C., N. Davidson (1966a) J. Molec. Biol 15: 111-123.

Wang, J.C., N Davidson (1966b) J. Molec. Biol. 19: 469-482.

Westmoreland, B., W. Szybalski. H. Ris (1968) Science 163: 1343-1348.

Wetmur, J.G., N. Davidson (1968) J. Molec. Biol. 31: 349-370

Wetmur, J.G. (1976) Ann. Rev. Bioph. Bioeng. 5: 337-361.

Wetmur. J.G. (1971) Biopolymers 10: 601-613.

Wiegand. R.D., G.N. Godson. C.M. Radding (1975) J. Biol. Chem. 250: 8848-8855.

Wood H.A. (1973) J. Gen. Virol. 20: 61-85.

Wrede. P., N.H Woo, A. Rich (1979) Proc. Natl Acad. Sci. 76: 3289-3293.

Wurst, R., J.N. Vournakis, A.M. Maxam (1978) Biochem. 17: 4493-4499.

Zimm, B.H. (1960) J. Chem. Phys. 33: 1349-1356.