High Resolution Thermal Denaturation of DNA: Thermalites of Bacteriophage DNA†

Douglas L. Vizard* and Allen T. Ansevin

ABSTRACT: High resolution thermal denaturation profiles are presented for the DNAs of bacteriophages λ and T7. It is concluded that the temperature increment in data gathering and the method of calculating results meet the requirements for quantitative recording of the large amount of information found in the thermal transitions of both DNAs. The high resolution derivative denaturation profiles of these bacteriophage DNAs demonstrate that individual subtransitions (thermalites) of natural DNA are Gaussian in form and have narrow transition widths. Curve resolution performed on these profiles indicates that the mean thermalite width (2σ) is 0.33 °C and that this breadth is relatively invariant. Transition widths are not influenced by the position of thermalites in the profile or by cation concentration in the range from 5 to 30 mM Na^+. However, the relative position of thermalites within a denaturation profile is a function of the solution ionic strength. The distribution of lengths of the DNA sequences which these thermalites represent is broad, with a number average length of 900 base pairs. Although we find an approximate similarity between the number of thermalites in the denaturation profile of T7 DNA and the number of looping regions in the electron microscopic partial denaturation map of Gomez and Lang ((1972), *J. Mol. Biol. 70,* 239–251) we conclude that free solution thermal denaturation experiments can be compared only superficially to the mapping results.

It has been apparent for some time that DNA hyperchromicity data can provide substantially more information than was revealed by early thermal denaturation experiments. Although there was some theoretical basis for expecting to find additional information (Crothers, 1968), only in very recent investigations (Reiss and Michel, 1974) has the extent of detail in a melting curve been appreciated. The study of Ansevin et al. (1975) discusses the thermal resolution and data treatment required to extract the information that is implicit in a faithful recording of absorbance changes during the thermal transition of DNA. By conforming to the principles of that study, it should be possible to characterize quantitatively the properties of the subtransitions (thermalites) which compose the thermal denaturation profiles of naturally occurring DNAs.

This research presents the derivative thermal denaturation profiles of the relatively simple DNAs of bacteriophage T7 and λ. These results illustrate the detail to be found in high resolution profiles for DNAs of relatively low genetic complexity: the T7 DNA molecule is 24×10^6 daltons or 36 000 base pairs (Gray and Hearst, 1968), and the λ DNA molecule is 31×10^6 daltons or 47 000 base pairs (Davidson and Szybalski, 1971). Because of the large number of thermalites present in natural DNAs, it will be seen that a full appreciation of the information content of denaturation profiles can be gained only by beginning with studies on DNA of this or even smaller size.

Among the investigations with which our experiments might be compared are denaturation results for λ DNA that range in complexity from a broad bi-modal transition of Hirschman et al. (1967), through the 3 or 4 components demonstrated by Falkow and Cowie (1968), to the very recent report of Reiss and Michel (1974) showing approximately 14 components. Our results for λ DNA imply a minimum of 34 thermal subcomponents. The melting curve of T7 DNA, in distinction to that of λ, is relatively narrow and was taken in the past as an example of a particularly simple transition. The T7 profile, with its overlapping subtransitions, serves as a suitable complement to the broader λ profile and provides an opportunity to compare our results for free solution, equilibrium denaturation conditions to the nonequilibrium denaturation map of T7 DNA derived from the electron microscopic studies of Gomez and Lang (1972).

Our results indicate that the thermal denaturation profiles of both DNAs are composed of a number of relatively discrete Gaussian-shaped thermalites having widths that are at least as narrow as those of simple paired homopolymers or copolymers. We find that the relative stabilities of some thermalites vary as a critical function of the salt concentration of the solution, which suggests that the stability of thermalites depends not only upon the average base composition, but also upon the nucleotide sequence. However, the thermal widths of the transitions do not appear to vary significantly with salt concentration. Finally, this study shows that the length distribution of the DNA sequences that can be derived from a quantitative evaluation of the mass fractions of thermalites appears to be a physical accounting of approximately gene-sized elements of these bacteriophage DNAs.

Materials and Methods

(*a*) *DNA Preparation.* Bacteriophages T7 and λ, kindly provided by Dr. R. R. Hewitt of this institution, were twice deproteinized by warm (35 °C) buffer-saturated phenol. The Tris salts of the DNAs were alcohol precipitated, alcohol washed, dried in vacuum, and redissolved in 10^{-3} Tris, 10^{-4} M EDTA, pH 8. The preparations were stored at −70 °C. Both phage DNA preparations were undegraded as assessed by analytical sedimentation at neutral pH. Alkaline sedimentation studies indicated that the single strands of λ were undegraded; one preparation of T7 DNA appeared to

† From the Department of Physics, The University of Texas System Cancer Center, M. D. Anderson Hospital and Tumor Institute, 6723 Bertner Avenue, Houston, Texas 77025. *Received August 11, 1975.* This is the second paper of a series.

have approximately 3 strand interruptions per double-stranded molecule.

(*b*) *Sample Preparation for Denaturation.* Solutions to be denatured were prepared by appropriate volume additions of DNA in the dilute Tris–EDTA buffer and NaCl solutions. This fact is important at lower ionic strengths (e.g., 5 mM) because there is a significant difference in final cation concentration between these constructed solutions and DNA solutions made by dialysis equilibrium. Even though Tris is not an ideal thermal buffer and, at 1 mM final concentration, has little capacity, it was routinely used because of its alcohol solubility and its insignificant perturbation of DNA stability in the presence of excess Na^+, which was the primary DNA counterion in these experiments. Calculations that consider the thermal coefficient of Tris buffers show that the initial pH of 8 (25 °C) is accurate; over the important temperature range, the actual pH falls between 6 and the initial value of 8 (at 25 °C). Experiments in this laboratory and elsewhere (Lewin and Pepper, 1965) have shown no important melting variation over the pH range of 5 to 10. In practice, we can discern no difference between profiles recorded in Tris buffer and those in dilute cacodylate buffer, for which the thermal coefficient of pH is quite small.

The DNA solutions, including the reference, were gently bubbled with helium and transferred to 1-cm Suprasil semimicro cuvettes (Precision Scientific). A minimum of 0.6 ml of solution was required by our beam geometry. The solutions were overlayed with washed Dow Corning 200 dimethylpolysiloxane and the cells were sealed with Teflon stoppers. No evaporation occurs in this system.

(*c*) *Thermal Denaturation Apparatus.* Techniques and equipment used in these experiments are slightly modified from those of Ansevin and Brown (1971), and the following is a brief summary of the more important features pertaining to this study.

A single beam monochromator with a Gilford Instruments photometer (Model 222) and sample positioner (Model 210D) was employed for these studies. Rapid readings of absorbance were made possible by operating in a mode in which the incident light was not adjusted to compensate for any changes in transmittance of the blank during the experiment. The linearity and absolute absorbance error of this photometer corresponds to 0.002 A, according to the manufacturer. These absolute errors result in negligible distortion of the experimental data. Of more immediate concern is the relative error or noise level of the measurements.

As previously described, the temperature was continuously increased in a linear fashion by driving the bath thermoregulator with a synchronous motor through an adjustable gear train interposed for heating rate selection. The rate of temperature rise within sample solutions in this study was 0.1 °C/min. Temperature was measured as described earlier (Ansevin and Brown, 1971; Ansevin et al., 1975) except that a quadratic temperature–voltage calibration was used to provide a more accurate estimate of the absolute temperature (about ±0.1 °C).

Photometer and thermistor-bridge voltages were digitized by a Systron Donnor (Model 7005) digital voltmeter interfaced with a Data Graphics controller (Datos 305). Output data were simultaneously punched on paper tape and typed by a Friden Flexowriter. The voltages corresponding to absorbance and temperature were digitized to five place readings; voltage sampling times were approximately 0.1 s. Although five-place precision is usually unnecessary for the absorbance measurement, since the standard deviation of the absorbance readings is about 1.4×10^{-4}, it is essential in recording the temperature (Ansevin et al., 1975), which has a standard deviation that is smaller by an order of magnitude. At the temperature increment that will be discussed (0.045 °C), the use of fewer than five digits generated large truncation errors in the calculation of the derivative.

The standard deviation for absorbance readings quoted above corresponds to the sum of all sources of error in observations on a nonabsorbing solution at a spectrophotometer slit opening of 0.34 mm (Ansevin et al., 1975). From beam energy measurements on our monochromator, we calculate that the variation in the absorbance approaches the level of photon statistical fluctuation. As expected, increased optical densities produce slightly higher noise levels; however, relatively high DNA concentrations ($A_{270} \approx 1$) serve to increase the precision of the data because a greater change in photometer output results when the sample denatures. For single wavelength studies of hyperchromism, high spectrophotometric resolution is not an absolute requirement; a band-pass of a few nanometers results in an insignificant distortion in the data. Therefore, large slit widths together with increased DNA concentrations can reduce noise levels.

From a number of observations, we conclude that our most significant source of noise is related to variations in light source and/or photometer output. Significant improvement upon these items, however, should be paralleled by appropriate precision in analogue to digital conversion to avoid truncation errors.

(*d*) *Data Processing.* The data punched on paper tape are compiled by a CDC 73 computer through a fast optical reader. Details of the program for evaluating the derivative of hyperchromicity have been discussed by Ansevin et al. (1975) and Ansevin and Brown (1971). Results are presented in the form of both tabulations and computer-drawn plots; the figures in this paper are the computer plots in which the points and lines have been broadened for clarity. Since calculating the voluminous data of our experiments is expensive, the computer program was modified to store the derivative results on disk, where they remain until the information has served all useful purposes. Of the few thousand profiles that we have calculated, numerous profiles may remain on the disk at any one time and are readily recalled for computer plotting in any desired form, size, scale, etc. Also, a number of similar profiles can be recalled for averaging and plotting. This is effected by weighting the individual profiles with an appropriate factor (e.g., an inverse function of the noise level in a profile), shifting the temperature scale of the individual profiles to compensate for minor differences of ionic strength or altered thermistor voltage, linearly interpolating the thermal derivatives of hyperchromicity to the temperature set of a selected profile, and adding and plotting the results.

Results

(*a*) *Recording and Analysis of Thermal Transitions.* The distinguishing feature of high resolution thermal denaturation is the narrow thermal increment at which measurements are recorded. This statement assumes, of course, that the equipment used for the experiment meets current, acceptable standards for precision and reliability. An example of the importance of temperature increment to the charac-

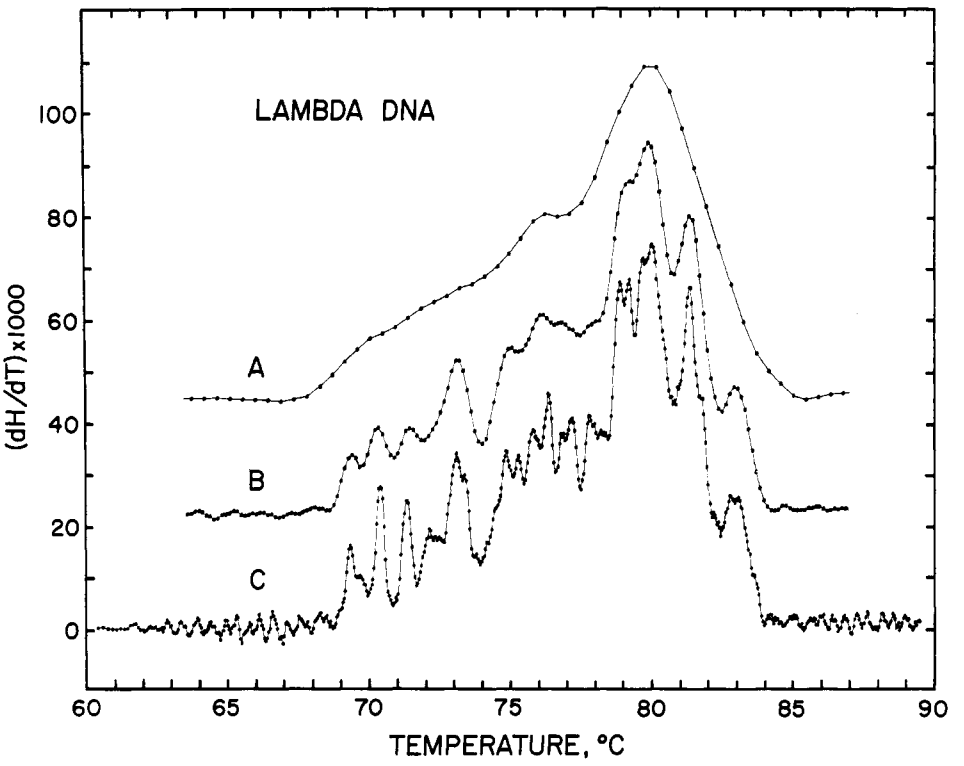

FIGURE 1. Thermal derivative of hyperchromicity at 270 nm vs. temperature for λ DNA in 30 mM NaCl. The derivatives were calculated from sequential 13-point cubic fits of a hyperchromic data set. (A) Every tenth point of the data set, a thermal increment of 0.45 °C. (B) Every third point of the data set, a thermal increment of 0.135 °C. (C) Complete data set, a thermal increment of 0.045 °C. Initial A_{270} = 0.797, final hyperchromicity was 1.52, heating rate was 0.1 °C/min.

ter of the derivative profile is provided by Figure 1, where data from a single experiment on λ DNA have been processed at three different thermal increments. In this figure, the derivative of hyperchromicity, calculated from sequential 13-point cubic polynomial fits of hyperchromicity, has been evaluated for each point, for every third point, and for every tenth point of the initial complete data set. A 13-point cubic fit was chosen for the analysis of these data to provide maximum smoothing of the profile for the full data set, with a negligible distortion of the features (see Ansevin et al., 1975). The profile corresponding to a temperature increment of 0.045 °C within the denaturation region shows a remarkable number of partially overlapping subtransitions, which, for convenience, we have termed thermalites (Ansevin et al., 1975). A subset of the same data, analyzed at a thermal increment of 0.135 °C, shows significantly broadened thermalites and a general lack of definition of the finer features in the derivative profile. The profile corresponding to a temperature interval of 0.45 °C, has been distorted beyond recognition simply by the decrease in data frequency with no change in accuracy of the basic measurements.

While it is apparent that rather remarkable detail can be observed in the profile at the closest data frequency of Figure 1, a fundamental question that must be asked is whether this fine structure is meaningful. One of the simplest ways to evaluate the significance of features in high resolution profiles is to test for reproducibility. This is demonstrated for T7 DNA in Figure 2. Curves A and B are independently measured profiles, while curve C is the average of the two profiles. Virtually all prominent features in the profiles, including maxima, minima, and points of inflection, are closely reproduced. Provided that DNA samples are free of significant contamination, such as strongly bound cations, basic proteins, or dust particles, we find that most deviations among profiles analyzed under the same conditions are at the level of the observable noise in the temperature regions outside of the transition.

Since, in effect, we are equating the term "high-resolution" to a thermal increment of 0.045 °C, we must also ask whether the data may contain finer details than those observed in the results presented here. It is not practical to increase substantially our thermal resolution (i.e., by an order of magnitude) to search for the much narrower transitions predicted by Eichinger and Fixman (1970), because the data recordings for high resolution experiments already involve 50 000 characters of information (for three profiles), and even more precision would have to be employed to achieve a higher resolution. The problem of finer detail was investigated, however, by using five-point fitting routines on the present data; such a procedure was shown to be capable of detecting even single-point discontinuities (Ansevin et al., 1975). We conclude from this analysis that thermalities narrower than those displayed in the present results do not appear in the data.

(*b*) *Interpretation of Derivative Profiles.* A straightforward interpretation of high-resolution derivative denaturation profiles involves the assumption that the area under different peaks is approximately proportional to the mass fraction of the DNA denaturing in that temperature interval. Unfortunately, this assumption is not valid for most thermal denaturation studies so far reported since the wavelength customarily chosen is 260 nm, the approximate absorption maximum of the DNA spectrum. At this wavelength, A·T base pairs are weighted much more heavily than G·C base pairs (Mahler et al., 1964). Any profile at 260 nm would, therefore, contain a strong bias against thermalite mass fraction in the higher temperature portion of profiles, which presumably represents DNA of higher G·C

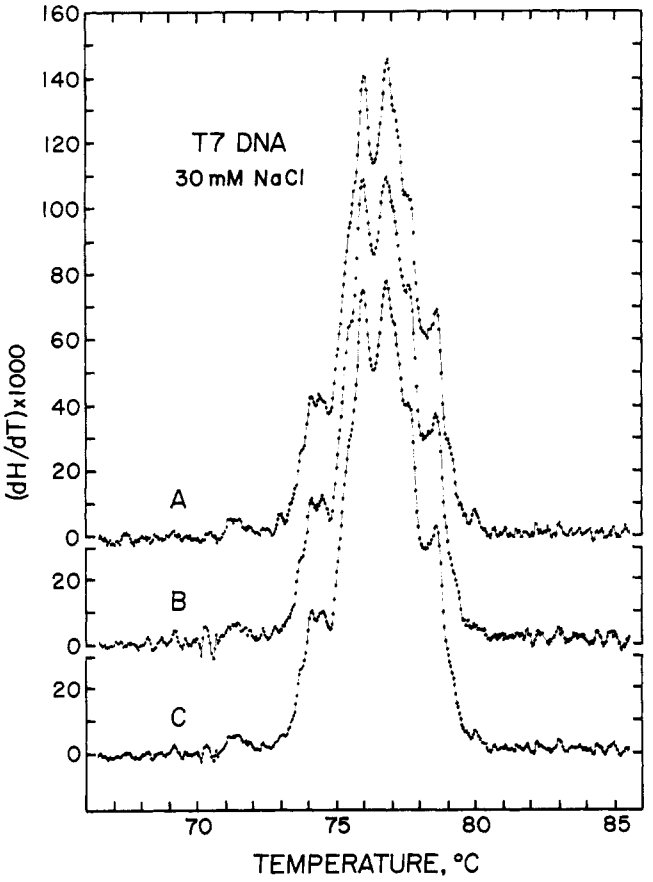

FIGURE 2. Thermal derivative of hyperchromicity at 270 nm vs. temperature for T7 DNA in 30 mM NaCl, calculated from 13-point cubic fits of hyperchromicity. Curves A and B are independent experiments and curve C is the interpolated average of A and B. Initial A_{270} = 0.937, final hyperchromicity was 1.52, and the heating rate was 0.1 °C/min for both experiments.

composition. Furthermore, this bias cannot be corrected accurately because, as will be seen below, the modal temperature of thermalites (T_{md}) is strongly influenced by factors other than the average base composition.

For this reason, we have chosen to record high-resolution experiments at an iso-extinction wavelength of the A·T and G·C hyperchromic spectra. Because there is disagreement as to the exact wavelength of this point (Mahler et al., 1964; Felsenfeld and Hirschman, 1965), we have taken 270 nm as a convenient approximation to the ideal wavelength at which the thermal denaturation profile accurately displays the mass fraction of DNA as a function of thermal stability. Fortunately, measurements can be made at 270 nm without loss of accuracy since the absolute change in absorbance upon denaturation for most natural DNAs is similar at 270 and 260 nm.

Another distortion which detracts from the interpretation of the derivative denaturation profile as a mass-fraction function of stability is the finite hyperchromism outside of the thermal transition region. This is attributed to pre-melting (Sarocchi and Guschlbauer, 1973) and unstacking (Brahms et al., 1966; Leng and Felsenfeld, 1966). Such hyperchromic increases lead to an elevated baseline in derivative plots but present little difficulty in the example of narrow bacteriophage transitions where the correction is relatively small; however, baseline correction can be more important for broader transitions. A further difficulty arises from the fact that hyperchromism is not a direct measure of hydrogen bond breaking but, rather, is a function of the changes in electronic transition moments of the bases that accompany base unstacking and solvation. A decoupling of the optical measurement from strand separation has been reported for nonequilibrium melting by Crothers and Spatz (1971). To what extent base unstacking and solvation are incomplete at the time of local strand separation remains unclear, but we have proceeded upon the assumption that any disparity is insignificant. It is then possible to relate the area under a derivative curve directly to the mass fraction of DNA undergoing a helix–coil transition in that thermal region.

Ideally, the DNA undergoing a helix–coil transition during thermal denaturation should be in a state of thermodynamic equilibrium at the time of each measurement if the information collected is to be readily interpreted. A sufficiently slow heating rate or very small temperature jumps with subsequent equilibration (Reiss and Michel, 1974; Hoff and Blok, 1970) may closely approximate an equilibrium condition. We have chosen to take advantage of the very smooth temperature rise obtainable by the technique of continuous heating.

That measurements made under the conditions of our experiments do reasonably approximate equilibrium results is indicated first by the fact that the profiles in this study are unaltered by doubling the usual heating rate of 0.1 °C/min. This observation also assures us that DNA degradation during the experiment is not a significant factor. In addition, we have tried to evaluate the effects of heating rate on the basis of existing theory concerning the time required to attain equilibrium following a rapid temperature rise. Although the initial denaturation step of hydrogen bond unzippering is known to be very fast (Pörschke, 1974), the overall transition for DNAs of the lengths in this study is kinetically limited by the much slower process of unwinding. Multiple extrapolations from the data of Crothers and Spatz (1971), assuming a linear length dependence (Massie and Zimm, 1969), and from the results of Hoff and Blok (1970) (whose data are more closely applicable to our experiments because of their smaller temperature jumps and similar ionic conditions), lead to the conclusion that any nonequilibrium distortion in our profiles is within the experimental noise levels. That is, the magnitude of theoretically calculated deviations from equilibrium indicate that there is not more than a few percent error in dH/dT, and that thermal lags are within the finest temperature increment of our data.

(*c*) *Noise Reduction and Profile Averages.* The ratio of signal to noise in the high resolution profiles depends upon a number of factors, including DNA concentration, light intensity, and the number of points simultaneously considered in the fitting routine. For genomes as simple as those of λ or T7 bacteriophage, the present noise level is certainly low enough to identify distinct thermalites. However, if thermalites are to be detected in a much more complex genome, or if a reliable quantitation of components is to be obtained for a simple DNA, some means for reducing the noise in a given recording is highly desirable. The optimization of sample concentration and incident light can have only a limited effect on the relative level of noise because instrument variations soon become limiting. Long integration times or repetition of readings at the time of data gathering also can contribute to the reduction of noise. In practice, we have found that the summation of profiles after completion of experiments is one of the most satisfactory ways to re-

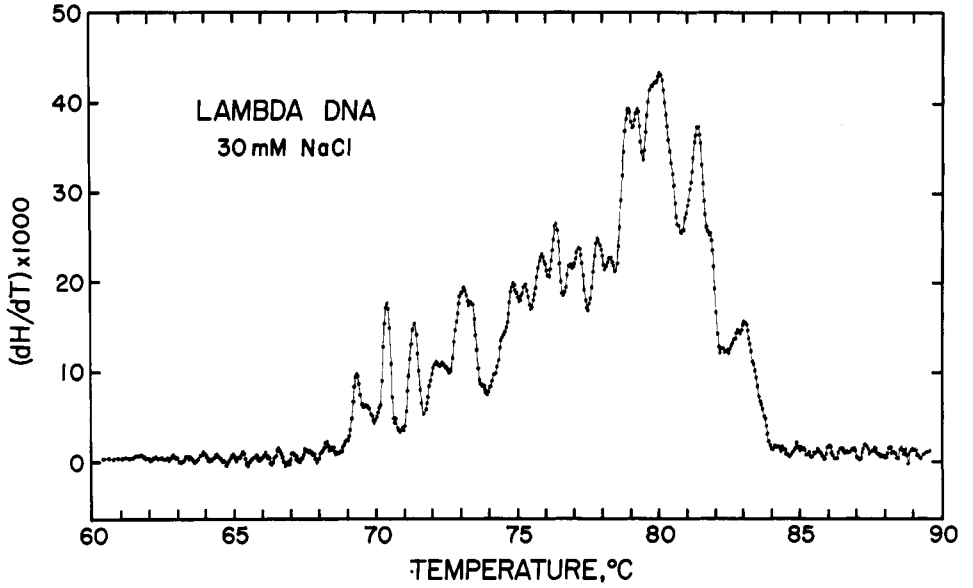

FIGURE 3. Thermal derivative of hyperchromicity at 270 nm vs. temperature for λ DNA in 30 mM NaCl. Interpolated average of two calculations as in Figure 1C, with the same initial absorptions, final hyperchromicities, and heating rates.

duce the noise level. Although the increase of signal to noise by means of averaging is limited by the square root of the number of profiles considered, we regard the results achieved by multiple profile comparison and averaging to be the most convincing demonstration of the physical reality of thermalites.

Examples of averaged profiles are shown in Figures 2–4. Curve C of Figure 2 displays the average of curves A and B for the denaturation of T7 DNA in 30 mM NaCl; Figure 3 is the average of two profiles of λ DNA in 30 mM NaCl, while Figure 4 shows the average of four profiles of T7 DNA in 5 mM NaCl. Somewhat greater noise is seen in Figure 3 than in Figure 2C because the λ average curve is derived from observations at lower DNA concentrations than for the 30 mM T7 DNA experiments. The progressive decrease of noise in going from a single recording to the average of two, and finally the average of four curves, is obvious in these figures.

(*d*) *Effect of Salt Concentration.* One of the most striking features of the T7 experiments (Figures 2 and 4) is the difference between profiles obtained in 5 and 30 mM NaCl. This and other information demonstrate that the pattern of thermalites in a profile is a strong function of cation concentration, particularly below 10^{-2} M salt (for weakly bound cations such as Na^+), and will be the subject of another communication (Vizard and Ansevin, in preparation). An examination of the difference between the profiles of Figures 2 and 4 demonstrates that many thermalites of T7 alter their position relative to one another as salt concentration is changed, in a manner not duplicated simply by a temperature scale distortion. Results obtained by resolving the two T7 profiles into components, as discussed below, suggest that the temperature difference between two thermalites can be modified by an amount that is greater than 10% of the mean temperature change of the thermalites between 5 and 30 mM salt.

(*e*) *Quantitation of Thermalites.* Several of our averaged high-resolution profiles were resolved into Gaussian components using a Du Pont electronic curve resolver. Such analogue curve resolution quantitates the three parameters of each Gaussian component: position, breadth, and area (or

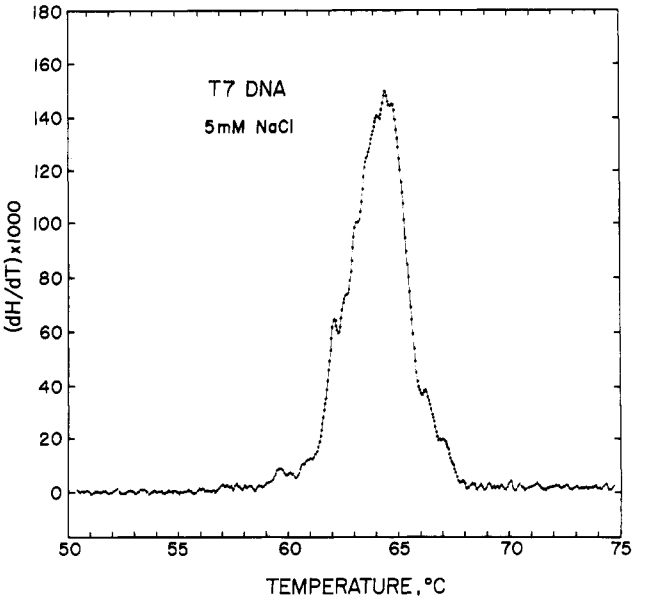

FIGURE 4. Thermal derivative of hyperchromicity at 270 nm vs. temperature for T7 DNA in 5 mM NaCl. The interpolated average of four 13-point cubic calculations. Initial A_{270} readings were 1.16 and 0.80 (two profiles each); final hyperchromicities were 1.52, and the heating rates were 0.1 °C/min.

mass fraction). From the area measurement one may calculate the length of a sequence associated with a given thermalite once the molecular weight of the whole DNA has been specified. A reasonably unambiguous resolution of a profile into Gaussian components, accurate to better than 5% of the total area, appears possible when peaks are well separated in the initial pattern. However, the determination of components in profiles exhibiting extensive overlap of peaks is considerably more difficult since important information must be gathered from points of inflection and the degree of skewing of curves. Even though the minimum number of components frequently is fairly well defined, the assignment of an exact size and width to elements within an overlapping profile must be viewed as tentative. In the case

Table I: The Thermalites of λ DNA[a] from Derivative Denaturation Profile Curve Resolution.

No.	T_{md} Modal Temp (°C)	Estimated Length Base Pairs	Estimated Breadth, 2σ (°C)	No.	T_{md} Modal Temp (°C)	Estimated Length Base Pairs	Estimated Breadth, 2σ (°C)	No.	T_{md} Modal Temp (°C)	Estimated Length Base Pairs	Estimated Breadth, 2σ (°C)
1	68.3	(150)[b]	0.37	14	73.4	1300	0.39	27	79.2	2900	0.38
2	68.9	(150)	0.32	15	73.8	450	0.55	28	79.5	900	0.29
3	69.2	490	0.26	16	74.1	450	0.32	29	79.9	4700	0.54
4	69.6	670	0.50	17	74.4	640	0.26	30	80.1	1100	0.37
5	70.0	(150)	0.25	18	74.8	1500	0.38	31	80.6	2400	0.52
6	70.4[c]	1200	0.30	19	75.3	1300	0.34	32	81.4	3600	0.52
7	70.9	340	0.50	20	75.9	2000	0.42	33	81.8	1700	0.40
8	71.4	1000	0.30	21	76.3	1700	0.33	34	82.3	830	0.35
9	71.6	(150)	0.22	22	76.7	900	0.32	35	82.6	750	0.33
10	72.0	830	0.36	23	77.2	2400	0.49	36	83.0	750	0.25
11	72.4	640	0.35	24	77.8	2000	0.41	37	83.2	710	0.38
12	72.8	980	0.35	25	78.3	1200	0.30	38	83.6	300	0.30
13	73.0	670	0.26	26	78.8	3200	0.43	39	84.2	(80)	0.28

[a] The λ DNA molecular weight assumed was 31×10^6 au or 47 000 base pairs as discussed by Davidson and Szybalski (1971). Inaccuracies in the sequence length and breadth for determinations range from 5 to 100%, depending upon the distinctness of the thermalite. [b] Thermalites under 200 base pairs, designated by parentheses, are close to the level of noise; they are, however, presented on the basis of reproducibility. [c] See Figure 6.

of the λ DNA profile, the correspondence of results in multiple attempts at curve resolution gives us reasonable confidence in the position and shape of the majority of thermalites; however, these attempts also reveal the fact that minor components can be estimated only to within a factor of about 2.

A summary of the resolved thermalites from the λ DNA profile is given in Table I. Similar analyses have been carried out for the T7 DNA profiles and, while any conclusions stated here appear to apply also to T7 DNA, the results for λ DNA present clearer arguments and are therefore elaborated. The less distinct T7 DNA profiles can be separated into approximately 22 components, compared with the 39 components suggested for λ DNA. This simple comparison supports the intuitive expectation that more thermalites should be found in a DNA of higher molecular weight, although it is likely that some components in both T7 DNA profiles remain unresolved because of overlap.

The thermalites listed in Table I show a broad distribution of chain lengths (listed in number of base pairs), but have a more restricted distribution of thermal breadths. Some of the larger and wider thermalites, such as the one at 81.4 °C (No. 32), are probably unresolved or overlapping transitions. This interpretation is supported by the histogram (Figure 5) of the frequency distribution of thermalite breadths for λ DNA. Since thermalite breadths are distributed without regard to stability (Table I), and since the wider thermalites appear to be separated from the majority in the frequency distribution (Figure 5), it would appear that the wider thermalites are artifacts of incomplete curve resolution. Excluding these wider thermalites ($2\sigma > 0.43$ °C), the mean width of the resolved λ DNA thermalites is 0.33 °C, and the standard deviation of the breadth distribution is 0.054 °C. This comparatively narrow distribution includes the inaccuracies of curve resolution, as well as any intrinsic variation of the thermalite breadth, and suggests that thermalite breadth is a relatively invariant quantity.

The shape of individual thermalites is similar to that of a normal curve of errors, as might be expected if helix-coil transitions are governed by a stochastic process involving many base pairs. A quantitative test of the shape, however, is desirable because current theories differ on this very point (Crothers, 1968; Eichinger and Fixman, 1970). Such a test has been applied to the most distinct of the thermalites in this study, the 70.4 °C thermalite of λ DNA. The data from this region of Figure 3 may be compared with the analytical Gaussian distribution of dH/dT with temperature, $dH/dT = Ae^{-(T-T_{md})^2/2\sigma^2}$, where A is the maximum ordinate of the distribution, T is the temperature, T_{md} is the modal temperature of the thermalite, and σ is the standard deviation of the distribution. For this test, interpolated base unstacking estimates were subtracted from dH/dT, for reasons previously discussed; these corrections were small, being less than 10%. Figure 6 is a plot of the corrected dH/dT values in the region of $T_{md} = 70.4$ °C (see Figure 3), in the form consistent with the analytical Gaussian expression. The thermalite data fit a straight line quite well over a range of about 0.5 °C, which represents over 85% of the transition; the slope of the line is consistent with a standard deviation of 0.15 °C ($2\sigma = 0.30$ °C). This rather close adherence of the data to a Gaussian form contrasts with predictions of current theories (Eichinger and Fixman, 1970; Applequist, 1969).

A major conclusion about the nature of thermalites in both types of bacteriophage DNA is that they lack any discernible trend in thermal width from the least to the most stable transitions, contrary to the theoretical prediction of Crothers (1968). Also noteworthy is the fact that curve resolution performed on the averaged T7 profiles disclosed that about the same number of components could be identified in the results at both salt concentrations when it was assumed that each thermalite was Gaussian in shape. Furthermore, the distribution of thermalite widths did not change between 5 and 30 mM salt even though the overall profile shapes were quite different. That transition breadths may not be a strong function of salt concentration is suggested also by the data of Blake (1973).

The distribution of sequence lengths derived from thermalite mass fractions is not easily evaluated because wherever there is a statistically large number of thermalites, the overlapping of components in the profile becomes a potential source of error. When the assumption is made that the larger, wider thermalites in the resolved profiles of the two bacteriophage DNAs studied here actually are multiple components, and therefore should be excluded from averages, as previously discussed, the remaining more resolved

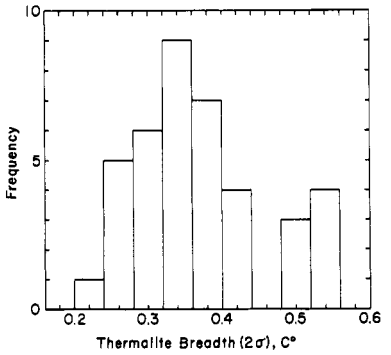

FIGURE 5. Frequency distributions of thermalite breadths (2σ) for the resolved derivative profile of λ DNA as shown in Table I. For the 32 of 39 thermalites whose $2\sigma < 0.43$ °C, the mean breadth is 0.33 °C with a standard deviation of 0.054 °C.

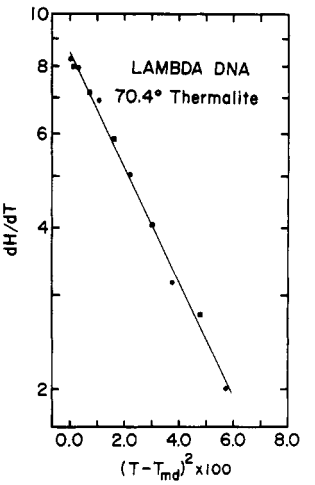

FIGURE 6. Semilogarithmic plot of the thermal derivative of hyperchromicity at 270 nm vs. $(T - T_{md})^2$ for the data in the region of the 70.4 °C thermalite of λ DNA (see Figure 3 and Table I). Squares (■) are data for $T < T_{md}$, and circles (●) are data for $T > T_{md}$. T_{md} (=70.38 °C) is the modal temperature of the thermalite estimated from the data symmetry. Values of the derivative were corrected for pre-melting and unstacking contributions by interpolation. The straight line corresponds to a Gaussian distribution of the form $dH/dT = Ae^{-(T-T_{md})^2/2\sigma^2}$, where the slope of the line indicates a standard deviation of 0.15 °C ($2\sigma = 0.30$ °C).

thermalites (51 examples) range in length from less than 200 to longer than 3000 base pairs. The apparent number average of the distribution is 900 base pairs, and the weight average is 1900 base pairs. Some thermalites smaller than 200 base pairs must be regarded as experimentally insignificant and were not considered in the averages, although their inclusion does not significantly change the numbers. It appears, therefore, that the average thermalite in λ or T7 DNA represents a gene-sized DNA sequence.

Discussion

(*a*) *Comparison of Conditions for High Resolution Melting.* The present investigation applies the technique of high resolution thermal denaturation to a study of the subtransitions (thermalites) of λ and T7 bacteriophage DNAs. Throughout the evolution and refinement of these methods we have attempted to achieve a fidelity in derivative profiles that would justify the quantitation of features observed in the profile. This has led us to explore a number of nontrivial problems which can significantly modify the final derivative denaturation profile. Several of the experimental approaches we have taken are different from those adopted by other investigators, who also have recently attempted to achieve high resolution in thermal denaturation experiments. In an effort to compare experimental approaches, we discuss below some of the distinctive features of the methods used here and their influence on the quantitation of results.

A notable difference in procedure among the few high resolution studies that so far have been conducted is the manner in which the temperature of samples is increased. We would suggest that the temperature-jump and the continuous-rise heating methods are thermodynamically equivalent if the rate of heating in each is sufficiently slow. Extrapolation from the previous relaxation studies of Crothers and Spatz (1971) and Hoff and Blok (1970) suggests that, under the conditions used here and in other recent experiments (Reiss and Michel, 1974), both techniques of temperature rise achieve reasonable approximations to a quasi-equilibrium denaturation of DNA. However, a small difference was reported by Reiss and Michel (1974) between profiles corresponding to temperature-jump heating rates of 0.025 and 0.1 °C/min. This apparent difference contrasts with an invariance of our profiles over the higher range of heating rates of 0.1 to 0.2 °C/min, where the heating was continuous. Since we calculate that a continuous heating rate of 0.1 °C/min should give results that are insignificantly different from equilibrium observations (see Results section), we conclude that other experimental variables may well account for the differences displayed by Reiss and Michel (1974). Among the experimental variables to be considered is the reduced salt concentration used in their study ($\approx$2 mM Na^+) which could slow the attainment of equilibrium (Crothers and Spatz, 1971). However, the lower molecular weight of their mitochondrial DNA (about 2×10^6 au) should more than compensate for the salt effect. Since it appears that the distinction between the Reiss and Michel (1974) profiles at different heating rates is an increasing function of temperature, it is possible that the disparity may have been caused by DNA degradation during their very lengthy experiment. In that case, the slower heating rate (0.025 °C/min), rather than being required for a better equilibrium approximation, may actually have been an experimental detriment.

One advantage of the continuous-heating method is that accurate temperature increments of less than 0.1 °C are achieved much more easily than with temperature-jump techniques. This is important because our results indicate that a thermal resolution of 0.05 °C or finer is necessary for quantitative profile analysis. With the temperature-jump method, it is advantageous to minimize the mass to be heated so that the equilibration can be rapid, but the maintenance of an assigned temperature in a small sample remote from the heating source is beset with problems caused by the large heat sinks that intervene and that surround the sample. The possible existence of thermal gradients within such a sample or a temperature fluctuation of the entire sample may explain the broader thermalites in temperature jump profiles (Reiss and Michel, 1974; Michel, 1974) relative to the breadths found in this study.

One of the more significant differences between our own studies and other experiments on high resolution thermal denaturation (Reiss and Michel, 1974; Vitek et al., 1974) is the way in which derivative profiles are calculated. We have made every attempt to ensure that the data frequency and the computational routines selected preserve the physical

realities of thermalites. To what extent the results in other studies are a function of the calculations applied we cannot estimate because of the fundamental differences among the various approaches to the problem and the lack of full details about other procedures.

(*b*) *High Resolution Denaturation Profiles of Bacteriophage DNA.* High resolution results for λ bacteriophage DNA, obtained in a medium containing 0.03 M sodium chloride (Figure 3 and Table I), indicate that at least 34 significant thermalites can be identified. Unfortunately, our profile cannot be compared directly with the highly detailed recent profile of Reiss and Michel (1974) since there is little correspondence in the positions of major thermalites in our respective results. This lack of agreement presumably can be taken as another example of the critical dependence of thermalite stability upon salt concentration. The increased breadths and lack of resolution in the latter profiles could be the result of a number of effects that we have already discussed. Nonetheless, it is apparent that high resolution profiles for λ DNA, at either 30 or 2 mM salt, are far more complex than those obtained by previously available thermal denaturation methods (Yabuki, 1971; Falkow and Cowie, 1968).

The DNA of T7 bacteriophage has been considered to have a particularly narrow transition, and we find that about 98% of the transition in 30 mM salt occurs in a 7 °C range. Despite the fact that this is far narrower than the 15 °C range of the λ DNA transition, it is quite apparent that the thermal transition of T7 DNA also consists of a series of thermalites; approximately 22 subtransitions could be detected by curve resolution. Some comparison of these results can be made to the free-solution derivative profile of T7 DNA shown by Gomez and Lang (1972) since the Na^+ concentration of that study (20 mM) was relatively close to our 30 mM solvent. The results agree surprisingly well considering the lower resolution of their experiment (a thermal increment of about 0.13 °C); of the 4 or 5 thermalites that can readily be distinguished in their profile, both mass fraction and relative modal temperatures are similar to our results.

(*c*) *Comparison to Partial Denaturation Maps.* The mapping of denatured segments of DNA by the direct observation of single-stranded loops with the electron microscope has been a most popular technique for identifying what are presumed to be A·T-rich regions in moderately simple genomes. There is great interest in comparing such results with our free-solution thermal denaturation profiles since both may reveal similar properties of DNA. Unfortunately, we must conclude that a correlation of electron microscope denaturation maps with high resolution solution denaturation profiles is at best very limited; in fact, it is not clear that a correlation should exist at all. The major difficulty stems from the fact that the electron microscope method involves a formaldehyde treatment (Inman, 1966) under distinctly nonequilibrium conditions. This treatment reveals a kinetic picture of loop formation at temperatures where little or no significant equilibrium denaturation would have occurred. It is perhaps reasonable to expect that the kinetic picture should correlate with "pre-melting" sequences (Sarocchi and Guschlbauer, 1973), but the relationship between the pre-melting phenomenon and the major thermal transition of DNA is not clear even in equilibrium solution studies. That a small sequence has a finite probability of looping at a temperature below the thermal range of equilibrium denaturation might be related to the ability of that sequence to nucleate a cooperative transition in the thermal range of equilibrium denaturation. However, the growth of a nucleating event into a major cooperative transition (a thermalite) will largely depend upon the stability of the adjoining sequences. This would necessarily limit the correlation between thermalites and denatured sites mapped by microscopy to the total numbers of each found in a genome; little useful information regarding mass fraction and stability should be available from the mapping experiment and no information regarding relative map positions from the thermalites of equilibrium denaturation.

Accordingly, some correlation might exist for the number of prominent denaturation regions between electron microscope denaturation mapping (either thermal or alkali-induced) and solution thermal denaturation. For instance, the 13 sites microscopically mapped on T7 DNA by Gomez and Lang (1972) might be tentatively correlated with a similar number of the more obvious peaks and inflections that we find in the T7 DNA profiles (Figure 2). Curve resolution of the derivative profile gives a larger number of thermalites (about 22); a more subjective resolution of the mapped site histograms can also give 20–22 sites. In view of the complexity of the λ DNA profile (Figure 3 and Table I), it is clear that a very large statistical sample of the complete partial denaturation map is needed to test a numerical correlation of the mapped sites with the λ DNA thermalites. Although complete mapping studies on λ are presently unavailable, the investigation of Inman and Schnös (1970) demonstrated the presence of a number of looping regions in the A·T-rich half of the λ genome. Clearly, the significance of the correlation between the number of thermalites and the number of looping regions in T7 DNA is tenuous. Proof of a good correlation between electron microscope mapping results and thermalites would require studies with a smaller genome having fewer and better resolved thermalites.

(*d*) *Relevance of Thermalites to Theories of DNA Helix-to-Coil Transition.* Since we believe that the details of thermalite shape are accurately preserved in the λ and T7 DNA profiles of Figures 2, 3, and 4, it is appropriate to compare the information that can be deduced about the nature of thermalites in these figures to existing theories of helix-to-coil transitions in DNA.

We have arrived at four conclusions about the thermal transitions of natural DNAs.

(1) Natural DNA helix–coil transitions are collections of discrete subtransitions, or thermalites, much as predicted by existing theories. (This is supported by considerable evidence in addition to the bacteriophage profiles presented here.) Extrapolation from results for the simple DNAs examined in this study suggests that the only broad transitions to be found in natural DNAs are those of pre-melting and post-transitional unstacking events.

(2) Thermalites appear to have a Gaussian shape and a relatively narrow distribution of widths, centered about a mean width (2σ) of 0.33 °C. Furthermore, thermalite width is not a function of position in the profile, nor is it strongly influenced by the salt concentration of the medium.

(3) The relative position of individual thermalites in a profile is not simply a function of salt concentration in the medium or base composition of the sequence.

(4) The number average sequence length for thermalites observed here is approximately 900 base pairs, and the length distribution is broad.

While the first of these conclusions is not particularly

surprising, the second is at variance with at least some of the models that have been presented for transitions in DNA of random sequence. The Gaussian form of the thermalite contrasts with the asymmetric form predicted by Eichinger and Fixman (1970); the precision to which the thermalite is Gaussian suggests that a large portion of the transition does not conform to the statistical picture of a one-dimensional phase transition (Applequist, 1969), where an enlarging coil makes an increasing entropic contribution to the coiled state. Furthermore, our observed transition breadths, which display little variance, are not a function of stability and therefore contrast with the corresponding prediction by Crothers (1968). On the basis of the Crothers model and our data, it would appear that internal looping and loop merging transitions are thermodynamically indistinguishable. That the transition breadth does not appear to be a strong function of ionic strength is a finding pertinent to relative base pair stabilities and ring weighting functions (Crothers, 1968; Eichinger and Fixman, 1970); however, the ionic strength variation in this study is not large enough to attribute great significance to this observation, even though it is consistent with a similar observation made by Blake (1973).

Another important result reported here is that the modal thermalite breadth is slightly narrower than that depicted by Crothers (1968), but significantly broader than that proposed by Eichinger and Fixman (1970). As pointed out by Crothers, absolute comparison of transition width between theory and experiment cannot be taken too seriously since few measurements were available to satisfy the empirical aspects of the theory. But there does appear to be a real discrepancy since adjustment of Crothers' theory to include a more realistic number of energetically distinguishable base pair doublets (Borer et al., 1974) should serve only to further broaden the predicted transitions. The possibility remains that the inclusion of the energetically distinguishable base pair doublets in nonrandom sequences could force convergence of the Eichinger and Fixman phase transition to the broader Gaussian form of the thermalite.

The observation that thermalite positions, relative to one another in the profile, change with salt concentration (conclusion No. 3) is evidence for sequence-specific information and will be the subject of a later communication containing more exhaustive data.

The fourth observation from the data of this paper is that the number average length of thermalites is about 900 base pairs. This presumably means that the more stable sequences that limit loop spreading are located at an average separation of 900 base pairs. That this length differs from the theoretical estimate of 300 base pairs (Crothers, 1968) is trivial since in a relatively small genome the average can differ considerably from that predicted for an infinite random sample, especially if nonrandom sequences happen to have biological utility.

(*e*) *Significance for Genomic Characterization.* The evidence presented here and elsewhere (Michel, 1974; Steinert and Van Assel, 1974) suggests that the thermal denaturation profiles of naturally occurring DNAs *in general* contain a great deal of information about the irregularities of base composition that are intrinsic to unique DNA sequences. It appears that a genome in the size range used in this study (50 000 base pairs) may be simple enough to permit a meaningful physical characterization of sequences in terms of their thermal stabilities. Of course, a more limited analysis might be possible for a larger genome if the distribution of thermalite T_{md} values is very broad. A complication that prevents a straightforward interpretation of thermalite data, however, is the fact that the modal temperature of a thermalite is not a simple function of base composition. However, there is the possibility that a limited sequence characterization can be achieved by analyzing the ionic strength dependence of a relatively simple profile. It is probable that the modal temperature of a thermalite characterizes the stability of the sequence that nucleates a particular transition and that the mass fraction of the thermalite is a measure of the distance between the more stable sequences terminating the cooperative transition. We expect salt concentration to influence both of these parameters for some thermalites.

Even without a detailed interpretation of thermalite stabilities, it is obvious that the type of characterization that is possible by high resolution thermal denaturation studies can be used to investigate genetic changes; indeed, this was recently demonstrated by Michel et al. (1974) for the simpler DNA of mutant yeast mitochondria. An interesting outcome of our results is that the distribution of thermalite mass fractions happens to be consistent with the approximate size range of genetic units. However, the mean size and size distribution of the sequences represented by thermalites may well be a consequence of the statistical stability distribution intrinsic to a quasi-random collection of base pairs and fortuitously similar to gene-sized segments.

The possible biological implications of a segmental heterogeneity in base composition were considered by Skalka et al. (1968). Although data available at that time suggested that the mean size of such regions in λ might be on the order of one-fourth the λ genome, or about 10 000 base pairs, our present limit of resolution implies that segmental heterogeneity along the DNA contour applies to a region less than 1/10th this size, or about 900 base pairs. The speculation that the gene-sized segments in this study of T7 and λ DNAs might be related to the 19 genes or more than 25 peptides of T7 (Studier and Maizel, 1969) and the 35 genes of λ (Szybalski and Herskowitz, 1971) is based only on the respective numerical correlations, which are tenuous. Clearly, a number of viable hypotheses remain to be tested.

Acknowledgments

This study was supported by the Robert A. Welch Foundation Grant G-290, and the Energy Research and Development Administration Contract AT-(40-1)-2832.

References

Ansevin, A., and Brown, B. (1971), *Biochemistry 10,* 1133–1142.

Ansevin, A., Vizard, D., Brown, B., and McConathy, J. (1975), *Biopolymers* (in press).

Applequist, J. (1969), *J. Chem. Phys. 50,* 600–609.

Blake, R. (1973), *Biophys. Chem. 1* , 24–34.

Borer, P., Dengler, B., and Tinoco, I. (1974), *J. Mol. Biol. 86,* 843–853.

Brahms, J., Michelson, A., and Van Holde, K. (1966), *J. Mol. Biol. 15,* 467–488.

Crothers, D. (1968), *Biopolymers 6,* 1391–1404.

Crothers, D., and Spatz, H. (1971), *Biopolymers 10,* 1949–1972.

Davidson, N., and Szybalski, W. (1971), in The Bacteriophage Lambda, Hershey, A. Ed., Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, pp. 45–54.

Eichinger, B., and Fixman, M. (1970), *Biopolymers 9,*

205-221.
Falkow, S., and Cowie, D. (1968), *J. Bacteriol. 96,* 777-784.
Felsenfeld, G., and Hirschman, S. (1965), *J. Mol. Biol. 13,* 407-427.
Gomez, B., and Lang, D. (1972), *J. Mol. Biol. 70,* 239-251.
Gray, H., and Hearst, J. (1968), *J. Mol. Biol. 35,* 111-129.
Hirschman, S., Gellert, M., Falkow, S., and Felsenfeld, G. (1967), *J. Mol. Biol. 28,* 469-477.
Hoff, A., and Blok, J. (1970), *Biopolymers 9,* 1349-1360.
Inman, R. (1966), *J. Mol. Biol. 18,* 464-476.
Inman, R., and Schnös, M. (1970), *J. Mol. Biol. 49,* 93-98.
Leng, M., and Felsenfeld, G. (1966), *J. Mol. Biol. 15,* 455-466.
Lewin, S., and Pepper, D. (1965), *Arch. Biochem. Biophys. 112,* 243-248.
Mahler, H., Kline, B., and Mehrotra, B. (1964), *J. Mol. Biol. 9,* 801-811.
Massie, H., and Zimm, B. (1969), *Biopolymers 7,* 475-493.
Michel, F. (1974), *J. Mol. Biol. 89,* 305-326.
Michel, F., Lazowska, J., Faye, G., Fukuhara, H., and Slonimski, P. (1974), *J. Mol. Biol. 85,* 411-431.
Pörschke, D. (1974), *Biophys. Chem. 2,* 97-101.
Reiss, C., and Michel, F. (1974), *Anal. Biochem. 62,* 499-508.
Sarocchi, M., and Guschlbauer, W. (1973), *Eur. J. Biochem. 34,* 232-240.
Skalka, A., Burgi, E., and Hershey, A. (1968), *J. Mol. Biol. 34,* 1-16.
Steinert, M., and Van Assel, S. (1974), *Biochem. Biophys. Res. Commun. 61,* 1249-1255.
Studier, F. W., and Maizel, J. (1969), *Virology 39,* 575-586.
Szybalski, W., and Herskowitz, I. (1971), in The Bacteriophage Lambda, Hershey, A., Ed., Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, p 778.
Vitek, A., Reddy, C., and Pivec, L. (1974), *Biochim. Biophys. Acta 353,* 385-391.
Yabuki, S., Fuke, M., and Wada, A. (1971), *J. Biochem. 69,* 191-207.

Subunit Structure of rDNA-Containing Chromatin†

Diane J. Mathis* and Martin A. Gorovsky

ABSTRACT: Recent studies indicate that chromatin has a repeating subunit structure. In an attempt to relate this organization to chromatin's role in selective gene transcription we have begun to examine the subunit structure of a specific gene. *Tetrahymena pyriformis* preferentially replicates the genes coding for rRNA (rDNA) during refeeding after prolonged starvation. By prelabeling cultures during exponential growth with [^{14}C]thymidine and pulse-labeling during refeeding with [^{3}H]thymidine, we have been able to differentially label bulk chromatin and rDNA-containing chromatin. Nuclei which contained at least 78% of their ^{3}H label in rDNA were digested with staphylococcal nuclease, and the DNA digestion products analyzed on agarose gels. Both the kinetics of digestion and the digestion products were similar for ^{14}C- and ^{3}H-labeled chromatin. In order to monitor protein exchange, digestions were also performed on partially purified rDNA-containing chromatin or free rDNA in the presence of nuclei. While the chromatin had a digestion pattern like nuclei, the rDNA was afforded no protection from digestion. Our conclusion is that the chromatin containing rDNA (a repeated, extrachromosomal gene in *Tetrahymena*) exhibits a particulate structure very similar to that of bulk chromatin. This organization does not exist in free rDNA and is not the result of protein exchange during the nuclease digestion.

Recent electron microscopic and nuclease digestion studies have provided evidence to support a particulate model of chromatin structure. Linear arrays of spherical particles have been observed in chromatin extruding from lysed nuclei (Olins and Olins, 1974) and in isolated chromatin depleted of histone F1 (Oudet et al., 1975). Arrays of nucleoprotein particles have also been detected after partial digestion of nuclei or chromatin with relatively nonspecific nucleases (Noll, 1974; Sahasrabuddhe and Van Holde, 1974; Sollner-Webb and Felsenfeld, 1975). To date, evidence of a repeating subunit structure has been found in chromatin from several vertebrates (Hewish and Burgoyne, 1973; Axel, 1975; Honda et al., 1975), a plant (McGhee and Engel, 1975), two lower eukaryotes (Lohr and Van Holde, 1975; Gorovsky and Keevert, 1975), and two animal viruses (Griffith, 1975; Louie, 1974).

Any model of chromatin structure must eventually be able to account for the transcription of selected DNA sequences. Thus, it is of interest to relate the apparent particulate structure of chromatin to proposed models of genome organization (Davidson and Britten, 1973). In this context, one can ask whether the subunit profile is similar for repeated vs. unique DNA sequences, expressed vs. repressed genes, or spacer vs. coding sequences. We have begun to examine these questions by comparing the subunit structure of bulk chromatin with that of chromatin containing a specific gene.

† From the Department of Biology, The University of Rochester, Rochester, New York 14627. *Received September 15, 1975.* This research was supported by National Institutes of Health Grant No. GM-21793. D.J.M. is a National Institutes of Health predoctoral trainee.