

DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: Correspondence with melting theory

(screening for mutation/gel electrophoresis/helix-coil transition/ λ phage/restriction fragment)

S. G. FISCHER AND L. S. LERMAN

Center for Biological Macromolecules, Department of Biological Sciences, State University of New York, Albany, New York 12222

Communicated by Rollin D. Hotchkiss, December 23, 1982

ABSTRACT DNA fragments 536 base pairs long differing by single base-pair substitutions were clearly separated in denaturing gradient gel electrophoresis. Transversions as well as transitions were detected. The correspondence between the gradient gel measurements and the sequence-specific statistical mechanical theory of melting shows that mutations affecting final gradient penetration lie within the first cooperatively melting sequence. Fragments carrying substitutions in domains melting at a higher temperature reach final gel positions indistinguishable from wild type. The gradient data and the sites of substitution bracket the boundary between the first domain and its neighboring higher-melting domain within eight base pairs or fewer, in agreement with the calculated boundary. The correspondence between the gradient displacement of the mutants and the calculated change in helix stability permits substantial inference as to the type of substitution. Excision of the lowest melting domain allows recognition of mutants in the next ranking domain.

Detection and localization of single base substitutions within long DNA sequences may be impractical by complete sequence determination and improbable on the basis of restriction endonuclease vulnerability. We present here the results of a procedure by which DNA molecules that have minimal sequence differences are separated and by which some conclusions can be drawn as to the nature of the change. A number of samples can conveniently be examined in a single slab gel; each DNA species is focused into a sharp band at a gel position determined by its sequence and composition. The physical separation of fragments of altered sequence provided by the denaturing gel makes possible further analysis and manipulation.

Our system makes an unconventional use of electrophoresis. Where DNA molecules migrate into a gradient of ascending concentration of denaturant, they undergo an abrupt decrease in mobility at a characteristic depth, resulting in positions and patterns that change little if application of the field is continued. The retardation depth in the gradient is determined by the least stable part of the molecule and is relatively insensitive to other parts of the sequence or to the overall length (1).

To understand the basis of the sensitivity of the system to single base substitutions, we have undertaken a close comparison of the gel results with those of a sequence-specific statistical mechanical theory of the stability of the double helix. Experimental studies on the helix-disorder transition, melting, have not yet provided a detailed test of the theory, which predicts intricate and interesting patterns for the progression of equilibria from full helicity to separated strands as the temperature increases for molecules of different sequence. In our system, the molecule is exposed to a gradual denaturation-pro-

moting change in the medium, linearly equivalent (2) to a gradual increase in temperature. The strong decrease in mobility as the helix unravels provides the basis for both sequence-determined separation and examination of the helix-disorder transition.

MATERIALS AND METHODS

DNA Preparation. λ strains were obtained from D. Wulff as *Sam7* derivatives substituted in the *y* region. Strain identifications and sequence determinations are from Wulff *et al.* (3) and D. Wulff and M. Rosenberg (personal communications). DNA from those strains was prepared as described. DNA was also prepared from λ strain *c1857* by lytic infection of *Escherichia coli* strain K-12 W3110 by standard methods. Plasmid pKM2 (K. McKenney and M. Rosenberg, personal communication), containing the wild-type λ sequence at position 38,989-40,291 inserted into the *Hind*III site of pBR322, was grown in *E. coli* HB101. Phage and plasmid DNAs were digested to completion with *Ava* I/*Bgl* II or *Alu* I/*Bgl* II under the conditions specified by the supplier (Bethesda Research Laboratories). Digestion was stopped by the addition of EDTA to a final concentration of 40 mM, glycerol to a final concentration of 10% (vol/vol), and a trace of bromphenol blue tracking dye.

Gel Electrophoresis. Nondenaturing polyacrylamide (65 mg/ml; acrylamide/bisacrylamide, 30:0.8) gels were run at 60°C in TAE buffer (40 mM Tris/20 mM NaOAc/1 mM EDTA, pH 8/HOAc), with 3-5 μ g of whole phage DNA or 0.5-1.0 μ g of plasmid DNA in each slot.

Denaturing Gradient Gels. The acrylamide concentration, 65 mg/ml, and the TAE buffer concentration were uniform throughout the gels. All gradients consisted of linearly increasing concentrations of urea/formamide at a constant ratio. Gels were poured from outgassed solutions containing ammonium persulfate at 0.1 mg/ml and 0.01% *N,N,N',N'*-tetramethylethylenediamine using a two-chamber gradient maker or the syringe gradient pump previously described (4, 5). The gels were submerged in a 5-liter aquarium that contained the anode electrolyte, TAE. The electrolyte was stirred and controlled at 60.0°C, and a field of 6 V/cm was applied.

RESULTS

Gradient Discrimination of Mutants. We have compared wild type and 16 strains of λ that have substitutions in the *y* region using fragments produced by cleavage with *Ava* I/*Bgl* II. The fragments contain 536 base pairs (bp) preceded by four unpaired bases from the *Bgl* II site. Numbering begins with the first paired base in the *Ava* I site. A map of the sites and substitutions for each of the mutants is shown in Fig. 1. Strain *cin-1 cnc-1* is a double mutation containing both *cin-1* and *cnc-1* substitutions.

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.

Abbreviations: bp, base pair(s); T_m , melting transition.

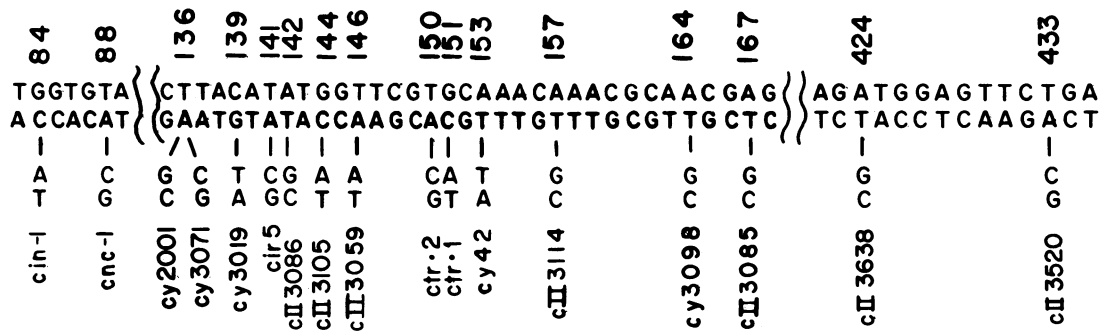


FIG. 1. Sites of λ region mutants. According to the convention by which the *EcoRI* site between the fragments whose lengths are 7,421 bp and 5,806 bp is numbered 40,000, the first base of this fragment, indicated here as 1, would be numbered 39,125.

The substitutions *ctr1* and *ctr2* are present only in strains also carrying the substitution *cII3059*, and strain *cir5* contains both *cir5* and *cII3086* substitutions.

DNA of each strain was digested with *Ava I/Bgl II* and separated by length in a nondenaturing polyacrylamide gel. The corresponding 536-bp fragments migrate to the same depth in each lane. The narrow strip carrying the bands was sealed across the top of a gradient gel in which denaturant concentration increased from top to bottom in a uniform polyacrylamide matrix. A field of 150 V was applied for 14.5 hr at 60°C.

As shown in Fig. 2, which presents a set of similar gradient gels with the wild-type fragment and 16 mutants, substitution of T·A for C·G at position 129, strain *cy3019*, results in retardation 0.9 cm higher in the gel than for wild type. The inverse substitution at position 136, strain *cy3071*, delays retardation beyond that of wild type. The shifts are not fully defined by the composition of the substitution; *cy3019* is displaced almost twice as far as *cII3105*, although both represent replacements of G·C by A·T. The difference between *cy2001* and *cy3071* is a simple transversion, the interchange of cytosine and guanine across strands, but there is a more-than-2-mm difference in gel positions. The doubly substituted strain, *cin-1 cnc-1*, identical to wild type in gross composition in that it carries both A·T → G·C and G·C → T·A substitutions, is also shifted. A single T·A → C·G substitution at position 142 in strain *cII3086* effects retardation below wild type while an additional A·T → G·C substitution in strain *cir5* delays retardation to an even greater depth. All of these substitutions occur within the first 144 bp adjacent to the *Ava I* site. The four mutants with substitutions at base 153 and above have no detectable effect on the gradient position. Strain *cII3059* and its derivative double mutants are discussed below.

Effect of Single Base Substitution on the Mobility Shift. The basis for separation is shown by comparison of the electrophoretic mobilities of two fragments in gels in which both fragments migrate through a constant concentration of the denaturing solvent (4). The variation in mobility over a substantial range of denaturant concentration is displayed by a denaturing gradient perpendicular to the electric field. In this procedure, 20 μ g of wild-type λ *Sam7* DNA and 10 μ g of DNA from strain λ *Sam7cy2001*, containing an A·T → G·C transition at position 136, were cleaved with *Ava I/Bgl II* and run into an agarose (15 mg/ml) slab gel from a single 13-cm-wide starting zone. An agarose strip containing the 536-bp band was cut from the ethidium-stained gel and sealed with hot agarose across the top of a polyacrylamide (65 mg/ml) gel containing a linear gradient from 1.4 M urea/8% (vol/vol) formamide on the left to 3.5 M urea/20% (vol/vol) formamide on the right. A field of 150 V perpendicular to the gradient was applied for 7 hr at 60°C (Fig. 3).

Fragments on the left migrate through a column of minimum denaturant concentration. From left to right, fragment mobility at first decreases gradually, then suffers a sharp reduction two-thirds of the way to the right at about 2.8 M urea/16% (vol/vol) formamide. At this point, the fragment from strain 2001 is distinguishable from wild type as a species that has half the staining intensity; its sharp mobility transition occurs at a very slightly higher denaturant concentration. The wild-type and mutant fragments appear to be superposed in both the high- and low-concentration regions where the mobility changes only slightly with increasing denaturant concentration. Fragment mobility decreases to less than one-fourth through the major transition zone.

The Calculated Melting Map. To examine the effect of a single base substitution on the calculated progression of the helix-

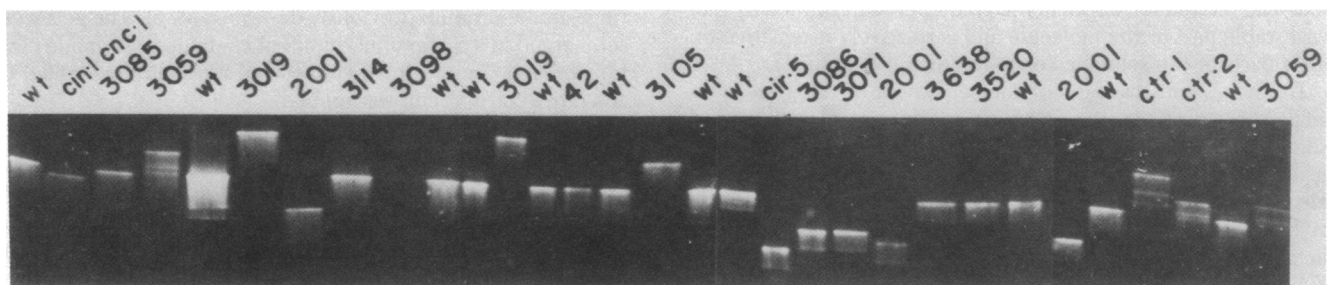


FIG. 2. Effect of mutation on fragment depth in denaturing gradient gels. Three to 5 μ g of λ *Sam7* DNA and derivatives substituted in the γ region (see Fig. 1) digested to completion with *Ava I/Bgl II* was electrophoretically separated in a nondenaturing polyacrylamide (65 mg/ml) gel. The ethidium-stained polyacrylamide strip containing the 536-bp γ -region fragment was sealed with hot agarose (10 mg/ml) in TAE buffer across the top of a polyacrylamide (65 mg/ml) gel containing a linear gradient of a urea/formamide mixture increasing, from 2.1 M urea/12% (vol/vol) formamide to 3.5 M urea/20% (vol/vol) formamide, in the same direction as the electric field. The figure is a composite of the central 2.5 cm of each of four gels, each containing one or more wild-type (wt) samples.

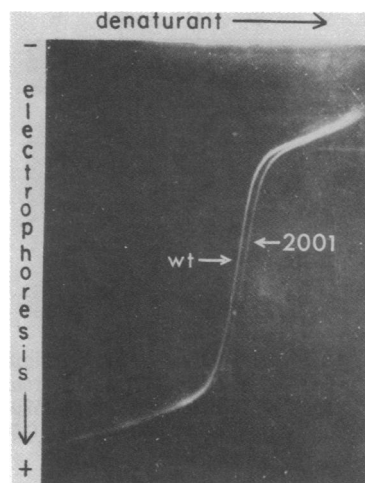


FIG. 3. Comparison of mobilities of wild-type and *cy2001* fragments in a denaturing gradient perpendicular to the electric field. A mixture of wild-type and *cy2001* fragments was electrophoresed in a gel in which the denaturant concentration was constant along the path of electrophoretic movement but increased linearly in the perpendicular direction. The sample was applied uniformly along the top from a strip of an agarose gel containing nominally the same total amount of both fragments at every point. The amount of mutant DNA was half that of wild-type DNA.

random chain equilibrium, we have used the Fixman and Friere (6) modification of the algorithm presented by Poland (7) for calculation of the equilibrium melting transition (T_m) probability as a function of sequence and temperature. We have replaced the 2-valued stability parameters for base pairs (usually given as T_{MAT} or T_{mGC}) by the set of 10 values for nearest-neighbor doublets suggested by Gotoh and Tagashira (8). The values are given in Fig. 4. A base-pair substitution changes the stability values in two adjacent positions, and a transversion without substitution results in a significantly different net value although the overall base composition remains constant. Since the Poland-Fixman-Friere calculation depends on at least two statistical parameters that are not precisely known (the cooperativity constant, σ , and a loop-closure exponent, α), we have compared results for several values and combinations of each. We have ascribed the variation in T_m of nearest-neighbor doublets to variation in ΔH , holding ΔS constant, but the converse assumption gives essentially the same results. The patterns of melting progression and the differences due to substitution show only insignificant differences so long as the parameters do not depart drastically from values previously published by others. All of the results shown here are based on $\sigma = 3.3 \times 10^{-5}$, the center of the range proposed by Amirikyan *et al.* (9), and $\alpha = 2.0$ (10).

The base sequence of the fragment enters the calculation of the melting progression as the sequence of nearest-neighbor stability values, represented as temperatures. The expected melting progression along the molecule calculated with standard parameters is shown in Fig. 4 as a melting map—the temperature at which each base will be at equilibrium with equal probability of helix or random chain configuration. At any temperature appreciably below the contour for that base pair, the pair can be regarded as helical and, at any temperature appreciably above the contour, the pair can be regarded as melted. As expected, a few base pairs at each end melt at lower temperatures; ends of the helix fray gradually prior to any large cooperative changes. The first domain to melt consists of the bases between 32 and 142. The uniformity of the ordinate value shows that all of the bases of this region melt as a block. A base at the

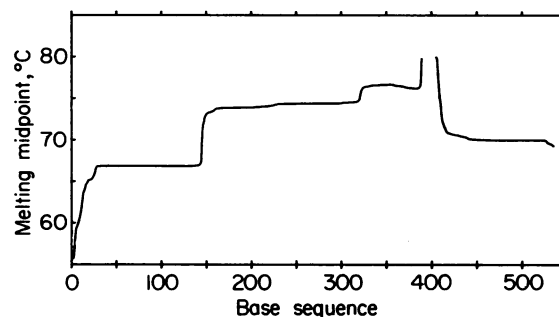


FIG. 4. Expected progression of melting of the wild-type λ fragment. The abscissa represents positions along the sequence between nearest-neighbor base pairs from the first pair following the *Ava*I scission position through the last pair at the *Bgl*II site. The distribution of each base pair between the helical and melted states was determined from the Fixman-Friere (6) algorithm at 0.1°C increments, and the temperature for 50:50 equilibrium was inferred by cubic interpolation. The function is not shown in the highest melting region between bases 389 and 405 where the properties depend on the probability of complete separation and bimolecular reassociation. $\sigma = 3.3 \times 10^{-5}$, $\delta(n) = \sigma n^{-2.0}$, approximated as a sum of nine exponentials; $\Delta S/R = 12.25$ mol. The Gotoh-Tagashira values for nearest-neighbor 5'-3' doublets in 19.5 mM Na^+ are as follows: T-A, 36.73°C; T-T, 54.50°C; T-G, 86.44°C; A-T, 57.02°C; A-G, 58.42°C; A-C, 97.73°C; C-G, 72.55°C; C-C, 85.97°C; G-C, 136.12°C (8).

center of the block progresses from 10% to 90% probability of unpairing and unstacking between 66.3°C and 67.3°C. The calculated contour rises 0.3°C between bases 143 and 144, 2.8°C between bases 144 and 145, 1.5°C between bases 145 and 146, and 0.4°C between bases 146 and 147, giving the appearance of a steep wall. The last melting necessary for strand separation is omitted; it is concentration dependent and unnecessary for the present analysis. The calculation shows, in general, that melting of a DNA molecule can be expected to proceed stepwise as the temperature is raised; the effect of cooperativity is strong enough that fairly long blocks of contiguous helix melt within temperature intervals much narrower than the temperature differences between entire blocks. The separation into blocks, or domains, follows from the algorithm, and no assumptions or *a priori* estimates of domain boundaries are required. The loci of domain boundaries are not discernable by inspection of the sequence.

We note that the lowest melting domain contains the *cyL* region, the sequence thought to determine polymerase recognition, and the *cyR* region falls into a higher melting domain beginning at about base 147 (3).

The changes in the melting progression effected by mutations between bases 83 and 168 are presented in Fig. 5, together with a section of the standard map. The calculations are shown in Fig. 5B and C as the differences in temperature for the 50% point in the melting equilibrium between the wild-type fragment and each mutant at each base pair. The difference maps were obtained by subtracting the melting map of the complete sequence of the wild-type fragment from the corresponding complete map of each mutant. As shown in Fig. 5B for four representative substitutions and a double substitution, substitutions that effect substantial displacements in the gels distinctly alter the melting temperature in the entire first domain. Except for slight shifts in the domain boundaries in this set (a shift appears as a spike in the melting map), melting is essentially identical with that of wild type at all other parts of the sequence. The effect of the double substitution, *cin-1 cnc-1*, and the comparison between *cy2001* and *cy3071* are of particular interest because corresponding differences are seen both in the gel positions and in the calculation between domains of identical

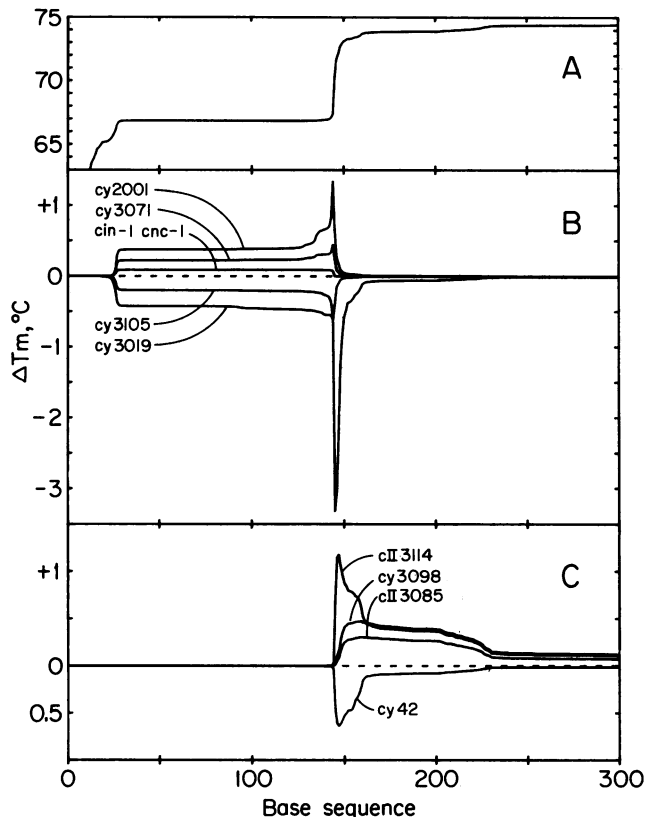


FIG. 5. Region of the first (lowest melting) domain and difference melting maps resulting from base substitutions. (A) The low-numbered region of the melting map of the wild-type fragment (Fig. 4) is shown with an expanded temperature scale. (B and C) T_m values for each base in the wild-type sequence have been subtracted from T_m values calculated for the mutant sequences.

base composition as a consequence of nearest-neighbor interactions.

The calculated melting difference maps for strains that reach retardation depths in the gel indistinguishable from the parental λ strain are shown in Fig. 5C. They have a negligible effect (less than 0.01°C) below base 145 in the first domain but depress the edge of the adjacent high-melting domain. The effects of these mutations tend to be significant from base 147 to about base 225. Substitutions at sites 424 in cII3638 and 433 in cII3520, show gradient positions indistinguishable from wild type and affect the melting map only near the high-numbered end.

Detection of Mutants Near the *Bgl* II End. The region within which substitutions in the 536-bp *Ava* I/*Bgl* II fragment are discernible in the gel corresponds to the lowest melting domain of the theoretical map. Excision of the low-numbered end of the sequence can be expected to promote the region of next higher T_m , extending from base 416 to the *Bgl* II end, to first melting, so that in the shortened fragment substitutions in this region should be recognizable. Gel positions of the truncated fragments of the wild type, cII3638, and cII3520, in which the same sequence was cleaved between bases 235 and 236 with *Alu* I, are shown in Fig. 6. Wild-type fragments from the recombinant plasmid pKM2 (lanes A and E) focused into a sharp band at the same depth as a band of the fragment from whole phage DNA (lane B). The mutant fragments, indistinguishable from wild-type fragments in the gradient in the original 536-bp molecule (Fig. 2) focus at greater depths, in good agreement with the calculated effect of the substitutions based on melting and mobility theory. Note that all of the fragments derived from whole phage DNA,

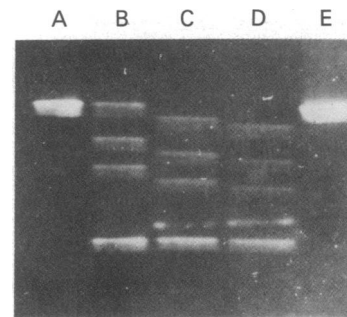


FIG. 6. Gel positions of truncated fragments substituted at the high-numbered end. The 301-bp *Alu* I/*Bgl* II fragments from pKM2 (lanes A and E), *Sam*7 (lane B), cII3638 (lane C), and cII3520 (lane D) were analyzed on a 42–60% denaturant/polyacrylamide (150 mg/ml) gel. The electric field was applied for 14 hr.

mutants and wild type, appear as triplet bands in which the substitutions uniformly shift all three members. A singlet lower in the gel, unaffected by the substitutions, provides a reference position. Tripletting of the 301-bp *Alu* I/*Bgl* II fragment was also obtained from wild-type phage grown lytically, while the plasmid-derived singlet was unaffected by mixed digestion with whole phage DNA. These results are consistent with the supposition that base modification during phage growth may be a source of the extra bands.

DISCUSSION

Because the gradient interval between any pair of fragments differing by a single base substitution in the determinant domain is larger than the width of the bands, samples can be recovered from the gels enriched for either component. The bands are narrower than those in simple length separations by constant-velocity electrophoresis because of the focusing due to the reduction in mobility as the determinant domain melts.

The correspondence between these results and properties calculated from the sequence by the Poland–Fixman–Friere algorithm provides more detailed support for this melting theory than has been available from hyperchromicity profiles. Because the six substitutions that alter gradient depth lie below position 145 and the six that do not alter gradient depth lie above position 152, it appears that the sequence above position 152 does not participate in determining the retardation depth. Following the explanation we have offered for identification of the decrease in mobility with partial melting (1, 11), we infer that the low-numbered region ending between bases 144 and 152 melts at a substantially lower temperature and independently of melting above base 152. That difference and the resulting decoupling emerges from the theoretical calculation as a distinct domain boundary within the limits specified by these mutants. The loss of mobility from melting of the first domain prevents the fragment from reaching the gradient depth necessary for further melting within the duration of the run.

Variation of the loop-entropy and cooperativity parameters can result in melting at a slightly lower temperature in the region between bases 83 and 131 than in the region near base 61, but there is no noticeable change in the maps shown in Fig. 5B nor in the boundary position.

There is a close proportionality between the calculated alteration, ΔT_m , of the plateau T_m of the first domain and the change in gradient depth reached by the fragments (Fig. 7). For sequences of different domain lengths, a simple comparison according to domain T_m is not appropriate and, hence, the sequence carrying the double substitution *cir*5–cII3086, which shifts the domain boundary about six bases to the *Ava* I end, is omit-

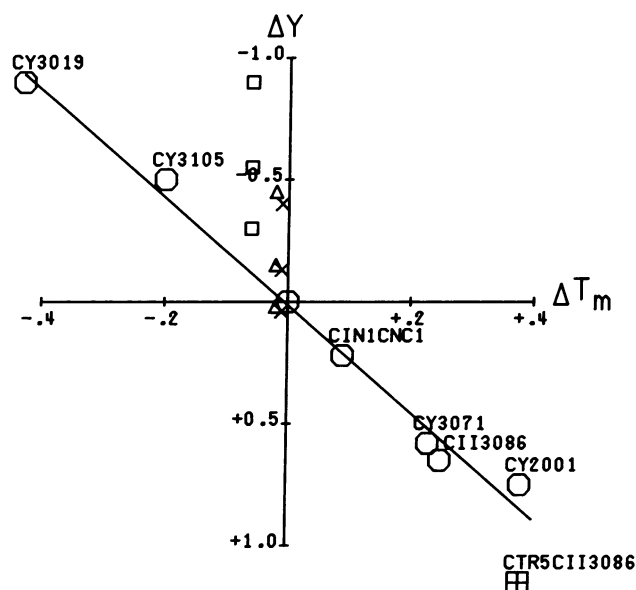


FIG. 7. Gel displacement and alteration in the T_m of the lowest melting domain. ΔT_m is the temperature difference between each mutant and wild type at the center of the first domain. ΔY is the increment in depth in the gradient referred to wild. The point at the origin represents wild type and seven strains carrying substitutions above base 152. The mutants that give triplet zones are as follows: Δ , cII3059; \square , cII3059 *ctr1*; \times , cII3059 *ctr2*; they have not been included in the calculation of the regression line.

ted from the regression line. The level of consistency between the calculated T_m of the first domain and the observed differences in gel depth support the relative assignments of nearest-neighbor stability values proposed by Gotoh and Tagashira (8) on the basis of less direct evidence. Our use of these values is somewhat arbitrary, in that they were proposed as applicable to conditions differing from those in the gels—0.02 M aqueous Na^+ , rather than 0.02 M Tris^+ and urea/formamide. However, the doublet stabilities always enter the calculation as the sum of nearest-neighbor values from both sides of each base pair, and the present results remain compatible with other doublet-stability assignments. Gel measurements with a large set of mutants may constitute an appropriate means to infer nearest-neighbor stability assignments.

Calculation of the melting map using only 2 stability values, one for G-C pairs and another for A-T, without consideration of neighbors results in a similar melting map. However, the 2-valued calculation cannot account for the gel displacement of the double mutant *cin-1 cnc-1*, which has a base composition identical to that of the wild type, nor for the difference due to transversion found in the comparison between *cy2001* and *cy3071*.

While most bands are accompanied to some extent by weak satellites slightly deeper in the gradient, the relative intensities of the satellites from cII3059 and its two derivatives are conspicuously greater. This property is retained through repeated plaque purification. Since the multiplicity also appears distinctly where the cII3059 fragment moves through a constant denaturant concentration, giving a pattern similar to that shown

in Fig. 3, the multiplicity is not generated by details of the gradient. The DNA in each band of the triplet behaves as a stable single component after isolation and migration into a new gradient. The calculated effect of the nominal base substitution in cII3059 appears in the difference melting map almost entirely as a perturbation beyond the boundary of the first domain. We have been unable to arrive at a significantly larger change in the first domain despite substantial *ad hoc* adjustment of many of the nearest-neighbor stability values by 5–20°, either singly or in combination, or by changes in other parameters. If the most displaced (highest) bands of the cII3059 mutants are related to the principal wild-type band, this effect could be construed as an influence on mobility originating from mutation about 6 bp pairs beyond the calculated boundary. If the least shifted (lowest) component is taken to indicate the effect of substitution, the displacements are compatible with the calculated domain boundary, as shown by the cluster of points near the origin in Fig. 7. The interval of 2 bp between cII3105 and cII3059 corresponds to the steepest part of the calculated boundary.

These results suggest that all substitutions from point mutations can be detected in denaturing gradient gels if they occur in the lowest melting domain of the molecule. Sensitivity will depend on the net stability change, which depends, in turn, on the specific context of the substitution, and on the length of the domain over which the effect is averaged. Perturbations at least as large as those from substitutions can be expected from single base pair insertions and deletions and from some base modifications. In the present example, the sensitive region initially constitutes 20% of the 536-bp fragment; it can be shifted to a different 20% by cleavage, which transfers highest melting priority to a different domain. Since the separations do not critically depend on the length of the molecules, random fragmentation may be as useful as restriction cleavage. By attachment of a higher melting section to small restriction fragments, nearly any sequence of appropriate length can be made to constitute the lowest melting domain in the resulting composite molecule, and nearly all substitutions may be made discernible.

We wish to thank Daniel Wulff and Michael Mahoney for strains from their collection of λ mutants. We are grateful to Karen Silverstein, Michele Soberman, Christine Cullen, Arthur Eisenberg, and Nelly Brown for assistance and to Marshall Fixman, Nadrian Seeman, and Richard Blake for helpful discussions. This work was supported by Grant GM-24030-06 from the National Institutes of Health.

1. Fischer, S. G. & Lerman, L. S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4420–4424.
2. Lerman, L. S., Fischer, S. G., Bregman, D. B. & Silverstein, K. J. (1981) in *Biomolecular Stereodynamics*, ed. Sarma, R. H. (Adenine, New York), pp. 459–470.
3. Wulff, D. L., Behr, M., Izumi, S., Beck, J., Shimatke, H., Brady, C., Court, D. & Rosenberg, M. (1980) *J. Mol. Biol.* 138, 209–230.
4. Fischer, S. G. & Lerman, L. S. (1979) *Cell* 16, 191–200.
5. Fischer, S. G. & Lerman, L. S. (1979) *Methods Enzymol.* 68, 183–191.
6. Fixman, M. & Friere, J. J. (1977) *Biopolymers* 16, 2693–2704.
7. Poland, D. (1974) *Biopolymers* 13, 1859–1871.
8. Gotoh, O. & Tagashira, Y. (1981) *Biopolymers* 20, 1033–1043.
9. Amiryanyan, B. R., Volodgodshii, A. V. & Lyubchenko, Y. L. (1981) *Nucleic Acids Res.* 9, 5469–5482.
10. Blake, R. D. & Fresco, J. R. (1973) *Biopolymers* 12, 775–789.
11. Lerman, L. S. & Frisch, H. L. (1982) *Biopolymers* 21, 995–997.