Proc. Natl. Acad. Sci. USA
Vol. 76, No. 1, pp. 101–105, January 1979
Biochemistry

DNA sequencing and melting curve

(specific fragmentation/fractioning/A+T content/bacteriophage φX174)

M. YA. AZBEL

Department of Physics and Astronomy, Tel-Aviv University, Tel-Aviv, Israel

Communicated by H. A. Scheraga, October 2, 1978

ABSTRACT **The dependence of DNA absorbance (for light at about 260 nm) on temperature is related to a specific DNA sequence structure in the vicinity of DNA thermal denaturation (the so-called DNA melting or coiling). A straightforward analysis of the experimental DNA melting curve allows us to determine the lengths, the A+T content, and the location in DNA of certain domains. In the case of a specific DNA fragmentation, the order of fragments in DNA can be learned from this analysis, nondestructively and quickly, without fractionating the fragments and other methods of fragmentation. If the DNA nucleotide sequence is known except for some sites and uncertain portions, the analysis determines these sites and the accuracy of the sequence at the portions. This information may complement exact methods of DNA sequencing. The proposed analysis is applied to bacteriophage φX174, whose melting curve is known. The results are compared to and found to be in an excellent agreement with the known φX174 nucleotide sequence.**

In recent years, powerful methods (1–3) of DNA sequencing have been developed. These methods provided, in particular, the complete sequencing of MS-2 (4–6), φX174 (7), simian virus 40 (8, 9), and bacteriophage FD. The procedure of the sequencing included specific fragmentations, then fractionation, and finally sequencing of each of the fragments. To determine the order of fragments in DNA, several different methods of fragmentation were used. Typically, after the sequencing was finished, the sequence at some portions remained less certain and some base pairs were undetermined.

This paper proposes a method that allows one (*i*) to learn the order of fragments in DNA from just one method of specific fragmentation without the fractionation of fragments; (*ii*) to determine the lengths, composition, and order of certain segments in a nonfragmented DNA; (*iii*) to verify the accuracy of the sequencing of uncertain portions; and (*iv*) to determine unknown base pairs. Of course, this information cannot compete with detailed sequence determination, but it may be important as a complementary source of information. The method is nondestructive, quick, and cheap. It is based on the measurement of the absorbance (or heat capacity) of DNA (or double-stranded RNA) at various temperatures in the vicinity of DNA thermal denaturation (which is usually denoted as DNA coiling or melting) and is related to the unbounding of base pairs and the separation of DNA strands. Because the light absorption in the region of 260 nm is less for bounded pairs than for unbounded ("helix") ones, an experiment directly determines (10–12) the number of unbounded ("melted") pairs (N_c). Herein it is shown qualitatively how $N_c(T)$ provides the above-mentioned information about DNA sequence structure. Then I present the quantitative approach to the problem, compare it to other approaches to the analysis of $N_c(T)$, apply it to the bacteriophage φX174 DNA (φXD), whose melting curve is known (13, 14), and compare the obtained results to the φXD nucleotide sequence (7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MAIN IDEA OF METHOD

The rigorous theory of DNA melting was published in 1973 (15) (see also refs. 16–21). The main statements of ref. 15 are as follows. DNA melting is local rather than continuous. It occurs step by step, domain after domain (Fig. 1). Each domain and its location in DNA is precisely determined for a given sequence by specific and rather tricky algorithm (15). The domain melts in a narrow temperature region (i.e., almost at a certain "melting" temperature). This discrete melting implies corresponding quasi jumps of N_c as a function of temperature (T). Quasi jumps imply narrow peaks in the differential melting curve (DMC) dN_c/dT (Fig. 2). Obviously, the height of each quasi jump, and therefore the area of each peak, approximately equals ℓ_m (ℓ_m is the "length" of the mth melting domain—i.e., the number of base pairs in it).

The described fine oscillatory structure of the DMC dN_c/dT was observed in numerous experiments (13, 14, 22–30) (e.g., see Fig. 3). All experiments demonstrate very narrow peaks (their half-width is typically about 0.5 K or 0.2% of the absolute temperature), thus experimentally verifying the discrete melting. The peaks are well reproduced in different experiments, thus verifying that the same domain melts at the same melting temperature—i.e., that the domains are specific rather than occasional.

The peaks at the differential melting curve are narrow but numerous. Therefore, they may overlap—i.e., a single experimentally observed peak may be related to the melting of several

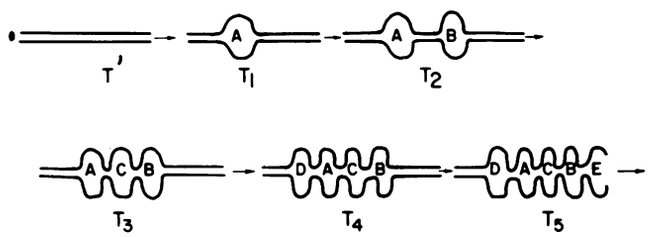

FIG. 1. Successive local melting of DNA ($T' < T_1 < T_2 < T_3 < T_4 < T_5 < \ldots$). The melting of domains A and B takes place in an unmelted helix environment. Each melting creates two phase boundaries ($n = 2$) between melted and helical regions. Domain C appears between two already melted domains and thus annihilates two phase boundaries ($n = -2$). The melting of domain D just displaces the phase boundary and does not change the number of phase boundaries ($n = 0$). The melting of the border segment E annihilates one phase boundary ($n = -1$). Vice versa, if $n = 2$ is experimentally observed, then it refers to a domain melting in a helical environment; $n = -2$ indicates a domain melting between already melted domains; $n = 0$ refers to a domain adjacent to a melted domain on one side and to a helical domain on the other side; $n = -1$ refers to a domain adjacent to a melted domain on one side and on a DNA border on the other side.

Abbreviations: DMC, differential melting curve; *φXD*, bacteriophage φX174 DNA; N_c, number of melted pairs; T_m, melting temperature in kelvins; ℓ_m, length of mth melting domain; n_m, number of phase boundaries that appear due to melting.

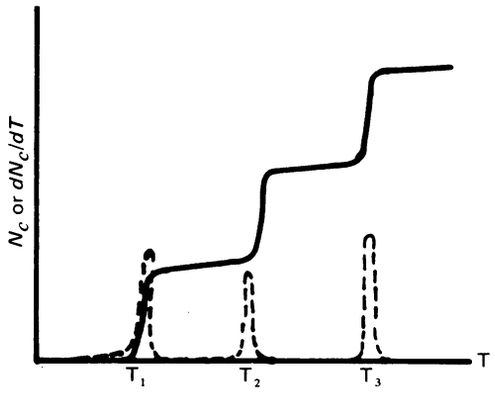

FIG. 2. N_c (solid line) and its derivative dN_c/dT (dashed line) versus temperature T. The height of each quasi jump and the area of the corresponding peak approximately equal the length (ℓ_m) of the melting domain.

domains (from different regions of DNA), with close melting temperatures. For the sake of vividness, in this, (and only in this) section I consider nonoverlapping domain peaks.

According to ref. 15, the melting temperature (T_m) of the mth melting domain depends on ℓ_m, A+T content (x_m), and number (n_m) of the phase boundaries [between melted and unmelted ("helix") states] which appear due to the domain melting (Fig. 1):

$$T_m = T_0 - x_m\tau + T_b n_m/\ell_m. \qquad [1]$$

(Later on, when it does not lead to any ambiguity, I omit the subscript m.)

The first two terms in Eq. 1 coincide with the Marmur–Doty formula (10–12), $T_{AB} = T_0 - \tau$ and T_0 being the homopoly-(A,T) ($x_m = 1$ in Eq. 1) and homopoly(G,C) ($x_m = 0$ in Eq. 1) T_ms, respectively. The last term represents the contribution of phase boundaries to the local coiling. T_0, τ, and T_b depend on the solvent. They can be determined (31, 32) from the melting curves (in the same solvent) of a polymer of the known composition.* If some characteristic of the solvent (e.g., its pH or its ion concentration, etc.) is slightly changed, then, by ref. 15, the melting domains and therefore their ℓ, x, and n remain unchanged and T_0, τ, and T_b change slightly. This implies the shift in the T_ms. For instance, when T_b increases, peaks with $n = 1,2$ increase their T_ms and move "to the right"; peaks with $n = -1,-2$ move to the left; peaks with $n = 0$ are not displaced. Because ℓ in Eq. 1 equals the area of the corresponding peak and thus can be considered known, the displacement of a peak allows us, by Eq. 1, to determine its n according to the equation

$$\delta[(T - T_0)/\tau] = (n/\ell)\delta[T_b/\tau]. \qquad [2]$$

Here, δ denotes the change of the quantities in square brackets. Because n may equal only 0, ±1, or ±2, even a low accuracy in Eq. 2 determines the precise n. With known n_m and ℓ_m, Eq. 1 determines x_m.

As we shall see in the next paragraph, the accuracy of the x thus determined may be high enough to provide the exact A+T and G+C contents of the domain. Then, if some base pairs at the domain are unknown in the DNA sequence, they can be determined. Also, the accuracy of the DNA sequence may be checked. Consequent ns determine the location of melting domains in DNA. This can be demonstrated with a simple example. Suppose the successive ns are 2,2,−2,0,−1, ... Then, $n = -2$ indicates that the third melting domain (C) is situated between the first (A) and the second (B) ones (Fig. 1). (The domain sequences ACB and BCA are identical because they coincide after the 180° DNA turn.) Now, $n = 0$ refers to the domain (D) that may be adjacent to A or to B (in both cases its melting just displaces the boundary). A change in the solvent clarifies the situation. Suppose it increases T_b. Then, by Eq. 1, T_ms of domains A and B (with $n = 2$) increase, those of C (with $n = -2$) and E (with $n = -1$) decrease (the decrease in C is larger than that in E), and the T_m of D (with $n = 0$) remains unchanged. Suppose the T_b increase is large enough to imply the melting domain sequence A-C-D-B-E with ns 2,0,0,0,−1 (the sequence of domains is determined according to their ℓ_m and x_m). Because now D has $n = 0$, it is adjacent to the melted A (it cannot be adjacent to C because C was determined to be adjacent to A and B). So the domain sequence is DACB. Now, $n = -1$ indicates the border domain (because it had only one phase boundary), which may be adjacent to D or B. Its location is determined by another change in T_b. If this change leads to the successive melting A-C-D-E-B with n = 2,0,0,1,−2, respectively, then B is situated between C and E (otherwise B would not change its n), and the sequence is DACBE. I will not discuss all possible cases of domain situation in DNA, but each time this situation is easily decoded. Different sequences of melting domains (related to different solvents) play the role of different fragmentations with respect to the determination of the domain order.

Another approach, somewhat less accurate but simpler, to the determination of the domain sequence is related to melting curves, which are slightly nonequilibrium (due to superheating or supercooling of certain domains). Of course, equilibrium would be achieved if one waited long enough. In the case of superheating, for example, at a temperature higher than the equilibrium domain T_m, a somewhat larger (than the equilibrium) domain may melt, thus shifting the boundary between adjacent domains. Eq. 1 is applicable to these domains with shifted boundaries. As one of the domains increases and another domain correspondingly decreases, we can determine the adjacent domains from slightly different nonequilibrium DMC. (This situation will be demonstrated in the example of ϕXD in the next section.)

All of this suggested analysis is nondestructive, if the DNA heating is stopped before the complete separation of strands. Suppose a specific fragmentation occurs—e.g., inside domain C in Fig. 1. This fragmentation affects only the former domain C, dividing it into two border domains. Their lengths and compositions are different from those of each other and from those of C. The melting of each of them may annihilate (or create) only one phase boundary (melting of C annihilated two phase boundaries). So, the melting of the parts of C occurs at "their" own T_ms. Thus, one peak at the DMC dN_c/dT after DNA fragmentation splits into two peaks at the DMC of two separated (but nonfractionated) fragments. This splitting is clearly seen in the ϕXD curve and the summary curve of its fragments, Y_1 and Y_2 in Fig. 3. The new peaks, whose summed area equals the area (i.e., the length) of the split peak, determine the fragmentized domain, its parts, and therefore the order of fragments in DNA. If, knowing the order of these fragments, we perform another fragmentation, we may determine the order of the new fragments. By performing successive fragmentations and each time determining the order of the resultant

* It may be, for instance, a periodic polymer with a period of ℓ A·T pairs followed by ℓ' G·C pairs. If ℓ and ℓ' are large enough, then, by Eq. 1 the peaks will be related to the successive melting of ℓ border A·T pairs ($n = 1$, $x = 1$); ℓ inside A·T pairs ($n = 2$, $x = 1$); ℓ inside G·C-pairs ($n = -2$, $x = 0$); and ℓ' border G·C pairs ($n = -1$, $x = 0$). The positions of the peaks allow us to determine, from Eq. 1, T_0, τ, and T_b.

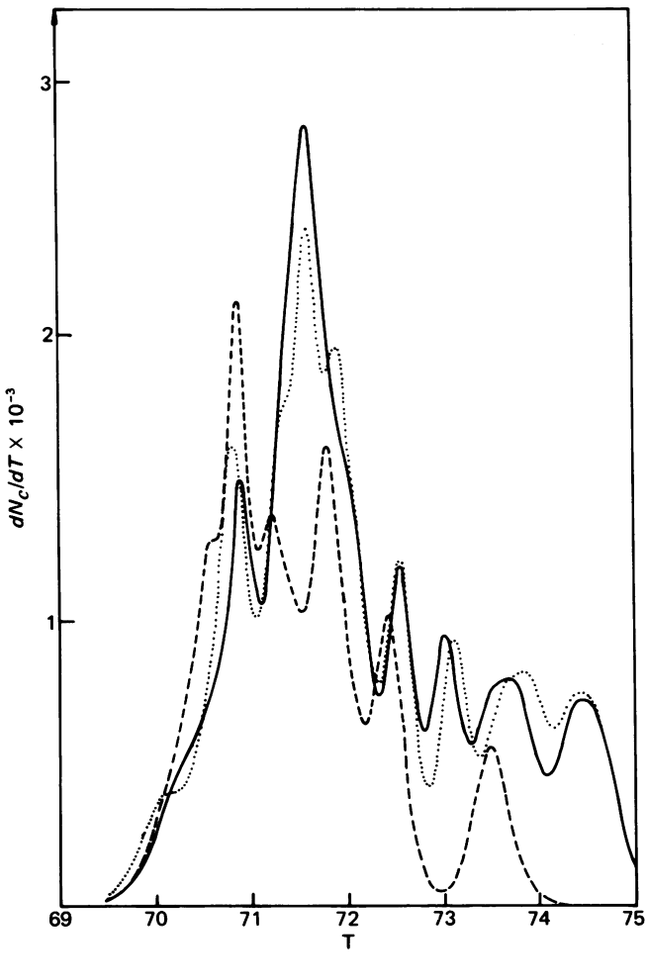

FIG. 3. Experimental DMCs in 15 mM NaCl/1.5 mM Na citrate solution of ϕXD (14) (solid line) and of the summary DMC (dashed line) of its fragments (13) Y_1 and Y_2, which form together the main part of ϕXD. The numerically computed (according to Eq. 3 with spectral data β, T_m, and ℓ_m from Table 1) curves coincided with experimental ones within the accuracy of the experiment. The dotted line represents the same computation for ϕXD but with changed T_ms changed according to the 20% decrease in T_b from Eq. 1. T_b (which is not even accounted for in the commonly used Marmur–Doty formula) bears the main responsibility for the change in the shape of the DMC due to the change in the solvent.

fragments, we finally determine the order of all fragments without fractionating them. Also, we do not destroy them if we stop DNA heating before a complete separation of strands.

QUANTITATIVE APPROACH AND ITS APPLICATION TO ϕXD

The quantitative formula for dN_c/dT, based on the accurate evaluation of the partition function (and, of course, accounting for excitations), is (19, 21):

$$\frac{dN_c}{dT} = \sum_n \frac{\beta \ell_m{}^2}{2 \cosh^2 [\beta \ell_m (T - T_m)]}. \qquad [3]$$

Here, ℓ_m is the length of the mth melting domain and T_m is its melting temperature, determined by Eq. 1. This formula (containing only one adjustable parameter, β, which can be determined by applying Eq. 3 to the melting of a polymer of a known structure—e.g., a periodic one) represents the explicit analytical description of the natural DNA melting. It coincides with the experiments (13, 14) within their accuracy (see below).

Decomposition of the experimental dN_c/dT into sum 3 allows us to determine accurate values of ℓ_m and T_m, even if domain peaks overlap, and to analyze these data.†

Eq. 3 is convenient for the spectral analysis. The maximal height of any peaks $(dN_c/dT)_{max}$ obviously exceeds $\beta \ell^2{}_{max}$, where ℓ_{max} is the maximal length among the lengths of domains essentially contributing to the peak. Thus, $\ell_{max} < [\beta^{-1}(dN_c/dT)_{max}]^{1/2}$. The minimal number of contributing domains is ℓ/ℓ_{max}, ℓ being the total area of the peak (i.e., the total length of these domains). The half-width of each domain peak approximately equals $1/(\beta \ell_m)$ and becomes larger as ℓ_m becomes smaller. So the half-width of the experimental peaks gives the estimate for the minimal length of the contributing domains. Experimental T_ms contributing to one peak are obviously within the width of the peak (i.e., relatively very close to each other). So the spectral analysis of Eq. 3 is easy and simple. I shall demonstrate it with ϕXD as the example. The DMCs in 15 mM NaCl/1.5 mM Na citrate solution of the entire DNA and its two fragments Y_1 and Y_2 were obtained by two different groups in Tokyo (13) and Moscow (14) (see Fig. 3; DMCs for separate fragments are not presented). Unfortunately, the experiments are not as accurate as I should like them to be and are slightly nonequilibrium (DMCs in two consecutive experiments differ somewhat). The authors of ref. 14 suggested contamination of DNA. The fragments in ref. 13 did not form an entire DNA. Only one solvent was used. Thus, the experiments are not too favorable for our purpose.

The parameters T_0 and τ were determined for the solvent 15 mM NaCl/1.5 mM Na citrate in another set of experiments (32) and were found to be $T_0 = 94.9°C$, $\tau = 42.4°C$. Yet, the results of the spectral analysis are quite demonstrative. Just two adjustable parameters ($\beta = 0.0106$ degree^{-1} and $T_b = 100°C$) are used for three DMCs (of ϕXD and two fragments). I obtain 50 parameters (ℓ_m and x_m) of 25 domains, which are in a surprisingly good agreement‡ (Table 1) with the corresponding values for the known (6) ϕXD nucleotide sequence. I replaced all unknown sites in DNA by A·T pairs. So, if the parameters T_0 and τ from ref. 32 and the parameter T_b are accurate, the coincidence between "our" results and the ϕXD data may indicate that: (*i*) three unknown base pairs at $1Y_2$ are A·T pairs, and elsewhere the $1Y_2$ sequence is precise; (*ii*) the unknown pair at $3Y_2$ is G·C; (*iii*) G+C content is underestimated at $4Y_2$, $4Y_1$, $5Y_1$, $6Y_1$, and $7Y_1$ and is overestimated at $1Y_2$. In general, the discrepancy between our data and those for the ϕXD sequence is about 1% where the sequence is precise and increases to 5–6% in two of the four cases in which the sequence is just as inaccurate.

Three sets of (ℓ_m, T_m) of ϕXD DMC are not included in Table 1 because they do not correspond to any domains in DNA and are related to the contamination indicated (but not specifically identified) in ref. 14. These sets are (348, 70.58°C), (344, 74.35°C), and (338, 74.64°C).

† The first proposal to apply the DNA melting curve to its sequencing was made in ref. 15. The Marmur–Doty formula has been applied to the analysis of the DMC (e.g., in refs. 13, 22–29) but none of the authors accounted for the important contribution (described by the last term in Eq. 1) to T_m of the phase boundaries and did not know the formula (3) for dN_c/dT which allows us to perform the spectral analysis. The spectral analysis is important because usually a single experimentally observed peak consists of several overlapping peaks from domains in quite different parts of the DNA.

‡ This agreement is related to a relatively small contribution to Eqs. 1 and 3 of other factors such as effects on T_m of a more detailed base sequence and of loop entropy (in particular, the influence of the adjoining loop regions on the T_m of a helical section). These effects are easily taken into account in our approach (15, 19, 21), but for natural DNA their contribution is less than the existing experimental accuracy. This is verified also by the agreement of Eq. 3 with experiments (13, 14).

Table 1. Spectral analysis of ϕXD and its fragments

ℓ_m (D)	ℓ_m^E (D)	ℓ_m (F)	ℓ_m^E (F)	x_m (D)	x_m^E (D)	x_m (F)	x_m^E (F)	δx_m	T_m (D)	T_m (F)	1st site* (D)	1st site* (F)	Fragment
291	289	287	289	0.599	0.599	0.599	0.599	±0.1	70.2	70.2	1180[c]	1180	$1Y_2$[†]
530	530	453	433	0.575	0.575	0.579	0.575	0	70.9	70.56	2361[d]	2361	$2Y_2$[‡]
256	220	215	220	0.577	0.591	0.574	0.591	±0.005	71.2	71.5	3932[a]	3932	$4Y_1$[§]
483	475	468	457	0.565	0.571	0.559	0.558	±0.002	71.36	71.2	4977[a]	4977[b]	$3Y_1$[‖]
380	369	354	348	0.553	0.557	0.553	0.557	+0.003 −0.006	71.44	71.44	1469[c]	1469	$3Y_2$[‖]
607	615	641	671	0.550	0.554	0.570	0.566	±0.002	71.6	70.87	77	59	$2Y_1$[‡‖]
486	523	520	544	0.536	0.545	0.536	0.544	±0.008	71.8	71.8	1838[c]	1817	$4Y_2$[‖]
411	412	233	239	0.550	0.546	0.585	0.582	+0.01 −0.1	72.06	70.5	3187[e]	3360	$1Y_1$[†‡]
340	333	333	333	0.526	0.571	0.526	0.571	+0.01 −0.05	72.06	72.0	3599[e]	3599	$5Y_1$[†]
516	490	481	490	0.528	0.559	0.530	0.559	±0.015	72.56	72.44	4152[a]	4152	$6Y_1$[†]
465	488	Extra fragment		0.505	0.502	Extra fragment		0	73.02		692		
368	335	374	335	0.493	0.501	0.493	0.501	±0.003	73.49	73.5	4642[a]	4642	$7Y_1$
374[§]	296	Extra fragment		0.485	0.527	Extra fragment		+0.02 −0.05	73.8		2891[e]		

Domain length (ℓ_m), melting temperature (T_m) and A+T content (x_m) according to spectral analysis; (D), for ϕXD and (F) for fragments indicated in the last column [e.g., $3Y_1$ is the third-melting domain of the fragment Y_1 (position 3360–5375–729); $2Y_2$ is the second-melting domain of the fragment Y_2 (position 1104–2793)]. Superscript E indicates the corresponding domains in the ϕXD sequence (6); their position in ϕXD (according to ref. 15) is determined by their first sites and their lengths. The uncertainty (6) in ϕXD sequence provides the inaccuracy δx_m in x_m (which is the same for DNA and its fragments). The first-melting domains create the first phase boundaries, so in Eq. 1 $n = 1$ for border domains $1Y_1$|and|$2Y_1$|and|$n = 2$ for|inner domain Y_2. The last-melting inner domains annihilate the last phase boundaries, so $n = -2$ for $7Y_1$ and $4Y_2$. The domain $3Y_1$, adjacent to the already melted $2Y_1$, has $n = 0$. Other ns are determined for ϕXD and its fragments according to ref. 15. Finally, subsequently melting domains have the following ns: 1, 1, 0, 2, −2, 0, and −2 for Y_1; 2, 1, 0, and −2 for Y_2; 2, 2, 2, 2, 2, 0, 0, −2, 2, −2, 0, −2, −2, and −2 for ϕXD.

* a, b, c, d, and e denote domains that approximately form genes A, B, F, G, and H (the positions of the genes are 3972–5375–133; 5064–5375–48; 1001–2275; 2362–2891; 2923–3906).

† Uncertain DNA sequence (6).

‡ Fragment boundary.

§ Contaminated domain.

‖ Domain boundary shift.

The domains $2Y_2$, $2Y_1$, and $1Y_1$, whose T_ms and lengths essentially differ for the fragment and for ϕXD, are border ones. (One of the Y_2 border domains is lost; it is very short—just 74 base pairs—and very refractory. It melts at ≈77.5°C, where experimental errors are of order of its own size.) A slight discrepancy between (ℓ_m, T_m) for a fragment and for ϕXD is related to a slight nonequilibrium shift of domain boundaries, which exposes the adjacency of the corresponding domains ($2Y_1$ and $3Y_1$; $3Y_2$ and $4Y_2$). An exception is $4Y_1$, which is probably contaminated.

It is remarkable that all three DMCs calculated according to Eq. 3 with the data for (ℓ_m, T_m) from Table 1 coincide with tht experimental ones within the experimental accuracy. This is related to the high sensitivity of Eq. 3 to any change in the parameters. Such a sensitivity arises, in particular, from the different shift of peaks due to a shift in T_b. This different shift explains also changes in DMC related to small changes in pH or ion concentration of the solvents (33). Fig. 3 demonstrates the change in the DMC of ϕXD due to a 20% change in T_b.

An essential change in T_b changes, according to ref. 15, the very melting domains. By Eq. 1, the less ℓ_m is, the more sensitive the T_m is to T_b. So, small domains may melt much earlier or much later than the rest of DNA.

The sensitivity of the DMC to the exact sequence and to the solvent makes each DMC very specific. Therefore, it can be used as an excellent "identification card" for DNA and, if the DNA sequence is known, for its contamination. It may be accurate enough even to identify a mutation or "cancer" DNA.

For an arbitrary wavelength the light absorption is a linear function of the numbers $N_c^{(1)}$ and $N_c^{(2)}$ of A·T and G·C melted pairs, respectively. The equations for $dN_c^{(1)}/dT$ and $dN_c^{(2)}/dT$ differ from Eq. 3 only by factors x_m and $(1 - x_m)$ in each term in the sum. Therefore, if one uses different wavelengths (21, 34, 35), one can determine ℓ_m, x_m, n_m from experiments for one solvent. The described approach is readily applicable to DNA interacting in the solvent with proteins, drugs, etc. (as well as to metalated DNA), thus increasing the information obtained and, in particular, allowing the determination of the location of the corresponding interactions with the DNA.

I am indebted to Professors A. Aharony, Y. Imry, and D. Bergman, whose help and support were invaluable to me. I am most grateful to Dr. H. Scheraga for stimulating discussions. Also, Dr. A. Golosovker's numerical computations are highly appreciated. My gratitude goes to Dr. M. Raman for his unfailing help and my sincere appreciation, to Professor Freeman Dyson for his hospitality at the Institute for Advanced Study, where this paper was revised and completed.

1. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94**, 441–448.
2. Sanger, F., Niclen, A. K. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
3. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
4. Min You, W., Haegman, G., Ysebaert, M. & Fiers, W. (1972) *Nature (London)* **237**, 82–88.
5. Fiers, W., Contreras, R., Duerinek, F., Haegman, G., Merregaert, J., Min Jou, W., Raeymakers, H., Volckaert, G., Ysebaert, M., Van de Kerckhove, J., Nolf, F. & Van Montagu, M. (1973) *Nature (London)* **256**, 273–278.
6. Fiers, W., Contreras, R., Duerinck, F., Haegemann, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van den Berghe, A., Volkaert, G. & Ysebaert, M. (1976) *Nature (London)* **260**, 500–507.
7. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hatchison, C. A., III, Sloconibe, P. M. & Smith, M. (1977) *Nature (London)* **265**, 687–695.

8. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978) *Science* **200,** 494–502.
9. Fiers, W., Coutreras, R., Haegeman, G., Rogiers, R., Van de Voozde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (London)* **273,** 113–120.
10. Poland, D. & Scheraga, H. A. (1970) *Theory of Helix-Coil Transition in Biopolymers* (Academic, New York).
11. Wartell, R. W. & Montroll, E. W. (1972) *Adv. Chem. Phys.* **22,** 129–203.
12. Vedenov, A. A., Dykhne, A. M. & Frank-Kamenetskii, M. D. (1972) *Sov. Phys. Usp.* **14,** 715–736.
13. Wada, A., Tachibana, H., Ueno, S., Husimi, U. & Machida, Y. (1977) *Nature (London)* **269,** 352–353.
14. Lyubchenko, Y. L., Vologodskii, A. V. & Frank-Kamienetskii, M. D. (1977) *Nature (London)* **271,** 28–31.
15. Azbel, M. Ya. (1973) *Phys. Rev. Lett.* **31,** 589–593; 1592E.
16. Lifshitz, I. M. (1974) *Sov. Phys. JETP* **38,** 545–549.
17. Azbel, M. Ya. (1972) *JETP Lett.* **16,** 128–131.
18. Azbel, M. Ya. (1973) *Biopolymers* **12,** 1591–1609.
19. Azbel, M. Ya. (1978) *Ann. Isr. Phys. Soc.* **2,** 25–36.
20. Azbel, M. Ya. (1975) *J. Chem. Phys.* **62,** 3635–3647.
21. Azbel, M. Ya. (1978) *J. Phys. A. Gen. Phys.* **12,** in press.
22. Steinert, M. & van Assel, S. (1974) *Biochem. Biophys. Res. Commun.* **61,** 1249–1255.
23. Yubuki, S., Gotoh, O. & Wada, A. (1975) *Biochim. Biophys. Acta* **395,** 258–273.
24. Vizard, D. L. & Ansevin, A. T. (1976) *Biochemistry* **15,** 741–750.
25. Ansevin, A. T., Vizard, D. L., Brown, B. W. & McConathy, J. (1976) *Biopolymers* **15,** 153–174.
26. Gotoh, O., Husimi, Y., Yabuki, S. & Wada, A. (1976) *Biopolymers* **15,** 655–670.
27. Akiyama, C., Gotoh, O. & Wada, A. (1977) *Biopolymers* **16,** 427–435.
28. Wada, A., Tachitana, H., Gotoh, O. & Takanami, M. (1976) *Nature (London)* **263,** 439–440.
29. Reiss, C. & Arpa-Gabarro, T. (1977) *Prog. Mol. Subcell. Biol.* **5,** 1–30.
30. Tachibana, H., Wada, A., Gotoh, O. & Takananui, M. (1978) *Biochim. Biophys. Acta* **517,** 319–328.
31. Poland, D. (1974) *Biopolymers* **13,** 1859–1871.
32. Lazurkin, Y. S., Lyubchenko, Y. L., Pavlov, V. M. & Frank-Kamenetskii, M. D. (1975) *Biopolymers* **14,** 1551–1562.
33. Blake, R. D. & Lefoley, S. G. (1978) *Biochim. Biophys. Acta* **518,** 233–246.
34. Hirschman, S. Z., Gellert, M., Falkow, S. & Felsenfeld, G. (1967) *J. Mol. Biol.* **28,** 469–477.
35. Hirschman, S. Z. & Felsenfeld, G. (1966) *J. Mol. Biol.* **16,** 347–358.