WEINER, Ronald Martin, 1942-
DETERMINATION OF INTRAMOLECULAR HETEROGENEITY
OF MICROBIAL DNA BY SPECTRAL ANALYSIS.
Iowa State University, Ph.D., 1970
Microbiology

University Microfilms, Inc., Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED
DETERMINATION OF INTRAMOLECULAR HETEROGENEITY OF MICROBIAL DNA BY SPECTRAL ANALYSIS

by

Ronald Martin Weiner

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of The Requirements for the Degree of DOCTOR OF PHILOSOPHY

Major Subject: Bacteriology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University Of Science and Technology Ames, Iowa

1970
TABLE OF CONTENTS

INTRODUCTION ........................................... 1
MATERIALS AND METHODS ................................. 7
RESULTS .................................................. 13
TABLES AND FIGURES .................................... 24
DISCUSSION ............................................... 47
LITERATURE CITED ....................................... 57
ACKNOWLEDGMENTS ....................................... 64
INTRODUCTION

Although the locations of the genes of many microorganisms are being mapped, relatively little is known of the inherent variation of base composition along the chromosome (or to use a phrase coined by Sueoka in 1964, of the intramolecular compositional heterogeneity of DNA molecules). Just as the units of function - cistrons - can be sequenced in their proper locations on the chromosome, so theoretically can the units of structure - the nucleotides - be mapped as a measure of regional guanine and cytosine (G+C) content. It was the purpose of this investigation to resolve into its component parts the chromosome of *Escherichia coli* at 50-52% overall G+C.

Although there are several methods available which determine overall G+C content, few permit analysis of intramolecular base composition. Of these, statistical analysis of thermal denaturation curves at 260 nm enables one to distinguish between different organisms possessing DNA molecules that are similar in overall G+C content. De Ley and Van Muylem (1963) were able to find differences among members of the genera *Pseudomonas*, *Xanthomonas*, *Chromobacterium*, and *Flavobacterium*; De Ley and Schell (1963) characterized several species of *Acetobacter* and *Gluconobacter*; De Ley and Rassel (1965) studied the genus *Rhizobium*; and Knittel *et al.* (1968) found heterogeneity in base composition among streptococci. Krieg and Lockhart (1968) employed a more rigorous statistical analysis of thermal transition curves than had been used previously. They measured not only the temperature midpoint and standard deviation of the thermal transition, but also its skewness and peakedness. While the
characteristic patterns obtained from these analyses are useful in taxonomic, diagnostic and evolutionary studies, they yield only a reflection of the internal base composition of DNA.

A parallel situation exists when one studies the unimodal band of DNA obtained in CsCl density gradient centrifugation. Since the band profile of a DNA sample is dependent on both its molecular weight and heterogeneity, an accurate estimation of heterogeneity is not easily obtainable. It has been shown by several of these studies, however, that mammalian DNA is more heterogeneous than bacterial DNA (Sueoka, 1964), and recent innovations in buoyant density determinations offer greater sensitivity and promise (Nandi, Wang and Davidson, 1965).

At a higher level of resolution, nearest neighbor base relationships may be determined with the aid of DNA polymerase (Josse et al., 1961; Berg et al., 1963) where 16 parameters are provided for the characterization of a given DNA. In this in vivo system, based on the specific modes of synthesis and degradation of DNA by specific enzymes, one can estimate the relative frequencies with which pairs of each of the 4 bases lie next to one another. This method has been used, successfully, for evolutionary studies with viral DNA (Bellett, 1967). But because it is limited in sensitivity and it is a technically difficult procedure to perform, nearest neighbor analysis has been employed infrequently.

Yet another method of investigation has been made possible by the discovery of Doty et al. (1960) and Marmur and Lane (1960) that the two strands of DNA may be separated and then reannealed with heterologous nucleotide strands. Because two strands of nucleic acids have more affinity for one another when they possess more base sequences in common,
the stability of such a heteroduplex becomes a measure of relatedness between the component strands. Thus, this method is useful in taxonomic and evolutionary studies and has been reviewed in this light by Britten and Kohne (1968) and McCarthy (1967). A main limitation of homology studies lies in their comparative nature, and that they thus fail to elucidate base sequences within any one given organism.

Additional evidence of compositional heterogeneity of DNA has come from salt fractionations, chromatographic analysis, and electron microscopy with base specific stains. Chargaff et al. (1953) separated calf thymus DNA into fractions containing 35 to 50% G+C; Brown and Watson (1953) report fractionation of calf thymus DNA from a histone kieselguhr column; and more recently, Miyazawa and Thomas (1965) were able to elute DNA fragments from hydroxyapatite columns. The separation of nucleotides by gel chromatography has also been used with some success (Lerner and Schepartz, 1969; Braun, 1967; Birnboim, 1966). Beer et al. (1966), combining chemical and visual techniques, were able to resolve concentrations of thymidine in electron micrographs of DNA polymers. It is unfortunate, however, that many of these methods are harsh and alter the DNA during analysis.

Because of the limitations of these methods, I chose to examine compositional heterogeneity of DNA by the technique of spectral analysis which was developed by Felsenfeld and Hirschman in 1965. The rationale for this technique is based on the different extinction coefficients of the four bases, the absorbance changes associated with the helix-coil transition during the thermal denaturation of double-stranded DNA, and on
the chemical and physical interactions between the different bases in the polymer. More complete discussions of this rationale and the parameters employed in calculations are presented by Felsenfeld (1968), Felsenfeld and Hirschman (1965), Votavova (1968), and Prouty (1970). Though the procedures involved are analogous to those used in measuring thermal transitions, one can not only determine overall base composition, but also base concentration and the extent of denaturation before analysis - and, most importantly, one can map the base composition of successively denaturing regions of DNA.

Using this method, Hirschman (1967) found intramolecular heterogeneity in the chromosome of Lambda bacteriophage and compositional heterogeneity among several Lambda mutants; Falkow and Cowie (1968) found that high intramolecular heterogeneity is a shared property among enteric phages; and Votavova et al. (1968) were able to characterize DNA from several higher plants and animals.

However, until the mechanism of denaturation is known, it will not be possible to extrapolate the thermal transition curve into a map of nucleotide concentrations along the chromosome. While the regions richer in A+T are known to denature initially (Inman, 1966) it is not known whether strand separation is initiated at one or both ends of the molecule, internally at single strand breaks, or preferentially near the areas of low regional G+C content.

The problem, then, becomes one of ordering on the chromosome the known lengths of known G+C content that denature during the course of spectral analysis.

To help accomplish this, it is possible to resort to techniques
analogous to those used in genetic mapping. One method of obtaining maps in bacteria is by doing a marker frequency analysis of the chromosome. This technique was devised by Sueoka and Yoshikowa (1965) with *Bacillus subtilis* and was employed by Stonehill and Hutchison (1966) with *Streptococcus faecalis* and by Altenbern (1968) with *Staphylococcus*. It is based on differences in the average state of chromosomal replication at different bacterial growth rates and is made possible by the consistent mode of DNA replication. The foundations or corollaries upon which marker frequency analysis rest have been experimentally substantiated. They are that DNA synthesis:

1- is semi-conservative (Watson and Crick, 1953; Meselson and Stahl, 1958; Cairns, 1963a; Cairns, 1963b; Bonhoeffer and Gierer, 1963)

2- is initiated at a fixed point on the chromosome and proceeds unidirectionally (Nagata, 1963; Yoshikawa and Sueoka, 1963; Lark et al., 1963; Donachie and Masters, 1966; Abe and Tomizawa, 1967; Berg and Caro, 1967; Cutler and Evans, 1967; Helmstetter, 1968; Caro and Berg, 1968; Caro and Berg, 1969)

3- is constant, taking about 41 minutes to replicate an entire bacterial chromosome assuming no substrate or energy limitations (Helmstetter, 1967; Cooper and Helmstetter, 1968)

4- is initiated at a fixed time during the cell division cycle with the time of onset of synthesis being regulated by (or regulating) the time between cell divisions (Lark, 1966a; Lark, 1966b; Clark and Maaløe, 1967; Helmstetter and Pierucci, 1968; Bleecken, 1969; Koch and Pachler, 1967; Helmstetter et al., 1968).

5- is initiated before completion of the preceding round of
replication in chromosomes of cells with generation times of less than 40 minutes, producing multiple replication points or forks (Oishi et al., 1964; Helmstetter and Cooper, 1968; Caro, 1970; Clark, 1968; Yoshikawa and Haas, 1968).

Thus, cells dividing every 25 minutes will have four copies of genetic loci near the origin to one of those near the terminus; cells dividing each 60 minutes will have a 2:1 ratio; cells with an infinite generation time will have a 1:1 ratio. Likewise, if there is heterogeneity of base composition along the chromosome, then this compositional heterogeneity should vary as the growth rate is varied.

Therefore, the problem of resolving the overall base content of the *E. coli* chromosome into segments can be approached by using spectral analysis to quantitate the G+C content of individual regions as they denature and by ordering these regions on the chromosome by a technique analogous to marker frequency analysis where one examines the structure, rather than the function, of chromosomal segments.

In this investigation, four organisms with overall G+C contents similar to that of *E. coli* were employed to evaluate the resolving capability of spectral analysis. The *E. coli* chromosome, in several distinct stages of replication, was then examined.
MATERIALS AND METHODS

Organisms  Escherichia coli, American Type Culture Collection (ATCC) strain 9637, Salmonella typhimurium, ATCC strain 13311, Proteus morganii, ATCC strain 8019, Neisseria sicca, strain Ne 12, obtained from Dr. B. W. Catlin at the Marquette School of Medicine, and Lambda bacteriophage B^C^ from the culture collection of Iowa State University Bacteriology Department were used in this investigation. Stock cultures were maintained on Brain Heart Infusion (BHI) agar slants at 5°C.

Media  E. coli was grown in nutrient broth (NB) and also in BHI in batch culture and in BHI plus 0.40 mg/L antifoam AF emulsion (Dow Corning Corporation, Midland, Michigan) in continuous culture. N. sicca was cultivated in medium containing 35 g BHI, 6 g NB, and 4 g yeast extract/L. P. morganii and S. typhimurium were grown in NB. All media were dissolved in distilled water and all incubation was at 37°C.

Propagation and purification of Lambda bacteriophage and DNA  Lambda was propagated, titrated and purified as described by Prouty (1970). DNA was purified by four phenol extractions. A solution was used that contained 50 g of freshly distilled phenol, 20 ml tris-HCl buffer, and 0.07 g 8-hydroxyquinoline. This solution was regulated to pH 7.0 and incubated overnight at 37°C in the dark. Residual phenol was removed with three washes of absolute ether.

Continuous culture system  The continuous culture system used in these studies has been described by Moore (1969) and Wright (1964). The system consists of a 15 L medium reservoir, a Beckman model 746 solution metering pump, a pre-warm coil, a custom-designed growth vessel of 510 ml
volume and several receiving vessels. Aeration, at a rate capable of supporting $10^{11}$ cells/ml of medium, was controlled by an airflow metering device (Ecker and Lockhart, 1959).

Five ml of *E. coli* from an overnight culture grown on a BHI agar slant were washed into the growth vessel containing 505 ml of pre-warmed medium. Growth was measured periodically in a Beckman-Spinco model 151 spectrocolorimeter at 540 nm and occasionally by plate count and Petroff-Hauser count. When the optical density was between 0.9 and 1.3 (about $10^9$ cells/ml), the continuous culture system was activated. Steady state growth rates of 0.40 and 2.23 generations/hr were achieved in accordance with the working equation of Herbert et al. (1956): specific growth rate equals the overflow rate (in ml/hr) divided by the volume of medium maintained in the culture vessel. At 2.23 generations/hr, the optical density of the culture was 0.80 and plate counts showed $2 \times 10^8$ cells/ml; at 0.40 generations/hr the optical density was 1.60 and there were $3 \times 10^9$ viable cells/ml. Cells were not harvested until they had been in steady state for at least four generations.

Cell harvest in continuous and batch culture DNA replication of harvested *E. coli* was inhibited with 60 µg of 2,4-dinitrophenol (DNP) per ml and crushed ice, or with 20 µg of naladixic acid (NA) per ml. The naladixic acid was a gift from S. Archer of the Sterling-Winthrop Research Institute. DNP is known to uncouple oxidative phosphorylation and prevent protein and DNA synthesis. NA acts directly on the chromosome (Deitz et al., 1966) to inhibit its synthesis (Goss et al., 1965) and after several hours exposure degrades it from replication fork to origin (Ramareddy and Reiter, 1969).
As the original experiments had been carried out with *E. coli* K12, the sensitivity of *E. coli* 9637 (W) to this antibiotic was tested. Cell division was inhibited within an hour and "typical" long forms were observed. The inhibitory effects of NA were reversible for three hours, and after this time a logarithmic death phase ensued. It was concluded that the sensitivity of *E. coli* 9637 is comparable to that of test strain *E. coli* K12 TAU.

Before bacteria in the early stationary phase of growth were harvested, crushed ice was added over a period of 20 minutes to retard nuclease action and to inhibit any possible new rounds of chromosomal replication. Between $10^{12}$ and $5 \times 10^{12}$ cells were collected for DNA extraction. The cells were washed twice with saline EDTA (0.15 M NaCl, 0.1 M ethylenediamine tetraacetic acid, pH 8.0) and resuspended in 120 ml of this solution. Cells were frozen until needed.

**Extraction of bacterial DNA** DNA was purified by a modification of the Marmur (1961) procedure. The organisms were lysed with sodium lauryl sulfate at 60°C. This was followed by at least 5 deproteinization steps with chloroform isoamyl alcohol and an RNA\(^ {\text{ase}}\) (Worthington) treatment. Nucleic acids were precipitated with 2 volumes of 95% ethyl alcohol. A phenol purification step was added when necessary. Initial purity was determined by Lowry, Biuret and orcinol methods. Neither RNA nor protein was detected within the limits of sensitivity of these tests. Typical ratios for absorption of the DNA at 260: 230: 280 nm corresponded to those reported by Marmur (1961), which were 1.0: 0.450: 0.515. It was concluded that protein and RNA contamination amounted to less than 0.5%.

Purified, spooled DNA was stored in 95% ethyl alcohol at 5°C until
needed, at which time it was dissolved in a phosphate salts buffer (0.0085M KH$_2$HPO$_4$, 0.0145M K$_2$HPO$_4$ and 0.100M NaCl plus deionized water) at pH 7.0.

This DNA solution was dialyzed for at least 72 hrs against 4 changes of the phosphate salts buffer at 8°C (500:1 buffer to DNA ratio). Immediately after dialysis, the DNA solution was diluted to 35 μg/ml ± 5 μg and spectral analysis was performed.

**Spectral analysis of DNA** 2.8 ml of the DNA sample was pipetted into 10 mm, matched, teflon stoppered, silica cuvettes. It made no difference whether these cuvettes had a working range from 180 nm to 2600 nm or 220 to 2600 nm. Spectral measurements were taken with a multiple sample absorbance recorder (Model 2000, Gilford Instrument Laboratories, Inc., Oberlin, Ohio) connected to a Beckman DU monochromometer. Temperature was regulated manually with a Haake Ultrathermostat (Poly Science Corp., Evanston, Illinois) and recorded through a platinum thermosensor located under the cuvette carriage. Control experiments showed no detectable difference between thermosensor readings and cuvette solution temperature.

Absorbance was read directly from the Detector-Indicator Housing, although it is also possible to calibrate a Gilford 209 absorbance meter to .001 accuracy and to take readings directly from this accessory.

The absorbance of the DNA samples at ambient temperature was measured at 230, 260, and 280 nm. The temperature of the cuvette chamber was then raised to about 10°C below the melting point of the DNA; the cuvettes were removed and opened to relieve pressure and eliminate bubbling, then the cuvettes were replaced and one hour was allowed for
equilibrium to be achieved. The absorbance at each of 15 wave lengths, 220 through 290 in 5 nm increments, was then measured after each temperature increase. Whenever possible, temperature intervals were varied, between 0.5 and 2.0°C, to maintain a more constant hyperchromic increase between intervals. Readings were taken at 11 to 25 temperature values, until increase in hyperchromicity was no longer noted.

**Buoyant density determinations**  The buoyant densities of DNA samples were determined using the Beckman-Spinco Model E analytical ultracentrifuge according to the procedures of Mandel, Schildkraut and Marmur (1968). The percent G+C was calculated from the equations of Schildkraut, Marmur and Doty (1962) and De Ley (1970). DNA from *Pseudomonas aeruginosa* (buoyant density 1.7266) was used as a reference.

**Characterization of the thermal transition curve and computations of spectral data**  The thermal transition curves were characterized by a statistical analysis (Krieg, 1968). The $T_m$, mean T, standard deviation, skewness and kurtosis for each curve was obtained to check and clarify spectral data. The $T_m$ and mean T are measures of the median and mean temperature of the thermal transition and are directly related to overall base composition; standard deviation yields an estimate of the range of heterogeneity of the DNA fragments; skewness is an indicator of regions that are very different from the overall base composition; and kurtosis measures homogeneity of base composition. These statistics were obtained for the thermal denaturation curves at 15 wave lengths, although results from 260 nm were considered to be most significant and representative.

The overall G+C content and concentration in moles/L of the DNA from
the denatured spectra, the overall G+C content and concentration in moles/L of the DNA from the hyperchromic spectra, the concentration ratio of hyperchromic to denatured DNA, and the fraction of A-T and G-C pairs denaturing at each temperature interval were calculated according to the equations and parameters of Felsenfeld (1968), Hirschman and Felsenfeld (1966), Felsenfeld and Hirschman (1965) and Hirschman, Gellert, Falkow and Felsenfeld (1967). The hyperchromic spectrum refers to those absorbance changes occurring while the DNA is progressing from the double stranded, helical configuration to the denatured or single stranded configuration. The denatured spectrum refers to the absorbance values of DNA in the single stranded configuration. The parameters of the denatured spectrum are used to measure the molar concentration and G+C content of all of the DNA in the sample. The parameters of the hyperchromic spectrum permit analysis of only that DNA which has denatured during a given temperature interval. Therefore, a comparison of the total molar concentration of DNA determined from each spectrum is indicative of the amount of DNA in the sample that is denatured before the initiation of analysis. Both a two term hyperchromic analysis (involving wave lengths 250, 260, 270, and 280 nm) and a three term analysis (involving wave lengths 220-290 in 5 nm increments) were done. The absorbance in all cases was corrected for the thermal expansion of the solvent.

All calculations were made with an IBM 360/65 computer by a program devised by W. R. Lockhart (personal communication).
RESULTS

The overall G+C contents of *Salmonella typhimurium*, Lambda bacteriophage, *Proteus morganii*, *Neisseria sicca* and *Escherichia coli* were obtained by buoyant density (De Ley, 1970) and thermal denaturation (Schildkraut and Lifson, 1965). The results of these studies along with those previously reported in the literature are presented in Table 1. (All tables and figures are presented in the following section). The values obtained during this investigation are within 2% G+C of those obtained in other laboratories.

There is also a correlation between results obtained from buoyant density and thermal denaturation experiments, although in this and other laboratories, the latter method yields a slightly higher G+C value when recently determined constants are employed (De Ley, 1970).

Either procedure, however, indicates that the variation of overall G+C content among the five organisms is less than 3%. Because of the near proximity of these overall G+C contents, there is a variation of less than 1.5°C in the mean melting temperatures. This overall compositional similarity notwithstanding, there is a clear difference in denaturation profiles which is indicative of intramolecular compositional heterogeneity. This is shown in Figure 1, where the change in absorbance at 260 nm (plots at wave lengths other than 260 nm show similar results) is plotted against temperature. If one looks at a single parameter, kurtosis (peakedness), it is evident that the denaturation curve of DNA from *E. coli* is most peaked, followed by those from *S. typhimurium*, *P. morganii* and *N. sicca*. This is an indication that the *E. coli*
chromosome possesses more regions of uniform base composition, while that of Neisseria comprises more regions of differing base composition. Furthermore, neither of these four bacterial chromosomes shows the multipeaked pattern characteristic of the sharp heterogeneic segmentation of temperate phage genomes (Falkow and Cowie, 1968).

Figure 2 illustrates change in absorption in DNA from E. coli cultivated and harvested under different conditions. A more peaked denaturation profile is seen in DNA isolated from logarithmically-growing E. coli (harvested with DNP) than in DNA isolated from E. coli either in the stationary phase (Figure 1) or growing at 0.40 or 2.23 g/hr (harvested with NA). This indicates that the chromosome of logarithmically grown E. coli has more regions of DNA at a given G+C content than E. coli harvested in the stationary phase of growth or than either of the cultures harvested with NA.

When one treats the thermal transition curves as cumulative frequency distributions, it is possible to quantitate the skewness, peakedness (or kurtosis) and standard deviation of each curve (Krieg, 1968). Table 2 shows these values for each of the ten types of experiments carried out during this investigation. All curves had a negative or left skewness value, indicative of more regions rich in A+T than those rich in G+C. Those of E. coli and S. typhimurium were most skewed, followed by P. morganii and N. sicca, with the Lambda ultraviolet absorbance curve being least skewed. E. coli had the most peaked values, followed by S. typhimurium, P. morganii, N. sicca and Lambda in decreasing order. The standard deviation is a measure of the melting range of DNA fragments.
A high standard deviation therefore is often an indicator of a broad thermal transition curve and consequently of a broad range of intramolecular G+C contents. A very high value would also appear in a multiply-peaked curve. The Lambda denaturation profile had the largest standard deviation, followed by *N. sicca, S. typhimurium, E. coli* and *P. morganii*.

The skewness, peakedness and standard deviation values of *S. typhimurium* are consistent with those previously reported by Krieg (1968), who obtained values of -0.9, 3.5, and 2.9 respectively.

Within *E. coli* samples, values for these parameters also vary. DNA from logarithmically grown, DNP-harvested *E. coli* is more skewed and more peaked than in other *E. coli* samples. The standard deviations are less than those of the stationary phase samples, which in turn are less than for samples inhibited with NA.

This comparative information suggests that the DNP-inhibited chromosomes of logarithmically grown *E. coli* have more regions of rich A+T content, larger (or more redundant) segments of uniform base composition, and a comparatively smaller percentage of regions with greatly varying G+C content than do the other *E. coli* samples. It also suggests that DNA from stationary phase *E. coli* has a smaller intramolecular variation of base content than DNA inhibited with NA. These results provided both a preliminary indication and a final check of results later obtained by spectral analysis.

It is also of interest to note that the skewness, peakedness and standard deviation values of DNA from logarithmically grown, DNP-inhibited
E. coli are comparable to those reported by Krieg (1968), which were 1.1, 4.1 and 2.6, respectively, for DNA extracted from E. coli 9637 during the late logarithmic phase of growth. The values for stationary-grown E. coli are comparable to those reported by Prouty (1970), which were 0.8, 3.6 and 2.7, respectively, for DNA extracted from slowly growing and stationary phase E. coli 9637.

Results obtained by spectral analysis yield a more quantitative measure of the findings. Lambda DNA was first characterized to standardize and check the technique. Figure 3 shows a plot of the fraction of all A-T pairs \( f_{AT} \) versus the fraction of all G-C pairs \( f_{GC} \) that denature at various temperature intervals during the melting of Lambda DNA. Although it is conceivable that a parabola can be drawn through the points, three linear regressions are also statistically satisfactory and yield more information about the Lambda chromosome. 19% of all the A-T pairs and 13% of all the G-C pairs denature in Region I, indicating that this region comprises 16% of the chromosome with an average G+C content of 41%. In Region II, 50% of the A-T pairs denature versus 47% of the G-C pairs, meaning that this region comprises 48.5% of the chromosome with an average G+C content of 49%. In Region III, 28% of the A-T pairs and 40% of the G-C pairs denature, indicating that 34% of the Lambda chromosome has a G+C content of 58%.

Although these spectral data do not show that Regions I, II and III are each largely single continuous portions of the DNA molecules, other evidence supports this premise. Hershey and Burgi (1965), doing density gradient centrifugation of sheared Lambda DNA, have shown that one region,
containing 10-20% of the DNA, has 41% G+C; a second region comprising 42% of the chromosome has a G+C content of 46%; and the third region, containing the remainder of the bases, has a G+C content of 56%.

Secondly, Inman's (1966) electron micrographs have shown that those parts of the molecule which denature first are located almost entirely in one portion of the chromosome. Thirdly, Szybalski et al. (1969) have shown that there are distinct regions containing 41, 47 and 57% G+C within the Lambda chromosome, and Falkow and Cowie (1968), doing spectral analysis, report regions of 44 and 58% G+C. As my results concur with these findings, both the validity and reproducibility of these spectral data are indicated.

Figure 4 shows the $f_{AT}$ - $f_{GC}$ plot for S. typhimurium and N. sicca DNA, and Figure 5 shows it for DNA from E. coli grown at 2.23 g/hr, harvested with DNP and also with NA. Each function is a composite of four experiments. This treatment of the data permits differentiation of bacteria and some qualitative estimate of intramolecular heterogeneity. However, it is evident that more subtle differences exist within bacterial DNA than Lambda DNA because bacterial DNA produces a parabola, indicating not three regions with very different average G+C content, but a large number of smaller areas, each of which differs slightly from any of the others in its nucleotide composition.

In order to characterize this smaller difference in the composition of successively melting regions, an additional calculation was performed. For each temperature interval of the thermal denaturation, the total fraction of DNA that was denatured and the G+C content of that fraction
were computed (Prouty, 1970). The fractions of DNA falling within a 3% G+C range were totaled for each individual experiment. Table 3 presents these representative data for four experiments with Lambda DNA. As shown in the last column of the table, the results of all experiments (computed by both a two-term and three-term analysis) were then averaged to estimate the mean percentage of total DNA contained in each fraction with a given mean molar G+C content.

Although, over a relatively broad compositional range, each individual determination mirrors the average of all determinations, it is clear that peaks are sometimes displaced, so that no two experiments yield identical results within the fine resolution of 3% G+C. This is attributed to differences from one analysis to another in the chosen temperature intervals and in the physical state of the DNA, and to minor fluctuations within the instrumentation. Nevertheless, when all of the results are averaged, each individual determination with DNA from Lambda or any of the other organisms reflects the summary values for a given class of DNA and can clearly be distinguished from results with another class of DNA. Furthermore, summary results from any five randomly chosen experiments with a given kind of DNA are reproduced by a summary of five different experiments with the same kind of DNA.

The ratio between the molar concentrations of DNA determined from the hyperchromic and denatured spectra, respectively, is a measure of the amount of denatured material present in the initial sample before analysis. This ratio ranged from 0.850 to 0.950 for most experiments. It is believed that there was less than 100% double-stranded DNA in each
sample before it was denatured thermally because strand separation had occurred during the extraction procedure and during storage in the phosphate buffer. It was noted, in fact, that the concentration ratio dropped about 0.015 for each day the DNA was dissolved in the 0.123M buffer. On the other hand, the DNA which was extracted from slowly growing cells harvested with NA yielded a slightly lower concentration ratio (0.790-0.900) than the other samples, and DNA from *E. coli* growing at 2.23 g/hr and harvested with NA yielded much lower concentration ratios (0.700-0.780). This finding is consistent with the hypothesis that NA acts directly on the DNA and results in its degradation. The presence of such degraded nucleotides or short fragments in the DNA sample could account for these observations.

Other computations and summaries like those shown in Table 3 provided the data for Figures 6 through 9. Each is based on from 6-12 determinations with DNA extracted under the conditions noted in column three of Table 1. These results are shown as histograms of the mean fraction of total DNA versus the mean molar percent G+C of each fraction. Figure 6 shows the patterns obtained for DNA from Lambda bacteriophage and stationary phase (NB) *E. coli*. The mole fractions shown for Lambda account for 97% of the total DNA and have a mean G+C content of 52.8%. This latter value is derived by multiplying each fraction of total DNA by its mean percent G+C and dividing by the sum of all the fractions of the total DNA. This value of 52.8% is close to the overall G+C values obtained by buoyant density (50.1%) and by thermal denaturation (51.4%) (Table 1). In all instances, these overall values, obtained by summing
the peaks, vary by less than 3% from the values obtained by other methods; this small variation may be attributed, in part, to the omission from the histograms of any fractions containing less than 0.5% of the total DNA. Note that the three major fractions of DNA appearing in the $f_{AT} - f_{GC}$ plot (Figure 3) are now resolved into smaller fractions of intramolecular heterogeneity in the Lambda bacteriophage histogram.

The fractions shown for stationary phase *E. coli* account for 90% of the total DNA and have a mean G+C content of 52.5%. 86% of all the "visible" genetic material falls within the relatively narrow range of 46 to 61% G+C, accounting for its aforementioned peakedness, while 12% of the chromosome falls within 31 to 43% G+C and 2% from 64 to 70% G+C, accounting for the large negative skewness. Lambda DNA, on the other hand, shown to be less skewed and less peaked, presents a pattern of more dispersed fractions. A similarly good correlation is found between all of these spectral patterns and all of the ultraviolet denaturation profiles.

Most importantly, it is shown in Figure 6 that the spectral patterns of a bacterium and a temperate phage are unique and distinct, as are patterns among bacteria with widely varying overall G+C contents (Prouty, 1970), as are patterns among "related" and unrelated bacteria with similar overall G+C contents (Figure 7).

The fractions of total DNA shown in the *S. typhimurium* histogram add up to 96% of the chromosome; those for *N. sicca*, 100%; and those for *P. morganii*, 96%. The remainder of the genetic material is found in fractions representing less than 1/2 of 1% of the total DNA. The overall G+C value, as determined from these histograms, is 54.0% for the
S. typhimurium, 54.2% for N. sicca, and 53.2% for P. morganii.

It can be seen that the patterns of the more closely related enteric bacteria (E. coli, S. typhimurium, P. morganii) are more similar to one another (though not nearly identical) than to those of N. sicca or Lambda. For example, although each of the enteric chromosomes has increasingly large fractions of DNA starting at 49% G+C, stationary phase cultures of E. coli have the largest peak (19% of the total DNA) at 58% G+C, those of S. typhimurium show a moderate decrease of genetic material at 58% G+C, while those of P. morganii show a sharp decrease of genetic material, with only an 8% fraction appearing at 58% G+C. These studies would indicate that spectral analysis provides fine enough resolution to conduct taxonomically oriented research and indicates possibilities of detecting intramolecular heterogeneity among DNA samples isolated from the same species of bacteria grown at different rates, under different conditions and inhibited by different chemicals.

Figure 8 shows the fraction of total DNA versus the mean G+C content for E. coli harvested with DNP while growing at 2.30 g/hr in a nutrient broth batch culture and for E. coli growing at 2.23 g/hr in continuous culture in brain heart infusion. Postulated average chromosomal replication positions, based on the statistical treatments of Sueoka and Yoshikawa (1965), are shown above each histogram. It should be pointed out here that, though it is possible to calculate the average replication positions of chromosomes of many cells, chromosomes from the individual cells may be of any configuration in these balanced cultures. Better resolution would seem possible through the use of synchronously growing
populations, where, theoretically, the positions of all replication forks coincide. This advantage is negated, however, when weighed against the difficulties of achieving high growth rates and large cell populations in synchronous systems. Accordingly, balanced systems were employed exclusively. Although culture conditions vary slightly, the two patterns in Figure 8 are nearly identical, as are the predicted chromosomal configurations. The histogram of the batch-grown *E. coli* accounts for 99% of the chromosome with a calculated overall G+C content of 51.6%, while that of the continuously grown *E. coli* shows 98% of the total fractions and has a calculated overall G+C content of 51.7%

Figure 9 shows the fraction of total DNA versus the mean percent G+C for *E. coli* grown in a batch BHI culture and harvested in stationary phase, and for *E. coli* grown in continuous culture at 0.40 g/hr and at 2.23 g/hr, harvested with NA. The fractions of total DNA shown in the stationary phase culture add up to 96% of the chromosome; those in the 0.40 g/hr culture, 97%; and those in the 2.23 g/hr culture, 100%. The mean G+C contents are 52.4, 53.3, and 53.5% respectively. The broken lines on the predicted average chromosome configuration indicate possible areas of degradation by naladixic acid.

The results in Figures 8 and 9 show that the *E. coli* DNA at the origin, where replication is initiated, is richer in A-T, while the terminus contains more G-C base pairs. As a result, the calculated overall G+C for *E. coli* increases as the ratio of terminus DNA to origin DNA increases.

In summation, the ten histograms can be more readily compared (Table
5) if one arbitrarily labels those peaks containing 7-11% of the total DNA as "small", those with 12-17% as "medium" and those with 18% or more as "large". These patterns of compositional intramolecular heterogeneity among the five different organisms are distinct and unique and can serve in their identification.

There are also differences within a single organism when it is grown at different rates or when a specific part of its chromosome is allowed to be degraded. When contrasted with stationary phase cultures, the increased concentrations of 52 and 55% G+C regions in the chromosomes of rapidly growing *E. coli* apparently reflect a four fold duplication of portions of their chromosomes. The "loss" of peaks at 43 and 46% in the 0.40 g/hr, NA inhibited cells is coincidental with a loss of specific regions of the chromosome, as is the absence of a large peak at 55% G+C in NA inhibited, rapidly growing cells.

It has therefore been shown with repeated determinations with DNA from *Escherichia*, *Salmonella*, *Neisseria*, *Proteus* and Lambda bacteriophage that the chromosome from a given organism has a characteristic distribution of portions with various nucleotide contents and that this distributional heterogeneity is a unique and distinct property which can be mapped.
TABLES AND FIGURES
Table 1. The overall base ratios (% G+C) of DNA isolated from the test organisms, as determined by buoyant density and thermal denaturation

<table>
<thead>
<tr>
<th>Organism</th>
<th>Previously reported % G+C</th>
<th>this study&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% G+C By mean T&lt;sub&gt;d&lt;/sub&gt;</th>
<th>% G+C By p&lt;sub&gt;c&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>By T&lt;sub&gt;m&lt;/sub&gt; Reference</td>
<td>By p&lt;sub&gt;c&lt;/sub&gt; Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Salmonella typhimurium</strong></td>
<td>54.6 De Ley (1970)</td>
<td>52.3 Schildkraut et al. (1962)</td>
<td>54.1</td>
<td>51.1</td>
</tr>
<tr>
<td></td>
<td>52.9 Krieg (1968)</td>
<td>51.1 Krieg (1968)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lambda bacteriophage</strong></td>
<td>49.7 De Ley (1970)</td>
<td>50.1 De Ley (1970)</td>
<td>51.4</td>
<td>50.1</td>
</tr>
<tr>
<td></td>
<td>49.7 Falkow &amp; Cowie (1968)</td>
<td>48.0 Hirschman et al. (1967)</td>
<td>49.2 Falkow &amp; Cowie (1968)</td>
<td></td>
</tr>
<tr>
<td><strong>Proteus morganii</strong></td>
<td>52.8 De Ley (1970)</td>
<td>50.1 De Ley (1970)</td>
<td>52.7</td>
<td>50.1</td>
</tr>
<tr>
<td><strong>Neisseria sicca</strong></td>
<td>50.7 De Ley (1970)</td>
<td>50.1 De Ley (1970)</td>
<td>54.3</td>
<td>50.1</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td>51.5 De Ley (1970)</td>
<td>50.1 De Ley (1970)</td>
<td>52.4</td>
<td>51.1</td>
</tr>
<tr>
<td></td>
<td>52.3 Krieg (1968)</td>
<td>52.4 Krieg (1968)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Each value reported represents an average of at least 6 determinations. Bacterial DNA for these determinations comes from cells harvested in the early stationary phase of growth.

<sup>b</sup>The median melting temperatures (T<sub>m</sub>) reported in the literature are converted to the mole percentage of guanine plus cytosine (% G+C) according to the formula: % G+C = (T<sub>m</sub>-69.37)/0.41 (De Ley, 1970).

<sup>c</sup>The relationship between p (buoyant density) and the % G+C is as follows: % G+C = 1038.47 (p-1.6616) (De Ley, 1970).

<sup>d</sup>The relationship between the mean T (mean melting temperature) and the % G+C is: % G+C = (mean T - 16.6 log M) - 81.5/0.41, where M is molar concentration of solvent (Schildkraut and Lifson, 1965).
Table 2. The mean skewness, mean kurtosis, and mean standard deviation for thermal transition curves (at 260 nm) of DNA isolated from the test organisms

<table>
<thead>
<tr>
<th>Organism</th>
<th>Culture conditions</th>
<th>No. of exp.</th>
<th>Skewness b</th>
<th>Kurtosis b</th>
<th>Standard deviation b</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>2.23 g/hr BHI NA</td>
<td>10</td>
<td>-0.81</td>
<td>3.60</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td>continuous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>0.43 g/hr BHI NA</td>
<td>10</td>
<td>-0.82</td>
<td>3.59</td>
<td>2.81</td>
</tr>
<tr>
<td></td>
<td>continuous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>Stationary NB</td>
<td>3</td>
<td>-0.83</td>
<td>3.56</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>Stationary BHI DNP</td>
<td>6</td>
<td>-0.79</td>
<td>3.59</td>
<td>2.69</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>2.23 g/hr BHI DNP</td>
<td>6</td>
<td>-0.91</td>
<td>3.91</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>continuous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>2.30 g/hr NB DNP</td>
<td>8</td>
<td>-1.01</td>
<td>3.98</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. typhimurium</td>
<td>Stationary NB DNP</td>
<td>6</td>
<td>-0.85</td>
<td>3.52</td>
<td>2.88</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. sicca</td>
<td>Stationary BHI + YE</td>
<td>6</td>
<td>-0.64</td>
<td>3.00</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. morganii</td>
<td>Stationary NB</td>
<td>6</td>
<td>-0.67</td>
<td>3.23</td>
<td>2.69</td>
</tr>
<tr>
<td></td>
<td>batch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambda</td>
<td></td>
<td>5</td>
<td>-0.39</td>
<td>2.43</td>
<td>3.26</td>
</tr>
</tbody>
</table>

aBHI, NB and YE refer to the growth media (brain heart infusion, nutrient broth and yeast extract) in which the cells were grown at the specific rate noted. In several of the experiments the bacteria were harvested with either naladixic acid (NA) or 2,4-dinitrophenol (DNP).

bThe skewness (3rd standard moment), kurtosis (4th standard moment), and standard deviation are defined by Krieg (1968).
Table 3. Base composition and percent fraction of the total DNA denatured during successive temperature increases, as determined by two-term and three-term hyperchromic spectral analyses in a series of experiments with DNA isolated from Lambda bacteriophage. The final column shows a composite percentage distribution of DNA in each fraction, obtained by averaging all two- and three-term results from the individual experiments.

<table>
<thead>
<tr>
<th>Range</th>
<th>Mean</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
<th>Experiment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% GC</td>
<td>Two-term % DNA</td>
<td>Three-term % DNA</td>
<td>Two-term % DNA</td>
<td>Three-term % DNA</td>
</tr>
<tr>
<td>30-32</td>
<td>31</td>
<td>0 2</td>
<td>3 2</td>
<td>5 8</td>
<td>8 0</td>
</tr>
<tr>
<td>33-35</td>
<td>34</td>
<td>0 0</td>
<td>0 6</td>
<td>0 0</td>
<td>0 8</td>
</tr>
<tr>
<td>36-38</td>
<td>37</td>
<td>0 0</td>
<td>0 0</td>
<td>0 13</td>
<td>2 17</td>
</tr>
<tr>
<td>39-41</td>
<td>40</td>
<td>7 0</td>
<td>0 0</td>
<td>0 5</td>
<td>0 0</td>
</tr>
<tr>
<td>42-44</td>
<td>43</td>
<td>16 0</td>
<td>7 0</td>
<td>0 13</td>
<td>0 13</td>
</tr>
<tr>
<td>45-47</td>
<td>46</td>
<td>0 9</td>
<td>17 8</td>
<td>10 0</td>
<td>5 0</td>
</tr>
<tr>
<td>48-50</td>
<td>49</td>
<td>0 11</td>
<td>8 0</td>
<td>12 0</td>
<td>22 0</td>
</tr>
<tr>
<td>51-53</td>
<td>52</td>
<td>21 0</td>
<td>11 12</td>
<td>6 3</td>
<td>0 10</td>
</tr>
<tr>
<td>54-56</td>
<td>55</td>
<td>0 12</td>
<td>14 11</td>
<td>9 48</td>
<td>0 25</td>
</tr>
<tr>
<td>57-59</td>
<td>58</td>
<td>34 13</td>
<td>13 13</td>
<td>13 0</td>
<td>13 12</td>
</tr>
<tr>
<td>60-62</td>
<td>61</td>
<td>13 0</td>
<td>16 0</td>
<td>13 0</td>
<td>21 5</td>
</tr>
<tr>
<td>63-65</td>
<td>64</td>
<td>9 30</td>
<td>0 30</td>
<td>0 0</td>
<td>10 11</td>
</tr>
<tr>
<td>66-68</td>
<td>67</td>
<td>0 9</td>
<td>0 0</td>
<td>8 8</td>
<td>5 0</td>
</tr>
<tr>
<td>69-71</td>
<td>70</td>
<td>0 0</td>
<td>9 12</td>
<td>0 13</td>
<td>0 8</td>
</tr>
</tbody>
</table>
### Table 4. Relative sizes of the fractions of total DNA of a given average mole composition (% G+C): a comparative study of 10 sets of experiments

<table>
<thead>
<tr>
<th>Organism</th>
<th>Culture conditions</th>
<th>Sizes of DNA fractions with avg mole composition (% G+C) of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambda</td>
<td></td>
<td>43 46 49 52 55 58 61 64 67</td>
</tr>
<tr>
<td>N. sicca</td>
<td>Stationary BHI + YE batch</td>
<td>S S S S S S S S S S</td>
</tr>
<tr>
<td>S. typhimurium</td>
<td>Stationary NB DNP batch</td>
<td>S M L M M M M M M M M</td>
</tr>
<tr>
<td>P. morganii</td>
<td>Stationary NB batch</td>
<td>S S S M L S S S S S</td>
</tr>
<tr>
<td>E. coli</td>
<td>2.30 g/hr NB DNP batch</td>
<td>S M L L M M M M</td>
</tr>
<tr>
<td>E. coli</td>
<td>2.23 g/hr BHI DNP continuous</td>
<td>S M L L M M M M M</td>
</tr>
<tr>
<td>E. coli</td>
<td>Stationary NP batch</td>
<td>S M M M M M L S</td>
</tr>
<tr>
<td>E. coli</td>
<td>Stationary BHI DNP batch</td>
<td>S M M M M M S</td>
</tr>
<tr>
<td>E. coli</td>
<td>0.40 g/hr BHI NA continuous</td>
<td>M M M M M M S</td>
</tr>
<tr>
<td>E. coli</td>
<td>2.23 g/hr BHI NA continuous</td>
<td>M M M M S</td>
</tr>
</tbody>
</table>

*BHI, NB and YE refer to the growth media (brain heart infusion, nutrient broth and yeast extract) in which the cells were grown at the specific rate noted. In several of the experiments the bacteria were harvested with either naladixic acid (NA) or 2,4-dinitrophenol (DNP).

*A small peak (S) represents 7-11% of the total DNA; a medium-sized peak (M) represents 12-17% of the total DNA; and a large peak (L) represents 18% or more of the total DNA.*
Figure 1. The absorbance change at 260 nm which occurred during each temperature interval during the denaturation of DNA isolated from stationary phase *Escherichia coli* (●), *Salmonella typhimurium* (■), *Neisseria sicca* (□), and *Proteus morganii* (○).
Figure 2. The absorbance change at 260 nm which occurred during each temperature interval during the denaturation of DNA isolated from *Escherichia coli* grown at 2.3 g/hr and harvested with 2,4-dinitrophenol (■), *E. coli* grown at 2.23 g/hr and harvested with naladixic acid (△), and *E. coli* grown at 0.40 g/hr and harvested with naladixic acid (●).
CHANGE IN ABSORBANCE 260 nm

TEMPERATURE °C

CHANGE IN ABSORBANCE 260 nm

0.08 0.16 0.24 0.32 0.40 0.48 0.56 0.64
Figure 3. Fractions of all adenine-thymine ($f_{AT}$) and guanine-cytosine ($f_{GC}$) pairs denatured at successive stages in the thermal transition of DNA isolated from Lambda bacteriophage. The three regions of internal heterogeneity of denaturation are identified by Roman numerals.
Figure 4. Fractions of all adenine-thymine ($f_{\text{AT}}$) and guanine-cytosine ($f_{\text{GC}}$) pairs denatured at successive stages in the thermal transition of DNA isolated from *Neisseria sicca* (□) and *Salmonella typhimurium* (■).
Figure 5. Fractions of all adenine-thymine ($f_{AT}$) and guanine-cytosine
($f_{GC}$) pairs denatured at successive stages in the thermal
transition of DNA isolated from *Escherichia coli* grown at
2.23 g/hr and harvested with 2,4-dinitrophenol (■) and
*E. coli* grown at 2.23 g/hr and harvested with naladixic
cacid ( □ )
Figure 6. Distribution of the fractions of total DNA according to their mean G+C content as determined by spectral analysis. 

Upper histogram shows DNA isolated from Lambda bacteriophage. Lower histogram shows DNA isolated from Escherichia coli which was grown in batch culture in nutrient broth and harvested in stationary phase.
Figure 7. Distribution of the fractions of total DNA according to their mean G+C content as determined by spectral analysis. Upper histogram shows DNA isolated from *Salmonella typhimurium*. Middle histogram shows DNA isolated from *Neisseria sicca*. Lower histogram shows DNA isolated from *Proteus morganii*
Figure 8. Distribution of the fractions of total DNA according to their mean G+C content as determined by spectral analysis. Upper histogram shows DNA isolated from *Escherichia coli* which was growing at 2.30 g/hr in nutrient broth in batch culture. Lower histogram shows DNA isolated from *E. coli* which was growing at 2.23 g/hr in brain heart infusion in continuous culture. Both cultures were harvested with 2,4-dinitrophenol. Postulated average chromosomal configurations are shown with each histogram.
Figure 9. Distribution of the fractions of total DNA according to their mean G+C content as determined by spectral analysis. Upper histogram shows DNA isolated from *Escherichia coli* in the stationary phase of growth in batch culture in brain heart infusion. Middle histogram shows DNA isolated from *E. coli* which was growing at 0.40 g/hr in brain heart infusion in continuous culture. Lower histogram shows DNA isolated from *E. coli* which was growing at 2.23 g/hr in brain heart infusion in continuous culture. The latter two cultures were harvested with naladixic acid. Postulated average chromosomal configurations are shown with each histogram. The broken lines in these configurations represent possible areas of degradation by naladixic acid.
DISCUSSION

The results of these experiments are in agreement with a number of findings and generally accepted theories. First of all, it is clear that NA not only inhibits replication of the *E. coli* chromosome, but also deletes portions of its base content. This is apparent when one compares spectral patterns of DNA from rapidly growing *E. coli* harvested with DNP on the one hand and harvested with NA on the other. As it has been substantiated that DNP does not alter chromosomal structure (Howland and Hughes, 1969), NA would therefore be the agent responsible for discrepancies between the two patterns. Ramareddy and Reiter (1969) have recently found that *Bacillus subtilis* cells degrade and lose from 20 to 30% of their DNA in 3.5 hr of exposure to 20 µg of naladixic acid per ml and that degradation appears to be initiated at the replication point and to proceed sequentially along the chromosome from the most recently synthesized DNA to older DNA, with both strands degraded equally. However, at this time, it is not clear whether the inhibited DNA is degraded by nuclease activity or directly by the antibiotic; it is not known whether degradation proceeds from one or all forks in a multiforked chromosomal structure; and the reaction kinetics of the degradative process have yet to be elucidated. From these spectral data, it appears that some portions of the chromosome near the origin of replication are not degraded in multiforked structures, because the 52% G+C peak increases sharply as the growth rate is increased and is not lost upon exposure to NA. This could be interpreted to mean that either degradation proceeds solely from the oldest fork (\[\text{\textbullet}---\text{\textbullet}\]) (although it is difficult to
envision a mechanism of such specificity) or that upon a 3.5 hr exposure to the antibiotic, degradation does not proceed to completion in the longest of the forks (\[ \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_]\). 

To choose between these alternatives one would work in synchronous culture, elucidate degradation kinetics, and study several more growth rates. Such experiments would be worthwhile, for controlled degradation of the chromosome would be an extremely useful tool in genetic mapping.

Secondly, in basing part of the experimental design of the investigation on the theory that bacterial DNA is multiply forked at growth rates above 1.5 g/hr, it is found that the results substantiate the premise. *E. coli* chromosomes of cells that are growing faster than 2.2 g/hr display very different patterns than the chromosomes of stationary phase cells, and preliminary studies of chromosomes of cells growing at 1.5 g/hr present a pattern that is intermediate between the former and the latter. The presence of dichotomous (forked) molecules at the highest rate would certainly account for these results.

Thirdly, and in a similar vein, it is thought that the average size and number of replication forks in bacteria is dependent on the time between cell divisions. This had previously been determined by a number of physical, chemical and genetic methods. Findings consistent with this hypothesis have been obtained in autoradiographic studies (Lark and Lark, 1965), tracer incorporation studies in synchronously growing cells (Helmstetter and Cooper, 1968; Cooper and Helmstetter, 1968; Helmstetter and Pierucci, 1968; Lark, 1966b), marker frequency analysis (Sueoka and Yoshikowa, 1965; Yoshikowa *et al.*, 1964), cell division studies in
replication-inhibited cells (Clark, 1968), and calculations of DNA content per nucleus of cells in balanced growth (Maaløe and Kjeldgaard, 1966; Bleecken, 1969). This study provides additional evidence, in showing that the intramolecular compositional heterogeneity of a bacterial chromosome varies as the growth rate is varied.

Fourthly, it is shown that DNA from *E. coli*, harvested at similar rates of growth, has similar intramolecular heterogeneity regardless of whether the culture conditions are continuous or batch, or whether the medium is nutrient broth or brain heart infusion. This is consistent with previous findings that at a given temperature, and assuming no substrate or energy limitations, the rate of DNA synthesis per replication point is constant irrespective of culture conditions or choice of medium (Maaløe and Kjeldgaard, 1966). One prediction of this hypothesis is that at similar growth rates, average chromosomal configurations are similar (Helmstetter and Cooper, 1968). This "one variable hypothesis" is supported by these experiments.

On the other hand, it has been shown that a nutrient-limited medium results in the slowing of DNA replication (Lark, 1966a). Prouty (1970), growing *E. coli* 9637 in continuous culture in carbon and nitrogen-limited medium at rates of 0.25 and 0.75 g/hr, found an increase of chromosomal material between 48-54% G+C at the faster growth rate. This is in agreement with my experiments where the fraction of total DNA within a similar G+C range also increases when
the cells are grown more rapidly. In fact, in both the 0.75 g/hr limited medium DNA sample and the 2.23 g/hr complete medium sample, 35% of all the *E. coli* chromosome has a G+C content between 49 and 54%. In stationary phase cells this figure falls below 25%. Whether or not multiple forking also occurs at 0.75 g/hr in carbon and nitrogen limited cells should depend on whether the time required for a round of replication is raised from 41 minutes to about 123 minutes by the substrate limitation. However, because the quantitative effect of substrate limitation on DNA synthesis is at this time conjecture, Prouty's average chromosomal configurations cannot be calculated, and a detailed comparison of the two studies is not yet fully possible. It is, however, suggested that in both the 0.75 g/hr nutrient-limited cells and the 2.23 g/hr non-limited cells there is a higher ratio of origin to terminus DNA and that the base content nearer the origin is richer in A+T.

It is possible to suggest an extension of this type of map by a further evaluation of the results obtained in Figures 8 and 9. In each given set of spectral analyses of *E. coli*, the fractions of total DNA display very little variation below 43% G+C and above 67% G+C. At 45-50% G+C, a significant variation in the fraction of total DNA is noted in the 0.40 g/hr NA sample. In the batch stationary BHI sample, for example, 24% of the total DNA is contained within this range, while in the sample where the origin fraction was degraded this value drops to 9%. Thus, based on the estimated chromosome configuration, approximately a sixth of the chromosome closest to the origin is rich in 45-50% G+C.

Secondly, the most rapidly growing cells contain 43% of their total
DNA within a range of 51-56% G+C, while those in stationary phase contain 30% of their total DNA within this range. This increase of 25% in the multiply-forked molecules over the non-forked molecules indicates that the 51-56% G+C region is located somewhere near the middle of the chromosome, composing about a third of it.

If this region were located near the origin, the discrepancy between the peak sizes would be greater than that determined, because of the 4:1 ratio in copies. If, on the other hand, the 51-56% G+C region were nearer the terminus, this fraction of the total DNA would be higher in the stationary phase cells than the logarithmically grown cells because a dichotomous molecule has a comparatively smaller percentage of its DNA at the terminus than do the non-forked structures.

This same rationale is applied to regions of between 57 and 64% G+C. 15% of the total DNA of the most rapidly growing cells is found between 57 and 64% G+C. The rest of these values are: 25% for stationary phase cells; 34% for the 0.40 g/hr, NA sample; and 36% for the 2.23 g/hr NA sample. Thus, it would appear that the terminus portion of the chromosome is partially composed of fragments with an average G+C content ranging from 57 to 64%.

It is also interesting to note that upon exposure to NA, the multiply forked molecule practically loses its 55% G+C region of intramolecular heterogeneity. There is a drop from a fraction consisting of 20% of the chromosome to a fraction consisting of 4% of the chromosome. It is not known whether this region is completely degraded or whether enough of it is degraded to change its overall G+C content sufficiently to displace the
peak in spectral studies.

It is furthermore likely that these variations in the intramolecular heterogeneity of DNA within a given organism are due to a change in average chromosomal configuration rather than to a gain or loss in extrachromosomal elements. Because plasmids are about 0.01 the size of a bacterial chromosome, they would have to be present in multiple copies to effect more than a 1% change in the spectral pattern. This strain of E. coli W is characterized as being F\(^{-}\), non-lysogenic and non-colicinogenic, and no plasmid with a different overall G+C content than the chromosome was isolated in buoyant density determinations.

In addition to the spectral patterns, the shapes of the denaturation profiles at 260 nm and the slight changes in the overall G+C contents among the various E. coli DNA samples are each indicative of a chromosome richer in A+T at the replication origin. For example, the denaturation profiles of DNA from logarithmically-grown E. coli are more negatively skewed than are any of the other samples, indicating more A+T rich regions in the multiply forked molecule. Also, as the growth rate is decreased or as NA is allowed to degrade portions of the chromosome near the origin, the overall G+C content increases by up to 2% over that of rapidly growing cells. It is interesting that, in the "trifragment" map presented here, the intramolecular G+C content increases as the terminus is approached.

A more exact, and perhaps finer map of intramolecular heterogeneity could be obtained by the analysis and comparisons of DNA from many other growth rates. By these techniques it is conceivable that the chromosome can eventually be sequentially divided into tenths (with each fragment
containing about $9 \times 10^5$ nucleotide pairs) and the overall base content of each ordered fragment identified.

The foundations of this technique rest on two premises: that the chromosome configuration can be predicted for each given set of cultural conditions and that there is intramolecular heterogeneity within the bacterial DNA under study. The first condition can be met by using any one of a number of mathematical models each of which is in essential agreement with the others (Sueoka and Yoshikawa, 1965; Cooper and Helmstetter, 1968). In each model there are essentially two constants: the time for a replication point to traverse the chromosome and the time between the end of a round of replication and cell division. The first constant is known for cells growing with no substrate limitation and, therefore, the second can be calculated for cells in a given state of balanced growth in continuous or batch culture (where an average chromosome configuration is determined), and more precisely perhaps in synchronously growing cultures where the chromosomes of each cell are presumably in the same replication state as those of its sister cells.

Each of these models rests on experimental data that DNA replication occupies a constant time and that the remainder of the bacterial life cycle of the cell is a "rest period" for DNA synthesis. A recent and lucid review of this model and the means by which one may determine chromosomal configurations is presented by Pierucci (1969).

As for the second premise, it has been shown that many species of DNA, including that of *E. coli* (Votavova, 1968; McCarthy, 1965), contain some degree of intramolecular heterogeneity. Miyazawa and Thomas (1965)
analyzed *E. coli* K12 DNA, sheared to $5 \times 10^5$ daltons, by fractional thermal elution from hydroxyapatite columns and found that individual fractions range from 44 to 57% G+C. Falkow and Cowie (1968) showed that one half of the *E. coli* K12 chromosome contains 47% G+C while the other half contains 55% G+C. Yamagishi (1970) sheared *E. coli* K12 DNA down to 1900 nucleotide pairs long ($1.3 \times 10^6$ daltons) and from the buoyant density of the mercury complexes found segments ranging from 39 to 56% G+C. He additionally found that segments containing an average G+C content of about 39% compose 3.3% of the total DNA.

Similar results are found in this investigation. The distribution patterns reflect an A+T rich area, as does the frequency distribution analysis where there are large negative skewness values. The histograms drawn from the spectral data further indicate that from 2 to 4% of the *E. coli* chromosome has an average G+C content of 39 to 41%.

In this investigation, however, about 15% of the DNA from stationary phase *E. coli* has an intramolecular base content beyond the limits observed by Votavova (35-56% G+C), Miyazowa (44-57% G+C) and Yamagishi (39-56% G+C). This apparent discrepancy could be due to a difference in *E. coli* strains but it is most likely to be a function of resolution. Rolfe and Meselson (1959) and Sueoka et al. (1959) were unable to find any intramolecular heterogeneity at all in *E. coli* DNA (sheared to $3 \times 10^8$ daltons) by equilibrium centrifugation in CsCl. Yamagishi (1970) also found none, working at a size of $5 \times 10^7$ daltons, but did find some intrachromosomal G+C variation among fragments of $2 \times 10^7$ daltons. Both Miyazawa and Yamagishi found a great deal more intramolecular heterogeneity among fragments in the $10^6$ dalton range. For *E. coli* K12 it would appear,
then, that fragments consisting of about $10^5$ base pairs are uniform in base content, while those of about $3 \times 10^4$ begin to vary, and that those consisting of about $5 \times 10^3$ base pairs are potentially very different from one another with respect to their base content.

It therefore becomes important to ascertain the limits of resolution that one can expect to obtain by spectral analysis. In the interval between any two temperatures it is possible to determine a 0.3% change in hyperchromicity. This could be interpreted to mean that 0.3% of all the molecules ($10^7$ molecular weight) denature completely within the given interval, or more likely that small regions of many molecules denature.

If, for example, 0.3% of each fragment denatures, then the resulting coil would represent a $3 \times 10^5$ dalton region and would contain less than 3000 nucleotide pairs. The average G+C content of these 3000 or so pairs would then be calculated by the parameters set forth by Felsenfeld and Hirschman (1965). It is, therefore, possible that the apparently greater resolution of spectral analysis over chromatographic and buoyant density methods accounts for the slightly wider range of intramolecular heterogeneity noted in this and in a previous study (Prouty, 1970).

It is interesting to note that intramolecular heterogeneity in *E. coli* first becomes apparent at chromosomal lengths just large enough to accommodate the size of an operon (Martin, 1963). If, indeed, boundaries between DNA segments of unlike composition coincide with genetic units of coordinated transcription, common ancestral operons might be envisioned.

A comparison of DNA from the taxonomically related *E. coli*, *P. morganii* and *S. typhimurium* reflects some similarities as well as the
differences already noted. *Proteus* and *Salmonella* both contain the largest fraction of total DNA at 54-56% G+C, as does the stationary phase sample of BHI-grown *E. coli*. In all three organisms at least 55% of their chromosomes have an intramolecular G+C content of between 48 and 59%. *Neisseria*, on the other hand, has 42% and Lambda 44% of their respective chromosomes within this range. Prouty (1970) shows these same values as 10% for *Pseudomonas aeruginosa* and less than 3% for *Staphylococcus aureus*. These and other comparisons of the patterns of different bacteria show that the results obtained by spectral analysis may be of taxonomic value. Similar patterns of intramolecular heterogeneity could be the mark of similar functional units common to related bacteria (Taylor and Trotter, 1967; Sanderson, 1967), and differences in these patterns could serve to "fingerprint" and thereby differentiate closely and distantly related microorganisms.
LITERATURE CITED


I wish to express gratitude to my major professor, Dr. William R. Lockhart, for his counsel and encouragement throughout this program; to the graduate faculty for their guidance; to Dr. Peter A. Pattee and David L. Berryhill for assisting with buoyant density determinations; to my wife, Marilyn, for her devotion; and to the National Science Foundation for their support of this project.