

To the Graduate Council:

I am submitting herewith a dissertation written by John W. Wills entitled "The Nuclear and Mitochondrial Genomes of *Candida albicans*: A Study of the Yeast Phase Cell." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Microbiology.

W. S. Riggsby
W. S. Riggsby, Major Professor

We have read this dissertation
and recommend its acceptance:

Dorothy M. Skinner
Jeffrey M. Belden
Ralph Smith

Accepted by the Council:

L. Evans
Vice Chancellor
Graduate Studies and Research

THE NUCLEAR AND MITOCHONDRIAL GENOMES
OF CANDIDA ALBICANS: A STUDY OF
THE YEAST PHASE CELL

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

John W. Wills

August 1982

3063280

Everything that doesn't destroy you
makes you stronger.

— Nietzsche

ACKNOWLEDGMENTS

I wish to give special thanks to a number of people who have been essential to the accomplishments represented by this dissertation. Most importantly, I thank Dr. Stuart Riggsby, my major professor, whose rigorous and quick thinking I hope have rubbed off on me a little. Without his assistance this dissertation would not have been nearly as much fun to produce and might not have happened at all. I thank Dr. Dorothy Skinner who made many trips between Oak Ridge and Knoxville to serve on my committee and who kindly made available her Model E centrifuge for my use. I thank Dr. Karl Sirotkin who introduced me to recombinant DNA techniques and who made them so easy to learn. I thank Dr. Jeffrey Becker for serving on my committee and for going out of his way to help make this happen. I thank all of the guys in the lab with whom I have spent countless, enjoyable hours working and talking: Luis Torres, Brent Lasker, Mark Mason (from whom I "borrowed" many milliliters of TE buffer), Tim Townes, and Ken Draper. I thank my parents and sisters for their morale, financial, and loving support. And finally, I thank my friend (and now wife), Becky, whose love made the tough times much easier.

ABSTRACT

The purpose of this study was to characterize the nuclear and mitochondrial genomes of the pathogenic, dimorphic fungus Candida albicans. Buoyant density measurements of whole cell DNA revealed two components: a major band of 32.8% G + C and a satellite band of 40.3% G + C. Cation-exchange chromatography of 5'-deoxyribonucleotides from whole cell DNA showed an average G + C content of 34.4%. Electrophoresis of restriction endonuclease EcoRI digested whole cell DNA revealed three distinct repetitive DNA bands (containing fragments 6.6, 3.3, and 2.6 kbp in length) which proved to be of nuclear DNA origin. Reassociation experiments using whole cell DNA and the HAP method of assay showed the presence of four kinetic components in randomly sheared DNA (300 bp, average length). The component composition of the DNA was found to be: 3.4% foldback, 2.5% fast repetitive ($K_{\text{pure}} = 742$), 11.0% slow repetitive ($K_{\text{pure}} = 13.1$), and 83.6% single-copy sequences ($K_{\text{pure}} = 0.063$). The total kinetic complexity (haploid genome size) was found to be 1.85×10^7 base pairs. Comparison with previously reported DNA content values indicates that Candida albicans is a diploid organism.

Methods were developed which enabled the successful purification of mitochondrial DNA (mtDNA) in large amounts from yeast phase cells of C. albicans strain H317. Buoyant density measurements revealed a G + C content of 38%. Digestion of mtDNA with restriction endonuclease EcoRI was found to produce six fragments, and digestion with PvuII was found to produce seven fragments. Based upon restriction endonuclease

analysis, the mtDNA was found to have a length of 41 kbp. The EcoRI fragments of this mtDNA (except for the largest fragment) were cloned in plasmid pBR322 without any detectable alterations. Hybridization experiments showed that each of the six EcoRI fragments contain sequences homologous to one or more of the other five fragments. Restriction endonuclease mapping of the EcoRI and PvuII sites proved that the mtDNA of this organism is circular and contains a large, inverted duplication. Comparison with EcoRI digests of the mtDNA of three other strains of C. albicans (ATCC, H350, and WD-18-4) suggests that all have the same restriction map. Thus, it appears that the map reported here is that of wild type C. albicans mtDNA.

PREFACE

Candida albicans is a very important organism. This dimorphic fungus is the most commonly encountered fungal pathogen of humans with infections having been observed in most tissues of the human body (Odds, 1979). In spite of its medical importance, very little is understood about the genetic factors responsible for the pathogenicity of this organism and for its ability to undergo conversion from the yeast to mycelial phase of growth. The primary reason for this paucity of information resides in the fact that C. albicans has no known sexual cycle and therefore has no convenient means for genetic analysis. Recently, some important steps towards the genetic analysis of C. albicans have been taken by using the parasexual cycle of protoplast fusion of auxotrophic strains (Poulter et al., 1981; Sarachek et al., 1981), but this promising approach has been hampered by difficulties in obtaining mutants of C. albicans (Whelan and Magee, 1981). Although conventional genetic approaches have had rather limited usefulness, techniques for molecular genetic analysis of the C. albicans genome have become very powerful. The purpose of the following dissertation is to present the results of my studies of the Candida albicans nuclear and mitochondrial genomes which have utilized molecular genetic techniques.

The dissertation is divided into two parts. Part I deals primarily with the organization of Candida albicans whole cell DNA as studied by reassociation kinetics. In this part, the sequences of

C. albicans are classified, according to their kinetic properties, into four groups: foldback, fast repetitive, slow repetitive, and single-copy sequences. These sequences are further characterized in terms of kinetic complexity and repetition frequency. The total haploid size of the C. albicans genome is calculated which allows the ploidy to be determined. In this part, the purity of nuclear DNA preparations is also explored. Problems with mitochondrial DNA contamination precluded investigations into the interspersions of the repetitive sequences using reassociation techniques. Nevertheless, preliminary interspersions information is presented from genomic blot experiments using a repetitive DNA probe.

The material in Part II of the dissertation arose because of the problems encountered in purifying nuclear DNA. Although nuclear DNA could not be isolated without mitochondrial DNA contamination, pure mitochondrial DNA has been obtained. In this part, the mitochondrial DNA is characterized. Experiments are presented which show how the DNA was cloned and how the size, restriction map, and organization were determined.

Part I and II were written with two different audiences in mind. Part I is intended not only for those interested in Candida albicans DNA but also for the beginning reassociationist. In particular, it is intended for the researcher who pursues the important interspersions problem once nuclear DNA can be obtained in pure form. With this in mind, much of the introductory chapter of this part (Chapter 1) was written to explain the important aspects of DNA reassociation experiments. Also, in the Appendix, important control experiments for the

hydroxyapatite method are presented both to justify my own experiments and to provide guidance for the beginner. Furthermore, a large portion of the results section (Chapter 3) concerns a rather detailed description of the statistical analysis of the reassociation data. Part II, on the other hand, represents a complete set of experiments which provide an initial characterization of C. albicans mitochondrial DNA. These experiments provide the foundation for extensive future studies. Thus, Part II was written with the hope that studies of the mitochondrial DNA will be carried to a higher level.

TABLE OF CONTENTS

CHAPTER		PAGE
PART I: REASSOCIATION KINETICS		
OF <u>CANDIDA ALBICANS</u> DNA		
I.	REVIEW OF DNA REASSOCIATION PRINCIPLES.	2
	Information Provided by DNA Reassociations.	2
	Mechanics of DNA Reassociation Experiments.	6
	Statistical Analysis of Reassociation Data	17
II.	MATERIALS AND METHODS	20
	Organisms	20
	Growth and Labeling of Cells.	20
	Preparation of Spheroplasts	22
	Purification of Nuclei.	23
	Purification of Mitochondria and mtDNA.	25
	Extraction of DNA from Spheroplasts and Nuclei.	25
	Extraction of DNA from <u>E. coli</u>	26
	CsCl Density Gradient Analysis of DNA	26
	Nucleotide Separation by Cation-Exchange Chromatography.	27
	Shearing of DNA	28
	Purification of Sheared DNA Fragments	29
	Sizing of DNA Fragments	30
	Measurement of Reassociation Kinetics Using Hydroxyapatite.	30
	Isolation of Repetitive DNA Sequences	32
	Statistical Analysis of Reassociation Data.	34
	Restriction Endonuclease Digestion and Gel Electrophoresis	34
	Transfer of DNA to Nitrocellulose	34
	Labeling of DNA by Nick Translation	34
	Blot Hybridizations	35
	Cloning of Mitochondria DNA	35
III.	EXPERIMENTAL FINDINGS	36
	Base Composition of <u>Candida albicans</u> DNA.	36
	Reassociation of <u>Candida albicans</u> Whole Cell DNA.	41
	Reassociation of <u>Escherichia coli</u> DNA	53
	Reassociation of Purified Repetitive DNA.	58
	Genomic Blot Hybridizations Using Repetitive DNA Probe	69
	Presence of mtDNA in Nuclear DNA Preparations	73

CHAPTER	PAGE
IV. DISCUSSION OF RESULTS.	80
Base Composition of <u>Candida albicans</u> DNA	80
Cation-exchange Chromatography	83
Component Composition of <u>Candida albicans</u> DNA.	85
Kinetic Complexity of the Components	87
Frequency of Repetition of the Repeated Sequences.	88
Genome Size and Ploidy of <u>Candida albicans</u>	90
Distribution of Repetitive Sequences	92
PART II: CHARACTERIZATION OF <u>CANDIDA</u>	
<u>ALBICANS</u> MITOCHONDRIAL DNA	
V. INTRODUCTION	96
The Functions of mtDNA	96
The Structure of mtDNAs.	100
The Size of mtDNAs	102
The Rearrangement of mtDNA	104
The Organization and Transcription of mtDNA.	109
Statement of Intent.	111
VI. MATERIALS AND METHODS.	112
Organisms.	112
Growth of <u>Candida albicans</u> Cells	112
Preparation of Spheroplasts.	113
Isolation of Mitochondria.	113
Isolation of Mitochondrial DNA	116
CsCl Density Gradient Analysis of mtDNA.	117
Restriction Endonuclease Digestion of DNA.	117
Polyacrylamide-Agarose Gel Electrophoresis	117
Agarose Gel Electrophoresis.	118
Transfer of DNA to Nitrocellulose.	118
Labeling of DNA by Nick Translation.	119
Hybridization to DNA on Nitrocellulose	120
Cloning of mtDNA	120
Screening of Transformants	122
Large Scale Purification of Recombinant DNA	
Plasmids	124
VII. RESULTS AND DISCUSSION	126
Isolation of Mitochondrial DNA	126
Base Composition of Mitochondrial DNA.	132

CHAPTER	PAGE
Single Digests of mtDNA.	135
Double Digestion of mtDNA.	137
Cloning of mtDNA	141
Mapping the PvuII Sites of the Cloned EcoRI Fragments.	143
Determining the Order and Orientation of the EcoRI Fragments.	147
Comments Concerning the Restriction Map.	157
LIST OF REFERENCES.	161
APPENDIX.	171
VITA.	177

LIST OF TABLES

TABLE	PAGE
1-1. Analytical CsCl Density Gradient Analysis of <u>Candida albicans</u> DNA.	38
1-2. Cation-exchange Separation of <u>Candida albicans</u> ³² P-labeled, 5'-deoxyribonucleotides.	40
1-3. Total <u>C. albicans</u> DNA Reassociation Experiment: Unconstrained Fits for One and Two Components	42
1-4. <u>C. albicans</u> DNA Uniqueness of Fit Analysis: $K_{\text{repetitive}}$ Fixed	47
1-5. <u>C. albicans</u> DNA Uniqueness of Fit Analysis: $K_{\text{single-copy}}$ Fixed.	49
1-6. <u>C. albicans</u> DNA Uniqueness of Fit Analysis: One Component Solutions Using Only Later Points	54
1-7. <u>E. coli</u> DNA Reassociation Experiment: Unconstrained Fits to One Component	57
1-8. <u>E. coli</u> DNA Reassociation Experiment: Uniqueness of Fit Analysis-- K_{obs} Fixed.	60
1-9. Minicot Experiment Using Repetitive DNA: Unconstrained Fits to Different Numbers of Components	65
1-10. Minicot Experiment Using Repetitive DNA: Uniqueness of Slow Component Fit Using Only Late Data Points	68
1-11. Repetitive EcoRI Fragments Detected by Autoradiography.	72
2-1. Fragment Sizes Produced by EcoRI and PvuII Digestion of mtDNA.	139
2-2. Sizes of Fragments Produced by PvuII Digesting the Cloned EcoRI Fragments.	146
2-3. Fragment Sizes of EcoRI Partial Digestion Products.	154
A-1. HAP Chromatography Using High Salt Elution.	172
A-2. Quenching of ³ H-cpm with Increasing NaP Concentration	174

TABLE

PAGE

A-3. Thermal Elution of dsDNA in the Absence of Carrier DNA.	174
A-4. HAP Chromatography Using Carrier DNA	176

LIST OF FIGURES

FIGURE	PAGE
1-1. Diagrammatic Representation of a DNA Reassociation Experiment to Determine Haploid Genome Size.	3
1-2. Ultraviolet Absorption Photographs of Analytical CsCl Gradients Containing <u>C. albicans</u> DNA.	37
1-3. Changes in the Least Squares Solution with Decreasing Values of Fixed $K_{S.C.}$	51
1-4. Changes in RMS Error with Increases and Decreases in Fixed K_{obs}	55
1-5. Kinetics of Reassociation of DNA of <u>Escherichia coli</u> and <u>Candida albicans</u> , Assayed by Hydroxyapatite Chromatography	59
1-6. Minicot Analysis of Rapidly Reassociating DNA of <u>Candida albicans</u>	67
1-7. Genomic Blot Hybridizations Using Repetitive DNA Probe.	70
1-8. Preparation of Nuclear DNAs from <u>C. albicans</u> Cells	74
1-9. Detection of mtDNA in EcoRI Digests of Nuclear DNAs.	75
1-10. Detection of mtDNA in PvuII Digests of Nuclear DNAs.	77
1-11. Hybridization of ^{32}P -labeled, Cloned Fragment E4 of mtDNA to PvuII Digested Nuclear DNAs.	79
2-1. Isolation of Mitochondria	114
2-2. Purity of mtDNA (lane a) As Demonstrated by EcoRI Digestion and Agarose Gel Electrophoresis.	128
2-3. Measurement of Base Composition of Purified mtDNA by Analytical CsCl Density Gradient Centrifugation	133
2-4. Digestion of <u>C. albicans</u> DNAs with the Isoschizomers, HpaII and MspI	134

FIGURE

PAGE

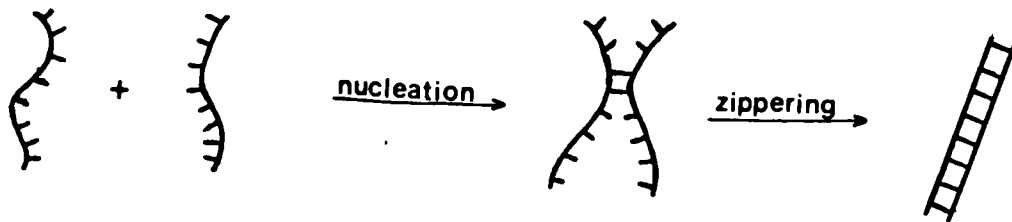
2-5.	PvuII and EcoRI Single and Double Digests of mtDNA Analyzed by Agarose Gel Electrophoresis.	136
2-6.	Polyacrylamide Gel Electrophoresis of Single and Double Digests of mtDNA and Cloned E6.	138
2-7.	EcoRI Digestion of mtDNA Containing, Recombinant DNA Plasmids from Maxipreps.	142
2-8.	EcoRI-PvuII Double Digestions of Recombinant DNA Plasmids Containing mtDNA.	144
2-9.	Location of PvuII Sites of Cloned EcoRI Fragments.	148
2-10.	Hybridization of ³² P-labeled Recombinant DNA Plasmids to EcoRI Digested and PvuII Digested mtDNA	149
2-11.	Partial Digestion of mtDNA with EcoRI.	153
2-12.	Circular Restriction Endonuclease Map of <u>Candida</u> <u>albicans</u> mtDNA	158

PART I: REASSOCIATION KINETICS
OF CANDIDA ALBICANS DNA

CHAPTER I

REVIEW OF DNA REASSOCIATION PRINCIPLES

The separation of the complementary strands of a DNA duplex, called DNA dissociation, is a process which can be reversed to allow duplex structures to reform. The reassociation of single stranded DNA is a two step reaction. The first step, which is rate limiting, involves the collision and stabilization of a few complementary bases. This is usually called the nucleation step. The second step, which is practically instantaneous, involves the pairing of complementary bases along the remaining lengths of the two strands. This is usually called the zippering step. These steps may be illustrated as follows.



Measurement of the rate at which reassociation of single stranded DNA occurs can yield a wealth of information about that DNA. The types of information obtainable are summarized in the following paragraphs.

Information Provided by DNA Reassociations

Genome size. The size of the haploid genome of an organism can be measured by reassociation kinetics. A simplified explanation as to why this is possible is illustrated in Figure 1-1. In this figure,

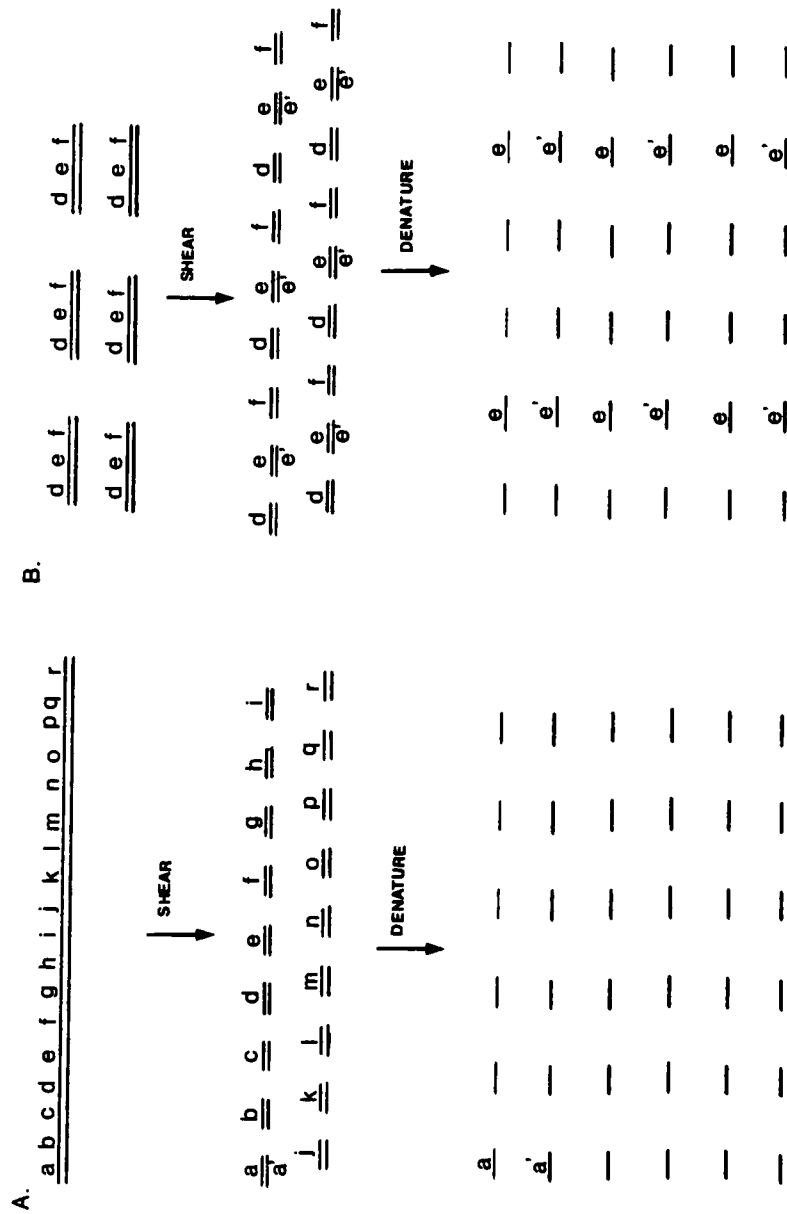


Figure 1-1. Diagrammatic representation of a DNA reassociation experiment to determine haploid genome size. The letters represent contiguous regions, each containing a unique sequence of nucleotides.

two types of double stranded DNA are diagramed, a large one (Figure 1-1a) which is six times longer than the other (Figure 1-1b). The diagram is intended to suggest that the two DNAs are contained in separate reaction vessels. Six molecules of the short DNA are presented to indicate that equal masses of each DNA are present and that the mass concentration of DNA is the same in each reaction vessel. The letters above each DNA serve as imaginary markers which allow identification of the different regions of each DNA. The sequences in each region are unique (i.e. no sequences are repeated in either DNA). To proceed with the reassociation experiment, both DNAs are sheared to the same size and dissociated so that the only difference between the two solutions is in the concentration and types of sequences present. Now, consider a specific region of each DNA, for example, region a in Figure 1-1a and region e in Figure 1-1b. Fragment a can only reassociate with fragment a', and fragment e can only reassociate with fragment e'. Since the relative concentration of e' is six times greater than that of a', the possibility of an e-e' reassociation is six times greater than an a-a' reassociation. Thus, it can be seen that the reassociation of the fragments from the small genome will proceed at a faster rate than the reassociation of the fragments from the large genome. In other words, the rate of reassociation is inversely proportional to the length of the unique sequences of a DNA.

Knowledge of the genome size of a DNA is useful for many purposes. For example, comparison of the haploid length of a DNA (measured by reassociation kinetics) with the total DNA content of the cell (measured by chemical methods) enables the ploidy of the organism to be determined (Lewin, 1980). Knowledge of the genome size also aids in

molecular cloning experiments by enabling an estimate to be made of the number of recombinant DNA molecules required to provide a complete representation of all the sequences of the genome.

Repetitive sequences. Sequences which are present in several copies in a given DNA will reassociate at faster rates than the remaining single copy sequences. This is easily envisioned with the aid of Figure 1-1 if one imagines that the two sets of diagramed sequences make up a single DNA with one sequence (def) repeated six times. After shearing and dissociation, the repetitive sequences will reassociate more quickly than the unique sequences (a-c and g-r).

Reassociation experiments can reveal many characteristics of sequences which are repeated. Measurements of reassociation rates enable one to estimate the lengths and repetition frequencies of repeated sequences (Lewin, 1980). These measurements also distinguish between sequences which are directly repeated and inversely repeated, since inverted repeats reassociate very rapidly in an intramolecular, concentration independent manner (Perlman et al., 1976). Furthermore, the arrangement or interspersal of repeated sequences among the single copy sequences can be determined by measuring the distance between repeats (Graham et al., 1974). Knowledge of these characteristics is useful in molecular cloning experiments since it provides predictions of sequence organization. Knowledge of the sequence organization of the repetitive elements may also be important to understanding the regulation of gene expression in eukaryotes (Davidson and Britten, 1979).

Mechanics of DNA Reassociation Experiments

DNA reassociation experiments all have in common several important features. In my experience, these features can be somewhat confusing when first encountered. The following comments, addressed to the newcomer to reassociation experiments, provide a rationale for the way in which these experiments are performed and, hopefully, stress the most important steps. Perhaps the best single reference for how to perform reassociation experiments is the classic paper by Britten, Graham, and Neufeld (1974).

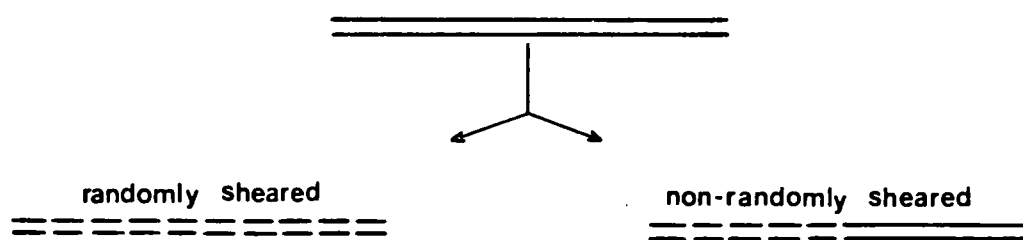
Purification of DNA. The method by which DNA is isolated for a reassociation experiment is not usually critical provided that the resulting preparation is free from 1) protein, 2) RNA, 3) divalent cations, and 4) mono- and oligonucleotides. Protein, RNA, and divalent cations may each interact with single stranded DNA in ways which alter the "normal" reassociation reaction. Protein and RNA are typically removed enzymatically as the first steps in DNA purification. Organic extractions using phenol and/or chloroform insure the removal of protein. Divalent cations are best removed (chelated) by passage through Chelex-100 resin. Chelation by adding EDTA to the DNA solution is less desirable since this chelator can destroy hydroxyapatite (which is used to separate the products of the reassociation) and absorbs light at 260 nm (the wavelength at which DNA concentration is measured). Mononucleotides need to be removed simply because their extinction coefficient at 260 nm is large (larger than that of DNA), and significant contamination may result in erroneous values when DNA concentration is measured spectrophotometrically. Mononucleotides are usually

removed by gel filtration using Sephadex G-50. In practice, mono-nucleotides and divalent cations are removed sequentially following shearing of the DNA, using a two layer column containing Chelex-100 and Sephadex G-50.

Shearing of DNA. As mentioned above, the rate-limiting step in a reassociation reaction is the nucleation event. If the DNA to be reassociated is first broken into small pieces, then a greater number of nucleation events will be required, and the rate of reassociation will be retarded. Thus, one reason for shearing DNA is to control the rate of reassociation. A second reason for shearing is to separate repetitive sequences from unique sequences. All eukaryotic DNAs contain repetitive sequences, and some contain a greater proportion of repetitive than unique sequences (Lewin, 1980). These repetitive sequences are often very short (a few hundred base pairs) and interspersed among the single copy sequences. If these DNAs are not sheared, then reassociated repetitive sequences will "carry along" long stretches of single stranded, unique sequences. If the reassociated repetitive sequences are then isolated on hydroxyapatite (see below), large amounts of unreassociated, single copy sequences will also be isolated. Thus, eukaryotic DNAs are usually fragmented to a very short length (a few hundred base pairs) to reduce the fraction of molecules containing both repetitive and unique sequences.

DNAs for reassociation experiments are almost always sheared in a random manner. Several satisfactory methods are available, but high speed blending (Davidson et al., 1973) is probably the best. This method provides great reproducibility with no chemical alteration of

the DNA. Methods which fragment DNA non-randomly (e.g. restriction endonuclease digestion) are not desirable since the resulting DNA reassociation may not follow second order kinetics (see section on statistical analysis of reassociation data). As an illustration of the effect of non-random shearing, consider two identical preparations of unsheared DNA. Now, suppose that one is randomly sheared and the other is non-randomly sheared. The resulting DNAs might be diagrammed as follows:



These diagrams are simplified. Random shearing does not produce DNA fragments of uniform size, but rather each sequence is contained in a set of fragments of various sizes. However, the size distribution and average length of each set of fragments are the same. Also, the fragments produced by endonuclease digestion are of discrete size and each contains specific sequences, but the pattern of digestion is not easily predicted. The asymmetrical pattern shown above was chosen to clarify the argument. If each preparation of fragmented DNA is dissociated and placed at reassociation conditions, nucleation and zippering reactions will begin. In the case of the randomly sheared DNA, the rate of reassociation will be uniform for all of the sequences (assuming that none of the sequences are repeated). This will not be the case with the non-randomly sheared DNA. Each nucleation involving the sequences in the "right-half" of the molecule will result in all of the sequences

in that half zippering to become double stranded. For the "left-half" of the molecule, many more nucleation and zippering reactions are required to give the same amount of double stranded DNA. In other words, the sequences in the "right-half" of the non-randomly sheared DNA molecule will reassociate more quickly than the sequences in the "left-half." Another example of the problems encountered by non-random shearing is provided by Singer (1979). He reported genome size measurements for a randomly sheared DNA which differed depending on whether a randomly sheared standard DNA or a standard DNA of discrete fragment size was used.

Dissociation of DNA. Before a reassociation reaction can be started, the DNA must be dissociated. The method of choice depends upon the nature of the experiment. For most experiments, one wishes to use dissociation conditions which can easily, rapidly, and reproducibly be shifted to renaturation conditions. For other experiments, one might sacrifice simplicity if it is desirable to allow a reassociation to proceed only briefly.

The simplest method of dissociating DNA is by heating. The amount of heat required depends upon 1) the G+C content of the DNA, 2) the length of the DNA, 3) the concentration of monovalent cations, and 4) the presence of chemical denaturants (e.g. formamide). The melting temperature, T_m , (i.e. the temperature at which 50% of the base pairs are dissociated) of a given DNA solution is best determined experimentally; however, theoretical equations are available which permit quite accurate estimates to be made (Mandel and Marmur, 1968). When shifting to the reassociation temperature, thermal equilibration

is rapidly achieved if the volume of the DNA solution is small. Solution with small volumes and large surface to volume ratios can be obtained by sealing aliquots of the DNA in siliconized, glass capillary tubes. (Siliconization prevents the DNA from sticking to the glass.) The capillary method also eliminates the problems of evaporation which are encountered when samples are removed from a large reaction during the course of a reassociation experiment.

Another simple method of dissociating DNA is by alkaline treatment. With this method one adds base (e.g. NaOH) to raise the pH to approximately 13. The dissociation should not be performed at elevated temperatures so that DNA degradation will be minimal. This method might be considered slightly more complicated than thermal dissociation because two steps are required to begin the reassociation: 1) neutralization with acid (e.g. HCl) and 2) equilibration to the reassociation temperature. One must be very careful to keep track of changes in the salt and DNA concentrations which result with the additions of base and acid. Furthermore, since the solution must be neutralized, the use of capillary tubes is not possible and evaporation may be a problem during the reassociation.

When it is desirable to obtain the reassociation products from a very brief reaction, the following procedure may be used. DNA dialyzed against H_2O is briefly heated (to insure dissociation) and then quickly cooled. Without salt, no reassociation can occur. The DNA is then equilibrated to the desired temperature, and salt is added to initiate the reaction. For the briefest reactions, the DNA is immediately passed through hydroxyapatite to separate single and double stranded DNA (see below for further explanation). This procedure

permits thermal equilibration to be achieved before any reassociation can occur and eliminates the time required to break open a capillary tube (see Perlman et al., 1976).

Conditions for reassociation. The progress of a reassociation reaction is expressed in terms of Cot. Cot is defined as the product of the DNA concentration (in moles of nucleotides per liter) and time (in seconds). It is appropriate to use the term Cot only when standard reassociation conditions are used. By convention, these conditions are (1) 0.12 M Na phosphate, pH 6.8 (= 0.18 M Na⁺) and (2) $T = T_m - 25^{\circ}\text{C}$ (i.e. 25^oC below the melting temperature). Other conditions can be used. Often the salt concentration is increased or decreased to speed up or slow down the rate of a reassociation. When non-standard salt conditions are used, a rate correction can be made to permit comparison with standard conditions (using Table I in Britten et al., 1974), but the appropriate term is then equivalent Cot (ECot). Similarly, if a non-standard temperature is used for reassociation, a correction can be made for comparison with the $T_m - 25^{\circ}\text{C}$ standard (Bonner et al., 1973). If both the temperature and salt conditions deviate from the standard conditions, it is not so easy to determine how the results would relate to the same experiment under standard conditions. In other words, performing reassociations with standard conditions, or conditions which can be easily normalized to standard conditions, is desirable since it allows comparisons of the results to be made with a large body of literature.

As discussed earlier, the fragment size also influences the rate of reassociation. There is no standard fragment size; however,

corrections can be made for differences in lengths so that results can be compared (Wetmur and Davidson, 1968). If possible, it is probably best to shear to the same size DNAs to be compared.

When very large Cot values are required, as is often the case with the very large and complex genomes of eukaryotes, reassociations must be allowed to proceed for days and sometimes weeks. Exposure to high temperatures for such extended periods of time can result in considerable degradation of the DNA. To minimize this degradation, reassociations are sometimes performed at lower temperatures in the presence of organic reagents which lower the melting temperature. The most widely used organic reagent is formamide (Hutton, 1977).

Termination of reassociation reactions. When measuring reassociation kinetics, it is imperative that one be able to stop a reassociation reaction nearly instantaneously; however, it is not always necessary to do so. The importance of stopping a reaction depends upon the speed of the reaction. When running a reaction containing a DNA of large sequence complexity to a very high Cot (which may take a week or more), a few minutes to either side of the projected termination time will not make any measurable difference in the amount of double stranded DNA obtained. On the other hand, reassociations of DNAs with low sequence complexity in high concentrations (e.g. repetitive sequences) may proceed very rapidly with the reaction being essentially complete in a matter of minutes. In this case, a few seconds to either side of a projected termination time can make considerable difference. When studying an unknown DNA, it is always wise to begin by running a reassociation using only a few widely separated Cot points in order to get a feel for the rate of reaction. One can

then go back and obtain additional Cot points in the regions of the kinetic curve where rapid reassociation occurs.

When it is necessary to stop a reaction rapidly, one has the choice of three parameters which can be adjusted quickly to discourage further reassociation: 1) temperature, 2) DNA concentration, and 3) salt concentration. If the reactions are carried out in capillary tubes, the best method is rapid freezing, usually by submersion in a dry ice-ethanol bath. (Slow freezing is to be avoided since it permits the concentration of DNA and salt in the fluid between slowly forming ice crystals, and this may result in further reassociation; Britten, et al., 1974.) The rapid change in temperature does not damage the capillaries, and even large volumes (e.g. 100 μ l) freeze very quickly. One must bear in mind, however, that fractionation of the reassociation products must be performed at higher temperatures (at the reassociation temperature in the case of hydroxyapatite chromatography), and steps must be taken to reduce further reassociation. This is simply accomplished by dilution of the frozen contents of the capillaries. The rate of reassociation is directly proportional to the square of the DNA concentration. Thus, a two fold dilution of the DNA concentration would result in a four fold reduction in the reassociation rate (see section on statistical analysis, below). Changes in the salt concentration can also dramatically change the reaction rate. For example, if the salt concentration is reduced from 0.12 M NaP to 0.06 M NaP (only a two fold dilution), then a 5.5 fold reduction in reassociation rate will be observed (see Table 1 in Britten et al., 1974).

Systematic errors are not easily dealt with in reassociation reactions. At first thought, one might suppose that as long as each Cot point is handled in an identical, reproducible manner and as long as the standard DNA is handled in a manner identical to the unknown DNA,

systematic errors would cancel out of calculations of sequence complexity. This is not the case. For example, assume that one is studying a previously unexamined eukaryotic DNA with DNA from E. coli being used as the reassociation standard. Furthermore, assume that one does not stop the reassociation reactions quite fast enough so that each reaction proceeds somewhat longer than realized. This systematic error will have less effect on the eukaryotic reaction since it will reassociate more slowly. In other words, the rate of the E. coli DNA reaction will be overestimated while the rate measured for the eukaryotic DNA reaction will be closer to its true value. However, since the calculation of the eukaryotic DNA complexity requires use of the measured reassociation rate of the E. coli DNA, the calculated eukaryotic complexity will also be in error (and overestimated).

Fractionation of reassociation products. The progress of reassociation reactions can be monitored by three techniques: 1) digestion of unreacted DNA by the single strand specific, S1 nuclease (Sutton, 1971; Smith et al., 1975), 2) measurement of optical hypochromicity (Subirana and Doty, 1966; DeLey et al., 1970; Gillis et al., 1970), and 3) hydroxyapatite separation of single and double stranded DNA (Kohne and Britten, 1971; Davidson et al., 1973; Britten et al., 1974). Each of these methods has its own advantages and disadvantages. In the experiments described in this dissertation only hydroxyapatite chromatography was used, and therefore, only this method will be discussed here.

Hydroxyapatite (HAP) is a calcium phosphate compound with a chemical composition of $\text{Ca}_3(\text{PO}_4)_2$. DNAs, both double stranded (ds)

and single stranded (ss), bind to the HAP when the phosphate concentration of the solution is low (e.g. 0.01 - 0.03 M NaP, pH 6.8). At moderate phosphate concentrations (e.g. 0.12 - 0.14 M NaP), ssDNA is eluted while dsDNA remains bound. At high phosphate concentrations (e.g. 0.4 - 0.5 M NaP) dsDNA is also eluted. Structural parameters other than simple strandedness influence the binding of DNA to HAP. For example, DNAs having a duplex structure shorter than 50 base pairs do not bind efficiently to HAP in moderate concentrations of phosphate (Wilson and Thomas, 1973). Also, dsDNAs with a duplex greater than 50 base pairs in length may not bind efficiently to HAP if they also have significant single stranded regions attached (Bell and Hardman, 1977). The mechanism by which HAP fractionation operates is not well understood. About all that can be said with certainty is that increases in base pairing and base stacking interactions, which lead to increased molecular order and decreased configurational entropy, lead to increases in affinity for HAP (Martinson, 1973).

When DNA from a reassociation reaction is to be fractionated on HAP, it is usually passed through a small HAP column (e.g. 1 ml) in phosphate buffer of moderate concentration at the reassociation temperature. The use of a moderate phosphate concentration permits the ssDNA to be washed through the column while the dsDNA binds to the HAP. Care must be taken not to change the stringency of the reassociation conditions during the fractionation procedure. For example, if high phosphate concentrations were used to reach high C_{ot} values, one must dilute the sample in order to permit the dsDNA to bind to the column. One must remember, however, that lowering the salt concentration also lowers the melting temperature and the temperature of

reassociation. More specifically, one must be careful to lower the temperature if the salt concentration is lowered. As another example, suppose that a reassociation reaction was performed at 55°C in 0.12 M NaP. If the HAP column was equilibrated at 60°C, some poorly matched duplexes might dissociate at the higher temperature, and this may or may not be important to the particular experiment. One must also keep in mind that DNA binds nonspecifically to HAP at temperatures below 50°C (Brenner et al., 1969). If an experiment requires that lower temperatures be used, formamide can be introduced in the system to prevent nonspecific binding (Monahan and Hall, 1975).

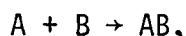
Two methods are available for the elution of dsDNA from HAP: 1) high phosphate washing and 2) thermal dissociation. Both methods are widely used, and the method of choice depends upon the nature of the experiment. If a preparative fractionation of DNA is to be performed, the high phosphate wash is best. Most preparative fractionations involve at least two cycles of thermal dissociation, reassociation, and HAP separation. Use of the high phosphate method of elution reduces the exposure of the DNA to high temperatures and therefore reduces the possibility of thermal degradation. Moreover, in some cases, one may want to recover the DNA in double stranded form. If an analytical fractionation of DNA is to be performed, thermal degradation is irrelevant since the eluted DNA will be discarded. In this case, the elution method of choice is largely a matter of personal preference. If the analysis is to employ assay of radioactivity, one must bear in mind the possibility of quenching problems. When scintillation fluids are used to measure radioactivity, increasing phosphate concentrations may cause decreasing values for cpm.

This problem can be totally avoided when ^{32}P -labeled DNA is used since measurements of radioactivity can be made by Cherenkov counting, but care must be taken when using ^3H -labeled DNA. Throughout the reassociation literature, very little mention is made concerning the quenching problem, but careful examination of the reported HAP methods indicates that experts in the field are well aware of it. When ^3H -DNA is used, so is the thermal elution method (Davidson et al., 1973) in order to maintain a constant phosphate concentration. If the high phosphate elution method is used, the column fractions are diluted to the same phosphate concentration before counting (Graham et al., 1974).

The separation of reassociated DNA from single stranded DNA by the hydroxyapatite method can be simple, reproducible, and reliable, but each investigator should perform careful control experiments to insure that problems do not arise. As an example, a detailed description of the hydroxyapatite control experiments that I performed can be found in the Appendix.

Statistical Analysis of Reassociation Data

After measuring the fraction of single stranded DNA remaining in a reaction mixture at various values of Cot , the kinetic curve which best describes the data can be found by computer assisted least squares analysis (Pearson et al., 1977). This analysis is based upon the observation that the reaction of randomly shear DNA follows second order kinetics when assayed by the hydroxyapatite method. Second order reactions have the general form



where A and B represent the randomly sheared, complementary sequences, and AB represents the double stranded reaction product. Mathematically, the second order reaction of single stranded DNA can be described as

$$\frac{dC}{dt} = -KC^2,$$

where C is the concentration of single stranded DNA and K is the second order rate constant. In words, the single stranded DNA concentration decreases with time at a rate proportional to K and the square of the concentration of single stranded DNA. (The value of K is dependent upon such things as salt concentration, temperature, fragment size, and the presence of non-reacting sequences.) Integration of the above equation yields

$$C/C_0 = \frac{1}{1 + KCot},$$

where C_0 is the initial single stranded DNA concentration, and C/C_0 is the fraction of single stranded DNA present at a given value of Cot.

The reassociation rate constant (K) can be calculated if the $(Cot)_{1/2}$ of the reaction (i.e. the Cot at which 50% of the sequences are measured as double stranded) is known. When C/C_0 equals 0.5, then K equals $1/(Cot)_{1/2}$. Proof of this follows.

$$\text{Given: } 0.5 = \frac{1}{1 + K(Cot)_{1/2}},$$

rearranging terms,

$$0.5 (1 + K (Cot)_{1/2}) = 1,$$

solving for K,

$$0.5 + 0.5 K(\text{Cot})_{1/2} = 1,$$

$$0.5 K(\text{Cot})_{1/2} = 0.5,$$

$$K(\text{Cot})_{1/2} = 1 ,$$

$$K = \frac{1}{(\text{Cot})_{1/2}} ,$$

or

$$(\text{Cot})_{1/2} = \frac{1}{K} .$$

If the DNA contains one or more sequences that are repeated, each of these will also react according to second order kinetics, but at a rate faster than that of the single copy sequences. The overall kinetic curve describing the reaction of the mixture of single copy and repeated sequences will represent the sum of all the kinetic components. Furthermore, the rate of reaction of each kinetic component in the mixture will be slower than if each component was reassociated by itself, because more collisions between non-reacting (i.e. non-complementary) sequences will occur. Thus, the rate constant observed for each kinetic component in the mixture (K_{obs}) will be smaller than the rate constants observed using each component by itself (K_{pure}). Estimates of K_{pure} for a particular component can be made by dividing K_{obs} by the fraction of the total DNA contained in that component (Lewin, 1980).

CHAPTER II

MATERIALS AND METHODS

Organisms

Candida albicans strain H317 (obtained from Dr. Jeffrey M. Becker, The University of Tennessee, Knoxville) was used throughout the experiments described. This strain is of interest because of its ability to undergo 100% conversion from the yeast to the mycelial phase of growth (Torres-Bauza, 1980). This strain was confirmed as C. albicans by standard tests (Silva-Hunter and Cooper, 1974). Stock cultures were maintained on Sabouraud agar Modified (Difco) at 4°C after about 48 hours of growth.

Escherichia coli strain K12, CSH42 (purchased from Cold Spring Harbor Laboratory) was used as a reference strain. This bacterium is Thr⁻Leu⁻Thy⁻Ilv⁻Bl⁻DNA_{ts} (Miller, 1972). Its DNA has a (G + C) content of 50% and a sequence complexity of 4.2×10^6 base pairs (Lewin, 1980).

Growth and Labeling of Cells

Low phosphate medium. Cells to be labeled with ³²P-phosphoric acid were grown in a phosphate depleted, rich medium. This medium was prepared by adding 10 ml of 1 M MgSO₄ and 10 ml of concentrated aqueous ammonia to 1 liter of 1% Bacto-yeast extract (Difco) and 2% Bacto-peptone (Difco). The phosphates (MgNH₄PO₄) were allowed to precipitate for 30 minutes at room temperature, and then they were removed by filtration through Whatman No. 1 filter paper. The filtrate was adjusted

to pH 5.8 with concentrated HCl and autoclaved. Sterile glucose was added to a final concentration of 2%. This procedure has been found to give a medium with a low inorganic phosphate concentration (on the order of 10^{-4} M) and a relatively high proportion of organic phosphates (Rubin, 1975). To label DNA with ^{32}P -phosphoric acid, a slight modification of the method of Rubin (1973) was followed. Cells were grown to exponential phase in low phosphate medium at room temperature, harvested, and resuspended in fresh medium at a concentration of 5×10^7 cells/ml. To each liter of cells, 50 mCi of inorganic ^{32}P -phosphate (Amersham) was added, and the culture was further incubated at 25°C for two generations. DNA isolated from cells labeled in this manner had a specific activity of about 5×10^5 cpm/ μg .

Yeast nitrogen base medium. Cells to be labeled with ^3H -uridine were grown in a minimal medium containing 1.34% Bacto-yeast nitrogen base without amino acids (Difco), 1% glucose, and 30 $\mu\text{g}/\text{ml}$ adenine (Hatzfeld, 1973). It has been reported that the addition of adenine to the medium gives better incorporation of label in Saccharomyces cerevisiae (Burns, 1966). C. albicans H317 does not require adenine for growth, and experiments to determine the efficiency of labeling in its absence were not performed. Cells to be labeled were grown to exponential phase in the above medium at room temperature, harvested, and resuspended in fresh medium at a concentration of 5×10^7 cells/ml. To each liter of cells, 5 mCi of either (5, 6 - ^3H)uridine or (6 - ^3H)uridine (Amersham) was added, and the culture was further incubated at 25°C for two generations. DNA isolated from cells labeled in this manner had a specific activity of $1 - 2.5 \times 10^4$ cpm/ μg .

YEPL medium. Cells used for studying mitochondrial DNA were grown in medium containing 1% Bacto-yeast extract, 2% Bacto-peptone, and 3% L-sodium lactate. Details about this medium can be found in Chapter 6 (Part II).

Bacterial DNA labeling medium. *E. coli* cells to be labeled with ^3H -thymine were grown at 30°C in a minimal medium containing (per liter): 10.5 g K_2HPO_4 , 4.5 g KH_2PO_4 , 1 g $(\text{NH}_4)_2\text{SO}_4$, 0.5 g Na Citrate $\cdot 2\text{H}_2\text{O}$, 0.2 g MgSO_4 , 0.5 ml of 1% thiamine, 10 ml of 4 mg/ml of each required amino acid (threonine, leucine, isoleucine, and valine), 10 ml of 20% glucose, and 50 $\mu\text{g/ml}$ thymine. Cells to be labeled were resuspended in fresh medium containing 25 $\mu\text{g/ml}$ thymine at a density of 0.8 A_{550} units per ml. To each liter of cells, 3.4 mCi of (6 - ^3H) thymine (Amersham) was added, and the culture was further incubated at 25°C for two generations. DNA isolated from cells labeled in this manner had a specific activity of 7×10^3 cpm/ μg .

Preparation of Spheroplasts

Osmotically sensitive cells were obtained by removing the cell walls from *C. albicans* following the method of Torres-Bauza and Riggsby (1980). Cells were harvested at 5000 rpm for 5 minutes at 4°C and washed twice with H_2O . The cells were then resuspended at a concentration of 0.2 grams/ml in freshly prepared pretreatment solution (0.77 grams dithiothreitol, 0.19 grams EDTA, 1.21 grams Tris, and 0.1 grams pronase per 100 ml of solution). After gentle shaking for 1.5 hours at 32°C , the cells were harvested and washed twice with osmotic buffer (3.5 grams K_2HPO_4 , 22.4 grams KCl, and pH adjusted to 6.0 per 500 ml of solution). For the treatment solution, the cells were resuspended in

osmotic buffer at a concentration of 0.2 grams/ml, and to each ml of suspension the following were added: 10,000 units of glucuronidase (i.e. glucuronidase), 0.2 mg of chitinase, and 0.5 μ l of mercaptoethanol. After 2 hours of gentle shaking at 32°C, a sample was checked for lysis in dH₂O. The osmotically sensitive spheroplasts were then harvested and washed three times with osmotic buffer and one time with 1 M Sorbitol, 0.1 M EDTA, pH 7.5.

Purification of Nuclei

Spheroplasts for isolation of nuclei were washed one time with 1M Sorbitol. Some nuclei were purified by the methods of Torres-Bauza (1980), and some were purified by modifications of those methods. A brief description of all the methods follows.

Crude nuclei method. The spheroplasts were resuspended at a concentration of about 0.2 g/ml of Buffer A (19% $\pm$ 0.5% (w/w) Ficoll, 0.5 mM MgCl₂, 20 mM KH₂PO₄, pH 6.5, containing 1 mM phenylmethylsulphonyl fluoride (PMSF)). After incubating 5 minutes at 0°C to allow the spheroplasts to lyse, an equal volume of Buffer B (19% $\pm$ 0.5% (w/w) Ficoll, 10 mM MgCl₂, 20 mM KH₂PO₄, pH 6.5, containing 1 mM PMSF) was added and mixed. Unbroken cells were removed by centrifugation at 5500 rpm for 10 minutes in a Sorvall SS34 rotor. The supernatant was transferred to a clean centrifuge tube and overlaid with 1/3 volume (relative to the total volume of Buffers A and B) of Buffer C (12% $\pm$ 0.5% (w/w) Ficoll, 5 mM MgCl₂, 20 mM KH₂PO₄, pH 6.5, containing 1 mM PMSF). The crude nuclei were then pelleted by centrifugation at 10,000 rpm for 35 minutes.

Discontinuous sucrose gradient method. Crude nuclei (prepared as described above) from about 1 gram of spheroplasts were resuspended in 8 ml of 0.25 M sucrose prepared in TKMCF buffer (50 mM Tris, pH 7.5, 25 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 6% Ficoll, containing 1 mM PMSF). This suspension was layered on a discontinuous gradient prepared with the following sucrose solutions (all in TKMCF buffer): 10 ml of 1.7 M, 10 ml 1.5 M, and 8 ml of 1.0 M sucrose. The gradients were prepared in polyallomer tubes and centrifuged at 27,000 rpm for 45 minutes at 4°C in a Beckman SW28 rotor. The purified nuclei formed a pellet at the bottom of the centrifuge tube.

Differential centrifugation method. Crude nuclei (prepared as described above) from about 1 gram of spheroplasts were resuspended in 8 ml of Buffer D. (For approximately 100 ml of Buffer D, the following components were mixed: 18.2 g Sorbitol, 25 ml 36% Ficoll, 30 g Glycerol, 10 ml 0.2 M KH_2PO_4 , 0.5 ml 1 M $MgCl_2$, and 80 ml H_2O . The pH was adjusted to 6.6 using 1 N KOH, and the refractive index was adjusted to 1.417 using H_2O .) The suspension was centrifuged at 5000 rpm for 5 minutes to remove unbroken cells, and the supernatants were transferred to clean centrifuge tubes. Additional unbroken cells were removed by repeating the low speed centrifugation one or two times. The purified nuclei were then pelleted by centrifugation at 10,000 rpm for 35 minutes.

Triton X-100 methods. This method was the same as the crude nuclei method (above) with the following modifications. After lysing the spheroplasts in Buffer A, an equal volume of Buffer B was added containing either 0.0%, 0.01%, 0.04%, or 0.2% Triton X-100 (w/v). After

removing unbroken cells and pelleting the nuclei, the nuclei were resuspended in a solution containing equal volumes of Buffers A and B and either 0.0%, 0.005%, 0.02%, or 0.1% Triton X-100. These suspensions were again overlaid with Buffer C, and centrifuged to pellet the nuclei.

Purification of Mitochondria and mtDNA

The methods for isolating mitochondria and mitochondrial DNA are described in Chapter 6 (Part II).

Extraction of DNA from Spheroplasts and Nuclei

The method was derived from the procedure of Cryer et al. (1975). Freshly prepared spheroplasts or nuclei were resuspended in 0.15 M NaCl, 0.1 M EDTA, pH 8.0 containing 100 µg/ml proteinase K. (Proteinase K stock solutions were prepared at a concentration of 5 mg/ml in 10 mM Tris, 5 mM CaCl₂, pH 8.6, self-digested for 30 minutes at 37°C, and stored at -20°C.) SDS was added to a final concentration of 1%, and the lysate was incubated at 37°C for 1-2 hours. The solution was then twice deproteinized by extraction with an equal volume of phenol-chloroform-isoamyl alcohol (50:50:1) and twice with chloroform-isoamyl alcohol (24:1). The nucleic acids were precipitated by the addition of 2 volumes of 95% ethanol and storage at -20°C overnight. The nucleic acids were collected by centrifugation at 2-3000 rpm for 5 minutes, dissolved in 0.1 X SSC (1 X SSC is 0.15 M NaCl, 0.015 M Na citrate, pH 7.0), made up to 0.3 M Na acetate, 0.1 mM EDTA, pH 7.0, and precipitated with 0.54 volumes of isopropanol. (This step not only differentially precipitates DNA but also eliminates a water soluble, yellow pigment which apparently is a component of the glucylase solution used

in preparing spheroplasts.) The isopropanol precipitate was again dissolved in 0.1 X SSC, made up to 1 X SSC, and digested with RNase A (25 µg/ml), TI RNase (100 units/ml), α-amylase (25 µg/ml), and β-amylase (25 µg/ml) at 37°C for 1 hour. The solution was again deproteinized, ethanol precipitated, isopropanol precipitated, enzyme treated, and deproteinized as above.

This method of extraction consistently yielded DNA with A_{260}/A_{280} of about 1.95 and A_{260}/A_{230} of about 2.25. No protein was detectable by the method of Lowry and no RNA was detectable by the orcinol method (Torres-Bauza, 1980). Furthermore, alkali treatment of purified radioactive DNA did not reduce the TCA precipitable cpm indicating the extraction method eliminates all labeled RNA (data not shown).

Extraction of DNA from *E. coli*

E. coli DNA was isolated using the protocol described above after harvesting the cells (6000 rpm for 5 minutes) and washing one time with 0.15 M NaCl, 0.1 M EDTA, pH 8.0.

CsCl Density Gradient Analysis of DNA

The buoyant density of DNA, which in CsCl gradients has a largely linear relationship with (G+C) content, was measured following the protocol of Mandel et al. (1968). Optical grade CsCl was purchased from Gallard and Schlesinger, and 65% (w/w) stock solutions were prepared in 20 mM Tris, pH 8.5. CsCl gradient purified *E. coli* DNA (50% G + C) and *M. lysodeikticus* DNA (72% G + C), purchased from Sigma, were used as reference DNAs. *C. albicans* DNA was dialyzed against 1 mM NaCl, 1 mM Tris, 1 mM EDTA, pH 7.0 prior to analysis. Samples to be

centrifuged were prepared by mixing 1) 0.84 ml of CsCl stock solution, 2) 0.5 μ g of reference DNA, 3) 3 μ g of *C. albicans* DNA, and 4) H_2O to bring the volume to 1.07 ml. The refractive index was adjusted to 1.4000 (i.e. density equals 1.71 g/cm³). Centrifugation was for 20 hours at 44,770 rpm at 25°C in a Spinco Model E analytical ultracentrifuge. Calculations of buoyant density were made using

$$P = P_0 + 0.0092 (r^2 - r_0^2),$$

where P_0 is the density of the reference DNA and r_0 is the distance of the reference DNA from the center of rotation. Corresponding values for mole fraction of (G + C) were calculated using

$$(G + C) = \frac{P - 1.660 \text{ g/ml}}{0.098}.$$

Nucleotide Separation by Cation-Exchange Chromatography

DNAs for chromatographic nucleotide analysis were first digested with DNaseI to oligonucleotides of an average length of four nucleotides (Matsuda and Ogoshi, 1966), and then digested with snake venom phosphodiesterase to 5'-mononucleotides (Hurst et al., 1951). Following the procedure of Miyazawa and Thomas (1965), DNA samples were dialyzed to 10 mM NaCl, 5 mM Tris, pH 7.3; the samples were then made 2 mM $MgCl_2$, and pancreatic DNase I (Worthington) was added to a final concentration of 10 μ g/ml. After incubation at 37°C for 2 hours, 0.05 volume of 1.0 M Tris, pH 9.0 was added and the Ca^{++} concentration adjusted to 5 mM. Venom phosphodiesterase (Worthington), treated to inactivate 5'-nucleotidase activity (Sulkowski and Laskowski, 1971), was added to

a final concentration of 6 units/ml, and incubation continued at 45°C for 2 hours.

The separation of 5'-mononucleotides was achieved using the method of Blattner and Erickson (1967). The cation-exchange resin was AG 50W-X4, minus 400 mesh (Bio-Rad), and the elution buffer was 0.1 M ammonium formate, adjusted to pH 3.2 with formic acid. The resin was suspended in H₂O and washed by decantation two times in 1N HCl, three times in concentrated NH₄OH, and three times in elution buffer. Columns 40 cm in length and 0.8 cm in diameter were packed using a peristaltic pump to maintain the flow rate at 20 ml/hour. Digestion mixtures (of 1 ml) to be chromatographed were mixed with an equal volume of elution buffer, loaded, and washed through the column at a rate of 20 ml/hour. The nucleotides (which elute in the order T, G, C, A) were collected in 50 fractions of 1 ml each; these were then pooled according to nucleotide. When unlabeled nucleotides were used, the pooled fractions were adjusted to the same volume, neutralized, and analyzed spectrophotometrically (using the extinction coefficients cited by Cohn, 1955) to determine the nucleotide concentrations. When ³²P-labeled nucleotides were used, the pooled fractions were adjusted to the same volume, and measured for radioactivity. Corrections for free ³²P were made by adding 0.5 ml of acid-washed Norit, filtering on Millipore filters, and measuring the filtrate for radioactivity.

Shearing of DNA

DNA was sheared to an average length of about 300 base pairs by the high speed blending method of Britten et al. (1974). DNA in 10 ml of low salt buffer (20 mM Na Acetate, pH 6.6 or 15 mM NaCl, 1.5 mM

Na Citrate, pH 7.0) was mixed with 20 ml of glycerol containing the same low salt buffer. The resulting 30 ml of 66% glycerol solution was homogenized for 30 minutes at 48,000 rpm in a Model 60K Virtis homogenizer using a 100 ml cup, a single blade, and cooling with an ethanol-dry ice bath. The sheared DNA was precipitated from the glycerol solution at -20° overnight after adding 0.1 volume of 3M Na Acetate, 1 mM EDTA, pH 7.0 and 2 volumes of 95% ethanol. The precipitate was collected by centrifugation at 25,000 rpm for 60 minutes in pollyallomer tubes in a Beckman SW27 rotor. The sheared DNA was then dissolved in low salt buffer and stored at 4°C to await further purification.

Purification of Sheared DNA Fragments

Sheared DNA in low salt buffer was chromatographed on a two resin column containing Chelex 100 (Bio-Rad) in the lower half and Sephadex G-50 (Pharmacia) in the upper half. Sephadex was prepared by swelling the beads overnight in 0.12 M NaP, pH 6.8 and then boiling for 10 minutes to inactivate nucleases. Chelex 100 was prepared following the method of Willard et al. (1970). 10 g of Chelex was suspended in 100 ml of concentrated NH_4OH at room temperature for 1 hour. It was then washed 10 times on a sintered glass funnel by repeated suspension in 150 ml of H_2O followed by filtration with suction. The Chelex was then suspended in 300 ml of 0.1 N HCl and filtered. The resin was washed again with water and then several times with 0.12 M NaP, pH 6.8. It was stored at 4°C in 0.12 M NaP. After construction of the two layer column, it was washed with 0.12 M NaP, pH 6.8 to insure equilibration. The DNA was then passed through the column to remove mononucleotides (Sephadex G-50) and divalent cations (Chelex 100).

DNA solutions requiring concentration were put in dialysis tubing, covered with dry Sephadex G-200, and placed at 4°C until the volume was reduced to the desired level. Prior to reassociation, all DNAs were extensively dialyzed against 0.12 M NaP, pH 6.8.

Sizing of DNA Fragments

Electrophoresis of sheared DNA fragments was conducted under non-denaturing conditions as described by Fisher and Dingman (1971) using 2.5% polyacrylamide-0.5% agarose gels and the electrophoresis buffer of Peacock and Dingman (1968). 10X Peacock's buffer contained (per liter): 108 grams Tris, 55 grams boric acid, and 9.3 grams EDTA. To prepare 100 ml of gel solution, 12.5 ml of Acrylamide-N, N'-methylene-bisacrylamide (19%: 1%), 10 ml of 10X Peacock's buffer, and 15 ml of H₂O were mixed with 50 ml of 1% agarose (in dH₂O; cooled to 45°C) and, lastly, 6.25 ml of 0.5% TEMED and 6.25 ml of 1.6% NH₄ persulfate. The solution was then rapidly poured into a 150 X 140 X 4 mm vertical gel apparatus and allowed 1 - 2 hours to polymerize. Pre-electrophoresis was for 30 minutes at 80 volts. Samples were loaded on the gel after mixing with 40% sucrose, 1 mM EDTA, 0.001% Bromophenol blue, pH 6.2 (approximately 1 : 4 dilution). Hinc II digests of ØX174 RF DNA (New England Biolabs) were used as markers. Electrophoresis was for 2 hours at 100 volts, constant voltage. DNA in gels was stained with 0.5 µg/ml ethidium bromide for 30 minutes and then visualized and photographed with u.v. light.

Measurement of Reassociation Kinetics Using Hydroxyapatite

Samples of sheared, purified DNA in 0.12 M NaP, pH 6.8 were sealed in sterile, siliconized, 20 - 100 µl capillary tubes, heated to

105°C for 5-10 minutes, and placed in a water bath set at the reassociation temperature (64.6°C for E. coli DNA and 58.9°C for C. albicans DNA). The reassociation temperatures were 25°C below the T_m (calculated using the equations of Mandel and Marmur, 1968). Capillaries were removed at various times, and the reactions were quickly stopped by freezing in an ethanol-dry ice bath. To discourage further reassociation in subsequent manipulations during HAP fractionation, the contents of each capillary were diluted in 1 ml of 0.12 M NaP, pH 6.8 containing 50 µg/ml of calf thymus DNA, and again frozen.

Hydroxapatite fractionations were performed in water-jacketed glass columns equilibrated at the reassociation temperature. A HAP slurry was prepared according to the manufacturer's instructions (BioRad) and was heated to 90 - 100°C for 30 minutes prior to use. Each HAP column was prepared by layering a volume of slurry equivalent to 1 ml of wet, packed HAP on top of a 1/2 inch bed of glass beads and pressure packing at 3 - 5 psi. Each HAP column was used for only one DNA sample; 1 N HCl was used to dissolve HAP.

Each 1 ml sample to be chromatographed was thawed, loaded onto the column, and allowed to equilibrate to the reassociation temperature (about 1.5 minutes). The sample was then run through the HAP under pressure until the meniscus reached the top of the bed, and the 1 ml eluate was collected for scintillation counting. Unreassociated DNA was further eluted from the column five times with 1 ml washes with preheated 0.12 M NaP, pH 6.8; the eluate from each wash was reserved for scintillation counting. To elute the bound, double stranded DNA, the column temperature was raised to 100°C, the HAP was washed six

times with 1 ml of 0.12 M NaP equilibrated to 100°C, and again the eluates were reserved for counting. The radioactivity of each eluate was measured by mixing with 9 ml of Aquasol-2 scintillation fluid (New England Nuclear) and counting in a Beckman scintillation counter. Care was taken throughout to correct for background cpm. Recovery of DNA from the HAP columns was greater than 90%. Control experiments were performed to check the reliability of the HAP method, and these are described in the Appendix.

The fraction of single stranded DNA in each sample was calculated as

$$\frac{(\text{cpm in low temp. wash})}{(\text{cpm in low temp. wash}) + (\text{cpm in high temp. wash})}.$$

The fraction of zero-time corrected, single stranded DNA in each sample was calculated as

$$\frac{(\text{uncorrected fraction ssDNA})}{1.0 - (\text{fraction bound at zero-time})}.$$

The progress of the reassociations was measured in terms of Cot, which was calculated for each sample using the equation of Britten et al. (1974) which is written as

$$\text{Cot} = \frac{(A_{260}/\text{ml}) (\text{hrs})}{2}.$$

Isolation of Repetitive DNA Sequences

C. albicans whole cell DNA was sheared, purified, and dialyzed against 0.12 M NaP, pH 6.8 as described above. The final DNA concentration of this starting material was found to be 0.616 A₂₆₀ units per

ml. Three samples (6 ml each) were sealed in sterile, siliconized, glass, screw capped test tubes ($3.7 A_{260}$ per tube). These tubes were heated to 100°C for 10 minutes and then transferred to a 59°C water bath for reassociation to a Cot of 1.59 (5.17 hours). At that point, the samples were frozen to await HAP fractionation.

To separate reassociated DNA from single stranded DNA, each 6 ml sample was thawed at 5°C and loaded on a 1.5 ml HAP column equilibrated to the reassociation temperature. From each column, the following sets of 1 ml eluates were collected: six from the sample loading, seven from the 0.12 M NaP washes, and seven from 0.48 M NaP washes. Thermal elution at 100°C was not used in order to reduce degradation of the DNA. Portions of each fraction were measured for radioactivity to determine which contained the double stranded DNA, and these (from all three columns) were pooled and dialyzed extensively against 0.12 M NaP.

To purify the repetitive sequences further, a second reassociation was performed. The DNA concentration of the once purified repetitive DNA was calculated (based upon the previously measured specific activity of the ^3H -DNA) and found to be equivalent to $0.18 A_{260}$ units per ml. This DNA, in 10.5 ml of 0.12 M NaP, was again put in a siliconized, sterile, glass, screw capped test tube, dissociated at 100°C , and transferred to the reassociation temperature. After reaching a Cot of 1.08 (12 hrs), the reaction was stopped by quick cooling. The sample was divided into two equal portions, and each was HAP fractionated as before. After pooling fractions containing the twice purified repetitive DNA, dialyzing against 0.12 M NaP, and determining the concentration, the repetitive DNA was then ready for kinetic analysis as described above.

Statistical Analysis of Reassociation Data

To compare the results of reassociation experiments to theoretical, second order kinetic curves, a computer program (Pearson et al., 1977) was used. This program assumes that DNA reassociations measured by the HAP method are described by

$$C/Co = \frac{1}{1 + KCot}$$

where C/Co is the fraction of DNA which is single stranded at any given Cot and K is the reassociation constant (Britten and Kohne, 1968). The program finds, by least squares analysis, the theoretical curve which most closely fits the input data points. Deviations of the best fit from the actual data points are expressed as the root mean square error (RMS error).

Restriction Endonuclease Digestion and Gel Electrophoresis

Restriction endonucleases were purchased from New England Biolabs and Bethesda Research Laboratories. Reaction conditions for the digestions were those of the manufacturer. DNA was electrophoresed in vertical agarose gels as described in Chapter 6 (Part II).

Transfer of DNA to Nitrocellulose

After electrophoresis, DNAs to be hybridized with nick translated probes were transferred from the gels to sheets of nitrocellulose by the blotting procedure described in Chapter 6 (Part II).

Labeling of DNA by Nick Translation

DNAs to be used as probes in blot hybridizations were labeled in vitro using ^{32}P -dCTP and E. coli DNA polymerase (Rigby et al., 1977).

The method for labeling mtDNA differed slightly from that used for labeling purified repetitive DNA.

Mitochondrial DNA probe. When mtDNA was labeled, nicks were introduced using DNase I following the procedure described in Chapter 6 (Part II).

Repetitive DNA probe. Prior to labeling, repetitive DNA (in 0.12 M NaP, pH 6.8) was dialyzed against 10 mM Tris, 1 mM EDTA, pH 7.5. Because randomly sheared DNA contains gaps and nicks following dissociation and reassociation, DNase I was omitted from the reaction mixture (Galau et al., 1976). All other conditions for labeling were the same as used for the mtDNA probes (Chapter 6, Part II).

Blot Hybridizations

The methods for prehybridization, hybridization, washing, and autoradiography of DNAs on nitrocellulose are described in Chapter 6 (Part II).

Cloning of Mitochondria DNA

EcoRI fragments of mtDNA were inserted into plasmid pBR322 and replicated in E. coli strain HB101 as described in Chapter 6 (Part II).

CHAPTER III

EXPERIMENTAL FINDINGS

Base Composition of Candida albicans DNA

The optimum temperature for DNA reassociation is about 25°C below the melting temperature (Bonner et al., 1973). For technical reasons, it was not possible to directly measure the melting temperature of C. albicans DNA. Because the melting temperature of a DNA is directly proportional to its (G + C) content (Marmur and Doty, 1959), it was possible to estimate the melting temperature from measurements of base composition of C. albicans DNA from strain H317 prior to performing reassociation experiments. These measurements were made using two methods.

CsCl buoyant density measurements. The buoyant density of DNA in CsCl increases linearly with its content of guanine plus cytosine (Szybalski and Szybalski, 1971). When samples of C. albicans DNA were centrifuged to equilibrium in CsCl gradients using a Model E, analytical ultracentrifuge, two distinct density components were observed. The u.v. photographs from two of these experiments are presented in Figure 1-2. Photo A in this figure shows the results of an experiment using whole cell DNA from spheroplasts. The majority of the DNA banded at a position corresponding to 32.6% (G + C), but a minor fraction banded at a more dense position of 40.4% (G + C). When an enriched nuclear DNA preparation was used (photo B), essentially the same result was obtained (32.7% and 40.5%). These and additional measurements are listed in Table 1-1. The results would seem to suggest that the satellite DNA is of nuclear origin rather than mitochondrial origin; however, such a

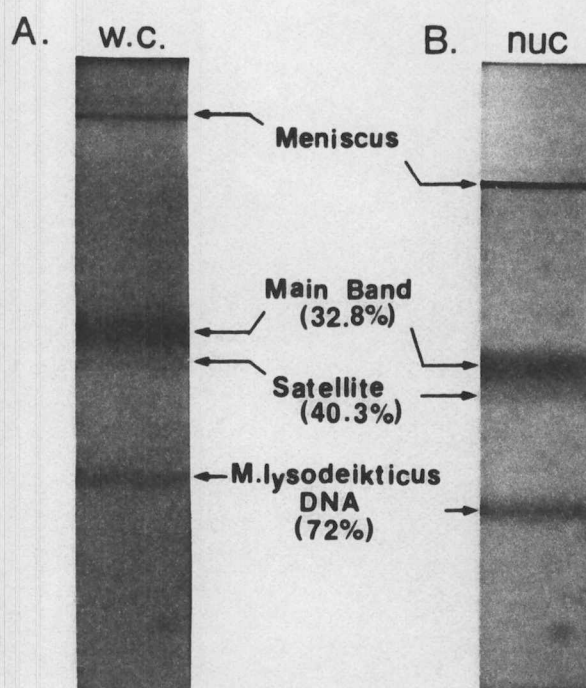


Figure 1-2. Ultraviolet absorption photographs of analytical CsCl gradients containing *C. albicans* DNA. A. Whole cell DNA. B. Enriched nuclear DNA. The numbers in parentheses are the averages of the measurements listed in Table 1-1.

Table 1-1. Analytical CsCl Density Gradient Analysis of Candida albicans DNA.

Source of DNA	% (G + C)	
	Main Band	Satellite
Spheroplasts	33.9	40.6
Spheroplasts	33.6	38.9
Spheroplasts	32.5	40.6
Spheroplasts	32.2	40.3
Spheroplasts	32.2	40.8
Spheroplasts	32.6	40.4
Nuclei ¹	32.7	40.5
Nuclei ¹	32.3	40.1
	Ave = 32.8	Ave = 40.3

¹Isolated by the crude nuclei method.

hypothesis assumes that the nuclear DNA was free from mitochondrial DNA contamination.

Cation-exchange chromatography. Sequential digestion of DNA with DNaseI and venom phosphodiesterase produces 5'-deoxyribonucleotides (Miyazawa and Thomas, 1965) which can then be separated by ion-exchange chromatography to enable the molar proportion of each species to be determined. C. albicans ^{32}P -labeled DNA was analyzed by this method, and the results are listed in Table 1-2. As shown, whole cell DNA and enriched nuclear DNA gave essentially identical values for the molar percentages of their nucleotides. The average percentage of guanine plus cytosine, based upon all four measurements, is 34.4%. It is interesting, however, that the pyrimidine containing nucleotides (T and C) were in excess over those containing purines (A and G). This result suggests that either the method of assay was inaccurate or the C. albicans DNA has unusual character (e.g. modified bases). As a control, an equal molar mixture of commercially available, unlabeled, 5'-nucleotides was prepared and analyzed. The results (bottom line of Table 1-2) show that the cation-exchange chromatography resolved the mixture without altering the nucleotide composition. More importantly, when unlabeled, E. coli DNA (50% G + C) was digested and analyzed (next to last line of the table), a value of 49% (G + C) was obtained. The slight deviations of the mole percentage of each nucleotide from the theoretical value of 25% is not unlike reported values (e.g. Davidson, 1976). These controls confirm that the method of assay was reliable.

Table 1-2. Cation-exchange Separation of Candida albicans ³²P-labeled, 5'-deoxyribonucleotides.

Source of DNA	Molar % of Nucleotide				% (C + C)
	A	T	G	C	
Spheroplasts	30.2	35.4	15.1	19.3	34.4
Spheroplasts	30.0	35.8	15.2	19.0	34.2
Nuclei ¹	30.5	34.8	15.1	19.6	34.7
Nuclei ¹	30.5	35.3	15.2	19.0	34.2
E. coli ²	23.9	27.1	24.4	24.6	49.0
Nucleotide mix ³	25.3	25.0	24.0	25.7	49.7

¹ Isolated by the crude nuclei method.

² Average of two measurements using unlabeled DNA.

³ Equal molar mixture of commercially prepared, unlabeled nucleotides.

Reassociation of Candida albicans Whole Cell DNA

In order to characterize the single copy and repetitive sequences of C. albicans DNA, I performed a reassociation experiment using ^3H -DNA from yeast phase cells. The DNA was sheared to an average length of 290 bp (data not shown), purified on a G-50-Chelex-100 column, dialyzed against 0.12 M NaP, and sealed in capillary tubes. Five of these capillary tubes were used to measure zero-time binding to HAP. For this purpose, the capillaries were heated to dissociate the DNA and then immediately frozen. The estimated Cot of these reactions was 3×10^{-4} or less. Additional capillaries (31) were incubated for various times at the optimum reassociation temperature. HAP chromatography was used to separate the reassociation products, and the fraction of single stranded DNA at each Cot value was calculated. Corrections for zero-time binding were made using the average of the values obtained from the five samples mentioned above (2.8%). The data were submitted to the computer for statistical analysis using the program of Pearson et al. (1977).

Unconstrained statistical analysis. The first step in the statistical analysis was to determine the minimum number of kinetic components needed to describe the data. To do this, the computer was instructed to fit one component and two component curves to the data. No constraints were placed upon the characteristics of the kinetic components. More specifically, the fraction of the DNA contained in each component, the complexity of each component, and the amount of unreactable DNA were allowed to vary as required to improve the fit. The solutions obtained from the computer are presented in Table 1-3.

Table 1-3. Total *C. albicans* DNA Reassociation Experiment¹: Unconstrained Fits for One and Two Components.

Data Set	Component Characteristics			Final Unreacted Fraction	Total DNA ⁴	RMS ⁵ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³			
zero-time ⁶ corrected	.957 (± 0.016)	.052 (± 0.000)	.054	-.018 (± 0.013)	.939	2.79%
not corrected	.128 (± 0.018)	1.491 (± 0.531)	11.690	.016 (± 0.007)	.967	0.75%
	.823 (± 0.017)	.052 ± 0.003	.063			
zero-time ⁶ corrected	.131 (± 0.018)	1.506 (± 0.529)	11.460	.016 (± 0.008)	.994	0.76%
	.847 (± 0.017)	.052 (± 0.003)	.061			

¹Using 31 Cot points.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the component by itself; $M^{-1}-1$

⁴Sum of the fraction of DNA in each component and the unreacted DNA.

⁵Root mean square deviation of the curve from the data points.

⁶The fraction bound at zero-time was 0.028.

The first line in Table 1-3 shows the results obtained from fitting a one component curve to zero-time corrected data. Two unusual observations can be made concerning the one component solution. First, the final, unreacted fraction of DNA is less than zero (-1.8%). [The presence of unreacted DNA can be attributed to three sources: 1) fragments too short to form stable duplexes, 2) duplexes too short to bind to HAP, and 3) additional components which reassociate at higher Cot values.] A negative unreacted fraction is obviously a physical impossibility. Second, the total amount of DNA is only 93.9%. A perfect fit would be expected to account for 100% of the DNA. Had the data not been corrected for zero-time binding, one might attribute the discrepancy to foldback DNA. [It should be noted here that the computer does not specifically describe foldback DNA because foldback DNA does not follow second order kinetics when it reassociates. As far as the computer analysis is concerned, the zero-time binding fraction is simply that portion of the DNA not accounted for by the kinetic components and the unreacted fraction.] Since a zero-time correction was made, it would seem that another rapidly reassociating component must be present in C. albicans DNA.

The solution obtained from fitting a two component curve to the zero-time corrected data is shown in the third line of Table 1-3. Interestingly, the K_{pure} value for the component containing the unique sequences (which includes about 85% of the DNA) is very close to the K_{pure} value from the one component solution (0.061 versus 0.054). This suggests that the computer indeed ignored the repetitive sequences when fitting the one component curve because they represented only a small fraction of the DNA. That the two component solution is better is shown

by three improvements in the fit. First, the amount of unreacted DNA is not less than zero (1.6%). Second, the total amount of DNA is nearly 100% (99.4%). Third, the deviation of the curve from the data points (i.e. the RMS error) is nearly four times smaller than with the one component curve (0.76% versus 2.79%). Fits with curves containing additional components could not be warranted due to the small amount of repetitive DNA (less than 15%) and small number of early points; however, minicot experiments (described later in this chapter) showed that the repetitive sequences contain at least two distinct kinetic components.

The second line in Table 1-3 shows the solution obtained when a two component curve was fitted to the data without prior correction for zero-time binding. These results are presented simply to demonstrate that the correction does not substantially alter the solution. If the measured value for the zero-time binding fraction (0.028) is added to the value for total DNA in the second line (0.967), the sum (0.995) essentially equals the value for total DNA in the third line of the table. Furthermore, the RMS error for each data set is essentially the same (0.75% versus 0.76%), and the K_{pure} values obtained for the two components (repetitive and unique) are essentially the same for each data set (11.69 and 0.063 versus 11.46 and 0.061). Lastly, it is clear that the differences between data sets in the fraction of DNA represented by each component directly reflects the zero-time correction. In other words, the fraction of the total DNA represented in each component is the same for both data sets ($.128/.967 = .131/.994 = .132$ and $.823/.967 = .847/.994 = .851$).

Uniqueness of fit analysis. Although the results from the unconstrained two component fit looked very good, further statistical analysis was required to find out whether other two component curves would equally as well describe the data. In other words, it was necessary to determine whether the unconstrained, two component solution was a unique solution. To understand the approach taken to search for other good solutions, it is essential to understand how the reassociation rate constant (K) relates to the Cot curve. As discussed in the introduction (Chapter 1), $Cot_{1/2}$ for a reassociation is equal to $1/K$. It follows that if K is large, then $Cot_{1/2}$ will be small, indicating that the reassociation proceeds rapidly. Conversely, if K is small, then $Cot_{1/2}$ will be large, indicating that the reassociation proceeds slowly. Now, suppose that the computer arrives at a particular value of K (K_{solution}) after fitting an unconstrained curve to a given set of data points. To look for alternative solutions to the right of the unconstrained Cot curve, one would instruct the computer to fit curves using the same data points but using fixed (i.e. constrained), smaller values of K_{solution} . Constrained values of K larger than K_{solution} would be used to force the computer to look for alternative solutions to the left of the unconstrained Cot curve. The point is that if the data points are not very good (e.g. contain a lot of scatter), then drawing the Cot curve to the right or to the left might not greatly affect the fit (e.g. the RMS error), and no conclusions should be drawn regarding the reliability of the solution parameters. If the unconstrained curve does provide a good description of the data points, then fixing K to either side of K_{solution} should cause the RMS error to increase significantly. Changes in RMS of 50 to 100% with changes in K are usually taken to indicate a

uniquely determined solution (Pearson et al., 1977). In the case of a two component curve, one must also consider the possibility of obtaining a good fit for one of the components but an arbitrary fit for the other.

Table 1-4 shows the results obtained when K_{observed} for the repetitive component (i.e. $K_{\text{repetitive}}$) of the unconstrained two component curve was fixed at values larger than, smaller than, and equal to K_{solution} . The first line of the table simply shows that when K_{rep} was fixed at the unconstrained solution value (1.506), the curve obtained was identical with the unconstrained curve (compare with the third line of Table 1-3). This result was expected, of course, and is only presented for comparison with the other solutions. The next three lines of the table show the results obtained when the computer was instructed to move the repetitive component to the left of the unconstrained repetitive component. K_{rep} was fixed at values two, three, and five times the K_{sol} value (3.012, 4.518, and 7.530, respectively), but all other parameters were free to be changed by the computer. Inspection of the results shows essentially no differences between the three solutions obtained other than the differences in K_{rep} . More specifically, shifting the repetitive component to the left had little effect upon:

1. the RMS error (at most a 20% increase),
2. the fraction of DNA in each component, and
3. the characteristics of the single-copy component.

Thus, it is clear that although a repetitive component is present in the C. albicans DNA, these data reveal little concerning its nature. This conclusion is further supported by attempts to move the repetitive component to the right of the unconstrained repetitive component.

Results obtained from fixing K_{rep} at values one half, one third, and

Table 1-4. C. albicans DNA Uniqueness of Fit Analysis¹: $K_{\text{repetitive}}$ Fixed.

Fixed Value of $K_{\text{repetitive}}$	Component Characteristics			Final Unreacted Fraction	RMS ⁴ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³		
K_{sol} ⁵	.13	1.506	11.46	.016	.76%
	.85	.052	.061		
2K	.11	3.012	26.55	.022	.80
	.87	.056	.065		
3K	.11	4.518	40.99	.025	.85
	.88	.058	.066		
5K	.11	7.530	66.38	.027	.92
	.88	.060	.068		
1/2K	.17	.753	4.42	.008	.82
	.81	.046	.058		
1/3K	.21	.502	2.41	.004	.91
	.77	.043	.055		
1/5K	.29	.301	1.05	.022	1.07
	.69	.037	.053		

¹Using 31 Cot points; corrected for 2.8% zero-time binding.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the purified component; $M^{-1}s^{-1}$.

⁴Root mean square deviation of the curve from the data points.

⁵Obtained from an unconstrained fit with two components.

one fifth the K_{SO1} value (0.753, 0.502, and 0.301, respectively) are presented in the final three lines of Table 1-4. Again, very little change in interpretation occurs as a result of the changes in K_{rep} . The somewhat poorer fits are probably due to interference with the single-copy component since the repetitive component was moved closer to it (i.e. to the right).

While the preceding uniqueness of fit analysis reveals ambiguity in the characteristics of the repetitive component, it also suggests that the single-copy component is quite well defined. The evidence for this resides in the observation that the computer always found essentially the same characteristics for the single-copy component in spite of the constraints imposed upon K_{rep} . To inspect more closely the fit of the single-copy component of the two component curve, attempts were made to find solutions to the data after fixing $K_{single-copy}$ at values larger than, smaller than, and equal to K_{SO1} . The results of this analysis are presented in Table 1-5, where once again the first line simply shows that fixing the unconstrained solution value without change ($K_{S.C.} = 0.052$) does not change the solution obtained (compare with Table 1-3).

The second, third, and fourth lines of Table 1-5 show the solutions obtained when attempts were made to move the single-copy component to the left of the unconstrained single-copy component. Several bizarre changes in the fit can be seen in these three solutions. When $K_{S.C.}$ was fixed at only twice K_{SO1} (.104), the computer responded by calculating a negative repetitive fraction of extremely large magnitude (-2×10^{27}) which, of course, has no physical meaning. Furthermore, this two-fold increase in fixed $K_{S.C.}$ also increased the RMS error by nearly 400%.

Table 1-5. *C. albicans* DNA Uniqueness of Fit Analysis¹: $K_{\text{single-copy}}$ Fixed.

Fixed Value of $K_{\text{single-copy}}$	Component Characteristics			Final Unreacted Fraction	RMS ⁴ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³		
K_{sol} ⁵	.13	1.506	11.46	.016	.76%
	.85	.052	.061		
2K	-2×10^{27}	1×10^{33}	-6×10^5	.108	2.90
	.87	.104	.120		
3K	-1.34	.225	-.168	.100	2.90
	+2.20	.156	.071		
5K	-2.90	.343	-.118	.142	4.36
	+3.70	.260	.070		
1/2K	.40	.235	.592	-.029	1.21
	.61	.026	.043		
1/3K	.55	.154	.279	-.044	1.35
	.46	.017	.037		
1/5K	.67	.121	.182	-.071	1.44
	.37	.010	.028		

¹Using 31 Cot points; corrected for 2.8% zero-time binding.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the purified component; $M^{-1}s^{-1}$.

⁴Root mean square deviation of the curve from the data points.

⁵Obtained from an unconstrained fit with two components.

Attempts to fix $K_{S.C.}$ at values three and five times the K_{SO1} value were also resisted by the computer. In both of these cases, the computer "refused" to allow the single copy component to be moved to the left and maintained K_{pure} for this component near the unconstrained value of 0.061 (0.071 and 0.070, respectively). These increases in fixed $K_{S.C.}$ were instead met with physically meaningless increases in the fraction of single-copy DNA (2.20 and 3.70, respectively). To maintain the total amount of DNA near 1.0, the computer compensated by calculating large but negative increases in the repetitive fraction (-1.34 and -2.90, respectively). The total amount of DNA in these two cases was .96 (i.e., $2.2 + .1 - 1.34$) and .942 (i.e., $3.7 + .142 - 2.9$). The lack of alternative curves in the region to the left of K_{SO1} is further exemplified by the increases in RMS error that result from fixing $K_{S.C.}$ at values three and five times K_{SO1} (about 400% and 600%, respectively).

The final three lines of Table 1-5 show the solutions obtained from attempts to move the single-copy component to the right of the unconstrained single-copy component. At first glance, it would seem that perhaps the data for the single-copy component were not all that good since the RMS error changes only slightly with decreasing fixed $K_{S.C.}$; however, closer inspection reveals otherwise. The effect of decreasing fixed $K_{S.C.}$ is illustrated in Figure 1-3. In this figure, four different solutions for the C. albicans data are presented. The first panel (solid line) shows the Cot curve representing the unconstrained, two component fit to the data and is described in the first line of Table 1-5. The second, third, and fourth panels (solid lines) show the Cot curves representing the fits obtained from fixing $K_{S.C.}$ at values one half, one third, and one fifth (respectively) of the

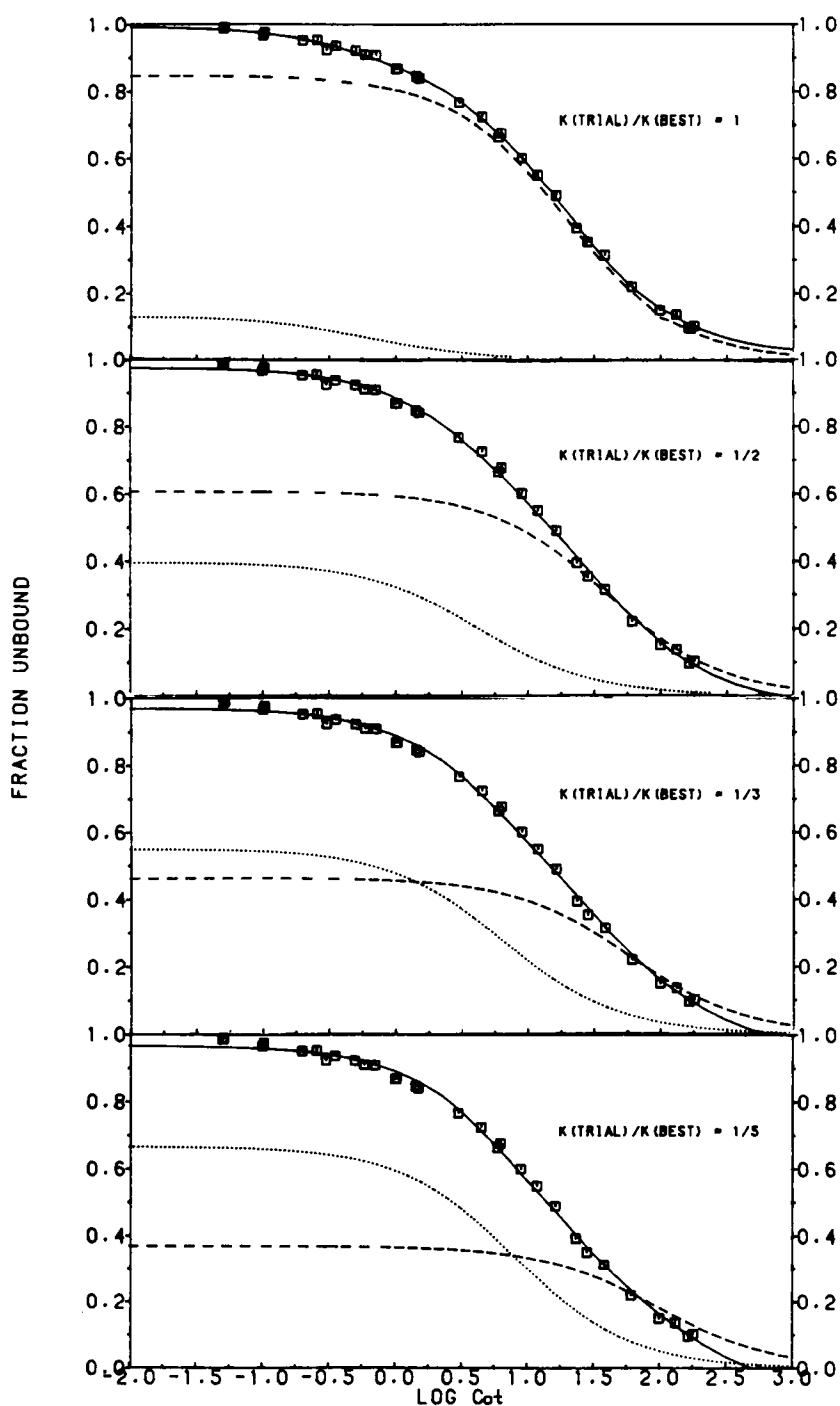


Figure 1-3. Changes in the least squares solution with decreasing values of fixed $K_{S.C.}$. The square symbols represent the zero-time corrected data points. The solid lines represent the best two component fits to the data under the constraint of fixed $K_{S.C.}$. The dashed and dotted lines portray the kinetics of the single-copy and the repetitive components, respectively.

unconstrained values. These are the curves described in the final three lines of Table 1-5. The square symbols in the figure represent the zero-time corrected data points which, of course, are the same in each panel. The dashed lines and the dotted lines portray the kinetics of the single-copy and the repetitive components, respectively. The fact that the solid curves are so similar reflects the fact that the RMS error of each curve is nearly the same. The curves are not identical, however, the most noticeable differences recurring at the x-axis intercepts. As shown in Table 1-5, the final unreacted fraction becomes increasingly negative as fixed $K_{S.C.}$ decreases.

The most striking differences between the Cot curves in Figure 1-3 are illustrated by the component curves. Specifically, as fixed $K_{S.C.}$ decreases, the repetitive fraction increases and the single-copy fraction decreases. This demonstrates that the computer resists moving the single-copy component to the right (and away from the data points). As the single-copy component is pushed to the right, the significance placed upon it by the computer is reduced. As the computer diminishes the single-copy component, it moves the repetitive component rightward to replace it. It must be concluded that the data points for the single-copy component are indeed quite good (i.e. contain little scatter). Although it was not tested, one can reasonably guess that with further decreases in fixed $K_{S.C.}$, the fraction of the single-copy component would continue to decrease (and eventually become negative) and the repetitive component would resist movement to the right of the data points (and its fraction would eventually exceed 1.0).

As a final demonstration of the quality of the data points and the unconstrained fit for the single-copy component, a third uniqueness

of fit analysis was performed using only those data points with Cot values greater than 1.5. [If it is assumed that the unconstrained, two component solution is correct, one can calculate that at Cot equals 1.5, 70% of the repetitive sequences and 7% of the single-copy sequences would have reassociated.] An unconstrained, one component fit was obtained, and attempts were then made to find solutions using values of fixed K_{obs} that were larger than, smaller than, and equal to K_{sol} . The results are presented in Table 1-6. The first line of this table shows the result of fixing K_{obs} at the unconstrained value (0.057). It is interesting to compare this solution with the one component fit obtained using all the data points (Table 1-3, page 42). By leaving out the early data points, the one component fit is greatly improved (i.e. the RMS error drops from 2.79% to 0.83%). This result confirms again the existence of a repetitive component. The remaining lines of the table show that increases and decreases in the fixed value of K_{obs} result in dramatically worse fits. A twofold increase or decrease of fixed K_{obs} increases the RMS error by more than 400%. Threefold changes in fixed K_{obs} increase the RMS error by more than 600%. Five-fold changes in fixed K_{obs} increase the RMS error by more than 800%. These changes in RMS error are plotted in Figure 1-4 (open triangles) where fixed K_{obs} is expressed as fixed $Cot_{1/2}$ (i.e. $1/K_{obs}$).

Reassociation of Escherichia coli DNA

Measurements of DNA reassociation kinetics provide unreliable and only approximate values for sequence complexity unless careful, parallel reassociation experiments are performed using a reference DNA of known size. For this purpose, I measured the kinetics of

Table 1-6. *C. albicans* DNA Uniqueness of Fit Analysis¹: One Component Solutions Using Only Later Points.

Fixed Value of K_{obs}	Component Characteristics			Unreacted Fraction	RMS ⁴ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³		
K_{sol} ⁵	.88	.057	.064	.021	0.83%
2K	.94	.113	.120	.101	3.72
3K	1.02	.170	.168	.140	5.74
5K	1.16	.283	.244	.182	8.12
1/2K	.89	.028	.032	-.087	3.59
1/3K	.93	.019	.020	-.173	5.40
1/5K	1.04	.011	.011	-.325	7.38

¹Using 16 data points (Cot greater than 1.5); corrected for 2.8% zero-time binding.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the purified component; $M^{-1}s^{-1}$.

⁴Root mean square deviation of the curve from the data points.

⁵Obtained from an unconstrained fit with one component.

UNIQUENESS OF LEAST SQUARES SOLUTION
FOR REASSOCIATION DATA
HYDROXYAPATITE ASSAY

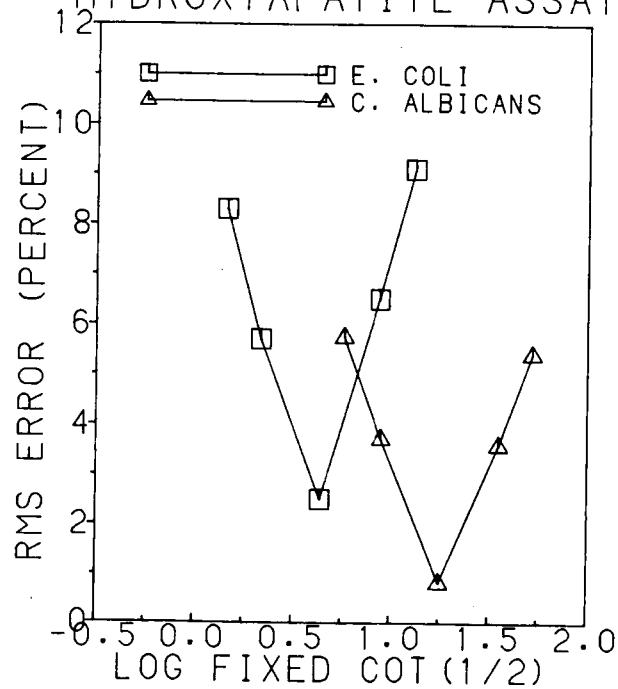


Figure 1-4. Changes in RMS error with increases and decreases in fixed K_{obs} . The values of fixed K_{obs} are expressed as $\log (1/K_{obs})$ (i.e. $\log Cot_{1/2}$). Open triangles represent C. albicans DNA. Open squares represent E. coli DNA.

reassociation of E. coli DNA using the HAP method of assay. All conditions were the same as for the C. albicans DNA experiment except the temperature of reassociation (because the melting temperatures of the two DNAs are different) which was set at the optimum (i.e. $T_m - 25^{\circ}\text{C}$) for E. coli DNA. More specifically, the DNA was sheared to an average length of 285 bp (data not shown), purified, dialyzed to 0.12 M NaP, and sealed in capillary tubes. Two zero-time binding measurements were made (estimated Cot of 2×10^{-4} or less), and 20 higher Cot points were obtained. The fraction of single stranded DNA for each Cot point was calculated, and corrections for zero-time binding were made using the averaged value (3.7%) from the two measurements. Statistical analysis of the data was accomplished using the computer program of Pearson et al. (1977).

Unconstrained statistical analysis. For measurements of sequence complexity, E. coli DNA can be assumed to contain no repeated sequences (Lewin, 1980). For this reason, only one component Cot curves were fitted to the data points. The solutions obtained when the computer was allowed to vary all reaction parameters are presented in Table 1-7. The first line in this table shows that when the data points were not corrected for zero-time binding, the total amount of DNA was only 97%. This suggests that 3% of the DNA might have been expected to bind at zero-time. The measured value for zero-time binding (3.7%), based upon only two samples, may therefore be somewhat overestimated. Consistent with this hypothesis, the total amount of DNA calculated when the corrected data points were used was greater than 100%, as shown in the second line of the table. Furthermore, the RMS error was found to be

Table 1-7. *E. coli* DNA Reassociation Experiment¹: Unconstrained Fits to One Component.

Data Set	Component Characteristics			Final Unreacted Fraction	Total ⁴ DNA	RMS ⁵ Error
	Fraction of DNA	² k_{obs}	³ k_{pure}			
not corrected	.946 ($\pm .017$)	.229 ($\pm .019$)	.242	.024 ($\pm .015$)	.970	2.43%
zero-time ⁶ corrected	.983 ($\pm .018$)	.229 ($\pm .019$)	.233	.025 ($\pm .016$)	1.008	2.52%

¹Using 20 Cot points.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the component by itself; $M^{-1}s^{-1}$.

⁴Sum of the fraction of DNA in each component and the unreacted DNA.

⁵Root mean square deviation of the curve from the data points.

⁶The fraction bound at zero-time was .037.

slightly greater when zero-time corrected data were used. These differences are only of minor importance, however, because the values obtained for K_{pure} are nearly identical. The unconstrained solution obtained using zero-time corrected data is illustrated in Figure 1-5 (square symbols). For comparison, the unconstrained, two component C. albicans DNA Cot curve is also illustrated (cross symbols). The E. coli data points clearly have more scatter about the curve than do the C. albicans DNA data points. This accounts for the larger RMS error value for the E. coli DNA curve: 2.52% versus 0.76% for the C. albicans DNA curve.

Uniqueness of fit analysis. The reliability of the E. coli DNA data points was tested by attempts to move the curve to the right and to the left of the unconstrained curve as discussed above for the C. albicans DNA Cot curves. The results of this analysis are listed in Table 1-8. The first line of this table simply shows that when K_{obs} was fixed at the unconstrained value ($K_{\text{sol}} = 0.229$), the solution obtained was not changed (compare with the second line of Table 1-7). When K_{obs} was fixed at values twice as large and half as large, the RMS error increased by more than 200%. Changing fixed K_{obs} by a factor of three increased the RMS error by more than 300%. These results are plotted in Figure 1-4 (page 55; open squares). From this analysis, it is clear that the E. coli DNA points are best described by using K_{sol} .

Reassociation of Purified Repetitive DNA

Reassociation of Candida albicans whole cell DNA revealed that only 13% of the DNA was repetitive. In order to study such a small portion of the cell genome, it was necessary to isolate the repetitive sequences away from the single-copy sequences. To do this, a second

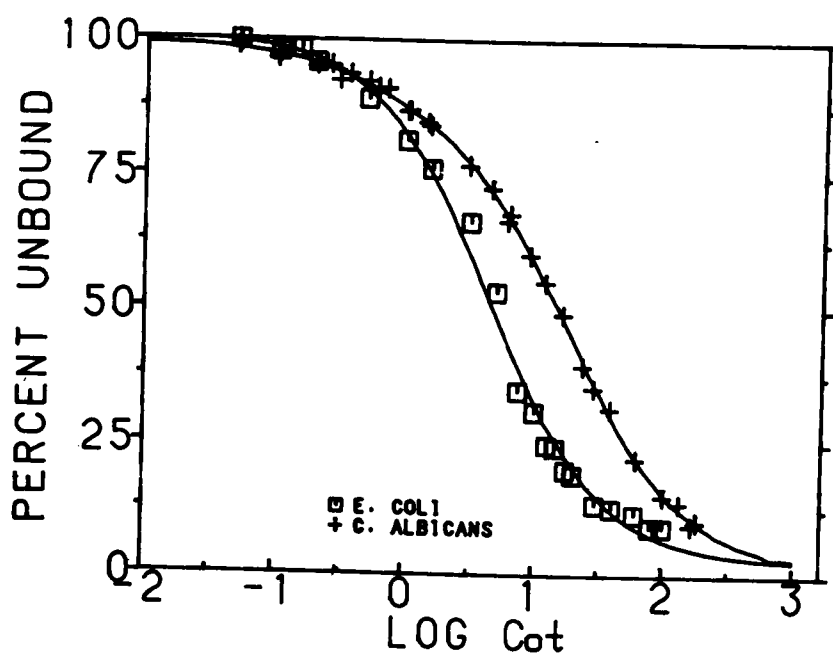


Figure 1-5. Kinetics of reassociation of DNA of *Escherichia coli* and *Candida albicans*, assayed by hydroxyapatite chromatography. The solid lines represent the best fits to the data, as discussed in the text.

Table 1-8. *E. coli* DNA Reassociation Experiment¹: Uniqueness of Fit Analysis-- K_{obs} Fixed.

Fixed Value of K_{obs} ²	Component Characteristics		Final Unreacted Fraction	Total ⁴ DNA	RMS ⁵ Error
	Fraction of DNA	K_{obs} ³			
K_{sol}	.983	.229	.025	1.008	2.52%
2K	.961	.458	.113	1.074	5.73
3K	.961	.687	.155	1.116	8.32
1/2K	1.032	.115	-.093	.939	6.05
1/3K	1.077	.076	-.179	.898	9.07

¹Using 20 Cot points; corrected for 3.7% zero-time binding.

²Solution for K_{obs} was obtained from an unconstrained fit.

³Observed rate constant; $M^{-1}s^{-1}$.

⁴Sum of the fraction of DNA in each component and the unreacted DNA.

⁵Root mean square deviation of the curve from the data points.

preparation of ^3H -labeled whole cell DNA was purified and sheared to an average length of 300 base pairs. Before isolating the repetitive sequences, samples of this DNA were used to perform a preliminary, test reassociation experiment to make sure that it reassociated with the same kinetics as the first DNA preparation, and this was found to be the case (data not shown). The DNA was then denatured and incubated to a Cot of 1.59 to permit the repetitive sequences to reassociate while leaving the single copy sequences largely single-stranded. After HAP fractionation, 19.4% of the DNA was recovered as dsDNA.

Using the results of the whole cell DNA reassociation experiment (Table 1-3, page 42), one can calculate the expected composition of the enriched repetitive DNA obtained. To do this, one first calculates the fraction of each kinetic component which would be expected to remain single-stranded at a Cot of 1.59, based upon the K_{obs} values and the second order equation. These calculations follow.

$$C/C_0 = \frac{1}{1 + K_{obs} Cot} = \text{fraction single stranded}$$

$$C/C_0 \text{ (foldback)} = 0.00 \text{ (or 100\% dsDNA)}$$

$$C/C_0 \text{ (repetitive)} = \frac{1}{1 + (1.491) (1.59)} = 0.297 \text{ (or 70.3\% dsDNA)}$$

$$C/C_0 \text{ (single-copy)} = \frac{1}{1 + (0.052) (1.59)} = 0.923 \text{ (or 7.7\% dsDNA)}$$

By calculating the fraction of the total genome that should be represented by each component in the form of double stranded DNA, one can then calculate the expected yield (i.e. the sum) of enriched repetitive DNA from the HAP chromatography.

$$\text{foldback: } (100\%) (0.028) = 2.8\%$$

$$\text{repetitive: } (70.3\%) (0.128) = 9.0\%$$

$$\text{single copy: } (7.7\%) (0.823) = \underline{6.3\%}$$

$$\text{expected yield} = 18.1\%$$

This value is very close to the measured value of 19.4%. The measured value may be higher than the theoretical value because of differences between analytical and preparative procedures. When performing analytical reassociation experiments, one uses very small volumes so that rapid changes in temperature can be achieved; moreover, the samples are diluted after the reaction to prevent further reassociation during HAP fractionation. During the preparation of the enriched repetitive DNA, however, relatively large amounts of DNA in large volumes of buffers were reassociated without dilution prior to the fractionation. Thus, it is likely that the actual Cot value reached was greater than 1.59. Assuming the above theoretical calculations to be correct, one can finally estimate the composition of the enriched repetitive DNA.

$$(2.8/18.1) (100) = 15.5\% \text{ foldback}$$

$$(9.0/18.1) (100) = 49.7\% \text{ repetitive}$$

$$(6.3/18.1) (100) = 34.8\% \text{ single copy}$$

In order to purify the repetitive DNA further, it was again dissociated and reassociated, this time to a Cot of 1.08. After fractionation on HAP, 60.3% of the DNA was recovered as double stranded DNA. To calculate the expected composition of this twice purified repetitive DNA, one must use the K_{pure} values estimated from the whole cell DNA reassociation experiment (Table 1-3, page 42) and the enriched

repetitive DNA composition information (calculated above) to estimate K_{obs} for each component. The values for K_{obs} are calculated below.

$$K_{obs} = (\text{fraction of DNA}) (K_{pure})$$

$$K_{obs} \text{ (repetitive)} = (.497) (11.69) = 5.810$$

$$K_{obs} \text{ (single copy)} = (.348) (0.063) = 0.022$$

Using these K_{obs} values, the fraction of each kinetic component expected to remain single-stranded, after incubation to a Cot of 1.08, can be calculated as follows.

$$C/C_0 \text{ (foldback)} = 0.00 \text{ (100\% dsDNA)}$$

$$C/C_0 \text{ (repetitive)} = \frac{1}{1 + (5.81) (1.08)} = 0.137 \text{ (86.3\% dsDNA)}$$

$$C/C_0 \text{ (single-copy)} = \frac{1}{1 + (0.022) (1.08)} = 0.977 \text{ (2.3\% dsDNA)}$$

The expected yield of double stranded DNA from the second purification procedure can now be calculated as shown below.

$$\text{foldback: } (100\%) (.155) = 15.5\%$$

$$\text{repetitive: } (86.3\%) (.497) = 42.9\%$$

$$\text{single-copy: } (2.3\%) (.348) = \underline{0.8\%}$$

$$\text{expected yield} = 59.2\%$$

This value is very close to the measured value of 60.3%. If these theoretical calculations are assumed to be correct, the composition of the twice purified repetitive DNA can be estimated as follows.

$$(15.5/59.2) (100) = 26.2\% \text{ foldback}$$

$$(42.9/59.2) (100) = 72.4\% \text{ repetitive}$$

$$(0.8/59.2) (100) = 1.4\% \text{ single-copy}$$

Having purified the majority of the repetitive sequences away from the majority of the single-copy sequences, samples of the twice purified repetitive DNA were dialyzed to 0.12M NaP, sealed in capillary tubes, denatured, and incubated to various Cot values. Two zero-time binding measurements were made, and the average of these was 20.0%, which is 6.2% lower than the expected value of 26.2% (calculated above). This difference may be due to breakage of strands containing foldback sequences, thus eliminating the intramolecular reassociation reaction (Perlman et al., 1976; Schmid et al., 1975; Cech and Hearst, 1975). Thirty-nine measurements were made at higher Cot values. These measurements were corrected for zero-time binding and analyzed using the computer program of Pearson et al. (1977).

Unconstrained statistical analysis. A reassociation kinetic curve obtained using an isolated fraction of the total DNA by itself is called a minicot curve (Britten et al., 1974). The first step in the statistical analysis of the C. albicans repetitive DNA reassociation experiment was to find the minicot curve with the fewest number of components required to describe the data points. To do this, the computer was instructed to fit one, two, or three component, unconstrained curves to the data. The results of this analysis are presented in Table 1-9.

Inspection of the three solutions listed in Table 1-9 reveals that the two component minicot solution (shown in the second line of

Table 1-9. Minicot Experiment Using Repetitive DNA¹: Unconstrained Fits to Different Numbers of Components.

No. of Components Specified	Component Characteristics			Final Unreacted Fraction	Total ⁴ DNA	RMS ⁵ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³			
1	.796 (± 0.012)	12.35 (± 0.77)	15.51	.163 (± 0.010)	.959	2.29%
2	.132 (± 0.125)	98.01 (± 171.0)	742.40	.148	.987	2.14
	.707 (± 0.130)	9.25 (± 2.68)	13.08	(± 0.016)		
3	.059 (± 0.276)	1221 ($\pm 24,600$)	20,540			
	.199 (± 0.933)	35.79 (± 232.0)	179.80	.145 (± 0.025)	1.015	2.14
	.612 (± 1.110)	8.14 (± 10.7)	13.29			

¹Using 39 Cot points; corrected for 20.0% zero-time binding.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the component by itself; $M^{-1}s^{-1}$.

⁴Sum of the fraction of DNA in each component and the unreacted DNA.

⁵Root mean square deviation of the curve from the data points.

the table) provides the best fit to the data. With this solution, the RMS error has its lowest value (2.14%), and the total DNA (0.987) is closest to the theoretical value of 1.0. Confidence in this solution is further provided by common features with the other two solutions. Specifically, all three solutions describe a slower component which comprises the majority of the repetitive DNA (.707 versus .796 and .612). Furthermore, the K_{pure} values for each of these slower components are very similar with each other (13.08 versus 15.51 and 13.29) and with the K_{pure} value obtained for the repetitive DNA from the whole cell DNA experiment (11.69; see Table 1-3, page 42). It can also be seen that when three components (rather than two) were specified, the computer responded by making adjustments primarily in the faster reassociating component (whose nature is uncertain) but resisted change in the slower, main component. This result suggests that the slower component has been well described by the data. Because the three component solution offers no better description of the data, its adoption is not justified (Pearson et al., 1977). The two component minicot curve is plotted in Figure 1-6.

Uniqueness of fit analysis. To confirm that the unconstrained minicot curve provides the best description of the slower repetitive component, the computer was instructed to fit one component curves to those data points having Cot values greater than 0.01 using fixed values of K_{obs} . (Using the component characteristics from the unconstrained two component solution, at Cot equals 0.01, 50% of the fast component and 8% of the slow component would have reassociated.) The results obtained from this analysis are presented in Table 1-10. Omission of the early data points improved the one component fit (compare the first

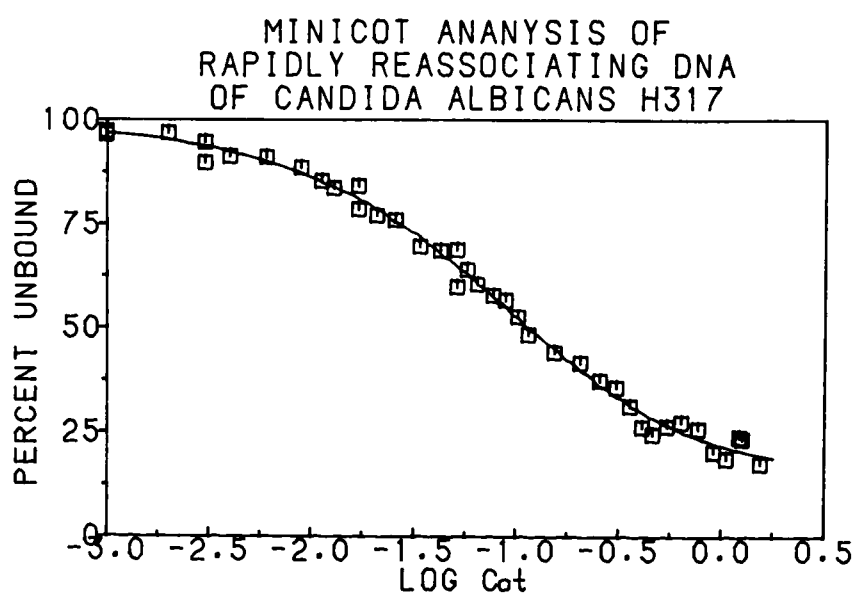


Figure 1-6. Minicot analysis of rapidly reassociating DNA of Candida albicans. The solid line represents the best, two component, fit to the data.

Table 1-10. Minicot Experiment Using Repetitive DNA¹: Uniqueness of Slow Component Fit Using Only Late Data Points.

Fixed Value of K_{obs}	Component Characteristics			Final Unreacted Fraction	RMS ⁴ Error
	Fraction of DNA	K_{obs} ²	K_{pure} ³		
K_{sol} ⁵	.773	10.74	13.90	.154	2.25%
2K	.854	21.48	25.14	.213	3.82
3K	.942	32.22	34.21	.241	5.28
5K	1.116	53.70	48.13	.270	7.01
1/2K	.757	5.37	7.10	.074	3.80
1/3K	.777	3.58	4.61	.012	5.23
1/5K	.841	2.15	2.55	-.094	7.01

¹Using 31 data points (Cot greater than 0.01); corrected for 20.0% zero-time binding.

²Observed rate constant; $M^{-1}s^{-1}$.

³Rate constant for the purified component; $M^{-1}s^{-1}$.

⁴Root mean square deviation of the curve from the data points.

⁵Obtained from an unconstrained, one component fit.

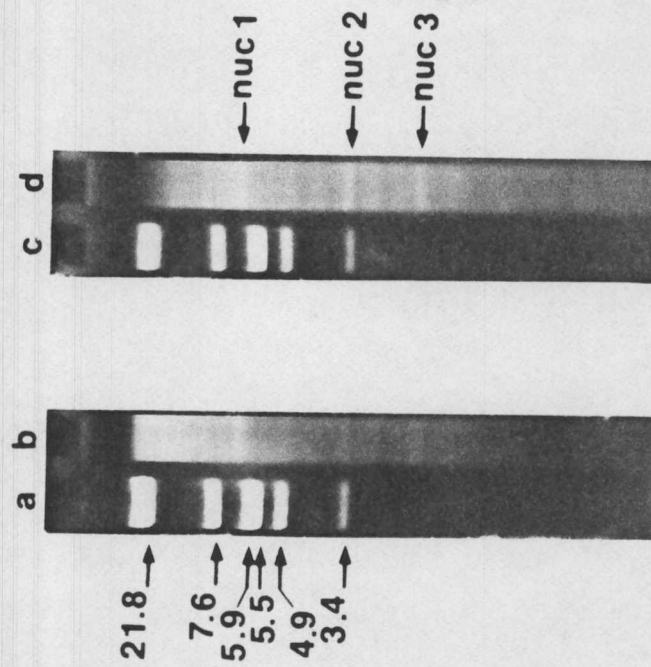
lines in Table 1-9 and Table 1-10). This result supports the contention that more than one component is contained in the repetitive DNA. Fixing K_{obs} at values that differed from K_{sol} by a factor of two, three, or five resulted in poorer solutions as judged by the increases in RMS error. For example, increasing and decreasing fixed K_{obs} by a factor of two increased the RMS error 70%.

Genomic Blot Hybridizations Using Repetitive DNA Probe

The repetitive DNA sequences of many higher eukaryotes are interspersed among the single-copy sequences, but in lower eukaryotes they appear to be clustered (Lewin, 1980). *C. albicans* whole cell DNA was shown above to contain about 13% repetitive DNA. In order to examine the arrangement of these repetitive sequences, whole cell DNA was EcoRI digested, agarose gel electrophoresed, and transferred to a sheet of nitrocellulose by the blotting procedure. Those EcoRI fragments containing sequences homologous to repetitive DNA sequence were detected by hybridization using twice purified, repetitive DNA which had been ^{32}P -labeled by in vitro nick translation (see Chapter 2). In order to identify repetitive DNA fragments derived from mitochondrial DNA, the experiment was also performed using EcoRI digested nuclear DNA (isolated from nuclei purified by the differential centrifugation method). The results of these experiments are shown in Figure 1-7.

In Figure 1-7a, the electrophoretic patterns of the EcoRI digested nuclear and whole cell DNAs are presented (lanes b and d, respectively). In both cases, several distinct bands can be seen. Three of these bands are considerably more intense than the others and are labeled nuc1-nuc3. These bands indicate the presence of non-equal molar concentrations of fragments due either to 1) repetitive sequences

A.



B.

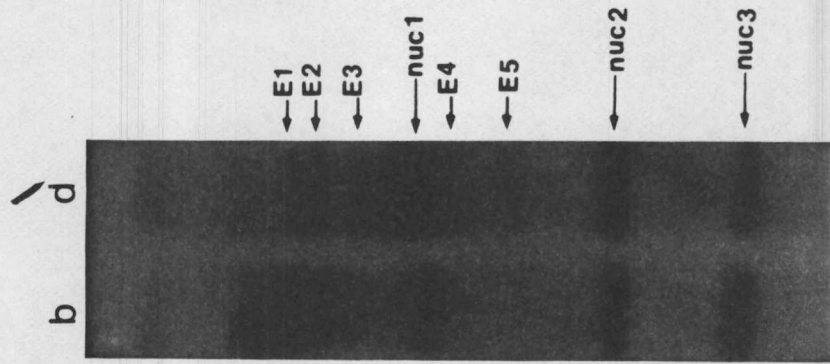


Figure 1-7. Genomic blot hybridizations using repetitive DNA probe. Nuclear (lane b) and whole cell (lane d) *C. albicans* DNAs were digested with *EcoRI*, electrophoresed, transferred to nitrocellulose, and individually hybridized. A. Electrophoretic patterns in agarose gel detected by U.V. fluorescence. B. Hybridization patterns detected by autoradiography. Markers are *EcoRI* digested Lambda DNA (lanes a and c).

which contain EcoRI sites, or 2) several single-copy sequences with similar distances between EcoRI sites. In the nuclear DNA digest, band nuc3 appears to be less distinct than in the whole cell DNA digest, but this is probably artifactual since the band was found in other samples from the same EcoRI digest (data not shown). In fact, bands nucl-nuc3 have been found with all the preparations of C. albicans nuclear and whole cell DNA that I have examined. (Additional examples are shown in Figure 1-9a, page 75.) This result is consistent with a hypothesis of nuclear (rather than mitochondrial) origin for the DNA in these three bands.

Figure 1-7b shows the autoradiography results obtained from the blot hybridization. The radioactive bands indicate the presence of EcoRI fragments containing repetitive DNA sequences. In the whole cell DNA digest (lane d), eight bands were easily observed, three of which were markedly more intense than the other five. These three prominent bands were found to correspond to bands nucl-nuc3 discussed above. The sizes of the fragments in each of the eight repetitive DNA containing bands are listed in Table 1-11. Using the nuclear DNA digest (lane b), three of the eight bands were found to have very low intensity. These bands (labeled E3, E4, and E5) presumably correspond to mitochondrial DNA EcoRI fragments which were purified away to some extent during nuclei isolation. Confirmation of this hypothesis was obtained by analyzing purified mitochondrial DNA. When purified mitochondrial DNA was EcoRI digested and electrophoresed, five bands were observed which correspond to repetitive bands E1-E5 (see Figure 1-9a, page 75 and Part II of this dissertation). The presence of the two largest mitochondrial DNA bands (E1 and E2) in the nuclear DNA digest suggests that

Table 1-11. Repetitive EcoRI Fragments Detected by Autoradiography.

Band	Fragment Size	Origin of DNA
E1	12.7 Kbp	mitochondrial DNA
E2	10.0	mitochondrial DNA
E3	8.6	mitochondrial DNA
nuc1	6.6	nuclear DNA
E4	5.9	mitochondrial DNA
E5	4.8	mitochondrial DNA
nuc2	3.3	nuclear DNA
nuc3	2.6	nuclear DNA

the nuclear DNA was contaminated with mitochondrial DNA. The total length of the mitochondrial DNA fragments is 42 Kbp which is very close to the actual length of C. albicans mitochondrial DNA (41 Kbp; see Part II). The total length of the nuclear DNA fragments is 12.5 Kbp. The sum of the lengths of all eight fragment sizes is 54.5 Kbp.

Presence of mtDNA in Nuclear DNA Preparations

The results from the genomic blot hybridizations using repetitive DNA probes suggested that the nuclear DNA was contaminated by mitochondrial DNA (mtDNA). To examine this problem in greater detail, nuclear DNAs were isolated by several methods from a common batch of C. albicans cells. Figure 1-8 shows the experimental design by which these nuclear DNAs were prepared. The cells were grown in YEPL medium (which contains lactic acid to encourage the presence of mitochondria), and spheroplasts were prepared. A portion of the spheroplasts was used for the isolation of whole cell DNA, while the remaining spheroplasts were used for the preparation of crude nuclei. A portion of the crude nuclei was then used for the isolation of DNA while the remaining nuclei were further purified by one of two methods: discontinuous sucrose gradient purification or differential centrifugation through a dense medium. Finally, DNA was also isolated from these latter two preparations of nuclei.

In order to assay for the presence of mtDNA sequences, the nuclear DNAs were digested with EcoRI or PvuII, electrophoresed, transferred to nitrocellulose, and hybridized with nick translated, ^{32}P -mtDNA. The results obtained using EcoRI digested DNAs are presented in Figure 1-9. The electrophoretic patterns of the digests (part A of

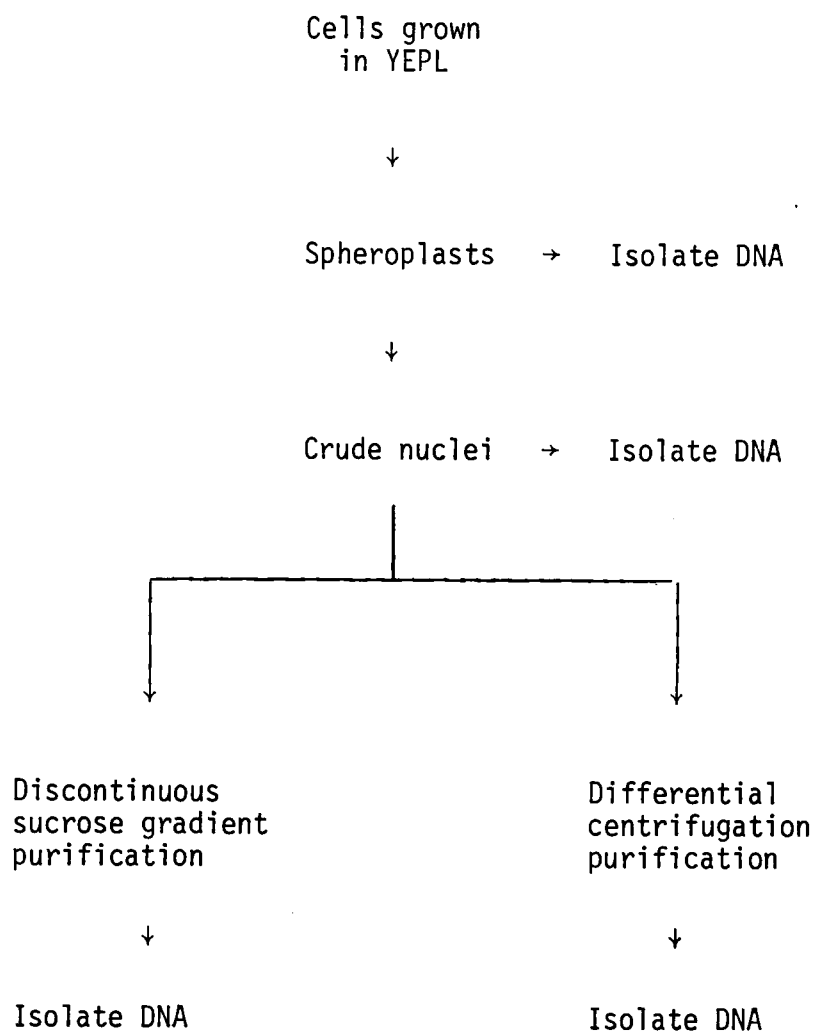
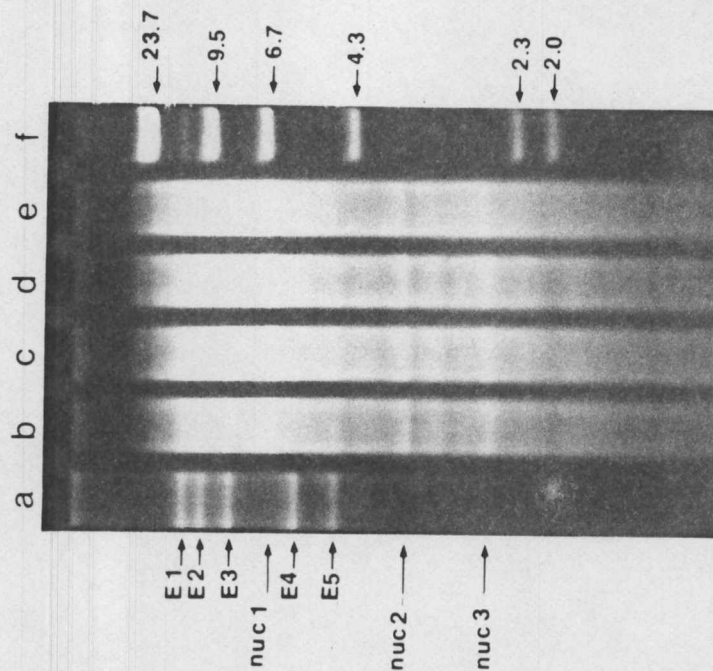


Figure 1-8. Preparation of nuclear DNAs from *C. albicans* cells. The procedures used at each step are described in Chapter 2.

A.



B.

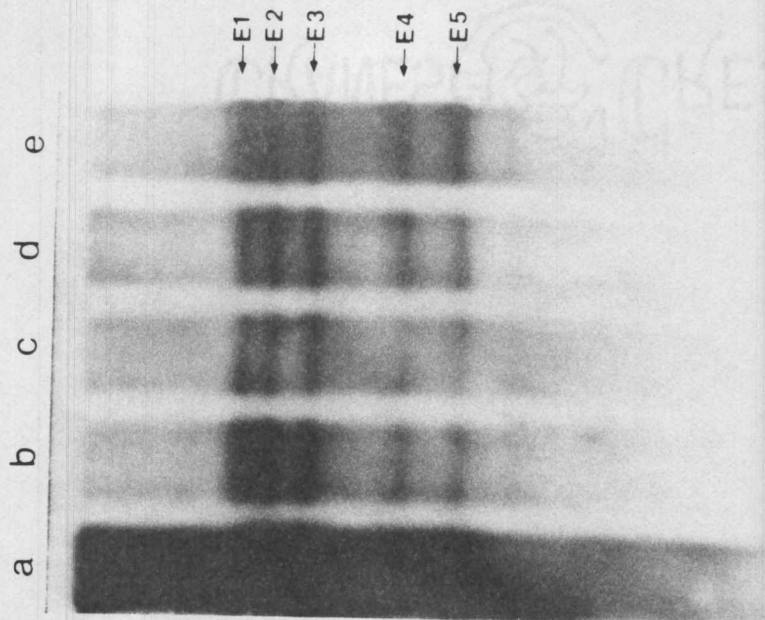


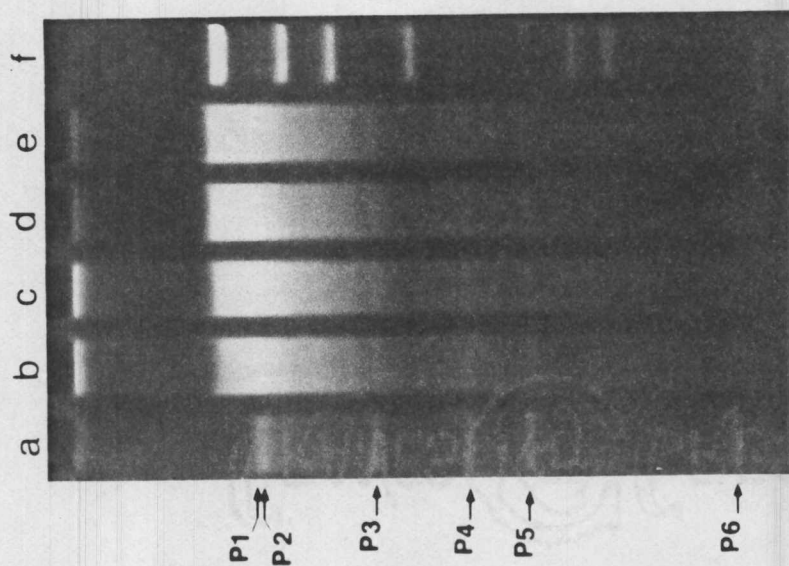
Figure 1-9. Detection of mtDNA in EcoRI digests of nuclear DNAs. A. Electrophoretic patterns in agarose gel detected by U.V. fluorescence. B. Hybridization pattern obtained using 32p-mtDNA and detected by autoradiography. Markers (lane f) are Hind III digested Lambda DNA. Lane contents: mtDNA (a); whole cell DNA (b); crude nuclear DNA (c); differential centrifugation purified nuclear DNA (d); gradient purified nuclear DNA (e).

the figure) are nearly identical with each other and with the previously described EcoRI digests of C. albicans DNA (Figure 1-7). The autoradiography results obtained from the hybridization (part B of Figure 1-9) show that the nuclear DNA digests (lanes b-c) each contained EcoRI fragments with electrophoretic mobilities identical to the five mtDNA fragments (lane a). The results obtained using PvuII digested DNAs are presented in Figure 1-10. This restriction endonuclease was found to cut the DNAs fewer times than EcoRI, and as a result, most of the fragments were large and remained near the top of the gel (lanes b-e, Figure 1-10a). Purified mtDNA was cut seven times (see Part II) and formed six well separated bands after electrophoresis (lane a). These six bands could also be seen in the nuclear DNA digests again suggesting that contaminating mtDNA was present. The autoradiography results (Figure 1-10b) confirmed the presence of mtDNA sequences in all of these bands.

Evidence that the sequences homologous to the mtDNA probe were not integrated in nuclear DNA was provided by two observations: First, the electrophoretic patterns of the radioactive bands in both the PvuII and the EcoRI digests of nuclear DNA were identical to the corresponding patterns of the purified mtDNA. That is, no other radioactive bands were seen to suggest the presence of flanking, nuclear DNA sequences. Second, the sum of the fragment lengths obtained from EcoRI digestion of mtDNA was the same as that obtained from PvuII digestion (41 kbp; see Part II).

Evidence that the hybridization results were not due to the presence of nuclear DNA contamination of the ^{32}P -mtDNA probe was also obtained. MtDNA fragments E2 and E4 were cloned (to purify away nuclear

A.



B.

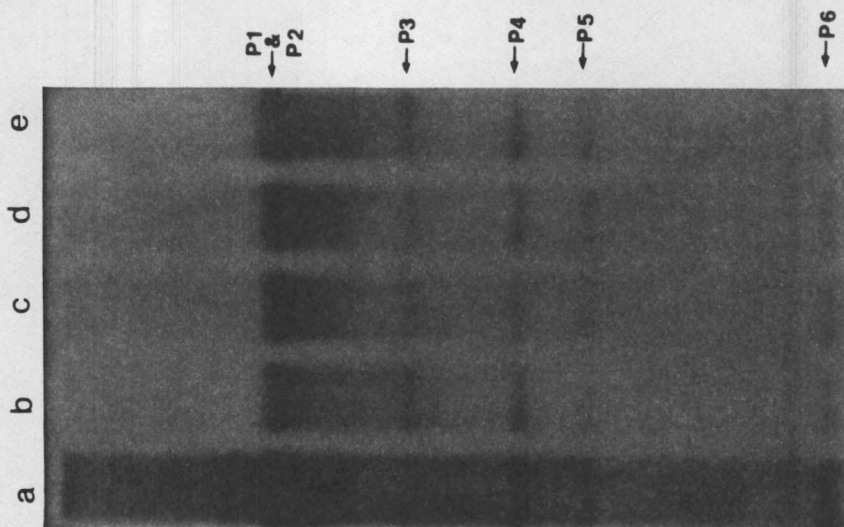
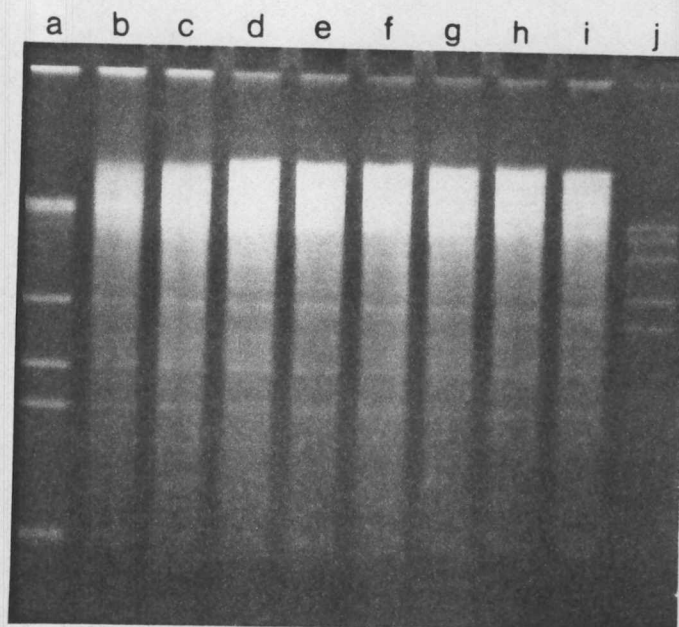


Figure 1-10. Detection of mtDNA in PruII digests of nuclear DNAs. The DNAs contained in each lane are the same as in Figure 1-9 (page 75).

DNA sequences), nick translated, and used as radioactive probes for mtDNA in PvuII digests of the nuclear DNAs. The results of the experiment using the cloned E2 DNA probe are shown in Figure 1-11. Lanes a and j contained PvuII and EcoRI digests of mtDNA, respectively. Lanes b-e contained the four C. albicans DNAs mentioned above. Hybridization with the PvuII digested nuclear DNAs resulted in the same electrophoretic pattern as obtained with the PvuII digested mtDNA. (Lane j is different because it contained EcoRI rather than PvuII digested mtDNA. As discussed in Part II, fragments E2 and E3 contain homologous sequences which overlap PvuII fragments P1, P2, P4, and P6.) When the experiment was performed using cloned E4 DNA as the probe, the electrophoretic patterns for the nuclear DNA and the mtDNA were again the same (data not shown).

The results above indicated that mtDNA was present in all the nuclear DNA preparations. Attempts to use Triton X-100 to selectively eliminate mitochondria also failed. The results are shown in Figure 1-11. Lanes f, g, h, and i contained PvuII digested nuclear DNAs isolated using 0%, 0.005%, 0.02%, and 0.1% Triton X-100, respectively. In the autoradiogram, mtDNA sequences were again detected. The fact that the radioactive signals obtained were somewhat less intense than with the other nuclear DNAs (lanes b-e) was probably due to the fact that a different batch of cells was used in the Triton experiment, and they probably contained a smaller amount of mtDNA.

A.



B.

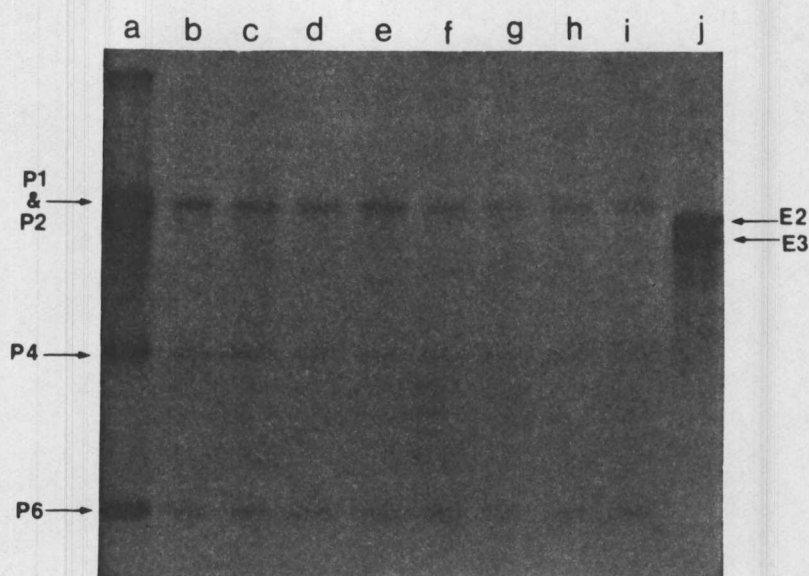


Figure 1-11. Hybridization of ^{32}P -labeled, cloned fragment E4 of mtDNA to PvuII digested nuclear DNAs. A. Electrophoretic patterns. B. Hybridization patterns. Lane contents: PvuII digested mtDNA (a); whole cell DNA (b); crude nuclear DNA (c); differential centrifugation purified nuclear DNA (d); gradient purified nuclear DNA (e); 0% Triton (f); 0.005% Triton (g); 0.02% Triton (h); 0.1% Triton (i); and EcoRI digested mtDNA (j).

CHAPTER IV

DISCUSSION OF RESULTS

Base Composition of *Candida albicans* DNA

The genus Candida is a very heterogeneous group of fungi. With respect to base composition, Candida DNAs have been found to have G+C contents ranging from 34.9% to 57.6% (for review, see Odds, 1979). C. albicans DNA has been reported to contain 35.1% G+C, based upon measurements of melting temperature (Stenderup and Leth Bak, 1968). I have measured the base composition of C. albicans strain H317 DNA by two additional methods: 1) equilibrium density gradient analysis in CsCl and 2) cation-exchange separation of 5'-mononucleotides.

CsCl buoyant density measurements. The relationship between G+C content and buoyant density in CsCl is linear in the interval from 20 to 80% G+C, provided that the DNA is not modified by methylation or glucosylation. Methylation and glucosylation both decrease the buoyant density of DNA (Szybalski and Szybalski, 1971). On the other hand, the relationship between G+C content and melting temperature is linear even if the DNA is methylated or glucosylated (Woodward and Lebowitz, 1980; Szybalski and Szybalski, 1971). When C. albicans strain H317 DNA was analyzed by the method of equilibrium density gradient centrifugation in CsCl, two density components were discovered. The majority of the DNA was found to band at a position corresponding to 32.8% G+C, but a small amount banded at a position corresponding to 40.3% G+C. It is interesting to compare these results with the thermal denaturation

measurements of C. albicans DNA reported by Stenderup and Leth Bak (1968). They detected only a single component, which suggests that the fraction of satellite DNA is too small to be distinguished by their method. Furthermore, the value they report for G+C content (35.1%) is 2.3% larger than the value reported here for the main band DNA (32.8%). This difference suggests that C. albicans DNA may contain modified bases, but measurements of melting temperature need to be obtained using strain H317 DNA.

The discovery of two density components in C. albicans DNA raises the question of satellite origin. Does the satellite contain nuclear DNA sequences which happen to have a distinctly richer G+C content, or is the satellite of cytoplasmic origin; for example, mitochondrial DNA? Analysis of whole cell DNA and nuclear DNA preparations showed they both contained the satellite component, and this is consistent with the hypothesis of a nuclear satellite. On the other hand, I have shown that all the nuclear DNA preparations were contaminated with mitochondrial DNA, a result consistent with the hypothesis of a cytoplasmic satellite. As shown in Part II, C. albicans mitochondrial DNA can be isolated in a highly purified form (i.e. free from nuclear DNA), and buoyant density measurements of the base composition yield an average of 38.2% G+C. If it is the case that mitochondrial DNA and satellite DNA are identical, one must assume that the difference in G+C content between the two was due to experimental error. The reproducibility of the CsCl method argues against such an error. Inspection of Table 1-1, page 38, reveals very little variation among measurements. The largest deviations (shown in the first and second lines of the

table) are from my earliest experiments during which various technical problems were encountered (e.g. poor u.v. photographs) which might explain the erratic values. All the other measurements differed by at most:

- 1) 0.5% G+C for the main band DNA,
- 2) 0.7% G+C for the satellite DNA, and
- 3) 0.6% G+C for the mtDNA (see Part II).

Differences between the average measurements of each of these DNAs are much greater. The difference between the average measurement of main band DNA and mtDNA is 5.4% G+C, and the average difference between mtDNA and satellite DNA is 2.1% G+C. Clearly, further investigation is warranted to resolve the satellite origin problem.

Three different approaches are available which may lead to the resolution of the satellite origin problem. First, it should be possible to better separate the main band DNA and the satellite DNA by using CsCl gradients which are less steep. This can be achieved simply by centrifuging the samples at slower speeds. With better separation, one could then examine the effect of adding purified mtDNA to a sample known to contain both density components. If the satellite band is of mitochondrial origin, then it should become increasingly prominent with increases in added mtDNA. If the mtDNA has a density intermediate to main band DNA and satellite DNA, then a third band should become increasingly prominent in the region between the two nuclear DNA bands.

The second approach to analyzing the satellite requires its isolation. Although the satellite band contains only a small proportion of the total DNA, it is well separated from the main band DNA

(Figure 1-2, page 37). It should be possible to run a series of preparative CsCl gradients to increasingly enrich and purify the satellite. After purification, the satellite DNA could then be analyzed with restriction endonucleases to determine if it is mitochondrial DNA (see part II for the restriction map of mtDNA).

The third approach to the satellite problem requires the purification of nuclear DNA. As shown above, all attempted methods for obtaining nuclear DNA, free from mtDNA, have not succeeded. As shown in Figure 1-10 (page 77), the degree of contamination is well above the detection limit of gel electrophoresis. The amount of mtDNA contamination can be determined by measuring the kinetics of reassociation using cloned mtDNA as a tracer. It may be the case that the existing levels of mtDNA contamination are below the level of detection of the analytical ultracentrifuge. If not, better purification methods might be developed, or strains of C. albicans might be isolated which do not contain mtDNA (as has been accomplished with S. cerevisiae; Locker et al., 1979). Given sufficiently pure nuclear DNA, the presence of a nuclear satellite DNA can be directly determined.

Cation-exchange Chromatography

The base composition results obtained by the cation-exchange method are consistent with the buoyant density measurements of the main band DNA (which represents the vast majority of the DNA). An average value of 34.4% G+C was obtained (Table 1-2, page 40), but it was reproducibly found that the purines were not in equal molar concentration with the pyrimidines. More specifically, the average difference

in molar proportion of thymine and adenine was 5.0%, and the average difference of cytosine and guanine was 4.0%. It is difficult to explain these observations without further investigation, but they may indicate the presence of modified bases. An explanation follows.

When DNA is digested with DNase I, the end products are polynucleotides with an average size of about four nucleotides (Matsuda and Ogoshi, 1966). With calf thymus DNA, less than 4% of the digestion products are mononucleotides (Sinsheimer and Koerner, 1951). To obtain 5'-mononucleotides, the DNase I digestion must be followed by digestion with phosphodiesterase. When these digestions are performed on DNA from T2r⁺ bacteriophage (which contains glycosylated, 5'-hydroxymethyl-deoxycytidylic acid), only 62% of the phosphorus of the DNA can be recovered in the form of mononucleotides; the remainder of the phosphorus is in the form of polynucleotides (Sinsheimer, 1954). Thus, base modification can result in incomplete digestion of DNA. When undigested DNA is run on the cation-exchange column, it is found to elute with the thymine containing nucleotides (data not shown). If Candida albicans DNA was not completely digested due to base modifications, the apparent molar proportion of thymine would be expected to be equal. In summary, it seems likely that C. albicans DNA contains modified bases but more detailed investigations are warranted. Nevertheless, the measurements of base composition presented here are consistent with one another and with other reported values.

Component Composition of Candida albicans DNA

The statistical analysis of the reassociation kinetics of whole cell Candida albicans DNA provides very strong evidence for the presence of at least four kinetic components. The results of the whole cell DNA reassociation (Table 1-3, page 42) show the composition of the whole cell DNA to be:

- 3.3% foldback sequences,
- 12.8% repetitive sequences, and
- 82.3% single-copy sequences,

with 1.6% non-reacting sequences. If it is assumed that the non-reacting sequences (e.g. very short fragments) are derived randomly from all the sequences of the DNA, then the composition can be corrected to give:

- 3.4% foldback,
- 13.0% repetitive, and
- 83.6% single-copy sequences.

As was shown by the minicot analysis (Table 1-9, page 65), the repetitive fraction actually contains at least two kinetic components. The fraction of repetitive DNA in each of these two components can be calculated as:

fraction in fast component = $.132/.839 = .157$, and

fraction in slow component = $.707/.839 = .843$.

Thus, the component composition of C. albicans whole cell DNA is found to be:

- 3.4% foldback,
- 2.0% = $(.157)(13.0)$ fast repetitive,

11.0% = (.843)(13.0) slow repetitive, and

83.6% single-copy sequences.

It is not possible, from the experiments presented here, to estimate the contribution of mitochondrial DNA to the component composition, but it may be significant. In Saccharomyces cerevisiae, 8% of the total whole cell DNA is mitochondrial DNA, and out of a total of 13% repetitive DNA, only 5% is of nuclear origin (Lauer et al., 1977). In Dictyostelium discoideum (a protozoan), 28% of whole cell DNA is mtDNA (Firtel and Bonner, 1972).

The component composition of C. albicans DNA is similar to that of other fungi, all of which have been found to contain relatively small amounts of repetitive DNA, in contrast to higher eukaryotes (Lewin, 1980). For example, Neurospora crassa nuclear DNA has been found to contain 2% foldback, 8% repetitive, and 90% single-copy sequences (Krumlauf and Marzluf, 1979). Aspergillus nidulans nuclear DNA has been found to contain 2-3% repetitive and 97-98% single-copy sequences (Timberlake, 1978). Saccharomyces cerevisiae whole cell DNA has been found to contain 1.7% fast repetitive, 11.4% slow repetitive, and 86.9% single-copy sequences (Lauer et al., 1977). And as a final example, Achyla bisexualis nuclear DNA has been found to contain 2% foldback, 2% fast repetitive, 14% slow repetitive, and 82% single-copy sequences (Hudspeth et al., 1977). Because of the small sizes of the genomes of these organisms (see below) and the small amounts of repetitive DNA they contain, it is generally thought that the repetitive sequences may mostly represent ribosomal and transfer RNA genes. This hypothesis requires further investigation.

Kinetic Complexity of the Components

Having determined that Candida albicans whole cell DNA contains two repetitive components, it is of interest to know how many base pairs are contained in the sequence which is repeated in each component. It is also of interest to determine the number of base pairs which are not repeated (i.e. contained in single-copy sequences). The length of a sequence, when measured by reassociation kinetics, is called the kinetic complexity (N) and can be calculated from the results for the K_{pure} values using the E. coli DNA results as a reference. The reassociation rate of purified E. coli DNA was found to be 0.233 (Table 1-7, page 57). The length of the E. coli genome can be reasonably assumed to be 4.2×10^6 base pairs (Lewin, 1980). Thus, for each component (except foldback),

$$N = (4.2 \times 10^6)(.233/K_{\text{pure}}).$$

Using this equation, the complexities of the components are found to be:

$$N_{\text{fast repetitive}} = 1.32 \times 10^3 \text{ base pairs,}$$

$$N_{\text{slow repetitive}} = 7.48 \times 10^4 \text{ base pairs, and}$$

$$N_{\text{single-copy}} = 1.55 \times 10^7 \text{ base pairs.}$$

The base pairs of a given kinetic component are not necessarily adjacent on the chromosome. For example, the fast repetitive component might consist of a sequence of about 1000 base pairs repeated many times in the genome, or it might consist of ten sequences, each about 100 base pairs in length, scattered throughout the genome with all ten sequences repeated about the same number of times. Furthermore, it is not possible to determine (from the experiments presented here) the fraction of base pairs in each component that is derived from mtDNA sequences.

As shown in Part II, the mitochondrial genome contains 4.1×10^4 base pairs, but some of the sequences have been found to be repeated within the mitochondrial genome. Thus, it is likely that portions of both repetitive components correspond to mtDNA sequences.

The kinetic complexities of the components of C. albicans DNA are very similar to those of other fungi. For example, in Neurospora crassa DNA, the component complexities have been found to be 1.53×10^4 for the repetitive component and 2.42×10^7 for the single-copy component (Krumlauf and Marzluf, 1979). In Aspergillus nidulans, the component complexities have been found to be 1.1×10^4 for the repetitive component and 2.5×10^7 for the single-copy component (Timberlake, 1978). In Saccharomyces cerevisiae, the fast repetitive component has a complexity of 3.2×10^3 ; the slow repetitive component has a complexity of 2.4×10^4 ; and the single-copy component has a complexity of 1.17×10^7 (Lauer et al., 1977). Lastly, Achyla bisexualis has component complexities of 3.8×10^2 , 7.6×10^4 , and 3.4×10^7 for the fast repetitive, slow repetitive, and single-copy components, respectively (Hudspeth et al., 1977). In summary, all fungi appear to have a repeated sequence of $2-8 \times 10^4$ base pairs and a single copy component containing $1-3 \times 10^7$ base pairs.

Frequency of Repetition of the Repeated Sequences

Each of the two repetitive components contains a sequence which is repeated several times in the genome. It is possible to calculate the approximate number of copies of each sequence in the Candida albicans genome from the reassociation results. The repetitive

frequency (R) of a repeated sequence is related to its complexity (N) and its fraction (f) of the total number of base pairs in the genome (G, see below):

$$R = \frac{f \cdot G}{N} .$$

The value of R can also be calculated from the rate of reassociation of each component, relative to the single-copy component as observed in unfractionated DNA:

$$R = \frac{K_{\text{obs, repetitive}}}{K_{\text{obs, single copy}}} .$$

The values of K_{obs} for each component (except foldback) can be calculated using

$$K_{\text{obs}} = (f)(K_{\text{pure}}),$$

so,

$$K_{\text{obs, fast repetitive}} = 14.84$$

$$K_{\text{obs, slow repetitive}} = 1.44$$

$$K_{\text{obs, single-copy}} = 0.053.$$

Thus, the number of copies of each repeated sequence can be estimated to be

$$R_{\text{fast}} = 280, \text{ and}$$

$$R_{\text{slow}} = 17.$$

These estimates are not intended to imply that the copies of each repeated sequence are identical but only that they have enough homology

to react under the conditions of the reassociation. The thermal stability of the repetitive reassociation products was not examined. Had more stringent reassociation conditions been used, the characteristics of the repetitive DNA (e.g. the fraction of DNA in each repetitive class) might have been much different. Furthermore, these estimates of copy number do not allow estimates to be made of the number of mitochondrial DNA molecules in the cell.

The repetition frequencies of C. albicans repetitive DNA are, however, typical of other fungi. For example, the repeated sequence of Neurospora crassa has been estimated to be present in 140 copies (Krumlauf and Marzluf, 1979). The repeated sequence of Aspergillus nidulans is present in 60 copies (Timberlake, 1978). In Saccharomyces cerevisiae, 70 copies of the fast repetitive sequence are present and 60 copies of the slow repetitive sequence are present (Lauer et al., 1977). Lastly, in Achyla bisexualis, the fast repetitive component consists of 2200 copies of a repeated sequence and the slow repetitive component consists of 80 copies of another sequence (Hudspeth et al., 1977). In summary, fungi tend to lack the highly repeated sequences found in higher eukaryotes (Lewin, 1980).

Genome Size and Ploidy of Candida albicans

Given the complexity (N) of a component of the DNA, its fraction (f), and its repetition frequency (R), the total number of base pairs in the haploid genome (G) can be calculated. Using the single-copy component, whose characteristics have been well defined in this study, the whole cell genome size of Candida albicans strain H317 can be calculated as:

$$\frac{N \cdot R}{f} = \frac{(1.55 \times 10^7)(1)}{(0.836)} = 1.85 \times 10^7 \text{ base pairs.}$$

Because the amount of mitochondrial DNA present in the whole cell DNA preparation is not known, it is not possible to calculate the size of the nuclear genome. However, based upon the amounts of mtDNA present in S. cerevisiae and D. discoideum (see above), the calculated value for haploid genome size should only be over-estimated by 10-20%.

Moreover, the genome size of Candida albicans is quite similar to that of other fungi. As examples, the genome size of Neurospora crassa is 2.7×10^7 bp (Krumlauf and Marzluf, 1979); Aspergillus nidulans has a size of 2.6×10^7 bp (Timberlake, 1978); Saccharomyces cerevisiae has a size of 1.35×10^7 bp (Lauer et al., 1977); and Achyla bisexualis has a size of 4.19×10^7 bp (Hudspeth et al., 1977). In summary, fungi genomes seem to range from three to ten times the size of E. coli DNA (0.42×10^7 bp).

One of the most controversial topics in the field of Candida albicans research concerns the ploidy of this fungus (for review, see Riggsby et al., 1982). One method of determining the ploidy of an organism is by determining the ratio of the total DNA content of the cell (as measured by chemical analysis) and the haploid genome size (as measured by DNA reassociation experiments). The correlation of chemical and reassociation measurements has been especially good for organisms with small amounts of repetitive DNA (Lewin, 1980). The DNA contents of several Candida albicans strains, including strain H317, have been measured by chemical methods and found to have values of about

38 fg (or 38×10^{-15} grams; Riggsby et al., 1982; Torres-Bauza, 1980).

The haploid genome size of strain H317 (1.85×10^7 base pairs; see above) is equal to 19.2 fg. Thus, the ratio of the two measurements equals 2, indicating that Candida albicans is a diploid organism.

This conclusion is supported by measurements of the haploid genome size using both S1 nuclease and optical methods of analyzing reassociation kinetics (data not shown). A more detailed description of these supporting results has been presented (Riggsby et al., 1982).

Distribution of Repetitive Sequences

In all higher eukaryotes, the repeated DNA sequences have been found to be arranged among the single-copy sequences in one of two ways. The first of these two interspersal patterns is the short-period interspersal pattern in which small (200-400 bp) repetitive sequences are distributed among the single-copy sequences at short intervals of less than 3000 bp. The second pattern is the long-period interspersal pattern in which long, repetitive sequences (about 6000 bp average) are contiguous with single-copy sequences of more than 10,000 bp (for review, see Lewin, 1980).

In most fungi, the repetitive DNA has been found to be clustered among the single-copy sequences with little evidence of interspersal. For example, in Achyla bisexualis it has been estimated that the average genomic lengths of repetitive and single-copy DNA are 27,000 bp and 135,000 bp respectively (Hudspeth et al., 1977). Electron microscopy studies of the Achyla DNA show no evidence for interspersal (Pellegrini et al., 1981). Similarly, in Neurospora crassa, the

repetitive sequences have been found to be organized in clusters 10,000 to 75,000 bp in length. Only 0.7% of the single-copy sequences have been found to be contiguous with repetitive sequences in fragments 6400 bp in length (Krumlauf and Marzluf, 1979). Because of the general lack of interspersion in fungi, it seems unlikely that repetitive DNA plans a regulatory function in these organisms.

Because of the mitochondrial DNA contamination of C. albicans nuclear DNA, interspersion experiments were not performed. Mitochondrial DNA would be expected to behave as clustered repetitive DNA and possibly obscure the nature of nuclear repetitive DNA. The genomic blot experiments, however, suggest that much of the nuclear repetitive DNA is clustered in C. albicans. In Figure 1-7 (page 70), three major bands of total fragment length of 12.5 kbp were detected which contained nuclear repetitive DNA (bands nuc 1-3). Because of the intensity of these bands, it is clear that they contain several EcoRI fragments from different portions of the genome. It may be the case that all of the fragments in each band are identical. It should be very easy to clone the fragments in these bands and then explore their arrangement in the C. albicans genome. It should be noted that it may be the case that only portions of the fragments in these three bands are repetitive. That is, it may be that the fragments contain single-copy sequences contiguous with repetitive sequences. Furthermore, other repetitive sequences interspersed within the genome may not have been detected on the

autoradiogram. In conclusion, much remains to be explored concerning the organization of C. albicans repetitive DNA.

PART II: CHARACTERIZATION OF CANDIDA
ALBICANS MITOCHONDRIAL DNA

CHAPTER V

INTRODUCTION

Mitochondria are organelles of eukaryotic cells; they contain the machinery for oxidative phosphorylation. They also contain characteristic DNA molecules (mtDNA). The purpose of Part II of this dissertation is to examine, for the first time, the structure and organization of the mitochondrial DNA of Candida albicans. As background, I first present a review of the information available concerning other mtDNA molecules.

The Functions of mtDNA

Mitochondrial DNA plays an essential role in the normal operation of the mitochondrial machinery. This DNA codes for many (but not all) of the proteins involved in oxidative phosphorylation, and these proteins are synthesized in the mitochondria (for a more detailed review, see Tzagoloff et al., 1979; Borst and Grivell, 1978). Several different mitochondrial systems have been examined, and from these studies, the following generalizations can be made.

1. Cytochrome oxidase. In all mitochondria examined, the three largest subunits of this enzyme (which consists of seven polypeptides in all) are specified by the mtDNA.
2. Coenzyme QH₂-cytochrome reductase. In all mitochondria examined, one polypeptide (cytochrome b) out of seven subunits is specified by the mtDNA.

3. ATPase. This intricate enzyme system is composed of ten polypeptide subunits. Each subunit is located in one of three fractions: 1) the water soluble fraction (five subunits), 2) the membrane bound, hydrophobic fraction (four subunits), or 3) the stalk fraction which holds the other two parts together (one subunit). In all mitochondria examined, only membrane bound subunits are specified by the mtDNA; however, differences exist among organisms with respect to which of the four subunits are specified by the mtDNA. Thus, a subunit gene located in mtDNA in one organism may be located in nuclear DNA in another organism.

These proteins represent most of the translational products of the mitochondria. While all of them are essential to mitochondrial function, they make up only about 5% of the protein in mitochondria; the other 95% is synthesized on cytoplasmic ribosomes and is specified by nuclear genes.

In addition to the messenger RNAs for the mitochondrial specified proteins, mtDNA also codes for all the other RNAs (ribosomal RNAs and transfer RNAs) necessary for mitochondrial protein synthesis. In all mitochondria examined, at least two (one large and one small) ribosomal genes have been found (Borst, 1977). In many species (e.g. Neurospora crassa, Xenopus laevis, Prosophilia melanogaster, Hansenula mrakii, Mus musculus, and Homo sapiens), the rRNA genes are adjacent on the mtDNA (Goldbach et al., 1978b; Wesolowski and Fukauhara, 1981; Bibb et al., 1981). In other species (e.g. Saccharomyces cerevisiae,

Tetrahymena pyriformis, Paramecium primaurelia, Aspergillus nidulans, and Kloeckera africans) the rRNA genes are separated by large distances (Goldbach et al., 1978b; Cummings and Laping, 1981; Macino et al., 1980; Clark-Walker et al., 1981). A few organisms have been found to contain "extra" ribosomal genes. In the two known cases (Tetrahymena pyriformis and Kloeckera africana), the duplicated sequences have been found to be contained in inverted repeats (Goldbach et al., 1978b; Clark-Walker et al., 1981). It should be mentioned that essentially all of the mitochondrial ribosomal proteins (50-80 proteins) are synthesized outside of the mitochondria (using nuclear DNA transcripts) and are transported into the mitochondria for assembly with the rRNAs (Borst and Grivell, 1978).

Perhaps the most unexpected recent discoveries concerning mitochondria have come from studies of mitochondrial transfer RNA (tRNA). It has been found that the genetic code of mitochondrial DNA is different from the putative universal genetic code. For example, in Saccharomyces and human mitochondria, the UGA codon is translated as a tryptophan codon rather than as a terminator codon. There are also differences between the genetic codes of different mitochondria (Barrell et al., 1979; Bonitz et al., 1980; Bibb et al., 1981). Furthermore, in all mitochondria examined, the numbers of different tRNAs have been found to be fewer than the theoretical minimum of 32 required by the wobble hypothesis (Crick, 1966). In the mouse, only 22 tRNAs have been found to be coded in the mtDNA (Bibb et al., 1981). In human mtDNA, only 23 tRNA genes have been found (Barrell et al., 1980). And, in Saccharomyces mitochondria, only 24 tRNA genes have been found (Bonitz et al., 1980). Moreover, studies of cytoplasmic tRNAs have

shown that none are transported into mitochondria, and no mitochondrial tRNAs hybridize to nuclear DNA (Aujame and Freeman, 1979). In order to function with a reduced number of tRNAs, mitochondria appear to have developed translational machinery which uses a "two out of three" decoding mechanism. In this system, a single tRNA may be used to recognize four codons which differ only in the third position (Bonitz et al., 1980).

In summary, mtDNA codes for a rather small number of functions. The majority of the mitochondrial protein is synthesized in the cytoplasm and transported across the mitochondrial membranes. On the other hand, nucleic acids are apparently unable to cross the mitochondrial membranes. The mitochondria synthesize all the RNAs (tRNA and rRNA, as well as mRNA) necessary to synthesize mitochondria specified proteins.

While it is clear that the cell requires mitochondria for the processes of oxidative phosphorylation, it is not obvious why mtDNA should be required. That is, it is not obvious why the few structural genes provided by mtDNA are not located in the nuclear DNA. Indeed, it seems that the cell has gone to great expense to maintain these genes in mtDNA. The nuclear DNA specified, mitochondrial proteins include: 1) ribosomal proteins, 2) aminoacyl synthetases, 3) DNA and RNA polymerases, 4) ligases, 5) RNA processing and modifying enzymes, 6) DNA recombination enzymes, and 7) others. In all cases examined, these imported proteins have amino acid sequences which are distinct from the corresponding cytoplasmic proteins. In other words, proteins that are imported into the mitochondria are not used by the cell outside of the mitochondria (Borst and Grivell, 1978). Thus, the nucleus provides all the information for all the proteins (at least 90) required by the mitochondria for the expression of only about seven

mitochondrial genes. It would seem to be simpler and more efficient to have all the oxidative phosphorylation genes located in the nucleus so that the protein subunits could be synthesized on the cytoplasmic ribosomes and then transported into the mitochondria.

Many hypotheses have been proposed to explain the existence of mtDNA (Borst, 1977). Until recently, the most attractive hypothesis has been that the polypeptides specified by mtDNA (which are very hydrophobic proteins) are more easily arranged in their appropriate sites in the mitochondrial membranes if they are synthesized inside the mitochondria. In other words, transport of these hydrophobic proteins would be difficult. This hypothesis is weakened by the above mentioned finding that the ATPase subunits specified in one organism by mtDNA genes are specified by nuclear DNA genes in another (Borst and Grivell, 1978). Perhaps the most popular hypothesis is centered around the belief that mitochondria have evolved from an endosymbiotic relationship between eukaryotic cells and bacterial cells. The bacterial symbionts are thought to have transferred, with time, many of their genes to the nuclear DNA, but the transfer has been incomplete. In different organisms, the extent of transfer has reached different levels. (For a review of the endosymbiont hypothesis, see Gray and Doolittle, 1982.)

The Structure of mtDNAs

Mitochondrial DNAs can be either circular or linear. Several methods have been used to determine the structure of mtDNAs. Direct observation with the electron microscope is sometimes useful; the observation of circular DNA molecules is good evidence for a circular structure. The observation is supported if the circular DNA is of the proper

size, as estimated by independent methods (e.g. by restriction endonuclease analysis or DNA reassociation analysis). Electron microscopic examination is often of little help in determining the structure of mtDNA because often it is difficult to isolate intact mtDNA. Thus, the observation of linear mtDNA by electron microscopy is not conclusive evidence for non-circularity. Another method of determining mtDNA structure is by restriction endonuclease analysis. If one restriction enzyme produces $\underline{m}$ unique fragments and another enzyme produces $\underline{n}$ unique fragments, then digestion with both enzymes should produce $\underline{m} + \underline{n}$ fragments if the DNA is circular and $\underline{m} + \underline{n} - 1$ fragments if the DNA is linear. Detailed restriction endonuclease mapping should confirm the DNA structure. To further confirm a suspected linear mtDNA, end labeling (at the 5' and/or 3' ends) experiments can be performed. By this method, limit restriction enzyme digestions should result in two fragments which correspond to the end pieces. Linear DNAs produced by random breakage of circular molecules would not be expected to produce specific labeled fragments after restriction endonuclease digestion.

The majority of mtDNAs examined have been found to be circular. A few exceptions have been found, however, which are linear. These exceptions have all been found in lower eukaryotes (but not all lower eukaryotes have linear mtDNAs). The species containing linear mtDNA are: Tetrahymena pyriformis (Goldbach et al., 1978a), Paramecium primaurelia (Goddard and Cummings, 1977), and Hansenula mrakii (Wesolowski and Fukuhara, 1981). There does not appear to be any unique feature among linear (or circular) mitochondrial DNAs. For example, replication proceeds bidirectionally in Tetrahymena, but unidirectionally in Paramecium (Cummings and Laping, 1981). Also,

Tetrahymena mtDNA contains an inverted repeat while Paramecium mtDNA does not (Goldbach et al., 1978b). Furthermore, the ribosomal RNA genes are adjacent in *Hansenula* but widely spaced in Tetrahymena and Paramecium (see above). Similarly, differences have been found among circular mtDNAs in 1) the mode of replication (Nass, 1980), 2) the presence of repeated sequences, and 3) the organization of the mtDNA genome (see above).

The Size of mtDNAs

The sizes of many mitochondrial DNAs have been determined. The smallest belong to the invertebrates and vertebrates. Among the organisms examined in this group are flatworms, sea urchins, amphibians, birds, insects, and mammals; and it has been found that the mtDNAs of these organisms are remarkably similar in size, ranging only from 15 to 18 kbp (Lewin, 1980; Borst, 1977; Suyama and Miura, 1968). The G + C contents of the mtDNAs in this group are more diverse, however, ranging from 31% to 52% (Gray and Doolittle, 1982). As implied above, all of these mtDNAs have been found to be circular.

Examination of the lower eukaryotes (Fungi and Protozoa) has revealed much larger mtDNAs with a greater range of sizes. Examples of this group are (Borst, 1977): 1) Saccharomyces (74 kbp), 2) Neurospora (57 kbp), 3) Tetrahymena (49 kbp), 4) Paramecium (41 kbp), 5) Kluyveromyces (33 kbp), and 6) Plasmodium (27 kbp). The G + C contents of these mtDNAs are also diverse and range from 18% to 47% (Gray and Doolittle, 1982). As mentioned above, a few of these mtDNAs are linear. Comparison of the mitochondrial DNAs of lower eukaryotes with those of higher eukaryotes suggests that evolution has selected mtDNAs of

smaller size; however, contradictory evidence is provided by examination of plant mtDNAs.

Plant mitochondrial DNAs are the largest known. Pea mtDNA is 110 kbp in length (Kolodner and Tewari, 1972); tobacco mtDNA has been found as large as 150 kbp (Sparks and Dale, 1980); and maize mtDNA has been reported to contain 480 kbp (Sederoff et al., 1981). Besides being large, plant mtDNAs differ from other mtDNAs in another important way; they are heterogeneous in size. For example, electron microscopic examination of mtDNA isolated from tobacco suspension culture cells has shown the presence of circular molecules of many sizes ranging from 1.8 kbp to 150 kbp with the most abundant size being 32 kbp. Curiously, restriction endonuclease analysis of this heterogeneous mixture yielded a size estimate of 270 kbp (Sparks and Dale, 1980). Restriction digests of mtDNA from plant tissue culture cells and mtDNA from whole plants have been found to be essentially identical (Dale et al., 1981). Investigations using mtDNAs isolated from whole plants and plant parts have also shown heterogeneous mixtures in soybeans, tobacco, cucumbers, corn, and potatoes (Sparks and Dale, 1980). An exception to the heterogeneity of plant mtDNAs is the mtDNA of pea which reportedly has a homogeneous size of 110 kbp (Kolodner and Tewari, 1972).

It is not clear why plant mitochondrial DNAs are heterogeneous in size; however, at least two hypotheses have been proposed (Dale et al., 1981). First, the genome of a plant mtDNA may actually be dispersed among the different size classes. In other words, the sequences of each size class of mtDNA might be different with ribosomal genes, for example, located on molecules of one size, ATPase subunit genes located on molecules of a different size, and so forth. Second, it

might be that a single, unique size class of mtDNA molecules was initially present and underwent extensive rearrangements and non-homologous recombinations (perhaps during development or perhaps during evolution) to produce a heterogeneous mixture of mtDNA molecules. Experiments using mtDNAs from several species of teosinte (an ancestor of maize) and maize have suggested that the evolution of plant mtDNAs has involved extensive rearrangements and recombinations of various sequences while conserving the structure of other sequences (Sederoff et al., 1981). Future studies of plant mtDNAs will undoubtedly be interesting.

The Rearrangement of mtDNA

Sequence rearrangements have been shown to occur frequently in the mitochondrial DNAs of several organisms other than plants. All of these organisms are lower eukaryotes. Two types of rearrangements will be described: 1) developmentally coordinated mtDNA sequence changes (Podospora anserina) and 2) spontaneous mtDNA sequence rearrangements involving massive deletions and amplifications (Saccharomyces cerevisiae and Neurospora crassa).

Saccharomyces cerevisiae. This organism is a facultative anaerobe. When yeast cells are plated on medium containing large amounts of a nonfermentable substrate (e.g. glycerol or lactic acid) and small amounts of a fermentable substrate (e.g. glucose), wild type cells form large colonies. Occasionally, small colonies are observed, called petite mutants, which have altered mtDNA and nonfunctional mitochondria. These petites cannot utilize the nonfermentable substrate, but they are still viable due to the presence of glycolytic enzymes.

(Petite mutants can also arise which have nuclear DNA lesions, but these will not be discussed here). Mitochondrial petite mutants can be grouped into two classes depending upon their behavior in crosses with wild type strains. Suppressive petites lead to predominantly petite phenotypes (i.e. small colonies) after mating with wild type. Neutral petites do not transmit the petite phenotype (i.e. after mating, only large colonies are observed).

Analysis of the mitochondrial DNAs from petite strains has revealed many interesting properties (for review, see Locker et al., 1979; Borst and Grivell, 1978).

1. The buoyant density of petite mtDNA depends upon the strain and may have values higher, lower, or the same as wild type mtDNA.
2. Petite mtDNA reassociates at a faster rate than wild type mtDNA, but different petite strains have different kinetic complexities (up to about 70% of wild type).
3. Hybridization of petite mtDNA to wild type mtDNA has shown that the former is derived from the latter.
4. Restriction enzyme analysis of petite mtDNA has shown that these DNAs result from the deletion of 20-100% of the wild type mtDNA. Individual petite strains may contain the genetic information from any part of the wild type mtDNA.
5. Electron microscopic examination has shown that petite mtDNA is about the same size as wild type mtDNA.

As a result of the many studies reported, the following understanding of petite mutants has emerged. Wild type mtDNA frequently and

spontaneously will undergo events (not yet understood) which result in massive deletions of mtDNA sequences. The remaining sequences are then amplified to give petite mtDNA molecules of the wild type size. It is thought that suppressive petites contain replicator sites with increased affinity for DNA polymerase. Thus, in crosses with wild type, the petite mtDNA would replicate more rapidly than wild type and result in predominantly petite phenotypes. On the other hand, neutral petites are thought to contain no mtDNA at all and thus have no means to transmit the petite phenotype.

Neurospora crassa. This organism is an obligate aerobe, and therefore, loss of mitochondrial function would result in loss of cellular viability. Comparison of restriction digest patterns of various N. crassa strains has shown that each contains one of two types of mtDNA (Mannella et al., 1979a). Recently, however, Mannella et al. (1979b) have characterized two variants of one type of Neurospora mtDNA and found sequence rearrangements. The first variant they discovered (called type IIa) produced an extra EcoRI band which contained fragments 2.1 kbp in length. Further analysis of these fragments showed that they derived from tandemly inserted reiterations of a specific 2.1 kbp sequence. In other words, the rearrangement involved placing copies of this sequence adjacent to one another. The repeats were found to be oriented head-to-tail with individual mtDNA molecules containing anywhere from 0 to 8 copies. Mannella et al. (1979b) also reported a second variant (called HI-10) which was derived from the first (i.e. IIa). This variant contained a much larger (27 kbp) amplified segment of the mtDNA. In both cases, mtDNA isolated from the variant strains

proved to be heterogeneous with significant numbers of mtDNA molecules containing no repeated sequences. (Mitochondrial DNA from type IIa cells contained 50% unaltered molecules.) It was not possible to measure the sizes of individual variant molecules to determine whether they were all the same (as with S. cerevisiae petites), because of partial degradation of the mtDNA during isolation. The researchers hypothesized, however, that the variant molecules (i.e. those containing tandem reiterations) were defective. When undigested type IIa mtDNA was analyzed on low concentration agarose gels, a single broad band was observed, but no novel bands of lower molecular weight were found. Thus, the evidence is consistent with a model of amplification in which the repeated sequences replace other sequences so that genetically incomplete (i.e. defective) molecules are produced. If this model is correct, then the variant cells would also be expected to maintain at least a few wild type mtDNAs to maintain cellular viability. Thus, it appears that similar mtDNA rearrangement mechanisms may be at work in S. cerevisiae and N. crassa (amplification accompanied by deletion) with heterogeneity differences due to physiological constraints.

It is tempting to extrapolate these ideas to explain the nature of another class of Neurospora mitochondrial mutants called stoppers. These interesting mutants have a phenotype characterized by an irregular pattern of variable periods of growth and no growth (Bertrand et al., 1976). Restriction enzyme analysis of the mtDNA of these mutants has revealed that deletions have occurred to produce defective mtDNAs (Manella et al., 1979b). It has been suggested that these defective mtDNAs replicate faster than the wild type mtDNA, which is required for

growth of the obligate aerobes. If so, it seems reasonable that the periods of no growth in stopper strains would result as the numbers of defective mtDNAs increased, and growth periods would resume after the replication of wild type mtDNAs "caught up." This hypothesis seems even more reasonable in the light of the next example.

Podospora anserina. This organism is also an obligate aerobe. It is an unusual fungus, however, in that it exhibits senescence (for review, see Smith and Rubenstein, 1973; Jamet-Vierny et al., 1980). After spore germination, all mycelia maintained in exponential growth (by continuous transfer to fresh medium) increase in size to a length called the incubation distance. This distance is characteristic for each strain and varies greatly (e.g. 10-100 cm) between strains. All members of a strain (or race) grow to at least this size. After reaching the incubation distance, the number of mycelia able to continue growth begins to decrease at an exponential rate which is also characteristic for each strain. All mycelia eventually reach a senescent state as their growth slows and then stops.

Senescence in P. anserina has long been known to be under cytoplasmic control, and the strain specific life span of a mycelium is known to be maternally inherited (see Jamet-Vierny, 1980). Recently Cummings et al. (1979) examined the mitochondrial DNA from young and senescent cultures. They found mtDNA from young cultures to be about 100 kbp in length and have a characteristic restriction enzyme pattern and buoyant density. Mitochondrial DNA from senescent cultures was different. Two buoyant density components were observed. Restriction enzyme analysis showed that the light mtDNA was missing several

fragments (as compared to young mtDNA). The heavy mtDNA (30 kbp) was found to be composed of a tandemly repeated sequence only 2.6 kbp in length. The study of senescent mtDNA was carried further by Jamet-Vierny et al. (1980). These researchers isolated mtDNA from senescent cultures and also found by restriction enzyme analysis, a tandemly repeated sequence (6.3 kbp in length). Electron microscopic examination of the senescent mtDNA revealed that the repeated sequence was contained in a multimeric set of circular molecules with a monomer size of 7.2 kbp. Hybridization experiments confirmed that the repeated sequences were derived from wild type (i.e. young) mtDNA. Thus, the sequence rearrangements of P. anserina mtDNA greatly resemble those of S. cerevisiae and N. crassa (discussed above).

It has been proposed (Jamet-Vierny et al., 1980) that senescence in P. anserina is caused by sequence rearrangements of the mitochondrial DNA. The relative amount of senescent mtDNA increases and the relative amount of wild type decreases as senescence progresses. This suggests that senescent mtDNA replicates preferentially to wild type mtDNA. Elimination of the wild type mtDNA would eliminate the capability for respiration and result in the death of this obligate aerobe.

The Organization and Transcription of mtDNA

All mitochondrial DNAs seem to carry nearly the same fundamental genetic functions. That is, all code for about the same oxidative phosphorylation subunits, two ribosomal RNAs, and 22-25 transfer RNAs (see above). Recent reports have shown, however, that the organization of this fundamental information varies greatly among mtDNAs. The most complete studies have been performed using HeLa cell mtDNA and

Saccharomyces cerevisiae mtDNA (for brief review, see Borst and Grivell, 1981; Attardi, 1981).

Many differences exist between the organizations of the yeast and the mammalian mtDNAs. Some of these differences are merely in the order of the coding sequences. For example, the two rRNA genes of yeast mtDNA are widely separated, but those of mammalian mtDNA are close together. Furthermore, the tRNA genes of yeast mtDNA are clustered, but in mammalian mtDNA, the tRNA genes have a scattered distribution among the other coding sequences. The most striking difference between yeast and mammalian mtDNAs, however, is in the compactness of the fundamental genetic information. For example, (1) yeast mtDNA is large (about 78 kbp), but the fundamental genetic information of mammalian mtDNA is contained in a small molecule (about 17 kbp). (2) The genetic information of yeast mtDNA is spaced apart by non-coding sequences, but the coding sequences for rRNA, tRNA, and mRNA of mammalian mtDNA are very close together. (3) The rRNA genes of yeast mtDNA are relatively large (21S and 15S), but those of mammalian mtDNA are small (16S and 12S). (4) Many of the coding sequences of yeast mtDNA are interrupted by one or more intervening sequence, but those of mammalian mtDNA have been found to be uninterrupted. In summary, mammalian mtDNAs seem to have evolved to a very small size by the elimination of the majority of the nonessential sequences.

Great differences have also been found concerning the transcription of yeast and mammalian mtDNAs. For example, (1) yeast mtDNA is transcribed asymmetrically using five or more promoter sites, but mammalian mtDNA is transcribed symmetrically and contains only two sites for the initiation of RNA synthesis (one on each strand). (2)

The mRNAs from yeast mtDNA have noncoding leader and trailing sequences, but the mRNAs from mammalian mtDNA contain little or no noncoding sequences. (3) Indeed, many mammalian mRNAs lack complete stop-translation codons until 3'-poly A tails are added. Yeast mtDNA codes for complete stop-translation codons on mRNAs which have very short or no poly A tails. Thus, even at the level of transcription, mammalian mtDNAs seem to have evolved to a very efficient degree. Future experiments will hopefully shed light on the nature of the selective pressures which have brought about the highly streamlined mtDNAs of mammals.

Statement of Intent

The preceding review illustrates many of the interesting and very diverse features of mitochondrial DNAs. The purpose of the research presented in the following two chapters was to characterize the mtDNA of Candida albicans. To my knowledge, no previous information has been reported concerning the mtDNA of this important, pathogenic fungus. The specific aims of this research were:

- 1) to establish methods for isolating large amounts of C. albicans mtDNA in pure form,
- 2) to determine the size and base composition of the mtDNA,
- 3) to identify restriction endonucleases useful for characterizing the mtDNA,
- 4) to clone the sequences of the mtDNA, and
- 5) to determine the structure of the mtDNA by restriction endonuclease mapping.

As a result of meeting these goals, a foundation for future studies of C. albicans mtDNA has been established.

CHAPTER VI

MATERIALS AND METHODS

Organisms

Candida albicans strain H317 was used as the source of mitochondrial DNA. This strain was described in Part I, Chapter II.

Escherichia coli strain HB101 (obtained from Dr. Karl Sirotkin, The University of Tennessee, Knoxville) was used in the cloning experiments. This strain carries the following markers: thr⁻, leu⁻, pro⁻, rec A⁻, thi⁻, hsd R⁻, and sup E⁻. The cloning vector was plasmid pBR322 (also obtained from Dr. Sirotkin) which confers resistance to ampicillin and tetracycline.

Growth of Candida albicans Cells

Yeast phase C. albicans cells for the isolation of mitochondria were grown in YEPL medium (Syverson, 1981; Zeman and Lusena, 1975). This medium contained 1% Bacto-yeast extract (Difco); 2% Bacto-peptone (Difco), and 6% D, L-sodium lactate (obtained from Sigma as a 60% syrup containing equal amounts of the D and L stereoisomers). Sterile sodium lactate was added after autoclaving the other components. To reduce foaming, approximately 1 ml of sterile antifoam B (J. T. Baker Chemicals) was added per liter of YEPL. To grow 8-12 liters of cells, a starter culture of 500 ml of YEPL was first heavily inoculated and incubated overnight at 25-30°C with good aeration. This culture was used to inoculate the large culture which was then incubated for about 48 hours at 25-30°C with vigorous bubbling (provided by forcing sterile air

through gas dispersion tubes). Approximately 7 grams (wet weight) of cells were obtained per liter of YEPL.

Preparation of Spheroplasts

The procedure for removing the cell walls from C. albicans to produce osmotically sensitive cells was described in Part I, Chapter II.

Isolation of Mitochondria

An outline of the procedure used for isolating mitochondria from C. albicans spheroplasts is shown in Figure 2-1. All steps were performed at 4°C.

To obtain a crude preparation of mitochondria, a modification of the procedure of Zeman and Lusena (1975) was used. Spheroplasts were osmotically lysed by resuspension in 0.25 M sucrose, 1 mM EDTA, 5 mM Tris-HCl, pH 7.5. The suspension was homogenized with several strokes in a dounce homogenizer and was then left to stand for 30 minutes to allow maximum lysis. The pH of the lysate was adjusted to 7.5 with 10 N NaOH and centrifuged with a Sorvall RC2-B centrifuge in a small Sorvall rotor (type SS-34) at 5000 rpm for 5 minutes. The supernatant (containing mitochondria) was transferred to a clean centrifuge tube to await further purification while the pellet was resuspended again in lysis buffer and dounce homogenized. The homogenate was centrifuged at 5000 rpm for 5 minutes, and the resulting supernatant was mixed with that from the first centrifugation. The pellet (containing nuclei and unbroken cells) was discarded, and the combined supernatants were centrifuged again (to get rid of additional nuclei) at 5000 rpm for 5 minutes. To pellet the enriched mitochondria, the supernatant

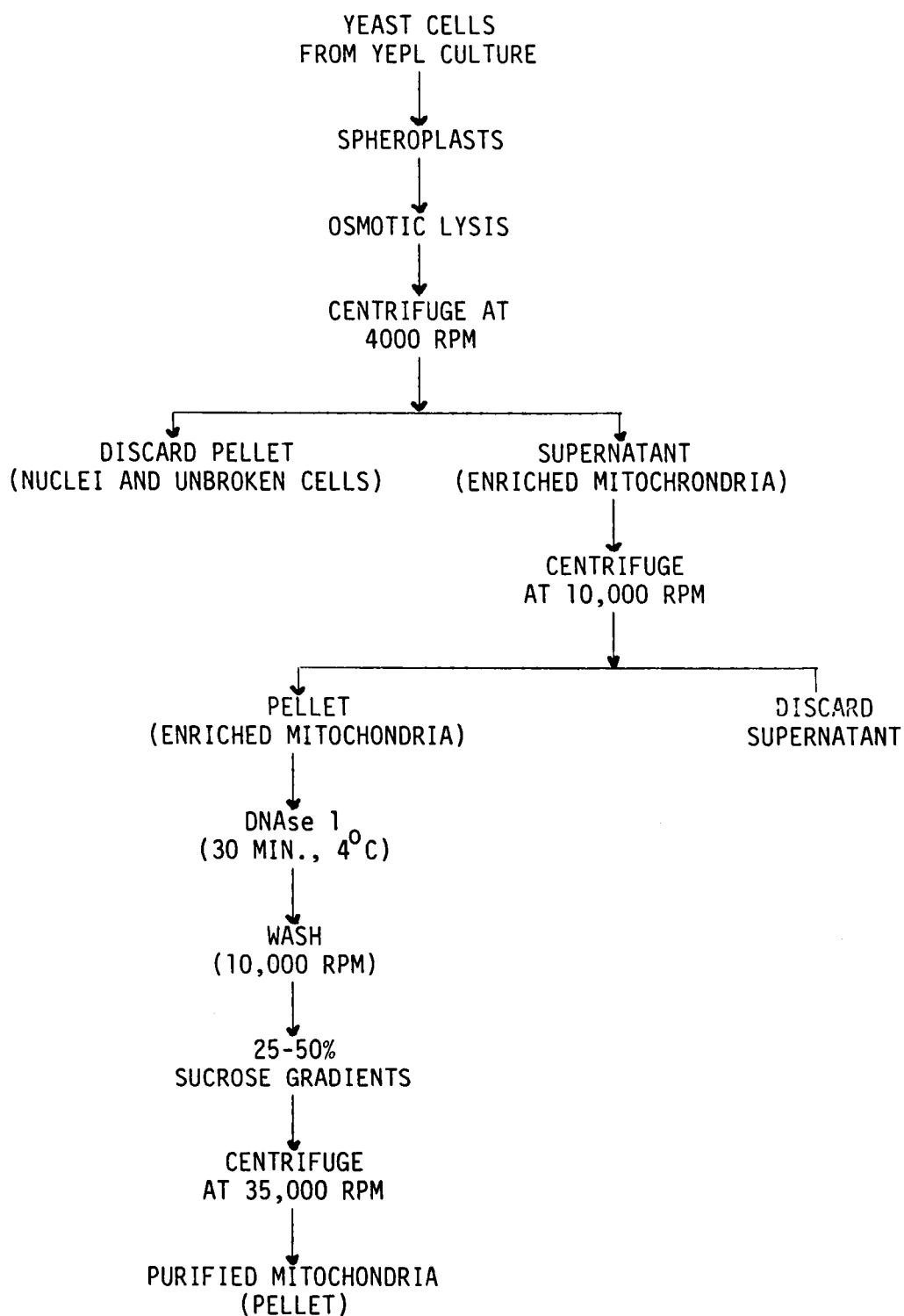


Figure 2-1. Isolation of mitochondria.

was transferred to a clean centrifuge tube and centrifuged at 10,000 rpm for 30 minutes.

To obtain highly purified mitochondria, it was necessary to eliminate nuclear DNA by digestion with DNase I (Zeman and Lusena, 1975). For this purpose, the crude mitochondrial pellet was resuspended in 0.3 M sucrose, 5 mM $MgCl_2$, 0.15% BSA, 20 mM Tris-HCl, pH 7.5. The suspension was homogenized to disperse clumps, and DNase I (Worthington) was added to a final concentration of 25 $\mu g/ml$. The reaction mixture was incubated at 4°C for 30 minutes. To stop the reaction, the mixture was diluted with 0.3 M sucrose, 1 mM EDTA, 0.15% BSA, 20 mM Tris-HCl, pH 7.5 and centrifuged at 10,000 rpm for 30 minutes. The mitochondria were washed 2 or 3 times by resuspension in the latter buffer followed by centrifugation.

To eliminate contaminating intact cells (which would resist DNase I treatment), the mitochondria were purified on 25-50% (w/v) linear sucrose gradients buffered with 10 mM Hepes, 1 mM EDTA, pH 7.0 (Linnane and Lukins, 1975). The gradients were centrifuged in a Beckman SW28 rotor and an L5-75 ultracentrifuge at 21,000 rpm for 90 minutes. Intact cells formed a pellet at the bottom of the tube while the mitochondria formed two diffuse bands approximately two-thirds of the way through the gradients. The nature of this mitochondria heterogeneity was not explored, but many have been due to 1) damage resulting from the isolation procedure or 2) heterogeneity of the mitochondria within the cells (Lloyd, 1974). After removal of both bands from the gradients, they were diluted about 1:1 with 25% sucrose gradient solution, and the highly purified mitochondria were collected by centrifugation at 35,000

rpm for 30 minutes in a Beckman 42.1 fixed angle rotor.

Isolation of Mitochondrial DNA

The mitochondria were lysed by resuspension in 0.1 M EDTA, 0.2 M NaCl, 0.1 M Tris, pH 8.0 containing 100-200 μ g proteinase K per ml followed by the addition of 25% SDS to a final concentration of 1%. The mitochondrial DNA (mtDNA) was isolated from the lysate by either 1) enzymatic purification (see Chapter II) or 2) CsCl-ethidium bromide gradient purification (Davis et al., 1980).

For the enzymatic method of purification, the lysate was extracted once with phenol : chloroform : isoamyl alcohol (25:24:1) and twice with chloroform : isoamyl alcohol (24:1). The nucleic acids were precipitated by the addition of two volumes of 95% ethanol and incubation at -20°C overnight. The precipitate was collected by centrifugation at 8000 rpm for 30 minutes and resuspended in 10 mM Tris, 1 mM EDTA, pH 7.5 (TE buffer). The solution was then brought to 1 X SSC (0.15 M NaCl, 0.015 M Na citrate, pH 7.0) using a 10 X stock solution, and enzymatically treated by adding per ml : 25 μ g RNase A, 50 units TI RNase, 25 μ g alpha-amylase, and 25 μ g beta-amylase. After incubation at 37°C for 45 minutes, the solution was extracted with phenol-chloroform-isoamyl alcohol, and was ethanol precipitated as before. The precipitate was collected and resuspended in TE buffer and isopropanol precipitated by adding 0.1 volume of 3M Na acetate, 1 mM EDTA, pH 7.0 and 0.54 volume of isopropanol. The precipitated DNA was immediately collected by centrifugation and resuspended in TE buffer.

For the CsCl-ethidium bromide gradient method of purification, the lysate was adjusted to 8.7 ml with TE buffer, and 8.3 grams of CsCl

and 0.9 ml of 10 mg/ml EtBr were added. The refractive index was adjusted (if necessary) to 1.390-1.396. The solution was centrifuged for 48 hours at 40,000 rpm at 20°C in a Beckman 50 Ti fixed angle rotor and L5-75 ultracentrifuge. Following centrifugation, the DNA (which formed a single band) was visualized by means of long wavelength, ultraviolet light (Ultraviolet Products, Inc.; model UVL-56). The band was removed by puncturing the side of the tube with a 22 gauge needle and syringe. The EtBr was removed from the mtDNA solution by extraction with isopropanol that had been saturated with 5M CsCl in TE buffer. The mtDNA was precipitated at -20°C overnight following the addition of two volumes of H₂O and six volumes of 95% ethanol. The precipitated DNA was collected by centrifugation at 8000 rpm for 30 minutes and resuspended in TE buffer.

CsCl Density Gradient Analysis of MtDNA

The (G+C) content of mtDNA was measured following the protocol of Mandel et al. (1968) described in Part I, Chapter II.

Restriction Endonuclease Digestion of DNA

Restriction endonucleases were purchased from New England Biolabs and Bethesda Research Laboratories. The reaction conditions for the digestions were those recommended by the supplier.

Polyacrylamide-Agarose Gel Electrophoresis

Small DNA fragments (i.e. less than 1000 bp) were sized using vertical, 2.5% polyacrylamide-0.5% agarose gels using the method described in Part I, Chapter II.

Agarose Gel Electrophoresis

Agarose solutions were prepared by dissolving Seakem agarose (Marine Colloids, Inc.) in electrophoresis buffer (40 mM Tris, 5 mM Na acetate, 1 mM EDTA adjusted to pH 7.8 using glacial acetic acid) by boiling over a bunsen burner. After cooling to about 50°C, vertical gels (150 X 140 X 4 mM; Watson Products) were poured. DNA samples to be electrophoresed were mixed with 1/4 volume of loading/dye solution (25% ficoll, 0.1% bromophenol blue, 0.1% xylene cyanole FF, 10 mM Tris-HCl, 1 mM EDTA, pH 7.8). Electrophoresis was performed using constant voltage conditions. Following electrophoresis, DNA was visualized by staining the gels for about 30 minutes in 0.5 µg/ml ethidium bromide and then using a Mineralight 302 nm, long wave ultraviolet light transilluminator (Fisher). Photographs were made using Polaroid type 57 film and a Kodak number 23A, Wratten gelatin filter.

Transfer of DNA to Nitrocellulose

After electrophoresis, DNA to be hybridized with radioactive probes was transferred from the gel to a nitrocellulose filter by the method of Southern (1979). Single stranded nicks were first introduced into the DNA by irradiating the EtBr stained gel with short wavelength, ultraviolet light. The DNA was then dissociated by soaking the gel (with gentle agitation) in 0.8 M NaCl, 0.4 M NaOH for 30 minutes. To neutralize the gel, it was then soaked for 30 minutes in 0.5 M Tris-HCl, 1.5 M NaCl, pH 7.6 (again with gentle agitation). The blotting apparatus was set up as suggested by Southern (1979) using Whatman 3 MM filter paper (Fisher), 10 X SSC, and nitrocellulose filters (type BA85 from Schleicher and Schuell). Blotting was carried out at room temperature

for 8-18 hours after which the nitrocellulose filters were washed with 3 X SSC, baked for two hours at 80°C in a vacuum oven, and stored at room temperature until used for hybridization reactions.

Labeling of DNA by Nick Translation

DNAs to be used as hybridization probes were labeled by the in vitro nick translation method (Davis et al., 1980). Each nick translation reaction contained 15 µl of DNA (33 µg/ml), 2.5 µl of 10 X buffer (0.5 M Tris, pH 7.8, 0.5 M MgCl₂, 0.5 mM dithiothreitol, and 50% glycerol), 1 µl of unlabeled deoxyribonucleotides (250 µM dATP, TTP, and dGTP), and 5 µl of 5'-[alpha-³²P]deoxycytidine triphosphate (approximately 50 µCi and 400 Ci/mmol; Amersham). After cooling the mixture to 0°C on ice, 0.25 µl of E. coli DNA polymerase (Boehringer Mannheim; 5000 units per ml) and 1 µl of 25 X DNaseI-BSA (25 ng/ml DNaseI and 1.25 mg/ml BSA) were added. The reaction mixture was then incubated at 15°C for about 3 hours. The reactions were stopped by adding 2 µl of 0.5 M EDTA, pH 8.0.

Unincorporated nucleotides were removed by Sephadex G-50 (Pharmacia) column chromatography. Dry Sephadex G-50 was hydrated in column buffer (0.25 M NaCl, 50 mM Tris, pH 7.9, 0.5% SDS, 2 mM EDTA). Samples to be purified were loaded on 4 ml, equilibrated columns and washed through with the above buffer. Fractions (200 µl) were collected, and those containing the excluded radioactivity were pooled (about 800 µl total). To determine the specific activity of the labeled DNA, 1-5 µl aliquots of the DNA was mixed with 1 ml of H₂O and 9 ml of Scintiverse (Fisher) and counted in a Beckman scintillation counter. Assuming total recovery of input DNA (i.e. 0.5 µg), specific activities

of 10^7 - 10^8 cpm/ μ g were routinely obtained.

Hybridization to DNA on Nitrocellulose

DNA on nitrocellulose filters were hybridized with radioactive probes using the procedure of Southern (1979). Prior to adding the probe, the filters were prehybridized to reduce nonspecific binding of radioactivity. Prehybridization solution contained 1 M NaCl, 50 mM Tris, pH 7.9, 1 mM EDTA, 2 mM $\text{Na}_4\text{P}_3\text{O}_7$, 0.1% sarkosyl, 0.2% ficoll, 0.2% BSA, 0.2% polyvinylpyrrolidone, and 100 μ g/ml single-stranded Salmon sperm DNA. (To prepare the salmon sperm DNA, 10 mg of lyophilized DNA was dissolved in 0.8 ml of 1N NaOH, boiled for 5 minutes, and then, 150 μ l of 2M Tris-HCl, 50 μ l of concentrated HCl, and 1 μ l of 0.5 M EDTA were added.) Each filter was sealed in a plastic bag with 5-20 ml of the above solution and incubated at 65°C for 5-15 hours.

The hybridizations were started by injecting G-50 purified, heat denatured probe (final concentration of 10^5 - 10^6 cpm/ml), resealing, and placing the bag again at 65°C. After 36-48 hours, the filters were removed from the bags and washed with 50 mM Tris, pH 7.8, 50 mM NaCl, 0.1% sarcosyl. Each filter was washed twice for 5 minutes at room temperature and twice for 30 minutes at 65°C. Radioactivity on the filters was located by autoradiography at -70°C using Kodak type XAR-5 x-ray film in film holders between two Dupont Cronex Lightning-Plus intensifier screens.

Cloning of mtDNA

EcoRI fragments of mtDNA were cloned in E. coli strain HB101 using plasmid pBR322 as the vector. pBR322 has a single EcoRI site

located in a nonessential region of the DNA outside of both the ampicillin and tetracycline resistance genes (Davis et al., 1980). mtDNA was inserted into this site by the following method.

EcoRI fragments from a limit digest of mtDNA (4 μ g in 107 μ l) were mixed with EcoRI digested pBR322 (0.19 μ g in 5 μ l), 10 μ l of 10 X ligase buffer (0.5 M Tris-HCl, pH 7.8, 0.1 M MgCl₂, 0.2 M DTT, 0.5 mg/ml BSA), and 10 μ l of 10 mM ATP. To start the ligation reaction, 40 units of T4 ligase (New England Biolabs) was added, and the mixture was incubated at 15°C for 2.5 hours. To encourage circularization of recombinant DNA molecules, the reaction was diluted by adding 400 μ l of 1 X ligation buffer and incubating for an additional 16 hours at 15°C (Dugaiczky et al., 1975).

E. coli HB101 was made competent for transformation by CaCl₂ treatment (Davis et al., 1980). The cells were grown to 0.6 A₆₀₀ units in L broth (0.5% NaCl, 0.5% Bacto-yeast extract, 1% Bacto-tryptone) and centrifuged at 5000 rpm for 5 minutes at 4°C. The collected cells were resuspended in 0.5 volume of sterile 0.05 M CaCl₂ and incubated on ice for 30 minutes. They were then centrifuged again and gently resuspended in 0.25 volume (relative to the original culture volume) of 0.5 M CaCl₂. To transform the cells, each 0.1 ml of ligated DNA was mixed with 0.2 ml of competent cells, incubated at 0°C for 45 minutes, and then incubated at 37°C for 10 minutes. The cells were plated directly on L-agar plates containing 10 μ g/ml tetracycline and 25 μ g/ml ampicillin and incubated at 37°C for about 36 hours.

Screening of Transformants

Colony hybridization. To identify transformants containing mtDNA sequences, a modification of the method of Grunstein and Hogness (1975) was used. Transformed colonies were randomly picked and transferred to 150 mm L-agar plates containing 20 µg/ml tetracycline and 50 µg/ml ampicillin. A streak about 1 cm long was made with each transferred colony, and about 50 colonies were arranged in a grid pattern on each plate. The plates were incubated at 37°C overnight. The cells were then replica plated onto sterile, 127 mm diameter, nitrocellulose filters (Millipore, HATF) which had been placed on a fresh L-agar plate containing ampicillin and tetracycline. [The filters were prepared by soaking in H₂O, placing between dry Whatman 3 MM filter paper, wrapping in aluminum foil, and autoclaving. Prior to replica plating, each filter was laid on a plate, peeled off, inverted, and again placed on the plate.] The cells were incubated at 37°C for about 12 hours to allow the cells to grow on the nitrocellulose. To amplify the plasmids in the cells, the nitrocellulose filters were then transferred to L-agar plates containing 500 µg/ml chloramphenicol and incubated at 37°C for about 16 hours (Hanahan and Meselson, 1980).

To lyse the cells, the filters were placed on Whatman 3 MM paper saturated with 0.5 M NaOH for about 5 minutes. The filters were then removed, blotted briefly on dry 3 MM paper, and the process repeated--once with 0.5 M NaOH, twice with 1 M Tris, pH 8.0, and once with 1.5 M NaCl, 1 M Tris, pH 8.0. The filters were then blotted dry and baked at 60°C for about 2 hours. Presumptive mtDNA containing

colonies were identified by hybridization with ^{32}P -mtDNA using the methods described above.

Minipreparation of plasmid DNA. In order to confirm the presence of mtDNA, plasmid DNA was extracted from presumptive colonies for restriction endonuclease analysis. For this purpose, the presumptive colonies were grown in 10 ml cultures of M9 medium (Norgard, 1981). Each liter of M9 medium contained the following sterile components: 100 ml of 10 X M9 salts (see below), 822 ml of H_2O , 10 ml of 0.1 M MgSO_4 , 10 ml of 0.01 M CaCl_2 , 25 ml of 20% (w/v) casamino acids (Difco), 20 ml of 20% glucose, 2 ml of 1 mg/ml thiamine (filter sterilized), 10 ml of 100 mg/ml uridine (filter sterilized), 1 ml of 10 mg/ml tetracycline, and 2.5 ml of 10 mg/ml ampicillin. The 10 X stock solution of M9 salts contained (per liter): 10 g NH_4Cl , 60 g Na_2HPO_4 , 30 g KH_2PO_4 , and 5 g NaCl . Uridine was added to promote plasmid amplification (Norgard, 1979). Tetracycline and ampicillin stock solutions were prepared using 100% methanol, and no further precautions were taken to sterilize these antibiotic solutions.

Plasmid DNA was isolated by the rapid, SDS-KAc method (Davis et al., 1980) from 9 ml of each culture after overnight incubation at 37°C . (The remaining 1 ml of each culture was refrigerated for subsequent use.) Following this method, the cells were collected by centrifugation, resuspended in 1.4 ml of 10 mM Tris, 1 mM EDTA, pH 8.5, and microfuged for 30 seconds. The cells were then resuspended in 0.5 ml of 50 mM Tris, pH 8.5, 50 mM EDTA, 15% sucrose, containing 1 mg/ml lysozyme (added just prior to use) and allowed to stand for 10 minutes at room temperature. To each tube, 1 μl of diethylpyrocarbonate (DEPC)

and 10 μ l of 10% SDS were added. After mixing, 50 μ l of 5M potassium acetate (KAc) was added, and the tubes were incubated on ice for 30 minutes. The K-dodecyl sulfate-protein precipitate was then sedimented by centrifugation in a microfuge for 15 minutes, and the supernatant was transferred to a clean microfuge tube. The nucleic acids in the supernatant were precipitated by adding about 1 ml of 95% ethanol and were collected by centrifugation in a microfuge for 5 minutes. The pellet was dissolved in 150 μ l of TE buffer. To purify the plasmid DNA further, 25 μ l of 5 M Kac and 25 μ l of 4 ml/ml yeast RNA were added to each tube; after mixing 300 μ l of 95% ethanol was added. The precipitate was collected by centrifugation for 10 minutes and resuspended in 25 μ l of 10 mM Tris, pH 7.4, 1 mM EDTA containing 10 μ g/ml RNase A.

In order to determine which EcoRI fragment of mtDNA was present in each recombinant DNA plasmid, 10 μ l of each preparation of plasmid was mixed with 10 μ l of 2 X EcoRI buffer and digested with restriction endonuclease. The digests were then electrophoresed in agarose gels using EcoRI digested mtDNA as marker. Those strains that were confirmed as containing mtDNA were grown in large batches (starting from the 1 ml, refrigerated samples) for the isolation of large amounts of each recombinant DNA plasmid.

Large Scale Purification of Recombinant DNA Plasmids

Recombinant pBR322-mitochondrial DNA plasmids were isolated free from chromosomal DNA contamination by a modification of the method of Haseltine et al. (1980). Transformed HB101 cells containing the plasmid to be purified were grown in 1 liter cultures of M9 medium (see above) at 37°C with vigorous shaking. When the cells reached a

density of $0.7 A_{550}$ units, the plasmids were amplified by adding 300 mg of chloramphenicol (in 2 ml of methanol) and incubating for an additional 16 hours. The cells were then harvested by centrifugation at 6000 rpm for 10 minutes, resuspended in 200 ml of 10 mM Tris, pH 8.5, 1 mM EDTA, and again collected by centrifugation. The cells were resuspended in 10 ml of 25% sucrose, 50 mM Tris, pH 8.0, 40 mM EDTA, placed in a 120 ml erlynmeyer flask, and kept on ice for the subsequent procedures. After all clumps of cells had been dispersed, 3 ml of 50 mM Tris, pH 8.0, 40 mM EDTA containing 10 mg/ml of lysozyme (freshly prepared) was added, and the suspension was left to stand on ice for 5 minutes with occasional gentle swirling. Three ml of 0.5 M EDTA, pH 8.0 was then added to the flask, and the suspension was again left to stand on ice for 5 minutes. To the flask, 20 ml of 0.1% Triton X-100, 50 mM Tris, pH 8.5, 50 mM EDTA was added and quickly swirled to mix. At this point, the mixture became viscous but not clear. After incubation on ice for 10 minutes, the chromosomal DNA was pelleted by centrifugation at 27,000 rpm and 4°C for 30 minutes in a Beckman SW28 rotor. The plasmid-containing supernatant was removed, and adjusted to 27 ml using 50 mM Tris, pH 8.0, 40 mM EDTA. To isolate supercoiled plasmid DNA, 26.5 g of CsCl and 2.7 ml of 10 mg/ml ethidium bromide were added (refractive index about 1.393) and the solution was centrifuged at 40,000 rpm for 48 hours. The (lower) plasmid band was removed by puncturing the side of the tube with a 22 gauge needle and syringe. The ethidium bromide was extracted, and the DNA was precipitated (as discussed above).

CHAPTER VII

RESULTS AND DISCUSSION

Isolation of Mitochondrial DNA

The method for isolating mitochondria from Candida albicans cells is outlined in Figure 2-1 (Chapter VI, page 114). The growth medium contained lactic acid, a nonfermentable carbon source, in order to encourage aerobic respiration and the synthesis of mitochondrial DNA (mtDNA). Lactic acid was chosen because the majority (if not all) of wildtype C. albicans strains can assimilate it (Syverson, 1981). In Saccharomyces cerevisiae, another facultative yeast, the synthesis of mitochondrial enzymes and mtDNA varies with the stage of growth and the culture conditions. When grown anaerobically using fermentable carbon sources (e.g. glucose), mitochondrial activity is found to be greatly repressed (Schatz and Kovac, 1974; Lewin, 1980). Preliminary experiments with C. albicans cells grown in glucose containing medium resulted in lower yields of mitochondria (but faster growth) than when lactic acid was used. The reason for the higher yield of mitochondria from lactic acid cultures was not investigated but was presumably due to increased numbers of mitochondria per cell.

Evidence that mitochondria were indeed isolated from C. albicans cells was provided by two observations. First, microscopic examination of the purified material revealed particles which had the appearance (e.g. size) of mitochondria. Only a very few nuclei could be seen in the purified material. Second, the purified material formed bands in

sucrose gradients at the position expected for mitochondria. When fractions from the gradients were assayed for succinic dehydrogenase activity (a mitochondria specific enzyme marker, Merkel et al., 1980), the peak of enzymatic activity was found to correspond to the position of the bands (B.A. Lasker, personal communication). Thus, the DNA isolated from the purified material is almost certainly mtDNA.

The most critical step in the mitochondrial DNA isolation procedure was found to be the DNase I treatment step. Although very few nuclei could be detected in all the mitochondria preparations, large amounts of nuclear DNA were found to be isolated without the DNase I treatment. Apparently, nuclear DNA binds in large amounts to the outside of the mitochondria. Because the mitochondria are impermeable to DNase I, nuclear DNA (but not mtDNA) is removed by the treatment (Zeman and Lusena, 1975). The evidence for this is presented in Figure 2-2. Lane a of this figure shows the electrophoretic pattern of EcoRI digested mtDNA isolated from DNaseI treated mitochondria. Six distinct bands can be seen. (As will be discussed below, the upper most band contains partially digested fragments.) Lane d of the figure shows the electrophoretic pattern of EcoRI digested whole cell DNA. The three prominent nuclear repetitive bands (discussed in Part I) are clearly visible, but the mtDNA bands are difficult to identify due to the abundance of nuclear DNA bands. Lanes b and c contained EcoRI digests of different mtDNA preparations isolated without DNaseI treatment. Both of these lanes clearly contained large amounts of contaminating nuclear DNA. The three repetitive bands of nuclear DNA are quite distinct as is a smear of bands having intermediate sizes which

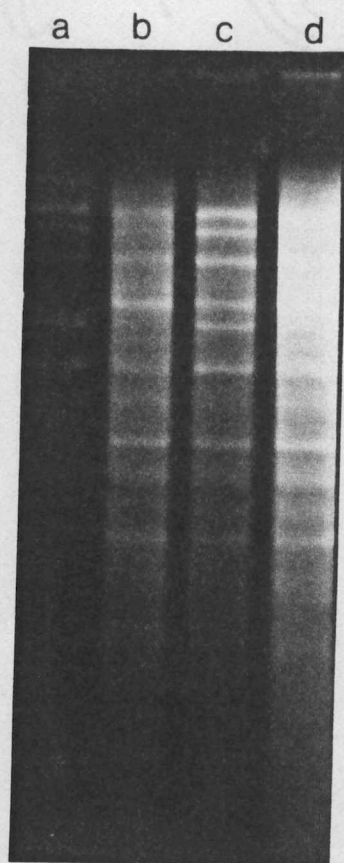


Figure 2-2. Purity of mtDNA (lane a) as demonstrated by EcoRI digestion and agarose gel electrophoresis. For comparison, enriched mtDNAs (isolated without the use of DNase I) are shown (lanes b and c) as is whole cell DNA (lane d).

together constitute a background of fluorescence.

A noticeable difference can be seen between lanes b and c of Figure 2-2. The mtDNA bands are much more intense in lane c, although the total amount of DNA was the same in both lanes. Little significance can be placed upon this observation since the two preparations of enriched mtDNA were isolated from different batches of cells, and care was not taken to handle the cells in an identical manner. It is possible, however, that the culture conditions were responsible for the different amounts of mtDNA in the two preparations. The DNA in lane b was isolated from cells grown in YEPL with good aeration (vigorous bubbling) for two days (8.5 grams of cells per liter). The DNA in lane c was isolated from cells grown in the same medium with poor aeration (stirring) for five days (2.5 grams of cells per liter). An alternative explanation is simply that different amounts of nuclear DNA were purified away from the mtDNA in the two preparations.

DNase I treatment of mitochondria is not required for the isolation of mitochondrial DNA from many organisms. In many cases, mtDNA can be easily separated from nuclear DNA on the basis of (1) differences in G+C content or (2) differences in structure. As an example of the first case, the mtDNA of Paramecium aurelia, which is a naturally occurring linear molecule, can be isolated in CsCl gradients due to the different densities of mtDNA and nuclear DNA (Cummings, et al., 1980). As an example of the second case, the small, circular mtDNA molecules of vertebrates (even those with base compositions similar to nuclear DNA) can be isolated in CsCl gradients containing dyes (e.g. ethidium bromide) which preferentially bind to supercoiled molecules and

increase their density relative to the nuclear DNA (e.g. hamster mtDNA; Nass, 1980). Supercoiled mtDNA from plant cells has also been purified by this method (Dale et al., 1981). As a final example, Saccharomyces cerevisiae mtDNA, which is a naturally occurring circular molecule, can be purified on the basis of buoyant density differences in CsCl (Lewin et al., 1978), but supercoiled molecules have not been isolated due to degradation problems (Locker et al., 1979).

Candida albicans mtDNA could not be purified by either of the two above methods. The density of mtDNA in CsCl gradients was found to be very close to that of nuclear DNA (see below). Likewise, when enriched mtDNA (i.e. isolated without the use of DNase I) was centrifuged in CsCl gradients containing ethidium bromide, only one band was observed. This band was found to contain both nuclear and mitochondrial DNA when analyzed with restriction endonucleases (data not shown). This suggests that the purified C. albicans mtDNA was composed of linear molecules. Since the restriction map of the mtDNA was found to be circular, it seems likely that degradation occurred during the isolation procedure. Additional evidence for the degradation of C. albicans mtDNA was provided by agarose gel electrophoresis. In gels of low agarose concentration, the mtDNA appeared as a smear as would be expected if linear molecules of many lengths were present. Such an observation has also been made with Neurospora crassa mtDNA which has not been isolated without degradation (Mannella et al., 1979) and with some plants mtDNAs (Dale et al., 1981).

Other reasonable explanations can also be proposed for the observed smear of C. albicans mtDNA on agarose gels.

1. Contamination with nuclear DNA fragments. When the smear of DNA was transferred to nitocellulose and hybridized with ^{32}P -labeled cloned fragments of mitochondrial DNA (characterized below), autoradiography also revealed a smear (data not shown). This argues that the smear contained only mtDNA; however, the same result might be expected if the smear contained nuclear DNA, and the probe contained sequences homologous to nuclear DNA.
2. Variable degrees of twisting of supercoiled mtDNA. This possibility does not seem likely because supercoiled molecules (as mentioned above) were not observed in CsCl-EtBr gradients. Also, mild treatment of the mtDNA with DNaseI to relax supercoiled molecules had no effect upon the smear (B. A. Lasker, personal communication).
3. Contamination with replicating mtDNA. This also seems unlikely since the smear contained fragments smaller than the size of non-replicating mtDNA (41 kbp; see below). No experiments were performed to test this hypothesis.
4. Size heterogeneity of mtDNA molecules. No evidence in favor or in disagreement with this possibility has been found. Precedence exists, however, with the heterogeneity of plant mtDNA and the sequence rearrangements of many fungi mtDNAs (see Chapter V).

Thus, at this time the nature of the smear obtained with C. albicans mtDNA has not been determined.

Base Composition of Mitochondrial DNA

As mentioned above, many mitochondrial DNAs have G+C contents that are significantly different from the nuclear DNA. Previous experiments using whole cell C. albicans DNA revealed the existence of two density components, a main band of 32.6% G+C and a satellite band of 40.0% (see Part I). Buoyant density measurements using purified mtDNA showed that its G+C content is very similar to those two components, Figure 2-3). The results of three measurements using two different mtDNA preparation were essentially identical (38.5, 38.2, and 37.9) with an average of 38.2% G+C. Thus, C. albicans mtDNA should be rather difficult to isolate from nuclear DNA on the basis of buoyant density.

As a side light, mtDNA, whole cell yeast DNA, and whole cell mycelial DNA were examined for methylation at CCGG sites. Conversion of C. albicans cells from yeast phase to mycelial phase was accomplished in amino acid synthetic medium (Lee et al., 1978) using the protocol of Torres-Bauza (1980). To detect methylation, portions of these three DNA preparations were digested with HpaII, and other portions were digested with MspI. These two enzymes are isoschizomers. HpaII does not cut at methylated recognition sites, but MspI (which recognizes the same CCGG site) is insensitive to methylation. The results of these digestions are presented in Figure 2-4. As can be seen, the electrophoretic patterns appear identical with both enzymes for all three DNAs. This result shows that methylation at this site occurs at most at low levels. It would be interesting to pursue the methylation question by transferring the DNA to nitrocellulose and using specific gene probes (e.g. labeled rRNA) to look for differences in methylation.

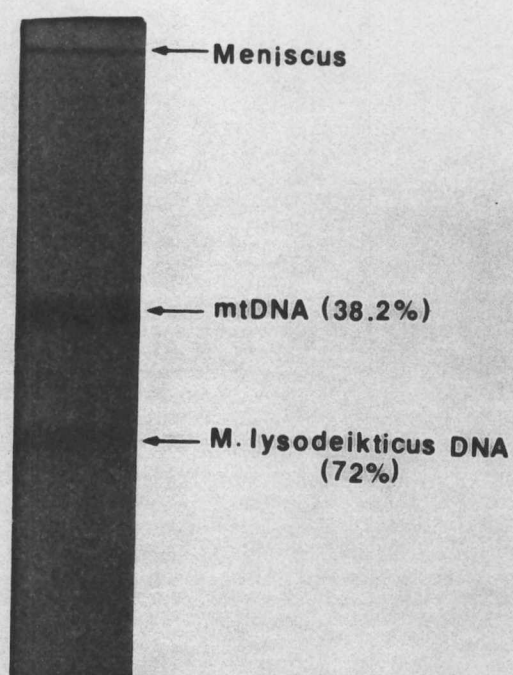


Figure 2-3. Measurement of base composition of purified mtDNA by analytical CsCl density gradient centrifugation. The value for the G + C content of mtDNA is based upon the average of three measurements.

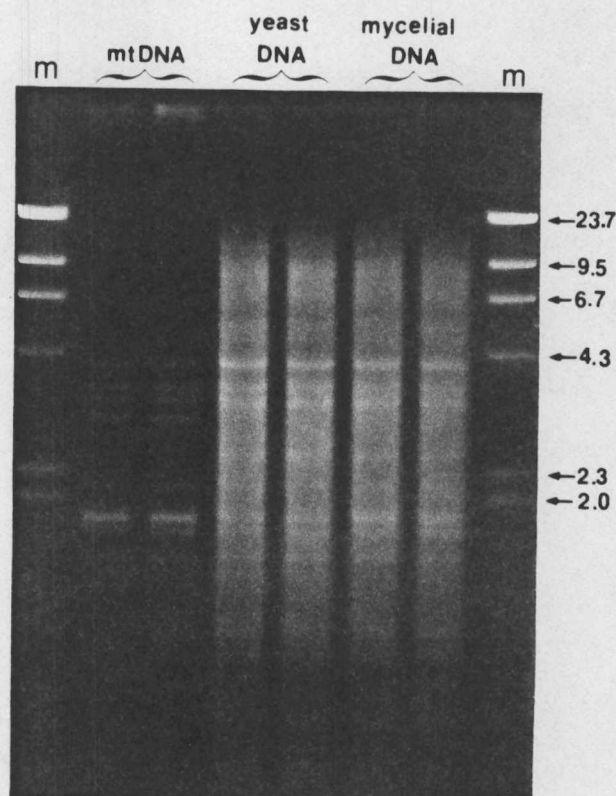


Figure 2-4. Digestion of *C. albicans* DNAs with the isoschizomers, HpaII and MspI. mtDNA was isolated from yeast phase cells. Whole cell yeast and mycelial DNA were isolated from ASM cultures. The left lane for each DNA contained the HpaII digest and the right lane contained the MspI digest. Hind III digested Lambda DNA was used as marker (lanes labeled m).

Single Digests of mtDNA

Preliminary studies of the digestion of Candida albicans mtDNA by single restriction endonucleases showed that EcoRI and PvuII should be useful enzymes for mapping. Each produced six well separated electrophoretic bands on agarose gels when digestion was allowed to proceed to completion. These bands are designated (starting with the largest) E1, E2, . . . E6 for the EcoRI digest and P1, P2, . . . , P6 for the PvuII digest.

Complete digestions were always more difficult to achieve with EcoRI than with PvuII. The reasons for this difference were not explored, but this was found to be the case with all of the mtDNA preparations. Figure 2-5 shows an example of a partial EcoRI digest (lane c). The largest visible band contains partial digestion fragments which disappear with extended digestion (compare with Figure 2-7, lane a, page 142). This particular partial fragment was the hardest to digest, presumably due to the sequences flanking the EcoRI site. The length of this partial digestion fragment (18.5 kbp) is equal to the sum of the lengths of fragments E2 and E3 (see Table 2-1) suggesting that these two fragments are adjacent on the map. This interpretation is supported by the fact that bands E2 and E3 in Figure 2-5c are of lower intensity than expected from the complete digestion pattern displayed in Figure 2-7a. Data presented below further confirm that E2 and E3 are adjacent.

The smallest EcoRI fragment, E6, is only 0.34 kbp in length and was therefore not usually seen in the agarose gels. Information concerning this fragment was obtained by electrophoresis in polyacrylamide

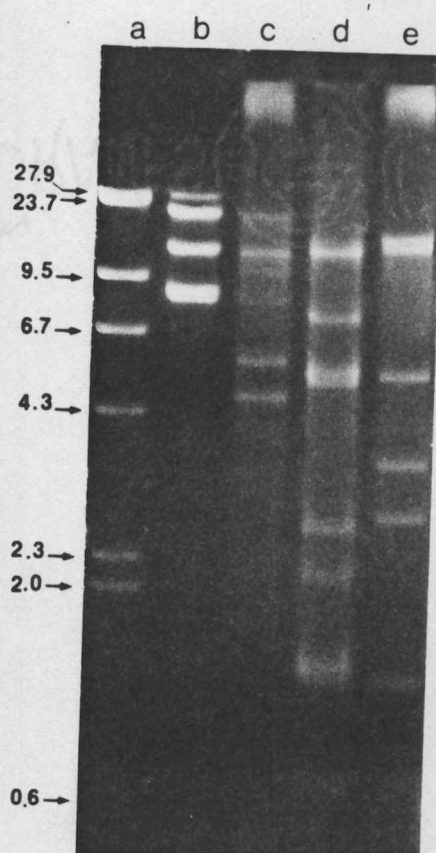


Figure 2-5. PvuII and EcoRI single and double digests of mtDNA analyzed by agarose gel electrophoresis. Lane contents: HindIII digested Lambda DNA markers (lane a), SmaI digested Lambda DNA markers (lane b), EcoRI digested mtDNA (lane c), EcoRI/PvuII digested mtDNA (lane d), and PvuII digested mtDNA (lane e).

gels (Figure 2-6, lane c). The sizes of the six fragments produced by complete EcoRI digestion are listed in Table 2-1 and their sum is 41.0 kbp.

Complete digestion by PvuII was routinely achieved (Figure 2-5, lane e). The band containing the smallest PvuII fragments, P6, is more intense than would be expected if these fragments were in equal molar concentration with the other PvuII fragments (see also, Figure 2-6, lane f). The fragments in this band were subsequently found to map to two distinct locations in the mtDNA; therefore the two fragments are designated P6a and P6b. As shown below, P6a and P6b appear to have very similar, if not identical, sequences. The sizes of the seven fragments produced by PvuII digestion are listed in Table 2-1; their sum, 41.6, is very close to that obtained from the EcoRI fragments.

Double Digestion of mtDNA

When mtDNA is digested with both EcoRI and PvuII, eleven bands are resolved by gel electrophoresis. These bands are designated EP1, EP2, . . . , EP11. Figure 2-5 (lane d) shows the results of one such double digest. In this figure, the three smallest fragments (EP9, EP10, and EP11) are not visible, and information concerning them was obtained from polyacrylamide gel electrophoresis (Figure 2-6, lane d). Bands EP7 and EP8 have intensities greater than would be expected if the fragments in these bands were in equal molar concentration with the other fragments. These two bands did not change with extended periods of digestion and therefore do not contain partial digestion products. Bands EP7 and EP8 were found to be "double bands." As shown below, the fragments in EP7 (designated EP7a and EP7b) have similar, if not

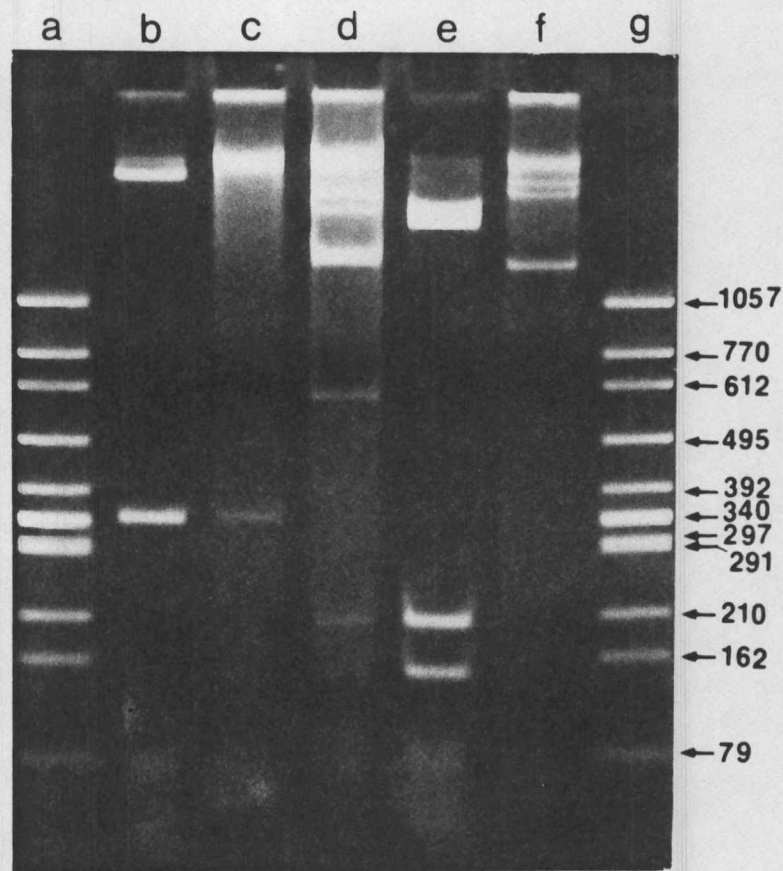


Figure 2-6. Polyacrylamide gel electrophoresis of single and double digests of mtDNA and cloned E6. Lane contents: HincII digested PhiX174 DNA markers (lanes a and g), EcoRI digested cloned E6 containing plasmid (lane b), EcoRI digested mtDNA (lane c), EcoRI/PvuII digested mtDNA (lane d), EcoRI/PvuII digested cloned E6 containing plasmid (lane e), and PvuII digested mtDNA (lane f).

Table 2-1. Fragment Sizes Produced by EcoRI and PvuII Digestion of mtDNA.

EcoRI		EcoRI and PvuII		PvuII	
Fragment	kbp	Fragment	kbp	Fragment	kbp
E1	12.2	EP1	12.2	P1	14.8
E2	10.0	EP2	7.4	P2	13.0
E3	8.1	EP3	5.4	P3	5.2
E4	5.7	EP4	5.1	P4	3.4
E5	4.7	EP5	2.6	P5	2.7
E6	<u>0.34</u>	EP6	2.1	P6a	1.27
		EP7a	1.36	P6b	<u>1.27</u>
		EP7b	1.36		
		EP8a	1.28		
		EP8b	1.28		
		EP9	0.60		
		EP10	0.20		
		EP11	<u>0.15</u>		
total = 41.01		total = 41.03		total = 41.64	

identical, sequences; the same is true of the fragments in EP8 (EP8a and EP8b). The lengths of the fragments contained in each of the double digest bands are listed in Table 2-1. The total length of the mtDNA obtained from the sum of all thirteen fragments is 41.0 kbp, which is the same as those obtained from the EcoRI and PvuII single digests. This further supports the contention that bands EP7 and EP8 are double bands.

Comparison of EcoRI-PvuII double digests with EcoRI single digests of mtDNA (Figure 2-5) showed that the fragments in bands E2, E3, E4, E5, and E6 all have one or more PvuII sites since these bands are not present in the double digest. E1 is the only fragment for which there are fragments of the same size (EP1) in the double digest. Since there are no EcoRI fragments larger than E1 from which EP1 could be derived, it can be concluded that E1 has no PvuII site.

Comparison of EcoRI-PvuII double digests with PvuII single digests of mtDNA (Figure 2-5) showed that the fragments in bands P1, P2, P3, P4, and P5 all have one or more EcoRI sites since bands of these sizes are not present in the double digest. P6 is the only PvuII band for which there is a comigrating double digestion band (EP8). This suggests that P6a and P6b do not contain EcoRI sites. Consistent with this hypothesis is the fact that the sum of the lengths of the fragments in bands EP9, EP10, and EP11 (which are all single bands) is only 0.98 kbp. If either (or both) P6a and P6b were digested by EcoRI, additional small bands would have been expected. To test the hypothesis further, the DNA fragments in bands P4 and P6 were extracted from a gel and digested with EcoRI. The

control (P4) was digested to give the two expected fragments, but P6 was not digested (data not shown).

Cloning of mtDNA

To aid in the restriction mapping, EcoRI fragments of the mtDNA were cloned. EcoRI fragments were chosen for cloning (in preference to PvuII fragments) for two reasons. First, each of the EcoRI fragments has a unique length. Had PvuII fragments been used, it would have been very difficult, if not impossible, to distinguish between P6a and P6b. Second, PvuII digestion produces two large fragments (P1 and P2) which would be difficult to insert into the plasmid, pBR322. All of the EcoRI fragments are smaller than the two largest PvuII fragments.

All of the EcoRI fragments were easily cloned in pBR322 with the exception of the largest, E1, which has a length of 12.2 kbp. Figure 2-7 show the results of EcoRI digestion and agarose gel electrophoresis of the isolated recombinant DNA plasmids (lanes b-f). Each of these lanes contains a common pBR322 band and a band which comigrates with one of the EcoRI fragments from mtDNA (lane a). (Results concerning cloned E6 were obtained from polyacrylamide electrophoresis, Figure 2-7, lane b.) These results suggest that the cloned EcoRI fragments are identical to the EcoRI fragments from mtDNA, and further characterization (below) supported this contention.

Experiments to clone fragment E1 were not successful, but only a few colonies were screened. For mapping purposes, however, it was not necessary to have this fragment cloned. All information concerning restriction sites in E1 can be deduced simply by comparing restriction

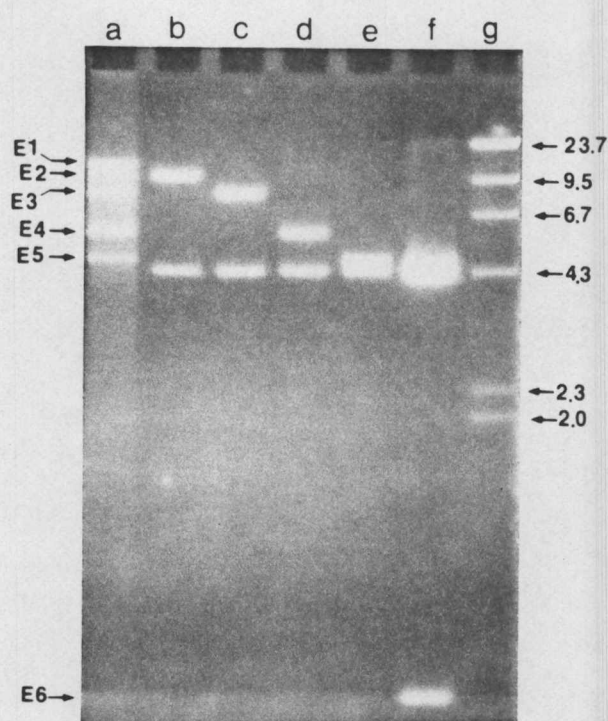


Figure 2-7. EcoRI digestion of mtDNA containing, recombinant DNA plasmids from maxipreps. Lane contents: EcoRI digested mtDNA (lane a), EcoRI digested plasmids containing fragments E2, E3, E4, E5, and E6 (lanes b, c, d, e, and f, respectively), and HindIII digested Lambda DNA markers (lane g).

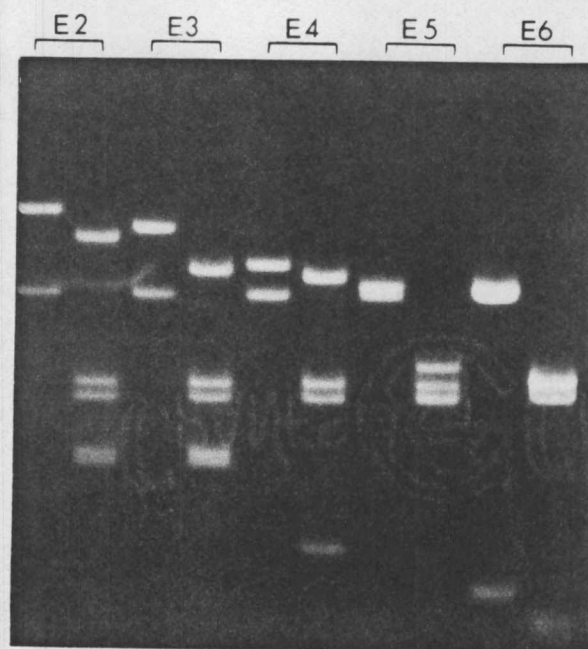
digests of intact mtDNA with digestions of the other cloned fragments. In the case of PvuII sites, the analysis is further simplified by the fact that E1 contains no PvuII sites.

Mapping the PvuII Sites of the Cloned EcoRI Fragments

To characterize further the cloned EcoRI fragments, the recombinant DNA plasmids were digested with EcoRI (to separate mtDNA from pBR322) and PvuII. The results of agarose gel electrophoresis of these digests are presented in Figure 2-8a. In this figure, an EcoRI single digest and an EcoRI-PvuII double digest are presented for each of the recombinant plasmids (left and right lanes, respectively). Since pBR322 contains only one PvuII site, the double digests produce two easily identified plasmid bands which are common to the right lane for each recombinant DNA. The figure clearly shows that each of the cloned EcoRI fragments (E2-E6) contains one or two PvuII sites, a result expected from the double digest of mtDNA (discussed above). (Figure 2-6, lane e, page 138, clearly shows the bands produced by cloned E6 in a polyacrylamide gel.) Only one "double band" is present in Figure 2-8a. Cloned E5 produces two fragments when digested with PvuII, one of which comigrates with the small pBR322 fragment. Thus, when the plasmid containing E5 was digested with EcoRI and PvuII, the band containing the smallest fragments was found to be more intense than either of the two bands containing the large fragments. In all, twelve fragments are produced by digesting the five cloned EcoRI fragments with PvuII. The lengths of these fragments are listed in Table 2-2. For each of these fragments there is a corresponding fragment of the same size produced by double digestion of the mtDNA (compare with Table 2-1).

Figure 2-8a also shows that cloned fragments E2 and E3 both

A.



B.

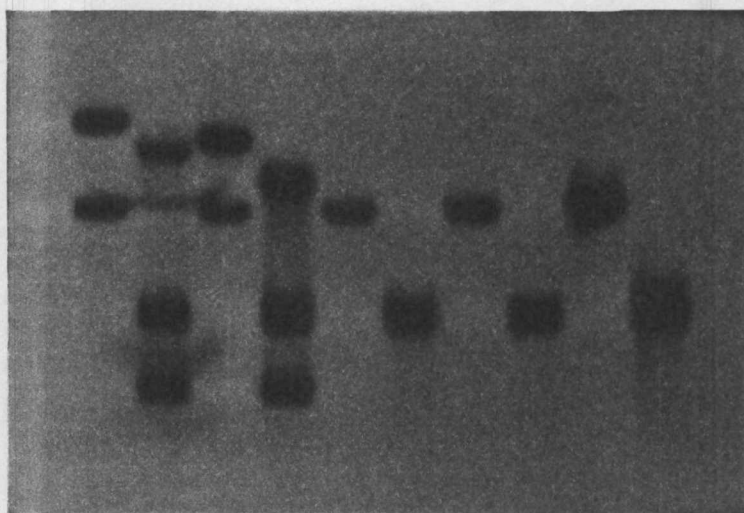


Figure 2-8. EcoRI-PvuII double digestions of recombinant DNA plasmids containing mtDNA. A. Recombinant DNA plasmids containing mtDNA fragments E2, E3, E4, E5, or E6 were single digested with EcoRI (left lane for each plasmid) and double digested with EcoRI and PvuII (right lane for each plasmid). B. Hybridization pattern obtained when the gel represented in part A was blotted and hybridized with ^{32}P -E3 containing plasmids.

produce a 1.35 and a 1.25 kbp fragment when digested by PvuII. This result is not surprising in view of the double bands of the same sizes found in the double digest of mtDNA (EP7 and EP8). When the DNA in the gel was transferred to nitrocellulose and hybridized with radioactive plasmid DNA containing E3, it was found that all the bands from both E2 and E3 hybridized with the probe (Figure 2-8b). This result indicates that the 1.35 kbp fragments from E2 and E3 are similar, if not identical; the same is true of the 1.25 kbp fragments. To explore the possibility that cloned E3 was an artifact derived by a deletion from E2, EcoRI-AvaI and EcoRI-KpnI double digests of cloned E2, E3 and mtDNA were compared (data not shown). In both cases, each of the fragments produced from E2 and E3 was of the same size as a fragment produced from the mtDNA. From these and the results above, it can be concluded that none of the cloned EcoRI fragments has been altered by the cloning procedure.

Mapping of the PvuII sites within each of the EcoRI fragments was trivial for E4, E5, and E6 since each of these yield only two fragments when digested by PvuII (Figure 2-8). Mapping of the PvuII sites within E2 and E3 (both of which yield three fragments when digested) was also straightforward. The maps of E2 and E3 follow from the fact that P6a and P6b (produced by PvuII digestion) contain no EcoRI sites (see above). For this reason, P6a and P6b are identical to EP8a and EP8b in the mtDNA double digest (Table 2-1, page 139). As shown in Table 2-2, EP8a and EP8b are derived from E2 and E3, respectively. Since EP8a has PvuII recognition sequences at both ends, this fragment must lie internal to the other two (larger) fragments produced by PvuII

Table 2-2. Sizes of Fragments Produced by PvuII Digesting the Cloned EcoRI Fragments.

Cloned EcoRI Fragment	EcoRI-PvuII Fragment Sizes	Corresponding mtDNA EcoRI-PvuII Fragments
E2	7.4 kbp 1.35 1.25 total = 10.0	EP2 EP7a EP8a
E3	5.4 1.35 1.25 total = 8.0	EP3 EP7b EP8b
E4	5.1 0.52 total = 5.62	EP4 EP9
E5	2.55 2.10 total = 4.65	EP5 EP6
E6	0.20 0.15 total = 0.35	EP10 EP11

digestion of E2 (which contain EcoRI sequences at one end and PvuII sequences at their other end). By a similar argument, the smallest fragment produced by PvuII digestion of E3 (EP8b) must be between the two larger fragments. The locations of all the PvuII sites in the EcoRI fragments are diagrammed in Figure 2-9.

Determining the Order and Orientation of the EcoRI Fragments

Having mapped the PvuII sites within the six EcoRI fragments, I next sought to deduce the arrangement of these fragments in the intact mtDNA. To explain how the restriction map of mtDNA (presented in Figure 2-12, page 158) was constructed, I arbitrarily begin the argument with fragment E4.

Hybridization of EcoRI fragments to PvuII fragments. When EcoRI-digested mtDNA was electrophoresed in agarose gels, transferred to nitrocellulose paper, and hybridized with ^{32}P -E4 DNA, the band containing E4 fragments was detected by autoradiography (Figure 2-10D, left lane). This expected result simply confirms that the cloned E4 DNA hybridizes with the E4 fragments from mtDNA. On the other hand, when PvuII-digested mtDNA on nitrocellulose was hybridized with ^{32}P -E4 DNA (right lane), bands P1 and P3 became radioactive indicating that E4 overlaps these two fragments (i.e. P1 and P3 are adjacent). Since the P3 band gave a very strong signal, it can be concluded that the 5.1 kbp PvuII fragment of E4 (Figure 2-9) lies within P3 and the 0.52 kbp PvuII fragment lies within P1.

We now move along the map in the direction of P3. P3 has a length of 5.2 kbp. This means that the next EcoRI fragment in this direction must have a PvuII site approximately 0.1 kbp from one end

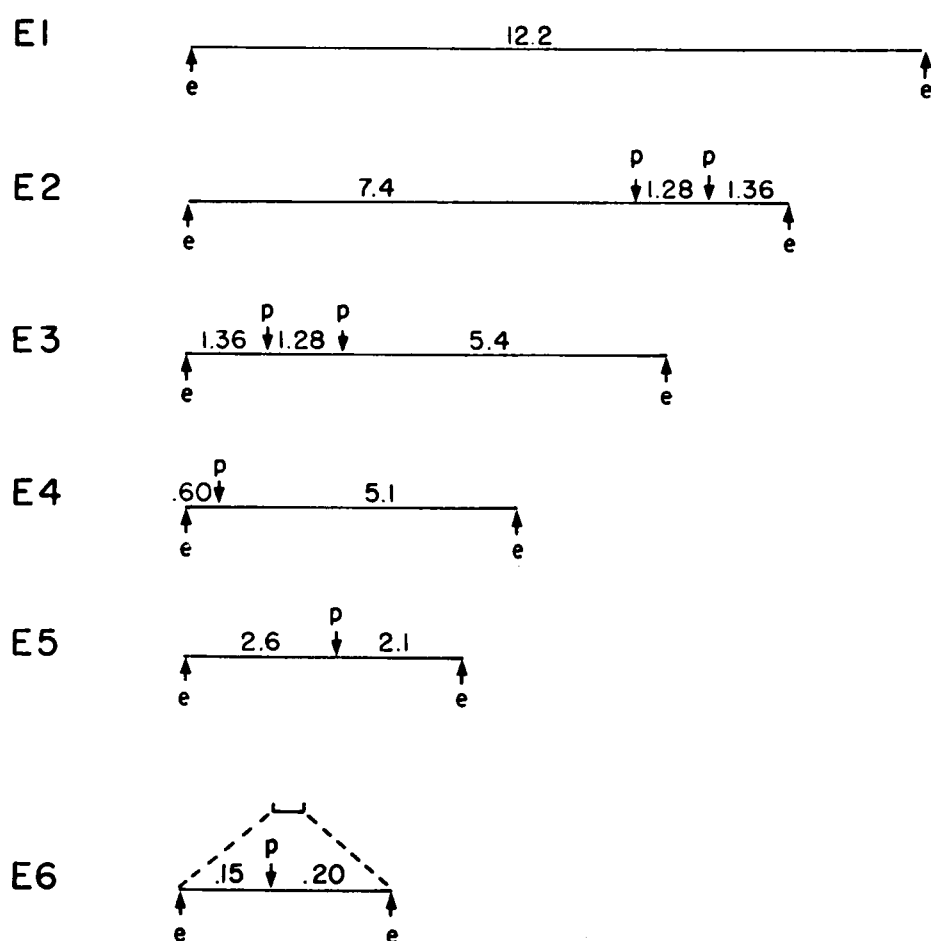


Figure 2-9. Location of PvuII sites of cloned EcoRI fragments. The arrows represent the EcoRI (e) and PvuII (p) sites. The numbers indicate the sizes of the fragments in kilobase pairs.

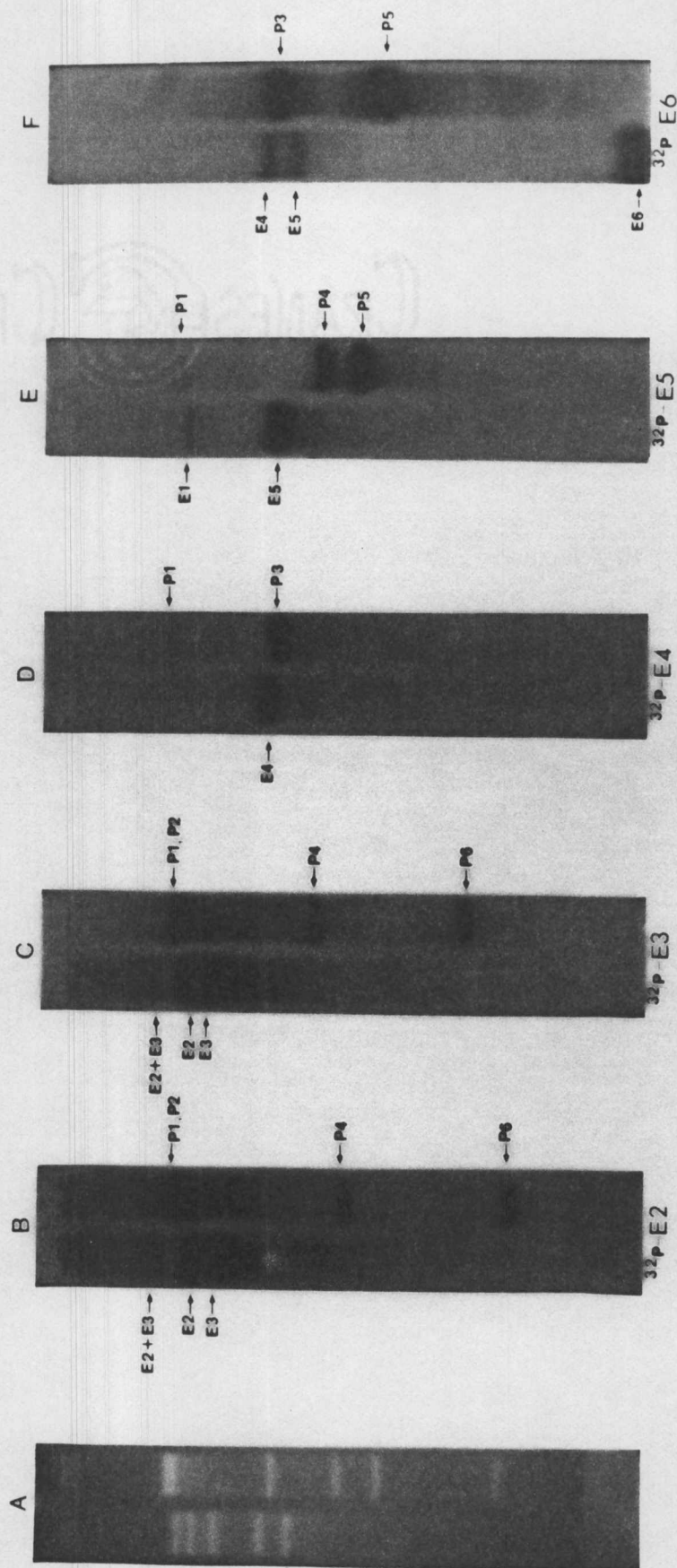


Figure 2-10. Hybridization of ^{32}P -labeled recombinant DNA plasmids to EcoRI digested and PvuII digested mtDNA. The left lane in each panel contained EcoRI digested mtDNA, and the right lane in each panel contained PvuII digested mtDNA. Panel A represents a typical electrophoretic pattern obtained using the digested mtDNA. Panels B-F show the hybridization patterns obtained when gels containing electrophoresed mtDNA were blotted and hybridized with the radioactive probes indicated below each panel.

(5.2 kbp-5.1 kbp = 0.1 kbp). E6 is the only possible choice. Because of its small size, E6 has not yet been oriented. When ^{32}P -E6 DNA was hybridized to EcoRI-digested mtDNA, band E6 became radioactive, but unexpected signals were also obtained from bands E4 and E5 (Figure 2-10F, left lane). This result establishes that E6 has sequences homologous to E4 and E5, but does not establish the order and orientation of these three fragments. However, when PvuII-digested mtDNA was hybridized with ^{32}P -E6 DNA, bands P3 and P5 became radioactive (right lane). This confirms that E6 is indeed adjacent to E4 and indicates that P5 is adjacent to P3.

Continuing to move along the map in the direction of P5, we next look for the EcoRI fragment that is adjacent to E6. Since P5 is 2.7 kbp in length, the next EcoRI fragment must have a PvuII site approximately 2.6 kbp from one end ($2.7 \text{ kbp} - 0.1 \text{ kbp} = 2.6 \text{ kbp}$). Once again, there is only one EcoRI fragment which will fit, E5. Furthermore, fragment E5 will fit in only one orientation. When ^{32}P -E5 DNA was hybridized to EcoRI-digested mtDNA, a very strong signal was obtained from band E5 (as expected) and a weak signal was obtained from E1 (unexpected). Again, this result does not establish the relative positions of E1 and E5, but establishes that these two fragments have sequences in common (Figure 2-10E, left lane). When ^{32}P -E5 DNA was hybridized to PvuII-digested mtDNA, very strong signals were obtained from bands P5 and P4, while a weak signal was obtained from band P1 (right lane). This result confirms that E5 is adjacent to E6 and suggests that P4 is adjacent to P5. I interpret the weak signal from

P1 as being due to repetitive sequences elsewhere in the mtDNA (discussed below).

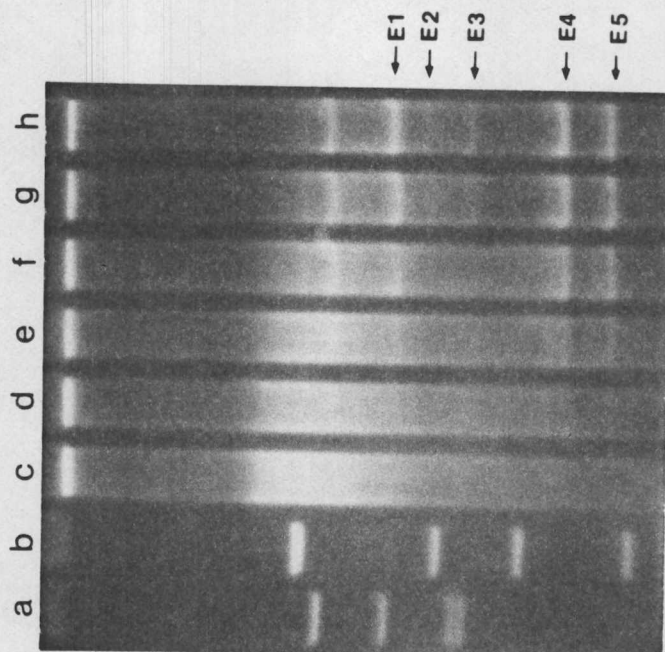
Moving along the map in the direction of P4, the EcoRI fragment adjacent to E5 must have a PvuII site approximately 1.3 kbp from one end since P4 has a length of 3.4 kbp ($3.4 \text{ kbp} - 2.1 \text{ kbp} = 1.3 \text{ kbp}$). At this point there exists two possibilities--E2 and E3. Hybridization of PvuII-digested mtDNA with ^{32}P -E2 DNA gave the same result as hybridization with ^{32}P -E3 DNA. Both radioactive probes hybridized to give strong signals from bands P2, P4, and P6 (Figures 2-10B and 2-10C). Both probes also hybridized to give weak signals from band P1 when the PvuII-digested mtDNA was electrophoresed on 0.4% agarose gels to obtain better separation of P1 and P2 (data not shown). These hybridization results are not surprising since experiments already discussed showed that E2 and E3 contain homologous sequences (Figure 2-8b). This was confirmed by hybridizing ^{32}P -E2 DNA and ^{32}P -E3 DNA to EcoRI-digested mtDNA. In both cases, bands E2 and E3 were made radioactive (Figure 2-10B and 2-10C). Since both E2 and E3 hybridize with band P4, an alternative method of ordering these two fragments was required.

Partial digestion of mtDNA with EcoRI. To determine if the EcoRI fragment adjacent to E5 is E2 or E3, the products of partial EcoRI digestions of mtDNA were examined. It was expected that either a 12.8 kbp fragment (E5 + E3) or a 14.7 kbp fragment (E5 + E2) would be found. Preliminary experiments, however, showed that most large partial digestion products were difficult to observe in EtBr stained agarose gels due to the presence of a smear of mtDNA (discussed above).

Figure 2-11a shows the results of an experiment in which samples of mtDNA were digested with EcoRI for various times and electrophoresed. This figure demonstrates the problems encountered in observing partial digest bands. A band was present in the region of the gel where 12.8 kbp fragments would migrate; however, E1 fragments (12.2 kbp) would also migrate to about the same position. A band could not be detected in the region where 14.7 kbp fragments would migrate, but a band there might have been obscured by the smear. To circumvent these problems, the DNA in the gel was transferred to nitrocellulose and hybridized with ^{32}P -E6 DNA. Figure 2-11b shows the autoradiogram obtained as a result of this experiment. Seven prominent bands were observed, and the lengths of the fragments in each are listed in Table 2-3. The band containing E6 fragments was run off the gel in order to resolve the larger fragments.

As shown above, fragments E4, E6 and E5 are adjacent on the map with E6 located in the middle. This arrangement was confirmed by the partial digestion experiment. Band 5 (5.7 kbp) and Band 7 (4.7 kbp) in the autoradiogram must contain E4 and E5 fragments, respectively, because none of the other EcoRI fragments (or combinations of fragments) have the length of the fragments in these two bands. Signals from bands E4 and E5 were expected since these fragments were shown to contain sequences homologous to E6 (Figure 2-10F). That E6 is adjacent to both E4 and E5 is shown by the presence of band 4 and band 6. Band 4 (6.1 kbp) contains fragments which could only be comprised of E4 and E6 since no other combination of EcoRI fragments gives a length of 6.1 kbp. By a similar argument, band 6 (5.0 kbp)

A.



B.

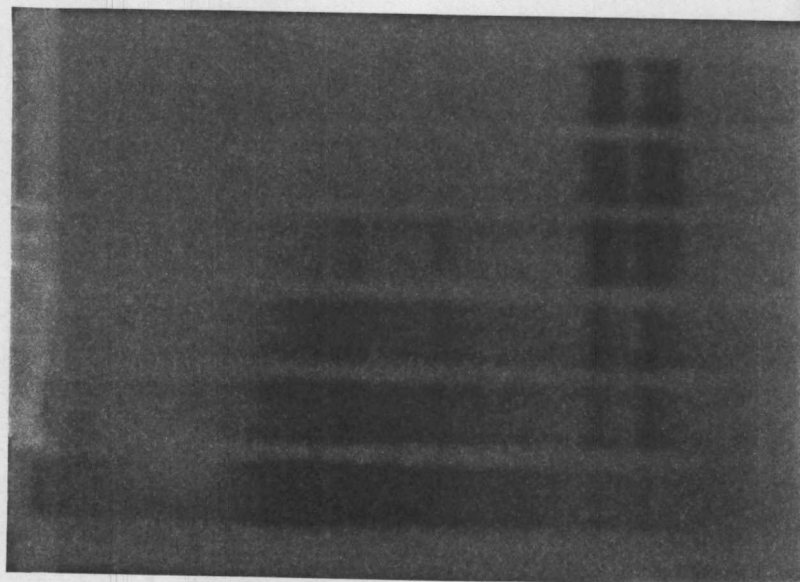


Figure 2-11.

Partial digestion of mtDNA with EcoRI. A. Electrophoretic patterns. 10 μ g of mtDNA in 100 μ l was mixed with 50 units of BRL EcoRI. 20 μ l of the reaction mixture was removed at various times for electrophoresis. Lane contents: SmaI digested Lambda markers (lane a), HindIII digested Lambda markers (lane b), mtDNA digested for 5 min (lane c), 10 min (lane d), 20 min (lane e), 45 min (lane f), 90 min (lane g), and 120 min (lane h). B. Hybridization patterns. The gel represented in part A was blotted and hybridized with 32p-E6.

Table 2-3. Fragment Sizes of EcoRI Partial Digestion Products.

Band on Autoradiogram	Length (kbp)	EcoRI Fragments in Band	Expected Length (kbp)
1	22.5	E1, E4, E6, E5	22.9
2	18.0	E4, E6, E5, E3	18.8
3	10.6	E4, E6, E5	10.7
4	6.1	E4, E6	6.0
5	5.7	E4	5.7
6	5.0	E6, E5	5.0
7	4.7	E5	4.7

must contain fragments comprised of E6 and E5. That only one copy of E6 is located between E4 and E5 is shown by the presence of band 3 (10.6 kbp). Only one other fragment (or combination of fragments) gives a length approximately equal to the sum of these three EcoRI fragments. E2 has a length of 10 kbp, but contains no sequences homologous to E6. Furthermore, hybridization of partial digests with ^{32}P -E4 and ^{32}P -E5 also reveal fragments of about 10.6 kbp (data not shown). Thus, the partial digestion results confirm the order of E4, E6, E5.

Because E4, E6, and E5 are adjacent and because bands E4 and E5 gave approximately equal signal intensities when hybridized with E6 (Figure 2-10F), it was expected that fragments containing all three of these would give stronger signals than fragments containing only one or two of the three. Band 2 in Figure 2-11b contained fragments 18.0 kbp in length. This is about the same length as a fragment comprised of E4, E6, E5, and E3 (18.8 kbp) suggesting that E5 and E3 are adjacent. If E2 and E5 were adjacent, a prominent band containing fragments of 20.7 kbp would be expected. Such a band was not detected. The only other combination of EcoRI fragments which would give a length of about 18.0 kbp and contain sequences homologous to E6 is that consisting of E1, E4, and E6. Since this set does not include E5, a strong signal from this combination was not expected. Consistent with this is fact that signals were not detected from fragments 13.4 kbp in length (containing E6, E5, and E3) nor from fragments 15 kbp in length (E6, E5, and E2). Thus, it can be concluded that E3 is the EcoRI fragment adjacent to E5. The only orientation possible for E5 is diagrammed in Figure 2-12 (page 158). Band 1 (22.5 kbp) on the autoradiogram suggests that E1 and E4 are adjacent on the map since this band contained fragments with a length about the same as fragments comprised of E1, E4,

E6, and E5. This is confirmed below.

The EcoRI partial digestion experiment was also performed using ^{32}P -E5 as a hybridization probe. After autoradiography, a complex set of major and minor bands were obtained (data not shown). This result was obtained due to the fact that E5 contains sequences homologous to two other (non-adjacent) EcoRI fragments--E1 and E6 (Figure 2-10E, page 149). Nevertheless, the results were consistent with E5 and E3 being adjacent (i.e. a minor band of 13 kbp was observed but a band of 14.7 kbp was not).

Completing the map. Having determined that E3 follows E5 on the map, we now look for the EcoRI fragment adjacent to E3. Figure 2-10C showed that ^{32}P -E3 hybridized with P2, P4, and P6. Because P4 overlaps E5 and E3, and because P6b is the internal PvuII fragment of E3, it can be concluded that P2 must overlap the next EcoRI fragment. Furthermore, since P2 has a length of 13 kbp, the next EcoRI fragment must have a PvuII site approximately 7.6 kbp from one end ($13 \text{ kbp} - 5.4 \text{ kbp} = 7.6 \text{ kbp}$). The only possible choice is E2, and E2 will only fit in one orientation. Additional evidence that E2 and E3 are adjacent has already been discussed (Figure 2-5, page 136). Further evidence was provided by hybridizing ^{32}P -E2 DNA to PvuII-digested mtDNA (Figure 2-10B, page 149). Bands P1, P2, P4, and P6 gave signals as mentioned earlier. It can be concluded from the P2 signal that E2 and E3 are indeed adjacent and are in the orientation diagrammed in Figure 2-12 (page 158). The signal from band P6 was expected since E2 contains the internal PvuII fragment, P6a. The signal from P4 is interpreted as being due to the repetitive sequences common to both E2 and E3 (Figure

2-8b). The signal from P1 suggests that P1 is adjacent to P6a. (This also suggests that E3 hybridizes with P1 due to the repetitive sequences common to E2 and E3.)

Moving in the direction of P1, we look for the EcoRI fragment adjacent to E2. Figure 2-10D (page149) showed that E4 also hybridized with P1. This observation suggests that P1 overlaps both E4 and E2, and that the map is circular. There must be an EcoRI fragment located between E4 and E2 because of the length of P1 (14.8 kbp). The EcoRI fragment must have a length of approximately 12.9 kbp (14.8 kbp kbp-0.5 kbp = 12.9 kbp) and contain no PvuII sites. Only E1 (12.2 kbp), the last remaining EcoRI fragment to be mapped, has these characteristics. This order is further confirmed by the partial digest mapping results as mentioned above (Figure 2-11b, Band 1, page153). The circular restriction map of the EcoRI and PvuII sites is presented in Figure 12.

Comments Concerning the Restriction Map

The results above have shown that the mitochondrial DNA of Candida albicans maps as a circular molecule of 41,000 base pairs. This size estimate is based upon three restriction enzyme digestions (EcoRI, PvuII, and the double digest). Each EcoRI fragment has been mapped using at least two pieces of evidence, partial digestion results and overlapping hybridization results. As further proof of the circular restriction map, it has been found that partial digestion results using PvuII are consistent with the map (W. S. Riggsby, personal communication). At this time, two EcoRI fragments remain unoriented. E1 has not been oriented simply because it contains no PvuII sites. E6 has not been

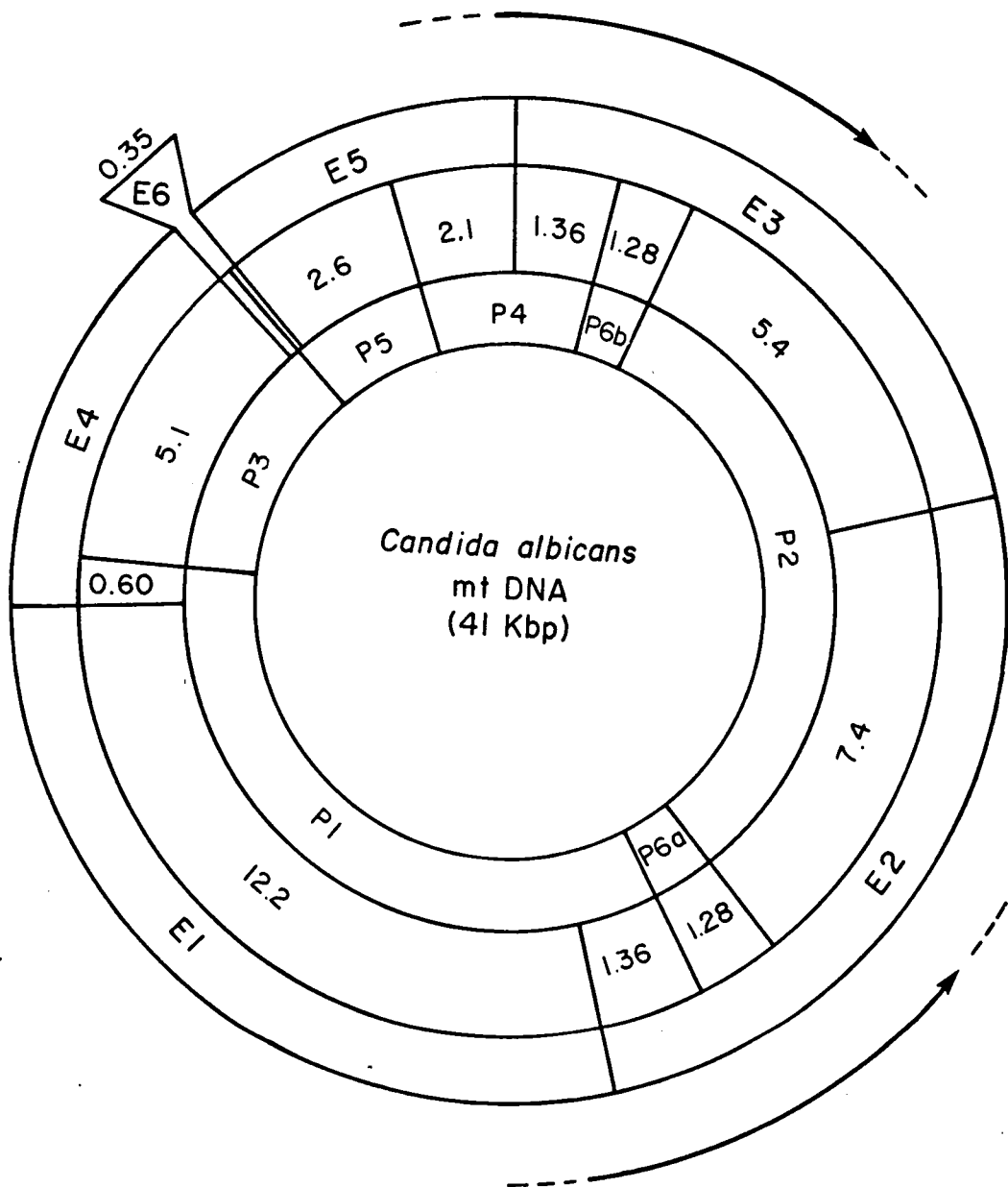


Figure 2-12. Circular restriction endonuclease map of *Candida albicans* mtDNA. The outer circle shows the positions of the *EcoRI* fragments and the inner circle shows the positions of the *PvuII* fragments. The middle circle indicates the sizes of the fragments in kilobase pairs. The arrows show the approximate position of the inverted repeated sequence.

oriented because of its small size. Nevertheless, a framework for future mapping has been established.

Examination of the mapping information reveals an interesting feature, namely a large inverted repeat. The evidence for this is as follows. First, EcoRI fragments E2 and E3 both contain two PvuII sites located at identical distances close to one end of each fragment (Figure 2-9, page 148). Thus, these two EcoRI fragments are asymmetrical with respect to the PvuII sites. Second, hybridization experiments have shown that the two smallest PvuII fragments of both E2 and E3 (i.e. the two adjacent fragments located close to one end of both E2 and E3) contain homologous sequences (Figure 2-8b, page 144). This establishes the existence and location of the repeated sequences. Third, the restriction mapping has shown that E2 and E3 are oriented in opposite directions relative to the PvuII sites (Figure 2-12, page 158). Thus, the repeated sequences are inverted. Intensive restriction mapping of E2 and E3 using additional endonucleases has confirmed the presence of the inverted repeat. All restriction sites mapping in the end of E2 which contains the PvuII sites have also been found to map in the same positions in E3 (B. Troutman, personal communication). At this time, however, the size of the repeat is not certain, but it appears to extend into E1 and E5. Evidence for this is provided by the fact that E5 hybridizes with E1 (Figure 2-10E, page 149).

The mtDNAs of two other organisms have been found to contain inverted repeats: Kloeckera africana and Tetrahymena pyriformis. K. africana is a yeast which contains a 27 kbp, circular mtDNA (Clark-Walker et al., 1981), and T. pyriformis is a protozoan

which contains a 49 kbp, linear mtDNA (Goldbach et al., 1978b). Thus, at this point it appears that inverted repeats are a feature only found in certain lower eukaryotic mtDNAs. In the above two organisms, the inverted repeats have been found to contain large ribosomal RNA gene sequences. In the case of K. africana, the duplication of the large rRNA gene is incomplete, and the duplication also contains sequences which hybridize to transfer RNAs. In I. pyriformis, the inverted duplication is terminally located on the linear molecule and contains complete large rRNA genes. Thus it should be interesting to determine if the inverted repeat of C. albicans mtDNA also contains rRNA genes.

The restriction map presented above does not appear to be unique to C. albicans strain H317. Whole cell DNAs from three other strains of C. albicans (ATCC 18804, H350, and WD 18-4) have been digested with EcoRI, electrophoresed, transferred to nitrocellulose, and hybridized with ^{32}P -labeled mtDNA from strain H317. Autoradiography showed that these three strains have the same electrophoretic pattern as the control DNA from strain H317 (M. Mason, personal communication). Thus, the information presented here should be of great importance to others interested in Candida albicans mitochondrial DNA.

LIST OF REFERENCES

LIST OF REFERENCES

- Attardi, G. 1981. Organization and expression of the mammalian mitochondrial genome: a lesson in economy. *Trends in Biological Sciences* 6: 100-103.
- Aujame, L. and K. B. Freeman. 1979. Mammalian mitochondrial transfer RNAs: chromatographic properties, size, and origin. *Nucleic Acid Research* 6: 455-470.
- Barrell, B. G., A. T. Bankier, and J. Drouin. 1979. A different genetic code in human mitochondria. *Nature* 282: 189-194.
- Barrell, B. G., S. Anderson, A. T. Bankier, M. H. L. DeBruijn, E. Chen, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden, and I. G. Young. 1980. Different pattern of codon recognition by mammalian mitochondria tRNAs. *Proc. Natl. Acad. Sci.* 77: 3164-3166.
- Bell, A. J. and N. Hardman. 1977. Characterization of foldback sequences in hamster DNA using electron microscopy. *Nucleic Acids Research* 4: 247-268.
- Bertrand, H., N. A. Szakacs, F. E. Nargang, C. A. Zagoneski, R. A. Collins, and J. C. Harrigan. 1976. The function of mitochondrial genes in *Neurospora crassa*. *Canadian Journal of Genetics and Cytology* 18: 397-409.
- Bibb, M. J., R. A. Van Etten, C. T. Wright, M. W. Walberg, and D. A. Clayton. 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell* 26: 167-180.
- Blattner, F. R. and H. P. Erickson. 1967. Rapid nucleotide separation by chromatography on cation-exchange columns. *Analytical Biochemistry* 18: 220-227.
- Bonitz, S. G., R. Berlan, G. Coruzzi, M. Li, G. Macino, F. G. Nobrega, M. P. Nobrega, B. F. Thalenfeld, and A. Tzagoloff. 1980. Codon recognition rules in yeast mitochondria. *Proc. Natl. Acad. Sci.* 77: 3167-3170.
- Bonner, T. I., D. J. Brenner, B. R. Neufeld, and F. J. Britten. 1973. Reduction in the rate of DNA reassociation by sequence divergence. *Journal of Molecular Biology* 81: 123-135.
- Borst, P. 1977. Structure and function of mitochondrial DNA. *Trends in Biological Sciences* 2: 31-34.
- Borst, P. and L. A. Grivell. 1978. The mitochondrial genome of yeast. *Cell* 15: 705-723.

- Borst, P. and L. A. Grivell. 1981. Small is beautiful--portrait of a mitochondrial genome. *Nature* 290: 443-444.
- Brenner, D. J., G. R. Fanning, A. V. Rake, and K. E. Johnson. 1969. Batch procedure for thermal elution of DNA from hydroxyapatite. *Analytical Biochemistry* 28: 447-459.
- Britten, R. J. and D. E. Kohne. 1968. Repeated sequences in DNA. *Science* 161: 529-540.
- Britten, R. J., D. E. Graham, and B. R. Neufeld. 1974. Analysis of repeating DNA sequences by reassociation. *Methods in Enzymology* 29E: 363-418.
- Burns, V. W. 1966. Regulation of pyrimidine biosynthesis and its strong coupling to the purine system. *Biophysical Journal* 6: 787-800.
- Clark-Walker, G. D., C. R. McArthur, and K. S. Sriprakash. 1981. Partial duplication of the large ribosomal RNA sequence in an inverted repeat in circular mitochondrial DNA from Kloeckera africana. *Journal of Molecular Biology* 147: 399-415.
- Cohn, W. E. 1955. Methods of isolation and characterization of mono- and polynucleotides by ion exchange chromatography. *Methods in Enzymology* 3: 724-743.
- Crick, F. H. C. 1966. Codon-anticodon pairing: the wobble hypothesis. *Journal of Molecular Biology* 19: 548-555.
- Cryer, D. R., R. Eccleshall, and J. Marmur. 1975. Isolation of yeast DNA. *Methods in Cell Biology* 12: 39-44.
- Cummings, D. J., L. Belcour, and C. Grandchamp. 1979. Mitochondrial DNA from Podospora anserina. II. Properties of mutant DNA and multimeric circular DNA from senescent cultures. *Molecular and General Genetics* 171: 239-250.
- Cummings, D. J. and J. L. Laping. 1981. Organization and cloning of mitochondrial deoxyribonucleic acid from Paramecium tetraurelia and Paramecium primaurelia. *Molecular and Cellular Biology* 11: 972-982.
- Cummings, D. J., R. A. Maki, P. J. Conlon, and J. Laping. 1980. Anatomy of mitochondrial DNA from Paramecium aurelia. *Molecular and General Genetics* 178: 499-510.
- Dale, R. M. K., J. H. Duesing, and D. Keene. 1981. Supercoiled mitochondrial DNAs from plant tissue culture cells. *Nucleic Acids Research* 9: 4583-4593.

- Davidson, E. H. and R. J. Britten. 1979. Regulation of gene expression: possible role of repetitive sequences. *Science* 204: 1052-1059.
- Davidson, E. H., B. R. Hough, C. S. Amenson, and R. J. Britten. 1973. General interspersion of repetitive with non-repetitive sequence elements in the DNA of Xenopus. *Journal of Molecular Biology* 77: 1-23.
- Davidson, J. N. 1975. *The Biochemistry of the Nucleic Acids*. New York: Academic Press.
- Davis, R. W., D. Botstein, and J. R. Roth. 1980. *Advanced Bacterial Genetics*. Cold Spring Harbor Laboratory, New York.
- De Ley, J., H. Cattoir, and A. Reynaerts. 1970. The quantitative measurement of DNA hybridization from renaturation rates. *European Journal of Biochemistry* 12: 133-142.
- Dugaiczyk, A., H. W. Boyer, and H. M. Goodman. 1975. Ligation of EcoRI endonuclease-generated DNA fragments into linear and circular structures. *Journal of Molecular Biology* 96: 171-184.
- Firtle, R. A. and J. Bonner. 1972. Characterization of the genome of the cellular slime mold Dictyostelium discoideum. *Journal of Molecular Biology* 66: 339-361.
- Fisher, M. P. and C. W. Dingman. 1971. Role of conformation in determining the electrophoretic properties of polynucleotides in agarose-acrylamide composite gels. *Biochemistry* 10: 1895-1899.
- Galau, G. A., W. H. Klein, M. M. Davis, B. J. Wold, R. J. Britten, and E. H. Davidson. 1976. Structural gene sets active in embryos and adult tissues of the sea urchin. *Cell* 7: 487-506.
- Gillis, M., J. De Ley, and M. De Cleene. 1970. The determination of molecular weight of bacterial genome DNA from renaturation rates. *European Journal of Biochemistry* 12: 143-153.
- Goddard, J. M. and D. J. Cummings. 1977. Mitochondrial DNA replication in Paramecium aurelia. Cross linking of the initiation end. *Journal of Molecular Biology* 109: 327-344.
- Goldbach, R. W., J. E. Bollen De Boer, E. F. J. VanBruggen, and P. Borst. 1978a. Replication of the linear mitochondrial DNA of Tetrahymena pyriformis. *Biochimica et Biophysica Acta* 521: 169-186.

- Goldbach, R. W. P., Borst, J. E. Bollen-De Boer, and E. F. J. Van Bruggen. 1978. The organization of ribosomal RNA genes in the mitochondrial DNA of Tetrahymena pyriformis strain ST. *Biochimica et Biophysica Acta* 521: 169-186.
- Graham, D. E., B. R. Neufeld, E. H. Davidson, and R. J. Britten. 1974. Interspersion of repetitive and nonrepetitive DNA sequences in the sea urchin genome. *Cell* 1: 127-138.
- Gray, M. W. and W. F. Doolittle. 1982. Has the endosymbiont hypothesis been proven? *Microbiological Reviews* 46: 1-42.
- Grunstein, M. and D. S. Hogness. 1975. Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci.* 72: 3961.
- Hanahan, D. and M. Meselson. 1980. Plasmid screening at high colony density. *Gene* 10: 63-67.
- Haseltine, W. A., C. P. Lindan, A. D. D'Andrea, and L. Johnsrud. 1980. The use of DNA fragments of defined sequence for the study of DNA damage and repair. *Methods in Enzymology* 65 (part 1): 235-248.
- Hatzfeld, J. 1973. DNA labelling and its assay in yeast. *Biochimica et Biophysica Acta* 299: 34-42.
- Hudspeth, M. E. S., W. E. Timberlake, and R. B. Goldberg. 1977. DNA sequence organization in the water mold Achyla: Proceedings of the National Academy of Science 74: 4332-4336.
- Hurst, R. O., J. A. Little, and G. C. Butler. 1951. The enzymatic degradation of thymonucleic acid. II. The hydrolysis of oligonucleotides. *Journal of Biological Chemistry* 188: 705-715.
- Hutton, J. R. 1977. Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucleic Acids Research* 4: 3537-3555.
- Jamet-Vierny, C., O. Begel, and L. Belcour. 1980. Senescence in Podospora anserina: amplification of a mitochondrial DNA sequence. *Cell* 21: 189-194.
- Kohne, D. E. and J. R. Britten. 1971. Hydroxyapatite techniques for nucleic acid reassociation. *Procedures in Nucleic Acid Research* 2: 500-512.

- Kolodner, R. and K. K. Tewari. 1972. Physiochemical characterization of mitochondrial DNA from pea leaves. *Proc. Natl. Acad. Sci.* 69: 1830-1834.
- Krumlauf, R. and G. A. Marzluf. 1979. Characterization of the sequence complexity and organization of the Neurospora crassa genome. *Biochemistry* 18: 3705-3713.
- Lauer, G. D., T. M. Roberts, and L. C. Klotz. 1977. Determination of the nuclear DNA content of Saccharomyces cerevisiae and implications for the organization of DNA in yeast chromosomes. *Journal of Molecular Biology* 114: 507-526.
- Lee, K. L., H. R. Buckley, and C. C. Campbell. 1975. An amino acid liquid synthetic medium for the development of mycelial and yeast forms of Candida albicans. *Sabouraudia* 13: 148-153.
- Lewis, A., R. Morimoto, and M. Rabinowitz. 1978. Restriction enzyme analysis of mitochondrial DNAs of petite mutants of yeast: classification of petites, and deletion mapping of mitochondrial genes. *Molecular and General Genetics* 163: 257-275.
- Lewin, B. 1980. *Gene Expression*, Vol. 2. New York: Wiley.
- Linnane, A. W. and H. B. Lukins. 1975. Isolation of mitochondria and techniques for studying mitochondrial biogenesis in yeasts. *Methods in Cell Biology* 12: 285-309.
- Lloyd, David. 1974. *The Mitochondria of Microorganisms*. New York: Academic Press.
- Locker, J., A. Lewin, and M. Rabinowitz. 1979. The structure and organization of mitochondrial DNA from petite yeast. *Plasmid* 2: 155-181.
- Macino, G., C. Scazzocchio, R. B. Waring, M. McPhail Berks, and R. W. Davies. 1980. Conservation and rearrangement of mitochondrial structural gene sequences. *Nature* 288: 404-406.
- Mandel, M. and J. Marmur. 1968. Use of ultraviolet absorbance-temperature profile for determining the guanine plus cytosine content of DNA. *Methods in Enzymology* 12B: 195-206.
- Mandel, M., C. L. Schildkraut, and J. Marmur. 1968. Use of CsCl density gradient analysis for determining the guanine plus cytosine content of DNA. *Methods in Enzymology* 12B: 184-195.
- Mannella, C. A., T. H. Pittenger, and A. M. Lambowitz. 1979a. Transmission of mitochondrial deoxyribonucleic acid in Neurospora crassa sexual crosses. *Journal of Bacteriology* 137: 1449-1451.

- Mannella, C. A., R. R. Goewert, and A. M. Lambowitz. 1979b. Characterization of variant Neurospora crassa mitochondrial DNAs which contain tandem reiterations. Cell 18: 1197-1207.
- Marmur, J. and P. Doty. 1959. Heterogeneity in DNA. I. Dependence on composition of the configurational stability of DNAs. Nature 183: 1427-1428.
- Martinson, H. G. 1973. The basis of fractionation of single-stranded nucleic acids in hydroxylapatite. Biochemistry 12: 2731-2736.
- Matsuda, M. and H. Ogoshi. 1966. Specificity of DNase I. Journal of Biochemistry 59: 230-235.
- Merkel, G. J., F. Naider, and J. M. Becker. 1980. Amino acid uptake by Saccharomyces cerevisiae plasma membrane vesicles. Biochimica et Biophysica Acta 595: 109-120.
- Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory.
- Miyazawa, Y. and C. A. Thomas. 1965. Nucleotide composition of DNA molecules. Journal of Molecular Biology 11: 223.
- Monahan, J. J. and R. H. Hall. 1975. Reduced effect of 50% formamide upon T_m of DNA absorbed on hydroxylapatite columns. Analytical Biochemistry 63: 620-623.
- Nass, M. M. K. 1980. Analysis of the two heavy and light strand origins and the direction of replication of mitochondrial DNA within a detailed physical map of this genome in transformed hamster cells. Journal of Molecular Biology 140: 231-256.
- Norgard, M. V. 1981. Rapid and simple removal of contaminating RNA from plasmid DNA without the use of RNase. Analytical Biochemistry 113: 34-42.
- Norgard, M. V., K. Emigholz, and J. J. Monahan. 1979. Increased amplification of pBR322 plasmid DNA in E. coli K-12 strains RR1 and X1776 grown in the presence of high concentrations of nucleoside. Journal of Bacteriology 138: 270-272.
- Odds, F. C. 1979. Candida and Candidosis. Baltimore: University Park Press.
- Peacock, A. C. and C. W. Dingman. 1968. Molecular weight estimation and separation of RNA by electrophoresis in agarose-polyacrylamide composite gels. Biochemistry 7: 668-674.
- Pearson, W. R., E. H. Davidson, and R. J. Britten. 1977. A program for least squares analysis of reassociation and hybridization data. Nucleic Acid Research 4: 1727-1737.

- Pellegrini, M., W. E. Timberlake, and R. B. Goldberg. 1981. Electron microscopy of Achyla deoxyribonucleic acid sequence organization. Molecular and Cellular Biology 1: 136-143.
- Perlman, S., C. Phillips, and J. O. Bishop. 1976. A study of foldback. Cell 8: 33-42.
- Poulter, R., K. Jeffrey, M. J. Hubbard, M. G. Shepherd, and P. A. Sullivan. 1981. Parasexual genetic analysis of Candida albicans by spheroplast fusion. Journal of Bacteriology 146: 833-840.
- Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. Journal of Molecular Biology 113: 237-251.
- Riggsby, W. S., L. J. Torres-Bauza, J. W. Wills, and T. M. Townes. 1982. DNA content, kinetic complexity and the ploidy question in Candida albicans. Molecular and Cellular Biology, in press.
- Rubin, G. M. 1973. The nucleotide sequence of Saccharomyces cerevisiae 5.8 S ribosomal ribonucleic acid. Journal of Biological Chemistry 248: 3860-3875.
- Rubin, G. M. 1975. Preparation of RNA and ribosomes from yeast. Methods in Cell Biology 12: 45-64.
- Sarachek, A., D. D. Rhoads, and R. H. Schwarzhoff. 1981. Hybridization of Candida albicans through fusion of protoplasts. Archives of Microbiology 129: 1-8.
- Schatz, G. and L. Kovac. 1974. Isolation of promitochondria from anaerobically grown Saccharomyces cerevisiae. Methods in Enzymology 48: 627-632.
- Sederoff, R. R., C. S. Levings, D. H. Timothy, and W. W. L. Hu. 1981. Evolution of DNA sequence organization in mitochondrial genomes of Zea. Proc. Natl. Acad. Sci. 78: 5953-5957.
- Silva-Hunter, M. and B. H. Cooper. 1974. Medically important yeast. In Manual of Clinical Microbiology (2nd edn), pp. 491-507. Edited by E. H. Lennette, E. H. Spaulding and J. P. Truant. Washington, D. C.: American Society of Microbiology.
- Singer, D. S. 1979. Arrangement of a highly repeated DNA sequence in the genome and chromatin of the African green monkey. Journal of Biological Chemistry 254: 5506-5514.
- Sinsheimer, R. L. 1954. Nucleotides from T2r⁺ Bacteriophage. Science 120: 551-553.

- Sinsheimer, R. L. and J. F. Koerner. 1951. Ion exchange separation of desoxyribonucleotides. *Science* 114: 42-43.
- Smith, J. R. and I. Rubenstein. 1973. The development of "senescence" in *Podospora anserina*. *Journal of General Microbiology* 76: 283-296.
- Smith, M. J., R. J. Britten, and E. H. Davidson. 1975. Studies on nucleic acid reassociation kinetics: reactivity of single-stranded tails in DNA-DNA renaturation. *Proc. Nat. Acad. Sci.* 72: 4805-4809.
- Southern, E. 1979. Gel electrophoresis of restriction fragments. *Methods in Enzymology* 68: 152-176.
- Sparks, R. B. and R. M. K. Dale. 1980. Characterization of ³H-labeled supercoiled mitochondrial DNA from tobacco suspension culture cells. *Molecular and General Genetics* 180: 351-355.
- Stenderup, A. and A. Leth Bak. 1968. Deoxyribonucleic acid base composition of some species within the genus *Candida*. *Journal of General Microbiology* 52: 231-236.
- Subirana, J. A. and P. Doty. 1966. Kinetics of renaturation of denatured DNA. I. Spectrophotometric results. *Biopolymers* 4: 171-187.
- Sulkowski, E. and M. Laskowski. 1971. Inactivation of 5'-nucleotidase in commercial preparations of venom exonuclease (phosphodiesterase). *Biochimica et Biophysica Acta* 240: 443-447.
- Sutton, W. 1971. A crude nuclease preparation suitable for use in DNA reassociation experiments. *Biochim. Biophys. Acta.* 240: 522-531.
- Suyama, Y. and K. Miura. 1968. Size and structural variations of mitochondrial DNA. *Proc. Natl. Acad. Sci.* 60: 235-242.
- Syverson, R. E. 1981. Variable assimilation of carbon compounds by *Candida albicans*. *Journal of Clinical Microbiology* 13: 163-166.
- Szybalski, W. and E. H. Szybalski. 1971. Equilibrium density gradient centrifugation. *Procedures in Nucleic Acid Research* 2: 311-354.
- Timberlake, W. E. 1978. Low repetitive DNA content in *Aspergillus nidulans*. *Science* 202: 973-975.
- Torres-Bauza, L. J. 1980. The nucleosome structure and the genome size of dimorphic fungus *Candida albicans*. Dissertation, University of Puerto Rico.

- Torres-Bauza, L. J. and W. S. Riggsby. 1980. Protoplasts from yeast and mycelial forms of Candida albicans. Journal of General Microbiology 119: 341-349.
- Tzagoloff, A., G. Macino, and W. Sebald. 1979. Mitochondrial genes and translation products. Annual Review of Biochemistry 48: 419-441.
- Wesolowski, M. and H. Fukuhara. 1981. Linear mitochondrial deoxyribonucleic acid from the yeast Hansenula mrakii. Molecular and Cellular Biology 5: 387-393.
- Wetmur, J. G. and N. Davidson. 1968. Kinetics of renaturation of DNA. Journal of Molecular Biology 31: 349-370.
- Whelan, W. L. and P. T. Magee. 1981. Natural heterozygosity in Candida albicans. Journal of Bacteriology 145: 896-903.
- Willard, J. M., J. J. Davis, H. G. Wood. 1969. Phosphoenolpyruvate Carboxytransphosphorylase. IV. Requirement for metal ions. Biochemistry 8: 3137-3144.
- Wilson, D. A. and C. A. Thomas. 1973. Hydroxyapatite chromatography of short double-helical DNA. Biochimica et Biophysica Acta 31: 333-340.
- Woodward, R. S. and J. Lebowitz. 1980. A revised equation relating DNA buoyant density to guanine plus cytosine content. Journal of Biochemical and Biophysical Methods 2: 307-309.
- Zeman, L. J. and C. V. Lusena. 1975. In vivo and in vitro synthesis of yeast mitochondrial DNA. Methods in Cell Biology 12: 273-283.

APPENDIX

APPENDIX

CHARACTERIZATION OF THE HYDROXYAPATITE METHOD

Before performing any reassociation experiments, it was necessary to make sure that the hydroxyapatite method of separating single and double stranded DNA would work in my hands. My initial plan was to use the method of high salt elution (0.48 M NaP) of dsDNA, after loading samples and washing ssDNA through with moderate salt (0.12 M NaP). To test this method, mixtures of denatured and undenatured, sheared, ³H-E. coli DNA were prepared in 0.12 M NaP and then chromatographed. The results from five test columns are presented in Table A-1.

Table A-1. HAP Chromatography Using High Salt Elution.

CPM Loaded on Column ¹		CPM Recovered		% Recovery ²	
Denatured	Native	0.12M NaP	0.48M NaP	ssDNA	dsDNA
1880	0	2445	35	130	(2) ³
2830	1000	3652	886	129	89
1880	2000	2226	1504	118	75
940	3000	1199	2270	128	76
0	2000	51	1301	(3) ³	65

¹Radioactivity of samples to be loaded was measured in 1 ml of 0.24 M NaP and 9 ml of Aquasol-2.

²Calculated relative to the cpm of loaded denatured DNA and native DNA, respectively.

³Numbers in parentheses were calculated using the total cpm loaded.

The results of the HAP columns suggested that a quenching problem existed. For example, only 65% of the 2000 cpm of dsDNA loaded on the fifth column was recovered in the 0.48 M wash (1301 cpm), yet only 3% of the loaded cpm were recovered as single-stranded DNA. On the other hand, 2445 cpm of ssDNA were recovered off the first column, but only 1880 cpm were loaded (130% recovery). For each column, the recovery of ssDNA was always much greater than the theoretical 100% value, and the recovery of dsDNA was always much lower than the theoretical 100% value. Further support for a quenching argument is provided by the fact that the radioactivity of the samples to be loaded was measured in a phosphate buffer intermediate in concentration to those used in the chromatography, 0.24 M NaP.

To test the hypothesis that ^3H -cpm decrease with increasing NaP concentrations (when using Aquasol-2 scintillation fluid), samples of ^3H -DNA were counted in a series of different buffers. As shown in Table A-2, the hypothesis was indeed found to be true. The number of cpm obtained in the presence of 0.24 M NaP (7876 cpm) was smaller than when an identical sample was counted in the presence of 0.12 M NaP (14,396 cpm) and larger than when 0.48 M NaP was used (5206 cpm). Not mentioned in the table is the observation that a white precipitate formed upon addition of Aquasol to each of the samples. The amount of precipitate increased with the phosphate concentration and presumably was responsible for the quenching. When Scintiverse (Fisher) was used in the place of Aquasol, the same effect was observed with an even greater degree of quenching.

Table A-2. Quenching of ^3H -cpm with Increasing NaP Concentration.

NaP Concentration	CPM ¹	% CPM ²
0.12 M	14396	100
0.24	7876	55
0.36	6431	45
0.48	5206	36

¹ Each value is the average of two measurements using 1 ml of aqueous solution and 9 ml Aquasol-2.

² Relative to 0.12 M NaP.

To circumvent the problems of high salt elution, I next tried the method of thermal elution of DNA bound to HAP. By raising the temperature of the column to 100°C, dsDNA is denatured and can be washed out of the column without increasing the NaP concentration. To test this method, 1 ml HAP columns were loaded with samples of dsDNA in 0.12 M NaP at 60°C, washed with 0.12 M NaP to remove nonbinding material, and then washed at 100°C with 0.12 M NaP to elute the bound DNA. The results of this experiment are presented in Table A-3.

Table A-3. Thermal Elution of dsDNA in the Absence of Carrier DNA.

DNA ¹ (μg)	% of Total CPM Loaded ²			Total Recovery
	60°C	100°C	Irreversible ³	
0.04	2.7	61.9	35.4	64.6
0.08	2.5	66.8	30.7	69.3
0.16	2.0	72.7	25.3	74.7

¹ Sheared, double stranded, ^3H -E. coli DNA.

² Measured in 1 ml of 0.12 N NaP and 9 ml of Aquasol.

³ 100% minus the total recovery.

Thermal elution was found to be inefficient for the chromatography of small amounts of DNA. As shown in Table A-3, irreversible binding was apparently a problem in the experiment resulting in total column recoveries of less than 75%. It can be seen, however, that the amount of irreversible binding was dependent upon the amount of DNA that was used. A four-fold increase in the amount of DNA loaded (from .04 μg to .16 μg) increased the column recovery 30% of the way towards the desired 100% mark (i.e. from 65% to 75%). This result suggested that the addition of unlabeled carrier DNA might eliminate the irreversible binding problem. To test this possibility, the thermal elution experiment was repeated, but this time the ^3H -DNA samples were mixed with 50 μg of native calf thymus DNA in 1 ml of 0.12 M NaP. In addition, the HAP was heated to 90-100°C for 30 minutes prior to its use in order to reduce nonspecific binding (Britten et al., 1974). The results of this experiment are presented in Table A-4.

When HAP chromatography was performed using the carrier DNA, the recovery from each column was greatly improved. Table A-4 shows that the amount of ^3H -E. coli DNA used did not influence the results; however, the column recoveries did fluctuate quite a bit with yields as high as 98% and as low as 82%. Because the fluctuations did not rise above 100% (as would be expected with random, experimental error), it seemed likely that they were due to another quenching problem. Evaporation during the 100°C, thermal elution would result in increased concentrations of NaP and reduced ^3H -cpm. To eliminate this possibility, I ran additional columns which were plugged with a rubber stopper during equilibrations to the high temperature. The results from these columns

are also found in Table A-4. As shown, stoppering of the columns seemed to eliminate the problem. Column recoveries obtained during the reassociation experiments described on the following pages were $100 \pm 10\%$ for the E. coli DNA experiment and $100 \pm 6\%$ for the C. albicans DNA experiment. Experts in the field report column recoveries of $100 \pm 8\%$ (Davidson et al., 1973).

Table A-4. HAP Chromatography Using Carrier DNA

³ H-DNA ¹ (μ g)	% of Total CPM Loaded ²			Total Recovery
	60°C	100°C	Irreversible ³	
0.04	6.4	75.7	17.9	82.1
0.08	3.5	93.5	3.0	97.0
0.16	4.4	84.5	11.1	88.9
0.16	4.6	93.3	2.1	97.9
0.16	3.2	83.5	13.3	86.7
0.16 ⁴	5.2	102.0	- 7.2	107.2
0.16 ⁴	5.2	91.2	3.6	96.4

¹Sheared, double stranded, ³H-E. coli DNA mixed with 50 μ g of unlabeled, native, calf DNA.

²Measured in 1 ml of 0.12 M NaP and 9 ml of Aquasol.

³100% minus the total recovery.

⁴Rubber stopper used during thermal elution.

VITA

John W. Wills was born on August 2, 1954, in Chicago, Illinois. He is the only son of Roland and Margaret Wills and has three sisters: Lucinda, Jennifer, and Jeannine. At age three, the author and his family moved to Memphis, Tennessee where he attended public schools and graduated from Fairley High School in May 1972. In August 1972 he began four enjoyable years of study at Grinnell College in Grinnell, Iowa where he majored in Biology. In May 1976 he graduated with honors and received a Bachelor of Arts degree. In September 1976 he accepted a teaching assistantship and began graduate studies at The University of Tennessee, Knoxville, in the Microbiology Department. Thanks to the influence of his major professor, Dr. W. S. Riggsby, he was able to become a molecular geneticist, and he graduated with a Doctor of Philosophy degree in August 1982.

The author is married to Dr. Rebecca C. Craven, who is sometimes referred to as Becky Wills. Both the author and his wife have accepted postdoctoral research positions at the University of Alabama in Birmingham where they now reside.