:PAGEversion9R;

K2 Compuscript [compuscript] - fot223 /data/journal/bioc bioc.b bioc2525.k root Thu Feb 11 12:03:25 1999 full terminal narrow nolist laser paper newhyph port cols:1 /data/K2/bin/seta ul:3000 priority:10

ELSEVIER

Biophysical Chemistry 76 (1999) 95–115

Biophysical Chemistry

Characterization of hairpin-duplex interconversion of DNA using polyacrylamide gel electrophoresis

Michael Shubsda, Jerry Goodisman*, James C. Dabrowiak

Department of Chemistry, Syracuse University, 1-014 Center for Science and Technology, Syracuse, New York 13244-4100, USA

Received 29 September 1998; received in revised form 30 October 1998; accepted 30 October 1998

Abstract

We show how polyacrylamide gel electrophoresis of radiolabeled DNA can be used to measure the hairpin-duplex equilibrium constant for DNA in solution. As an aid to the interpretation of the experiments, the differential equations associated with diffusion, migration and chemical reaction of the DNA forms are solved and intensity patterns generated. Two kinds of experiments were performed on several DNA 12-mers: in the first, electrophoresis time was constant while DNA concentration varied; in the other, concentration was constant while time varied (a 'load-and-run' gel). The observed patterns depended on the gel temperature and not the temperature at which the DNA was equilibrated before loading in the well, because reequilibration occurs before the DNA leaves the well to enter the gel proper. During this time, mixing also occurs, changing the concentration and ionic strength of the sample. A method of calculating the true DNA concentration, including the unmeasured concentration added with the radiolabel, is given. When the intensity pattern consists mainly of monomer and dimer peaks, the equilibrium constant K is easily calculated from peak intensities. However, when there is significant intensity between the peaks (which the calculations show results from monomer–dimer interconversion in the gel), K will be inaccurate. An accurate value of K may be determined from a load-and-run gel by extrapolating back to time 0. When the intensity pattern consists of a single broad peak (from rapid monomer-dimer interconversion in the gel), K cannot be calculated without additional information. The rate of interconversion increases with temperature. Estimated rates in the gel are more than an order of magnitude smaller than in bulk solution at the same temperature. Derived values of K for several DNAs are compared with literature values.

Keywords: DNA; Dimerization; Hairpin; Duplex; Electrophoresis

* Corresponding author.

0301-4622/99/$ - see front matter
PII: S0301-4622(98)00217-8

1. Introduction

It is well known that supercoiled DNA under high torsional stress can form extrusions or cruciforms which serve as sites of drug or protein binding in the cell [1–4]. The models typically used to study the mechanism of cruciform formation are small DNA hairpins which can dimerize to form duplexes. Like DNA, RNA can also be involved in a monomer–dimer equilibrium. For example, the genomes of all retroviruses, including the human immunodeficiency virus type 1 (HIV-1), contain two copies of genomic RNA which are linked through Watson–Crick base pairing near their 5′-ends [5,6]. Since the dimer is important to a number of viral functions, many studies have focused on finding the sequences which initiate the dimerization and identifying the mechanism by which the monomer is converted into a dimer [7–17].

Many techniques, including equilibrium centrifugation [18], UV-melting and non-denaturing polyacrylamide gel electrophoresis, PAGE [19–25], NMR [23,26,27], Raman and circular dichroism [24], have been used to study monomer/dimer interconversions involving DNA and RNA. Despite the fact that polyacrylamide gel electrophoresis (PAGE) is a convenient way to separate and quantify nucleic acid forms, the technique is not often used for measuring thermodynamic and kinetic parameters associated with hairpin-duplex interconversions. Since the time scale for electrophoresis is relatively slow (hours), and the amounts of the species can change in the gel as the separation takes place, calculation of equilibrium constants from spot intensities in the gel is not straightforward.

In this report we show how the autoradiographic spot intensities from native polyacrylamide gels can be used to measure the hairpin-duplex equilibrium constant for DNA in solution. In order to understand the effects of monomer–dimer interconversion, we solve the differential equations associated with diffusion, migration and chemical reaction of the DNA forms. By comparing the calculated with the observed intensity patterns, the latter can be interpreted to yield equilibrium and kinetic data for the monomer/dimer interconversion. The differential equations associated with migration of an interconverting dimer–monomer system, and a dimer–monomer–tetramer system, have previously been solved numerically by Yang et al. [28] with the same purpose. Kleinschmidt et al. [29] performed computer simulations to obtain gel patterns for binding of a protein to two targets on a DNA fragment and used them to interpret gel mobility shift experiments. Other work along these lines has been reviewed by Cann [30].

2. Theory

2.1. Equilibrium constants from gel data

The goal of this study was to determine if the band intensities observed in native polyacrylamide gels for monomer (hairpin) and dimer (duplex) DNA could be used to determine the monomer–dimer equilibrium constant for the system in solution. Because electrophoresis causes the monomer and dimer to separate and interconversion in the gel may take place during migration, spot intensities for the monomer and dimer in the gel may not accurately reflect the amounts of the two forms present in solution before electrophoresis took place. Since the M/D equilibrium constants for the 12-mers, WC, 6,7 GA and 5,8 AG, were earlier measured by Ross et al. [18] using equilibrium centrifugation, these oligomers were ideal candidates for the gel studies.

Let the concentrations of monomeric (hairpin) and dimeric (duplex) DNA be $[M]$ and $[D]$, respectively. The equilibrium constant for the dimerization reaction

$$M + M \rightleftharpoons D \tag{1}$$

is written:

$$K = \frac{[D]}{[M]^2} \tag{2}$$

To obtain the concentrations of monomeric and dimeric DNA from the intensities of the corresponding spots on an autoradiogram of a non-de-

naturing PAGE gel, one first notes that the intensities come from radiolabeled DNA, which constitutes a small fraction, say f, of the total DNA. The ratio of the concentration of radiolabeled monomer to the total concentration of monomer is f, and the ratio of the concentration of single-radiolabeled dimer to the total concentration of dimer is $2f$. The ratio of doubly-labeled dimer to the total is f^2, which may be neglected since f is small: all the labeled dimers may be considered to have a single label.

The measured intensity is proportional to the concentration of radiolabeled DNA, so the intensity of the monomer spot is $\theta_M = cf[M]$ and the intensity of the dimer spot is $\theta_D = 2cf[D]$, where c is the constant of proportionality. Therefore

$$K = \frac{1}{2}\frac{\theta_D/cf}{(\theta_M/cf)^2} \tag{3}$$

Since the value of c is not known, we multiply both sides of the equation by C, the total concentration of oligonucleotides, expressed as monomers; C is equal to θ_T/cf, where θ_T is the total intensity. Now

$$KC = \frac{1}{2}\frac{\theta_D\theta_T}{(\theta_M)^2} \tag{4}$$

which does not depend on c or f, and involves measurable spot intensities only. (These spot intensities are areas under peaks in a plot of intensity vs. position in the electrophoresis direction.) Another reason to use Eq. (4) is that the right hand side is independent of the amount of DNA, so it automatically 'normalizes', correcting for any variation in the amounts of oligomer loaded into the lanes of the gel.

2.2. Differential equations

As pointed out by Cann [30], the interpretation of an electrophoresis gel for a reacting system, such as this one, is made difficult by the fact that reactions may take place as the material migrates down the gel. Depending on the rate constants and the migration velocities, peaks in addition to those from reactants and products may appear, or the peaks for reactants and products may coalesce to form a single peak. To understand the patterns seen in our gels and to help in their interpretation, we have solved numerically the differential equations governing this system to produce simulated electrophoresis results: intensity as a function of electrophoresis distance. These plots were analyzed as if they were experimental results.

The differential equations take into account diffusion, electrophoresis, and chemical reaction. The local concentrations of the two species, monomer and dimer, are represented by $\rho_M(x,t)$ and $\rho_D(x,t)$, respectively, where x is the electrophoresis direction and t is the time. The chemical reactions are (1) dimerization (duplex formation), assumed second-order:

$$M + M \underset{k_f}{\rightarrow} D$$

and (2) dissociation of dimer to monomers, assumed first-order:

$$D \underset{k_r}{\rightarrow} M + M$$

The dimerization equilibrium constant in the gel is equal to k_f/k_r, but this may differ from K, the equilibrium constant in bulk solution. At each position (value of x), the chemical reactions change ρ_D and ρ_M according to

$$\left(\frac{\partial\rho_D(x,t)}{\partial t}\right)_x = k_f\rho_M(x,t)^2 - k_r\rho_D(x,t) \tag{5}$$

and $\partial\rho_M/\partial t = -2\partial\rho_D/\partial t$.

The differential equations give $\partial\rho_D/\partial t$ and $\partial\rho_M/\partial t$ as a sum of three contributions (diffusion, migration, and reaction):

$$\left(\frac{\partial\rho_D(x,t)}{\partial t}\right)_x = D_D\left(\frac{\partial^2\rho_D(x,t)}{\partial x^2}\right)_t - \nu_D\left(\frac{\partial\rho_D(x,t)}{\partial x}\right)_t + k_f\rho_M(x,t)^2 - k_r\rho_D(x,t) \tag{6}$$

and

$$\left(\frac{\partial \rho_M(x,t)}{\partial t}\right)_x = D_M\left(\frac{\partial^2 \rho_M(x,t)}{\partial x^2}\right)_t - \nu_M\left(\frac{\partial \rho_M(x,t)}{\partial x}\right)_t - 2k_f \rho_M(x,t)^2 + 2k_r \rho_D(x,t) \quad (7)$$

Here, D_D and D_M are the diffusion constants for dimer and monomer; ν_D and ν_M are the migration velocities for dimer and monomer (in the absence of diffusion and chemical reaction, dimer would move at velocity ν_D and monomer at velocity ν_M). The migration velocities are proportional to the local electric field (assumed independent of x). The ratio D_D/D_M is expected to be equal to ν_D/ν_M. It is assumed that, at $t=0$, both ρ_M and ρ_D are Gaussians centered at some small value of x.

The equations simplify in two limits. If k_f and k_r are very small, monomer–dimer interconversion can be neglected and the density of each species is a broadening Gaussian moving to positive x at the appropriate migration velocity. If k_f and k_r are very large, the reaction comes to equilibrium instantly, so $\rho_D/\rho_M^2 = k_f/k_r$ for all x. In general, a numerical solution is required. This is obtained by the Runge–Kutta method [31].

2.3. Results of simulations

Having obtained $\rho_D(x,t)$ and $\rho_M(x,t)$ for a desired time t, we calculate the apparent total intensity or optical density, $fc[2\rho_D(x,t) + \rho_M(x,t)] = fc\,\rho_T(x,t)$. Since we are interested only in the shape of the intensity profile, and the units of intensity are arbitrary, we take fc equal to unity. In the results of the numerical solution shown in Figs. 1–4, apparent total intensity is actually the concentration of oligonucleotide strands in μM. Our results, and the plots of total intensity, are very similar to those of Yang et al. [28], except that the leading and trailing edges of their calculated intensity patterns are much sharper than ours. This is because they considered gradient electrophoresis, for which migration velocities depend on position.

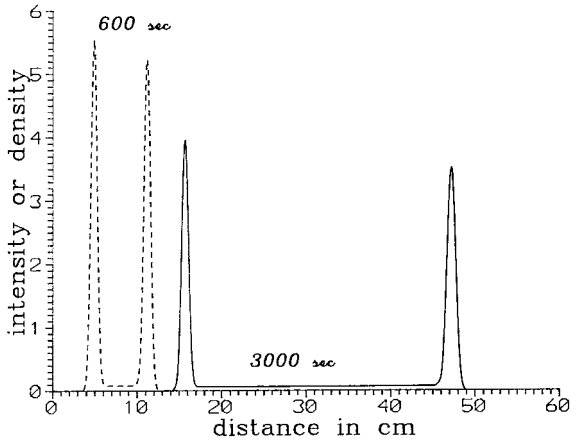

Fig. 1. Intensity vs. distance at 600 and 3000 sec, from solution of differential Eqs. (6) and (7) with $\nu_M = 0.015$ cm s^{-1}, $\nu_D = 0.0045$ cm s^{-1}, $D_M = 2\times10^{-5}$ cm^2 s^{-1}, $D_D = 6\times10^{-6}$ cm^{2}s^{-1}, $k_f = 8$ M^{-1} s^{-1}, $k_r = 9.6\times10^{-5}$ s^{-1}. Initial dimer and monomer concentrations: 6 μM and 3 μM.

In all the simulations, monomer and dimer densities at time $t=0$ were assumed to be Gaussians of equal width centered at $x = 2.1$ cm, corresponding to equilibration in the well of the gel. The relative heights then represented the equilibrium ratio of concentrations. For the first few

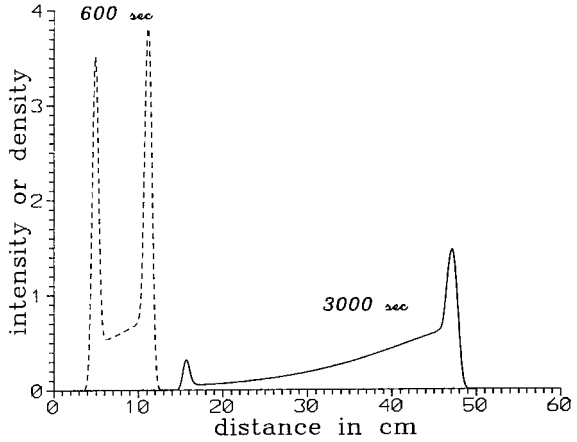

Fig. 2. Intensity vs. distance at 600 and 3000 sec, from solution of differential Eqs. (6) and (7) with $\nu_M = 0.015$ cm s^{-1}, $\nu_D = 0.0045$ cm s^{-1}, $D_M = 2\times10^{-5}$ cm^2 s^{-1}, $D_D = 6\times10^{-6}$ cm^2 s^{-1}, $k_f = 80$ M^{-1} s^{-1}, $k_r = 9.6\times10^{-4}$ s^{-1}. Initial dimer and monomer concentrations: 6 μM and 3 μM.

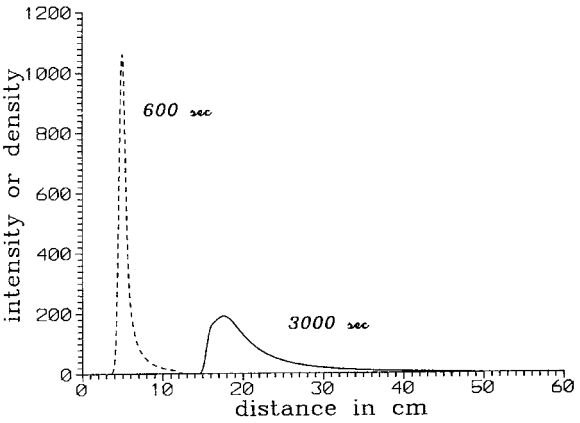

Fig. 3. Intensity vs. distance at 600 and 3000 sec, from solution of differential Eqs. (6) and (7) with $\nu_M = 0.015$ cm s^{-1}, $\nu_D = 0.0045$ cm s^{-1}, $D_M = 2 \times 10^{-5}$ cm^2 s^{-1}, $D_D = 6 \times 10^6$ cm^2 s^{-1}, $k_f = 80$ M^{-1} s^{-1}, $k_r = 9.6 \times 10^{-4}$ s^{-1}. Initial dimer and monomer concentrations: 768 μM and 96 μM.

solutions (Figs. 1–3), the migration velocities of monomer and dimer were $\nu_M = 0.015$ cm s^{-1} and $\nu_D = 0.0045$ cm s^{-1}, respectively, and the diffusion coefficients were taken to be $D_M = 2 \times 10^{-5}$ cm^2 s^{-1} and $D_D = 6 \times 10^{-6}$ cm^2 s^{-1}, respectively. The equations were integrated to a time of 3000 s, so that the monomer would migrate a distance of 45 cm and the position of the undisturbed monomer peak would be at 47.1 cm.

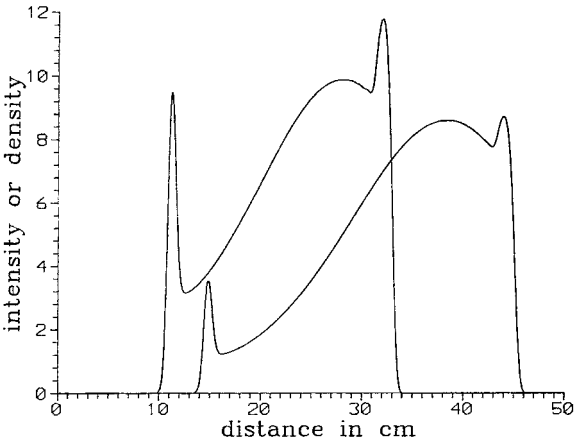

Fig. 4. Intensity vs. distance at 2000 and 2800 sec, from solution of differential Eqs. (6) and (7) with $\nu_M = 0.015$ cm s^{-1}, $\nu_D = 0.0045$ cm s^{-1}, $D_M = 2 \times 10^{-5}$ cm^2 s^{-1}, $D_D = 6 \times 10^{-6}$ cm^2 s^{-1}, $k_f = 20$ M^{-1} s^{-1}, $k_r = 0.0012$ s^{-1}. Initial dimer and monomer concentrations: 96 μM and 48 μM.

In the first simulation (Fig. 1), the rate constants for the forward and reverse reactions were taken to be $k_f = 8$ M^{-1} s^{-1} and $k_r = 9.6 \times 10^{-5}$ s^{-1}, respectively, making the equilibrium constant in the gel 8.33×10^4 M^{-1}. The maximum densities of dimer and monomer in the well (at small x) were 6 μM and 3 μM, respectively: $[D]/[M]^2$ or ρ_D/ρ_M^2 for the homogeneous solution was thus $(3 \times 10^{-6}$ M$)/(6 \times 10^{-6}$ M$)^2 = 8.33 \times 10^4$ M^{-1}. Thus the value of K, the equilibrium constant in bulk solution, is the same as that in the gel. In the second simulation, both k_f and k_r were multiplied by 10, increasing forward and reverse reaction rates but maintaining the ratio constant. The maximum densities of dimer and monomer in the well were again 6 μM and 3 μM. The third simulation used the same values for rate constants as the second, but higher initial concentrations of dimer and monomer: 768 μM and 96 μM, respectively. Note that 768×10^{-6} M $= (8.33 \times 10^4$ M$^{-1})$ $(96 \times 10^{-6}$ M$)^2$, so that there is still equilibrium for the same value of K, but the dimer concentration is eight times the monomer concentration, instead of two times as previously.

The results of each simulation were analyzed as if they were experimentally measured density profiles: the total density was fitted to a sum of Gaussians and a linear background, and the areas of the Gaussians were obtained. One peak was ascribed to monomer and one to dimer, and the total intensity was calculated as the sum of the areas of all peaks. Then KC was computed according to Eq. (4). The equilibrium constant K was obtained by dividing KC by the known total concentration of DNA strands.

The first simulation is a very favorable case: the intensity profiles consist of two Gaussians, with little intensity between them, because the rates of interconversion of monomer and dimer are slow compared to their migration rates. The quantity KC can be well calculated from the data at any t. In the second simulation, Fig. 2, much more reaction takes place as seen from the rate at which monomer and dimer peaks decrease with time, and from the large intensity between the peaks.

In the simulation (although not in a real experiment) one can determine that this intensity comes

mostly from monomers formed by dissociation of the dimer. This monomer migrates at velocity ν_M but, since it forms from dimer migrating at ν_D, it lags behind the main monomer peak. The lag is smaller for the monomer formed earlier, and the intensity of earlier-formed monomer is higher because it was formed when there was more dimer. This gives the appearance of a streak extending from the monomer toward the dimer, Fig. 2. In this situation, one may obtain the initial amounts of dimer and monomer (and determine KC) by measuring peak intensities for several migration times and extrapolating back to time 0 (when the monomer and dimer peaks coincide). This is the purpose of a load-and-run gel. The intensities of Fig. 2 were fitted to three peaks (including a broad peak centered between the monomer and dimer peaks) plus a linear baseline. For evaluation of KC, the area of the broad peak was included in θ_T, and values of KC, extrapolated to $t = 0$, agreed well with the theoretical value.

The third simulation, Fig. 3, assumes much higher DNA concentration, so that one starts with mostly dimer. The numerical intensity values are much higher than for the previous simulations, but it should be realized that, in an experiment, a much smaller fraction of the DNA would be radiolabeled, and the intensities would be numerically about the same as in other experiments. Since only one peak appears, even at 600 s, there is no obvious way to obtain dimer and monomer concentrations. The calculations show that a substantial fraction of the intensity of this peak is due a to monomer in equilibrium with a dimer. The migration velocity of the peak, 0.0051 cm/s, is between ν_D (0.0045 cm/s) and ν_M (0.015 cm/s). In an experiment, it might be possible to estimate the relative amounts of dimer and monomer from the migration velocity, if the migration velocities of dimer and monomer were known.

The results of Fig. 4, obtained using initial monomer and dimer concentrations of 96 and 48 nm, respectively, show that it is possible to obtain more than two peaks from a monomer–dimer equilibrium. The large broad peak between the monomer and dimer peaks is a combination of the shoulders or plumes attached to the monomer and dimer peaks (monomer formed from dimer and dimer formed from monomer, respectively). In fact, the three-peak pattern is an intermediate stage between the initial two-peak pattern and the pattern of Fig. 3. With increasing time, the monomer and dimer peaks decrease in size and the middle peak grows; eventually only the middle peak remains. From an experimental pattern which resembles Fig. 4, it is impossible to reliably obtain a dimerization constant without an assumption about how the area of the middle peak is to be partitioned between monomer and dimer. A load-and-run experiment is recommended.

3. Results and analysis

3.1. Equilibration in solution and in the gel

In a series of experiments, solutions containing various concentrations of the dodecamers were allowed to equilibrate for different times (up to 18 h) and temperatures (0–90°C), before being loaded into an electrophoretic gel. A photograph of an autoradiogram of a 5°C electrophoresis gel using 6,7 GA equilibrated at various temperatures is given as Fig. 5. The intensity patterns obtained from this gel are shown in Fig. 6. If the same samples are electrophoresed in a gel at approximately 35°C, the intensity patterns are different from those of Fig. 6, but again independent of the equilibration temperature.

In every case, it was found that the intensity patterns after electrophoresis depended on the gel running temperature, but not on the incubation times and temperatures outside the gel. This indicates that monomer–dimer equilibrium is established at the running temperature of the gel, before separation of monomer and dimer. This must occur in the loading well of the gel, before the DNA forms enter the gel matrix, since monomerization and dimerization rates in the gel are slow (see below).

Apparently, 10 or 15 min are required for the DNA to move out of the gel well and enter the gel. This is seen in Fig. 7, an autoradiogram of a 'load-and-run' gel for WC DNA, produced by loading 5 μl of sample (WC DNA concentration

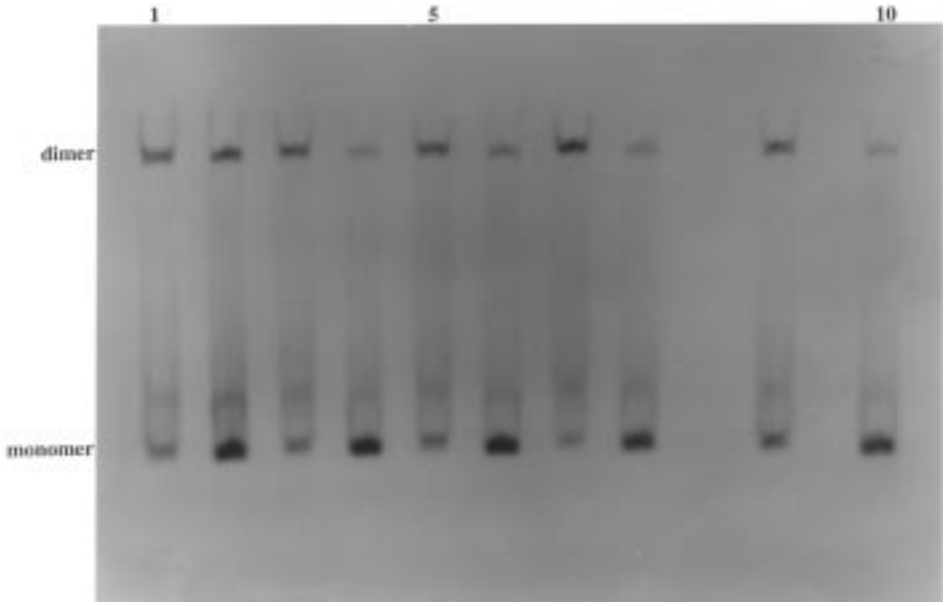

Fig. 5. Autoradiogram of electrophoresis gel for 6,7 GA, pre-equilibrated at temperatures of 90°C (lanes 1–2), 56°C (lanes 3–4), 37°C (lanes 5–6), 25°C (lanes 7–8), and 5°C (lanes 9–10). The odd numbered lanes have ionic strength of 190 mM and even numbered lanes have $I = 97$ mM. Since the direction of migration is from top to bottom, the lower spot in each lane is monomer, and the higher is dimer.

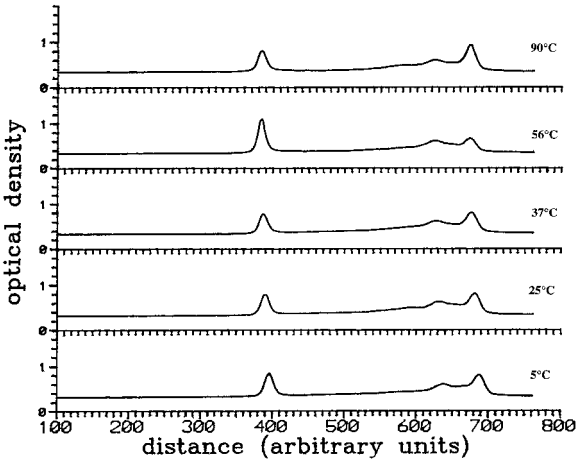

Fig. 6. Optical density as a function of electrophoresis distance for 6,7 GA equilibrated at various temperatures before loading into 5°C gel. Ionic strength = 190 mM. From top to bottom, equilibration temperatures are: 90°C, 56°C, 37°C, 25°C, 5°C.

0.75 μM, NaCl concentration 100 or 500 mM) into different lanes at different times while electrophoresis continues at 5°C. Intensity plots from this gel are given in Fig. 8. At $t = 0.25$ h there is only one, somewhat broadened peak, near the position of the well.

In the gel proper at 5°C, equilibration occurs much more slowly than in the well. This is seen clearly from load-and-run gels and from constant-time gels, such as those giving rise to Figs. 5–9. One sees distinct bands for the monomer and dimer with only a minor broad component between the peaks. As is evident from the simulated patterns in Figs. 2 and 4, the broad band between the peaks is due to some of the dimer decomposing to monomer, and some of the monomer forming dimer, while migration in the gel takes place. If equilibration occurred rapidly in the gel, there would be no distinct bands for monomer and dimer.

After 5 h of electrophoresis, the areas of the monomer and dimer peaks in Fig. 8 are 20.89 and 5.63, respectively while the areas of the shoulders, representing monomer formed from dimer and dimer formed from monomer, are each 2.02. Thus only a small fraction of dimer has dissociated to

Fig. 7. Autoradiogram of load-and-run gel for WC DNA at 0.287 μM, ionic strength 190 mM in odd numbered lanes and 100 mM in even numbered lanes. In each lane, lower spot is monomer, upper spot is dimer. Electrophoresis times in h: 5 (lanes 1–2), 4.5 (3–4), 3.5 (5–6), 2.5 (7–8), 1.5 (9–10), 0.75 (11–12), 0.25 (13–14).

monomer, and only a small fraction of monomer has associated to dimer, because equilibration is much slower than in solution.

Electrophoresis of WC DNA at approximately 35°C of monomer–dimer mixtures, previously equilibrated at temperatures above or below room temperature, exhibit only a single broad band, at a position roughly midway between the positions of the WC monomer and dimer spots in gels run at 5°C (data not shown). It is clear that this spot is an equilibrating mixture of monomer and dimer, with a migration velocity between that of monomer and that of dimer, as was seen in the simulated separations shown in Fig. 3. As expected, at a gel temperature of approximately 35°C the rate constants k_f and k_r are considerably higher than at 5°C, making it impossible to measure amounts of monomer and dimer. A single band was also observed for 6,7 GA and 5,8 AG at a running temperature of 35°C (data not shown). However, as discussed below, the band for 5,8 AG appears to be monomer, rather than an interconverting mixture as for the other DNAs.

From these experiments it follows that: (1) the kinetics of interconversion in *solution* are rapid even at the lowest temperature (5°C), so that, in the time it takes for the forms to enter the gel from the loading well, the system comes to equilibrium at the temperature in the well of the gel; and (2) the kinetics of interconversion in the gel at 5°C, but not above room temperature, are slow enough so that discrete spots due to dimer and monomer can be seen and measured. Since 5,8 AG yielded a single band under all conditions tried, no information on the kinetics of interconversion of the monomeric and dimeric forms of this 12-mer was available.

3.2. Experimental correction for dilution in the well of the gel

Since the 10 minutes the sample remains in the

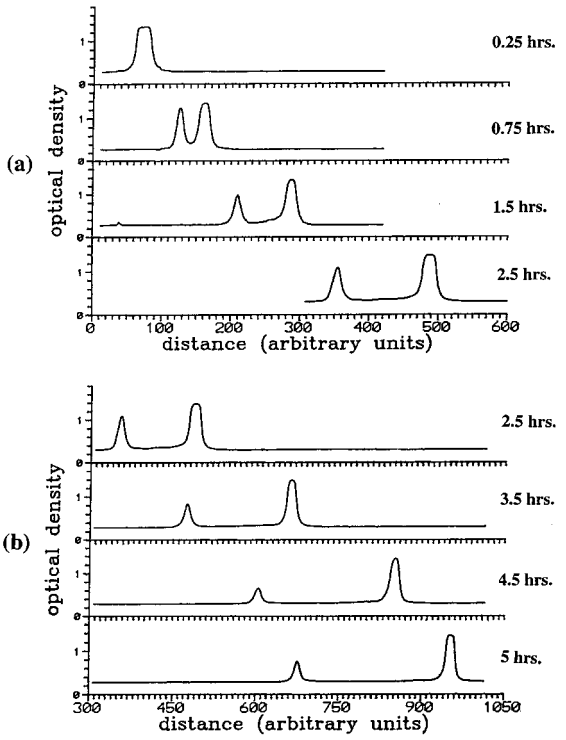

Fig. 8. Optical density as a function of electrophoresis distance for load-and-run gel for WC DNA; (a) running times, in h, are (top to bottom): 0.25, 0.75, 1.5, 2.5. (b) running times, in h, are (top to bottom): 2.5, 3.5, 4.5, 5.

gel well is enough time for monomer–dimer equilibration, dilution of the sample also occurs during this time. A volume V_i (5–8 μl in our experiments) of solution containing monomer and dimer is loaded into the well, of total volume 21.6 μl, which contains some buffer solution. It is observed that a tracking dye fills its entire well quickly, and that the boundary between the sample and the buffer disappears within a few minutes. This implies that the volume of sample increases from V_i to some V_{eff}, where V_{eff} is close to 21.6 μl, before passing into the gel. This reduces the DNA concentration by a factor of V_i/V_{eff} and changes the ionic strength, on which K depends, if the running buffer's ionic strength is different from that of the sample. The value of V_{eff} cannot exceed 21.6 μl; in the absence of further information, we will assume V_{eff} equals this volume.

It is also necessary to ascertain the true DNA concentration present in the sample before loading into the gel well. This is a problem because the amount of DNA introduced into solution from the sample containing *labeled* DNA is not known. Very few of the molecules in this sample actually carry a radiolabel, but all of them contribute to the total DNA concentration. The concentration of DNA recovered in the labeling procedure is typically too low to measure optically [33].

3.3. Determining the true DNA concentration

By the concentration of 'labeled DNA' we mean the concentration of the DNA from the radiolabeled sample which is added to the unlabeled DNA. For WC DNA, this concentration was obtained by analysis of the 5°C gel experiment whose autoradiogram is shown in Fig. 9. The samples, as loaded into the gel, had WC concentrations (uncorrected for dilution) between 0 and 1.875 μM and ionic strengths 100 or 500 μM; electrophoresis time was 5 h. The intensity profiles were of the 3-peak form of Fig. 4, except that the area of the central broad peak was always smaller than the sum of the other peak areas. Somewhat arbitrarily, the central peak area was assigned half to dimer and half to monomer.

This gave the areas in Table 1 (for lower DNA concentrations, no dimer spot was seen); KC was calculated according to Eq. (4) and divided by the DNA concentration, as in Table 1. The apparent values of K increase markedly with decreasing [WC], because the radiolabeled DNA concentration has not been taken into account. Thus, KC is really $K(U+L)$, where U is the known concentration of unlabeled DNA and L the unknown concentration of labeled DNA, so a plot of KC vs. U should be linear with K as the slope and $-L$ as the x-intercept. Fig. 10 shows such plots, made from the data of Table 1.

From the 500 mM salt data we obtain $K = 1.50 \pm 0.09$ $(\mu\text{M})^{-1}$ and $L = 0.14 \pm 0.09$ μM; from the 100 mM salt data we obtain $K = 0.48 \pm 0.04$ μM^{-1} and $L = 0.28 \pm 0.12$ μM. The absolute values of K are not significant because the dilution correction has not yet been made, but the values for L are. Averaging, we find the concen-

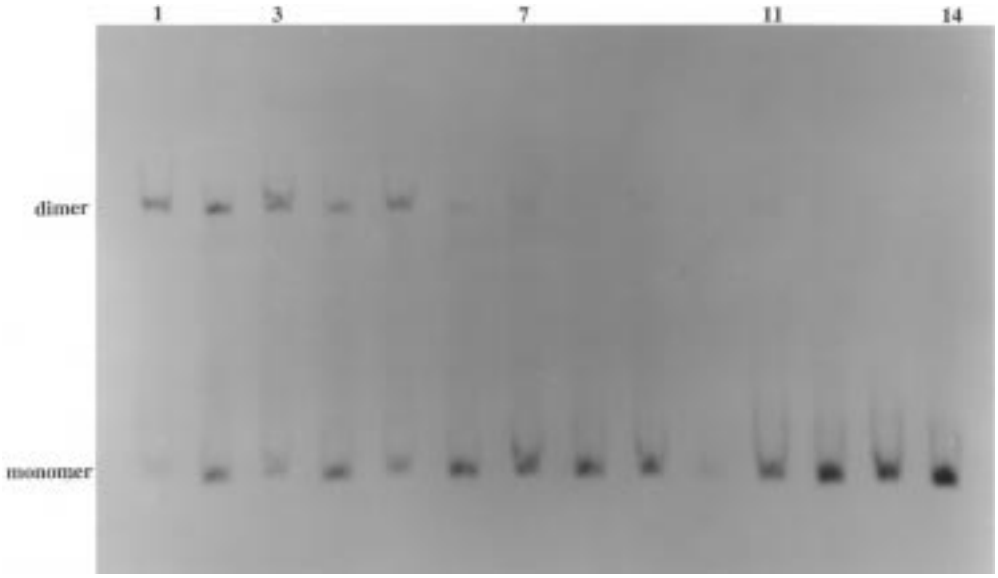

Fig. 9. Autoradiogram of WC DNA gel. The concentrations of the unlabeled DNA, before correction for dilution in the well, are 1.875 μM (lanes 1–2), 0.75 μM (lanes 3–4), 0.375 μM (lanes 5–6), 0.075 μM (lanes 7–8), 0.052 μM (lane 9–10), 0.030 μM (lanes 11–12), and 0 μM (lanes 13–14). Electrophoresis time = 5 h, gel temperature = 5°C. Alternate lanes are for ionic strengths 97 and 190 mM (right to left).

tration of radiolabeled DNA to be $L = 0.21$ μM, which must be added to the concentrations of WC used in subsequent experiments.

A similar correction must be made for other DNAs, unless the dimerization equilibrium constant is small. Then, a high concentration of unlabeled DNA would be used in the experiments, so that the amount from the labeled sample could be neglected. However, for a large equilibrium constant, the concentration of DNA in solution would be small, and it would be important to determine the amount added from the labeled sample.

3.4. WC DNA

The results of the load-and-run experiment with 0.75 μM WC DNA at 5°C, shown in Figs. 7 and 8,

Table 1
Calculation of apparent K for determination of labeled DNA concentration

[WC], μM	[NaCl], mM	Monomer area[a]	Dimer area[a]	KC	K (μM)$^{-1}$
1.875	500	4.17	8.41	3.042	1.622
1.875	100	16.71	16.91	1.018	0.543
0.750	500	17.21	19.43	1.202	1.602
0.750	100	16.08	10.67	0.552	0.736
0.375	500	11.76	11.50	0.967	2.579
0.375	100	15.81	5.90	0.256	0.683
0.075	500	17.20	6.99	0.286	3.808
0.075	100	17.85	5.02	0.180	2.402
0.052	500	18.75	6.78	0.246	4.686

[a] Integrated areas of spots, in units of optical density times pixel number.

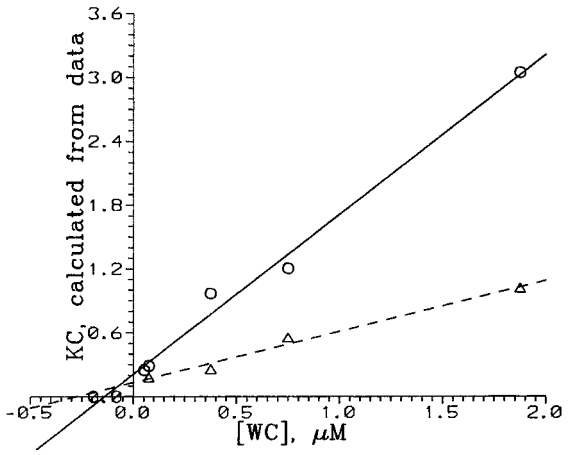

Fig. 10. Measured values of *KC* plotted vs. DNA concentration in μM and linear fits. The slope gives the true value of K and the x-intercept the negative of the concentration of DNA added with the radiolabeled DNA.

were used. For each running time (except 0.25 h, for which there is only one peak), the intensity was fit to a sum of three Gaussians and a linear background. The central Gaussian was very broad and of very low intensity, so hardly contributed to the total area. From the peak areas, *KC* was calculated according to Eq. (4). Then linear extrapolation back to time 0 gave $KC = 0.552$ for 500 mM salt. Corrected for dilution in the well and for the labeled DNA, the DNA concentration was $(6/21.6)(0.75 + 0.21)\mu\text{M} = 0.267\ \mu\text{M}$, so that $K = 2.48 \times 10^6\ \text{M}^{-1}$. A similar treatment of the data for 100 mM salt gives $KC = 0.183$ or $K = 8.2 \times 10^5\ \text{M}^{-1}$.

In another load-and-run experiment with 500 mM NaCl and original DNA concentration 1.0 μM, the value of *KC*, extrapolated to zero time, was 0.461. Since the corrected (actual) DNA concentration is $5(1.21)/21.8 = 0.278\ \mu\text{M}$, the calculated value of K is $1.66 \times 10^6\ \text{M}^{-1}$. For 100 mM NaCl and 1.0 μM WC DNA, the value of *KC*, extrapolated to zero time, was 0.183, giving $K = 6.6 \times 10^5$. Combining these results with those of the previous paragraph, we find $K = 2.07 \pm 0.41 \times 10^6\ \text{M}^{-1}$ and $7.4 \pm 0.8 \times 10^5\ \text{M}^{-1}$ for the two ionic strengths.

Because of mixing, the actual ionic strength of the sample in the well may differ from the original ionic strength in the sample as loaded. The ionic strength of the buffer solution in the gel well is 0.096 mM, so that the dilution of the sample from 5 μl to V_{eff} changes the ionic strength to

$$I = \frac{5}{V_{eff}} J + \frac{V_{eff}^{-5}}{V_{eff}} (96\ \text{mM}) \tag{8}$$

where J is the original ionic strength, 100 or 500 mM. Therefore the ionic strength in the well was actually 97 mM in the former case and 190 mM in the latter.

In order to compare our results with those of Ross et al. [18], we interpolate, assuming that ln K is linear in the square root of the ionic strength I (this is important only for the 500 mM case). The results of Ross et al. [18] for this DNA at 5°C give $\ln K = 11.913 + 4.74\sqrt{I}$, so that $K = 6.53 \times 10^5\ \text{M}^{-1}$ for the lower ionic strength and $1.178 \times 10^6\ \text{M}^{-1}$ for the higher. Agreement with our results is good, less so for the higher ionic strength, possibly because of the assumed interpolation rule.

3.5. GA DNA

Several gels were run with 6,7 GA at 0.867 μM and ionic strength either 0.097 M or 0.190 M (concentrations after mixing in well), with solutions incubated at various temperatures. One such gel was shown in Fig. 5, with intensity patterns for 0.190 M ionic strength shown in Fig. 6. As discussed above, these demonstrate that monomer–dimer equilibrium is established during the time the sample remains in the gel well (at 5°C), and that rate constants for monomer–dimer interconversion are much smaller in the gel proper than in the well. However, the existence of significant density between the two peaks for monomer and dimer in Figs. 5 and 6 shows that these rate constants in the gel are much larger for 6,7 GA than for WC (see Figs. 7–9).

The patterns for 6,7 GA differ in another way from those for WC, which show a dimer peak, a monomer peak, and a broad peak between them due to monomer–dimer interconversion. In Fig.

6, there are two fast-migrating peaks instead of one. The extra peak could be due to a second form of monomeric DNA, migrating slightly more slowly than the first, or it could be part of the equilibrating monomer–dimer mixture. On fitting the intensity plots to a sum of Gaussian peaks plus a linear background, the peak areas given in Table 2 were obtained; peak 2 is the broad monomer–dimer peak and peak 3 is the new peak. The last two columns of Table 2 give values of *KC* calculated from the two interpretations. In each calculation, the intensity ascribed to the monomer–dimer equilibrium (peaks 2 and 3 in the one-monomer model, peak 2 in the two-monomer model) is assigned half to monomer and half to dimer.

Experiments at 5°C were also carried out on 6,7 GA with varying DNA concentrations and ionic strengths of 0.097 M and 0.190 M. The intensity patterns were similar to those of Fig. 6, except the central broad spot showed some structure, and could be better fit to two Gaussians than to one. This favors interpreting it as arising from monomer being formed from dimer and dimer being formed from monomer, and assigning the second sharp peak to a second form of monomer. Furthermore, the calculations show that one cannot get four peaks from a simple monomer–dimer system. Using a single Gaussian to fit the central peak, we obtain the measured spot intensities given in Table 3. The values of *KC* calculated according to the two-monomer model, with monomer intensity equal to the sum of peaks 3 and 4, are also given. We plot *KC* vs. *U* and obtain *K* as the slope. The result is $K = 0.241 \pm 0.026 \times 10^6\ M^{-1}$ for high salt and $0.142 \pm 0.006 \times 10^6\ M^{-1}$ for low salt (compare the averaged results of Table 2 $0.256 \times 10^6\ M^{-1}$ and $0.110 \times 10^6\ M^{-1}$); errors are uncertainties from the linear fits. The *x*-intercept is $-L$, $L = 0.36 \pm 0.32\ \mu M$ and $0.29 \pm 0.13\ \mu M$ for high and low salt.

The average, 0.325 μM, is added to the concentration used in a load-and-run experiment, 0.231 μM (after mixing in the gel well). Experiments were carried out at two ionic strengths, with running times 0.25, 0.75, 1.5, 2.5, 3.5, 4.5, and 5 h. The intensity patterns for the 6,7 GA load-and-run gel were similar to those shown for WC DNA in Figs. 7 and 8. For the shortest running time, the monomer and dimer peaks could not be resolved; for all other running times, the intensity pattern consisted of three peaks with some intensity between them. Each such pattern was fitted to four Gaussian peaks, with areas given in Table 4. The quantity *KC* was calculated according to Eq. (4) for each running time, the central peak contributing to the total intensity only. Then the six results were linearly extrapolated to time 0 to

Table 2
Peak areas[a] for 6,7 GA, and *KC* values calculated from two models

Incubation temperature	Ionic strength	Peak 1 (dimer)	Peak 2 (central)	Peak 3 (sharp)	Peak 4 (monomer)	*KC* (M^{-1}) model 1[b]	*KC* (M^{-1}) model 2[c]
90°C	0.40	7.88	20.96	10.73	8.16	1.111	0.508
56°C	0.40	7.02	20.55	12.21	6.82	1.284	0.469
37°C	0.40	6.38	25.31	13.56	6.67	0.322	0.457
25°C	0.40	9.81	22.04	13.14	4.57	1.616	0.625
5°C	0.40	7.23	13.24	9.46	8.50	1.017	0.440
90°C	0.10	11.28	45.69	23.67	23.11	0.896	0.365
56°C	0.10	3.17	23.92	15.12	18.84	0.660	0.219
37°C	0.10	4.10	18.65	13.90	17.68	0.656	0.218
25°C	0.10	3.03	5.06	11.14	14.95	0.527	0.116
5°C	0.10	3.40	7.68	9.99	15.09	0.520	0.157

[a]Areas for peaks 1–4 are in units of optical density times number of pixels.
[b]Model 1 assumes only one form of monomer, with peak 3 being part of the broad peak arising from the monomer–dimer mixture.
[c]Model 2 assumes the monomer–dimer mixture is peak 2 only, and that peaks 3 and 4 arise from two forms of monomer.

Table 3
Peak areas[a] for 6,7 GA run at various concentrations and calculated values for *KC* and *K*

[6,7 GA][b]	*I*[c]	Peak 1	Peak 2	Peak 3	Peak 4	*KC*[d]	$K/10^6(M^{-1})$
5.76	0.40	8.21	5.00	3.05	2.33	1.603	0.278
5.76	0.10	5.48	27.38	5.02	2.40	0.866	0.150
2.88	0.40	6.08	24.36	12.40	0.76	0.620	0.215
2.88	0.10	5.18	18.42	8.18	5.22	0.521	0.181
1.44	0.40	4.22	28.90	11.22	0.89	0.598	0.415
1.44	0.10	3.05	16.58	15.12	6.22	0.264	0.183
0.72	0.40	2.10	18.37	13.20	1.26	0.353	0.490
0.72	0.10	1.34	7.08	8.97	7.21	0.154	0.214
0.28	0.40	1.60	5.81	7.39	3.24	0.222	0.771
0.28	0.10	0.62	5.56	5.63	7.28	0.132	0.458
0.043	0.40	0.70	4.60	5.53	5.94	0.133	3.093
0.043	0.10	0.21	3.32	6.72	11.04	0.053	1.233
0[b]	0.40	0.00	0.00	9.62	16.92	–	–
0[b]	0.10	0.00	0.00	7.56	13.84	–	–

[a]Areas of peaks 1–4 given in units of optical density times number of pixels.
[b]Concentration of unlabeled DNA in μM, corrected for dilution in the well but not for addition of labeled DNA.
[c]Ionic strength in M.
[d]Calculated according to model 1, which assigns peaks 3 and 4 to two forms of monomer; monomer intensity is the sum of the areas of these peaks.

obtain the initial value of *KC*, and hence *K*. For ionic strengths of 0.190 M and 0.097 M, *K* was $2.91 \pm 0.86 \times 10^6$ M^{-1} and $6.21 \pm 2.11 \times 10^5$ M^{-1} (stated errors are uncertainties in the calculated slopes). The values of *KC* for the two highest running times were subject to large errors because the areas of the dimer peaks were very low. Interpolating in the values given by Ross et al. [18], assuming ln *K* is linear in the square root of ionic strength, yields values of 8.51×10^5 M^{-1} and 4.83×10^5 M^{-1} for ionic strengths of 0.190 M and 0.097 M. The second, but not the first, value agree with ours within our stated errors. Results from a second load-and-run gel on 6,7 GA are shown in Table 5. The total DNA concentration (including DNA added with the radiolabel) was 0.385 μM, and ionic strengths were 0.190 M and 0.097 M. Analysis similar to that for the first gel led to $K = (2.65 \pm 0.56) \times 10^6$ and $(6.65 \pm 2.85) \times 10^5$ M^{-1}, respectively, in good agreement with our previous values. As for the WC results, it may be that the interpolation, assuming ln *K* is linear in ionic strength, is not valid.

Table 4
Load and run results for 6,7GA at 5°C

Run time, h	Peak intensities for $I = 0.19$			Peak intensities for $I = 0.10$		
	Monomer[a]	Central	Dimer	Monomer[a]	Central	Dimer
0.75	9.10	6.30	6.54	18.75	3.25	6.04
1.50	7.04	9.51	6.40	14.61	9.05	3.31
2.50	5.91	12.66	4.94	9.84	12.32	3.13
3.50	7.09	11.98	3.10	11.91	9.09	1.00
4.50	8.32	10.55	1.48	15.26	10.42	0.77
5.00	8.80	14.70	1.75	17.06	10.75	0.80

[a]Sum of areas of two monomer peaks.

Table 5
Load and run results for 6,7 GA at 0.49 μM, 5°C

I[a]	Run time, h	Dimer[b]	Broad peak[b]	Sharp peak[b]	Monomer[b]	KC[c]
0.40	0.75	15.21	3.66	3.66	11.06	1.179
0.40	1.50	7.12	4.79	6.88	6.50	0.503
0.40	2.40	4.55	5.34	11.05	6.70	0.200
0.40	3.50	2.21	16.51	9.81	8.21	0.125
0.40	4.50	1.06	9.00	15.18	8.42	0.032
0.40	5.00	0.58	5.14	19.78	7.71	0.013
0.10	0.75	6.91	4.40	8.09	16.25	0.208
0.10	1.50	4.84	8.75	3.60	14.63	0.232
0.10	2.40	3.46	7.14	12.42	17.18	0.079
0.10	3.50	1.86	15.34	6.41	9.46	0.122
0.10	4.50	0.70	5.31	17.23	17.01	0.012
0.10	5.00	1.00	11.90	7.87	16.64	0.031

[a] Ionic strength, M.
[b] Peak areas in units of optical density times pixel number.
[c] Calculating using sum of monomer and sharp peaks for monomer intensity.

Measurements were also made on 6,7 GA at a DNA concentration of 4.32 μM. It was expected from previous results that the concentration of monomer would be too small to measure accurately; only a single broad dimer spot was observed on the gel. It proved to consist of three closely spaced peaks, which we fitted as Gaussians. The average fractional intensities were: 0.192 ± 0.040, 0.628 ± 0.042, and 0.180 ± 0.013 for ionic strength 0.190 M; and 0.338 ± 0.025, 0.543 ± 0.032, and 0.119 ± 0.027 for ionic strength 0.097 M. If two forms of the monomer, say a and b, can combine unchanged to form a dimer, there would be three forms of the dimer: aa, ab, and bb. Their relative amounts can be calculated in terms of the relative amounts of a and b and compared with measured values (see Discussion).

3.6. AG DNA

For the 5,8 AG DNA, both a variable-concentration and a load-and-run experiment were performed at 5°C. The concentrations used in the former (aside from the radiolabeled DNA, of unknown concentration) ranged from 0 to 2.66 μM, and the ionic strengths were 0.097 and 0. 190 M. In every lane, only a single spot appeared on the gel, always at the same distance from the loading well. This spot could represent monomer, dimer, or an interconverting mixture of the two. Since the spot appeared even at nominal DNA concentration zero (only labeled DNA present), it is highly unlikely that it represents dimer. The fact that the position of the spot is independent of DNA concentration tends to rule out the third possibility, since the simulations show that the migration velocity of an interconverting monomer–dimer mixture depends on the total concentration. The fact that the distance of the spot from the well is close to that for the monomer spot for other DNAs makes it highly likely that the spot is due to monomer.

A load-and-run experiment, with nominal DNA concentration of 0.22 μM, was carried out to confirm this interpretation. At every electrophoresis time from 0.25 to 5 h, a single spot was seen. The distance from the spot to the loading well was proportional to the electrophoresis time, showing constant migration velocity. The shape of the spot did not change; there was no evidence of separate peaks disappearing to leave a growing broad spot between them, as in the simulations. The conclusion is that dimeric 5,8 AG DNA is never seen in our experiments, but only monomer. This means that the equilibrium constant for dimerization of this DNA is too small for a dimer to be formed at the concentrations used.

4. Discussion and conclusions

In this study, we have shown how gel electrophoresis can be used to determine hairpin-to-duplex dimerization constants for small pieces of DNA, and pointed out the problems associated with such measurements. Two kinds of experiments were conducted on three hairpin DNAs at two different ionic strengths. In one kind, DNA concentration was varied with a fixed electrophoresis time, and, in the other, the concentration was fixed and the electrophoresis time varied (the 'load-and-run' experiment). From numerical solutions of the differential equations governing simultaneous chemical reaction, diffusion, and migration in a gel, we generated plots of DNA concentrations as a function of distance and time, to be compared with experimental intensity plots. We showed that, depending on the relative rates of these processes, the intensity profile could consist of two peaks, two peaks with shoulders, a single peak, or three peaks (the center peak being broad). Such intensity profiles are exactly what we observe. Unlike the experiments, the calculations show which species contribute to the different peaks in these profiles, and we use this information to interpret the peaks observed experimentally.

In the most favorable case, one sees two peaks in the intensity pattern, corresponding to monomer and dimer. It is easy to measure their areas by fitting them to Gaussians, as we have done. Commonly, one or both of these peaks is accompanied by a shoulder. The calculations show that shoulders arise from monomer formed by dissociation of dimer and dimer formed by association of monomers. If large enough, the two shoulders combine to give a broad peak between the monomer and dimer peaks.

If the area of the shoulders is not much smaller than the areas of the monomer and dimer peaks, the latter will not yield a good estimate of the monomer–dimer ratio. The remedy is to carry out a load-and-run gel, and extrapolate monomer and dimer peak intensities, which should both decrease with time, back to time 0. From the relative intensities at time 0, one can calculate the equilibrium constant in the gel well, which corresponds to the bulk solution (this equilibrium constant may not be equal to the equilibrium constant in the gel). Gels run for WC DNA at 5°C, which show two spots with shoulders, were analyzed in this way.

The solutions to the differential equations show that the three-peak pattern becomes a single peak at longer electrophoresis times, as the monomer and dimer peaks disappear and the central peak builds up. The single-peak pattern by itself cannot give the monomer-to-dimer ratio. It is necessary to find an electrophoresis time short enough so that monomer and dimer peaks are still present, but long enough so that they are separated. This may not be possible if k_r and k_f are large in the gel, as seems to be the case at elevated temperatures.

A rough estimate of the rate constants in a gel at 5°C can be obtained from peak areas. For example, the 5-h results of Fig. 8 (for Wc GA) give the following areas: monomer, 5.63; dimer, 20.89; dimer from monomer, 2.02; monomer from dimer, 2.02. The original concentrations of monomer and dimer are proportional to $5.63 + 2.02 = 7.65$ and $\frac{1}{2}(20.89 \pm 2.02) = 11.45$, respectively. If the dimer dissociates by a first-order process with rate constant k_r,

$$\ln(5.63/7.65) = -k_r(1.8 \times 10^4 \mathrm{s})$$

and $k_r = 1.7 \times 10^{-5}\ \mathrm{s}^{-1}$. If the dimerization is a second-order process, and $[D]_o$ is the original dimer concentration,

$$(20.89[D]_o/22.91)^{-1} - [D]_o^{-1} = k_f(1.8 \times 10^4 \mathrm{s})$$

so $k_f\ [D]_o = 5.4 \times 10^{-6}\ \mathrm{s}^{-1}$. Previous calculations gave $K = 2.5 \times 10^5\ \mathrm{M}^{-1}$ and total DNA concentration in the gel well = 0.32 μM. Let $[M]_o$ be the original monomer concentration so $[D]_o$ $\frac{1}{2}(0.32\ \mu\mathrm{M} - [M]_o)$. The equilibrium expression

$$2.5 \times 10^5\ \mathrm{M}^{-1} = \frac{1}{2}\,\frac{0.32\mu\mathrm{M} - [M]_0}{[M]_0^2}$$

yields $[M]_o = 0.28\ \mu M$, $[D]_o = 0.02\ \mu M$ and $k_f = 270\ M^{-1}\ s^{-1}$. For 6,7 GA, k_f and k_r in the gel are larger than for WC, as evidenced by the significant intensity between the monomer and dimer peaks for GA.

The rate constants k_f and k_r in bulk solution are more than an order of magnitude smaller than in the gel. This is shown from a series of experiments in which DNA samples were held at various temperatures before being loaded into a gel for electrophoresis. The resulting intensity pattern did not depend on the equilibration temperature. Thus for 6,7 GA at ionic strength 0.50 M, linear fits of KC (calculated from the spot intensities) vs. Celsius temperature gave slopes of 0.0003 ± 0.0085 or 0.0001 ± 0.0013, assuming one or two forms of monomer were present (see below). Since monomer–dimer equilibrium is established during the 10 min the DNA remains in the gel well, the electrophoresis experiment can measure the equilibrium constant only at the gel temperature, unless the interconversion rates in bulk solution are much slower than for the DNAs considered here.

The fact that equilibrium is established in the well in 10 min at 5°C also places approximate lower limits on the rate constants in bulk solution. Thus k_r, the rate constant for dissociation of the WC dimer, probably exceeds $1.6 \times 10^{-3}\ s^{-1}$ and k_f, the rate constant for dimerization, probably exceeds $1.6 \times 10^{3}\ M^{-1}\ s^{-1}$ (1 μM DNA assumed). If the rate constants in the gel proper were not much smaller, one would not see separate spots for monomer and dimer but, as shown by the solutions to the differential equations, only a single spot, corresponding to an equilibrating monomer–dimer mixture. In fact, gels run above 35°C show this one-spot pattern, as is expected since rate constants increase with temperature. From a single-peak pattern alone, one cannot find the monomer-to-dimer ratio. (It may be possible to estimate it if the migration velocity of the spot is measured, and the migration velocities of dimer and monomer are known from other experiments.)

A reason for the low rate constants in the gel may be that a monomer in a 'cell' of the gel [32] cannot form a dimer unless another monomer is present in the same cell, and a dimer dissociating in a cell will reform unless one of the monomer products moves out of the cell. The decrease may also be related to the orientation of the molecules in the gel during electrophoresis. If the dimer forms from two monomers through a 'kissing' structure involving the loops of the monomers [15,16], the orientation of the monomers in the field may reduce the possibility of the 'kissing dimer' from forming. Similarly, decomposition of a duplex dimer would appear to require reorientation of the 'kissing' or cruciform intermediate [23] in the electric field, making decomposition to the monomer more difficult in the gel than in bulk solution.

If the rate constants are small enough to give measurable spots for monomer and dimer, their areas can be interpreted to yield the equilibrium constant for dimerization in homogeneous solution (which may differ from the equilibrium constant in the gel). However, two issues must be dealt with: dilution of the DNA in the gel well and the effect of adding radiolabeled DNA on the overall concentration of DNA.

In experiments using a standard DNA sequencing gel, a volume V_i of sample is loaded into a gel well of volume V, and the remaining volume, $V - V_i$, is filled with buffer solution. There is visual evidence that mixing occurs in the well: a tracking dye fills the well volume and, when the sample contains glycerol, the meniscus between the sample and the buffer disappears within a few minutes. Since much of the sample remains in the well for 10 min or more, the concentration of the oligonucleotide is decreased. We have assumed the dilution factor is V/V_i. In addition, the mixing changes the ionic strength, on which the dimerization constant depends strongly.

The second issue arises because detection of DNA in polyacrylamide gels usually requires that the oligomer be radiolabeled. Some DNA is labeled with ^{32}P and added to DNA of known concentration. Since the fraction of DNA molecules having a radiolabel is unknown, the total concentration of DNA in solution is uncertain. It may be possible to use concentrations of unlabeled DNA far exceeding the concentration of labeled DNA. If not, one requires a series of electrophoresis experiments at fixed elec-

trophoresis time and with different concentrations of unlabeled DNA. The concentration of DNA contributed by the labeled sample, L, can be found by fitting the measured values of KC to a linear function of the known (unlabeled) DNA concentration assuming constant values for K and L.

After correcting for mixing and for the DNA added with the radiolabeled sample, one can derive the relative monomer and dimer concentrations and calculate a value of KC according to Eq. (3). The ratio of concentrations of monomer and dimer is twice the ratio of spot intensities because a dimer is twice as likely to be radiolabeled as a monomer. For WC DNA, we derived K from a single-time electrophoresis experiment, obtaining 2.07×10^6 M^{-1} for ionic strength 190 mM and 7.4×10^5 M^{-1} for ionic strength 97 mM. These compare with the values derived from those of Ross et al. [18] by interpolation, 1.18×10^6 M^{-1} and 6.5×10^5 M^{-1}.

For 6,7 GA DNA, the monomer and dimer spots were clearly apparent, but the intensity in the region of the gel between the spots was comparable to the intensity in the spots themselves. The solutions to the differential equations show that the shoulders on the monomer and dimer peaks grow with electrophoresis time, eventually producing a broad peak between the monomer and the dimer peaks. Further assumptions are required to estimate the relative amounts of monomer and dimer before electrophoresis. One could assign the intensity of the central peak half to monomer and half to dimer, or, if it is much closer to the dimer peak than to the monomer peak, assume it to represent dimer formed from monomer in the gel, and assign its intensity to monomer. The best solution is to produce a load-and-run gel from which peak areas can be extrapolated back to time 0. For 6,7 GA, this procedure produced equilibrium constants in agreement with literature values.

However, the intensity pattern for 6,7 GA consisted of four peaks; in order of electrophoresis distance, they were the dimer peak, a broad peak, a sharp peak, and the monomer peak. This intensity pattern is also seen for small RNAs (to be published). The sharp peak could be due to a second form of monomer, or to the equilibrating monomer–dimer mixture. We believe the former to be correct because of the sharpness of this peak, because this model gives less scatter in the calculated values of K, and because we found it impossible to obtain such a peak from solutions to the differential equations. It should be noted that, in a denaturing gel, 6,7 GA runs as a single peak, so the two forms differ only in secondary and tertiary structure. In analyzing our experiments, the areas of the two monomer peaks were combined to give the monomer intensity. Then KC was calculated, using Eq. (3), as $(2.91 \pm 0.86) \times 10^6$ M^{-1} and $(6.2 \pm 2.1) \times 10^5$ M^{-1} for ionic strengths of 0.19 M and 0.097 M, which compare with interpolated literature values: 2.06×10^6 M^{-1} and 4.82×10^5 M^{-1}.

The 6,7 GA results show an advantage of electrophoresis over other methods of measuring dimerization constants. For instance, using ultracentrifugation [18], it is necessary to assume that only two species (monomer and dimer) are present, and find their relative amounts by fitting the observed density profile to the model. As long as the fit is adequate, the presence of more than two species can not be inferred, whereas the species are seen directly when gel electrophoresis is used.

At a concentration of 7.18 μM, 6,7 GA shows a single peak with structure, due to the dimer. It was fit to three Gaussians, the areas of which were measured and divided by the total to give fractional intensities. Average values of these are: for ionic strength 0.19 M, 0.192 ± 0.040, 0.628 ± 0.042, 0.180 ± 0.013; for ionic strength 0.097 M, 0.338 ± 0.025, 0.543 ± 0.032, $0.119 + 0.027$. If the two different forms of the monomer (see Table 4) were maintained on dimerization, the dimer would exist in three forms and the fractional intensities would be calculable: if r is the ratio of Peak 3 to Peak 4 in Table 3, the relative populations of the three dimer forms should be: r^2, $2r$, and 1.

The value of r at ionic strength 0.19 M is 1.825 (average of ratios of Peak 3 to Peak 4 in Table 3), so the relative populations of the three dimer forms would be $(1.825)^2$, 2(1.825), and 1 (homo-, hetero-, and homo-dimers), and the fractional intensities would be 0.417, 0.457, and 0.125, which is not consistent with the measured values. For

ionic strength 0.097 M, r is 0.804, making the fractional populations of the three dimer forms 0.199, 0.494, and 0.307, again not consistent with the measured values. We conclude that the three parts of the dimer peak are not three dimers formed from two non-interconvertible forms of monomer. This is consistent with the fact that only one form of the monomer appears in a denaturing gel.

For 5,8 AG DNA, no evidence of dimer was seen in either the concentration-dependent or the load-and-run gels. The equilibrium expression, Eq. (2), may be rearranged to:

$$K = \frac{[D]}{(C - 2[D])^2}$$

where C is the total DNA concentration and $[D]$ the concentration of dimer. If $[D]$ is small compared to C, the solution is $[D] = KC^2/1(1 + 4KC)$. One can measure spot intensities as small as 0.5 units in a total intensity of 25 units, which would correspond to $[D]/C = 1/100$. Since the dimer spot is not observed, $KC/(1 + 4KC)$ must be less than 0.01, or KC less than 0.0104. The highest concentration of DNA used was 7.1 μM (including the concentration added with the labeled DNA sample). Thus the fact that no dimer was observed in our experiments implies that K is less than approximately 1.5×10^3 M^{-1}.

This is much smaller than the values cited by Ross et al. [18] for 5,8 AG DNA at 5°C (2.5×10^5 M^{-1} at 100 mM salt, 7.8×10^5 M^{-1} at 500 mM salt). These authors in fact remarked on the unexpectedly high values that they found for this oligomer. It is conceivable that the reason for our inability to observe the dimer is that the rate of monomerization of this DNA in the gel is very fast, and the rate of dimerization unusually slow. Then the dimer originally present in solution would be gone after 0.75 h (at which two peaks are observed for other DNAs), and no dimer would have formed from monomer. There is no reason for the kinetics of this DNA to be radically different from the kinetics of others, however.

In summary, this study has shown when and how dimerization constants in solution can be obtained from electrophoresis measurements on radiolabeled DNA dodecamers. The method should be applicable to any system composed of equilibrating forms which can be separated using gel electrophoresis, which requires kinetics in the gel to be slow. In addition to dimer/monomer equilibria involving DNA, dimer/monomer and complexation equilibria involving RNA have been studied (to be published). When only two peaks are observed, one can obtain a value of KC (K = dimerization constant, C = total concentration of oligonucleotide, expressed as monomers) from the integrated intensities of the monomer and dimer spots. In order to obtain K, the value of C must be calculated taking into account mixing in the gel well and the unlabeled DNA which is added with the radiolabel.

For the systems studied, the temperature at which the DNA was equilibrated before loading into the gel had no effect on the intensity patterns observed. This is because the forward and reverse rate constants for the dimerization in solution are large, and the DNA comes to equilibrium in the gel well before it can move into the gel proper. The fact that individual spots for monomer and dimer can be observed in a gel shows that the rate constants at 5°C are much smaller in the gel than in homogeneous solution. At 35°C, rate constants are larger, leading to a single spot, corresponding to an equilibrating monomer–dimer mixture, being observed.

In cases where the monomer and dimer peaks are observed, but streaking or other features are also found, the intensities for the monomer and dimer peaks cannot be used to accurately calculate the equilibrium constant. On the basis of solutions to the differential equations describing migration, diffusion, and reaction, we have interpreted shoulders and a broad peak between the monomer and dimer peaks, and discussed how their intensities may be assigned to monomer or dimer. The solution to this problem is to produce a load-and-run gel, from which peak intensities can be obtained as a function of electrophoresis time. These intensities, or the calculated values of

KC, can then be extrapolated back to zero electrophoresis time, which is the solution at equilibrium in the gel well.

5. Materials and methods

5.1. DNA oligomers

The DNA oligonueleotides, 5′-GGTACGCG-TACC-3′ (WC), 5′-GGTACGAGTACC-3′ (6,7 GA), purchased from Midland Certified Reagent Company as 'GF Grade', were further purified by polyacrylamide gel electrophoresis [33], followed by desalting using a Centricon concentrator (Amicon), molecular weight cut-off 3000 Da. The oligomer 5′-GGTAAGCGTACC-3′ (5,8 AG), purchased as 'AE Grade' (Midland), was desalted and used without further purification. Concentrations of the oligomers were determined optically at 260 nm in water using published values of the extinction coefficients (M^{-1} cm^{-1}): 106 200 (WC), 103 440 (6,7 GA), and 118 920 (5,8 AG) [34]. The concentration of DNA is stated in strands.

5.2. End-labeling

The oligonucleotide was 5′-end labeled in the presence of [γ-^{32}P]ATP (Amersham) and T4 polynucleotide kinase (Promega). The labeling solution contained: 15 μM DNA, 50 mCi of [γ-^{32}P]ATP and 10 units of polynucleotide kinase in a total volume of 20 μl in 1 × kinase reaction buffer. After incubation for 1 h, the labeled oligomer was precipitated by addition of cold ethanol and purified using 20% denaturing PAGE in 1 × TBE running buffer (90 mM tris-borate, 1 mM EDTA, pH 8). After visualization of the labeled DNA by autoradiography, the portion of the gel containing the oligomer was excised and the labeled oligomer recovered by suspending the crushed gel segment in 1 ml of 1 × TBE buffer [33]. A 5-μl sample of oligomer recovered in this manner yielded 60 000 cpm on scintillation counting.

In order to determine that the various forms observed using non-denaturing PAGE conditions were in fact due to a single-length oligomer, the gel-purified, end-labeled DNA was electrophoresed a second time using denaturing PAGE conditions. In all cases, autoradiography showed the presence of only one band indicating the presence of a single length radiolabeled oligomer.

5.3. Monomer–dimer equilibria

The reactions involving WC, 6,7-GA, and 5,8-AG were carried out in a total volume of 25 μl. They contained sufficient labeled DNA for autoradiography in 1–2 days (20 000–60 000 cpm), DNA of known concentration and 100 or 500 mM NaCl in the buffer, 0.01 M sodium cacodylate, pH 7. The solutions containing DNA were held at 100°C for 10 min and rapidly cooled in ice to remove all DNA secondary structure. The solutions were equilibrated at a specific temperature for at least 2 h after which time a 6 μl portion was removed and 2 μl of a 1:1 glycerol/water mixture added, followed by brief vortexing. A 5-μl portion of this sample was loaded into the wells of a 20% native polyacrylamide gel having a running buffer of 1 × TBE, 90 mM tris–borate, 1 mM EDTA, pH 8.0.

As outlined in the Results and Analysis section, introduction of the sample into the loading well of the gel caused dilution of the sample. The concentrations, ionic strengths and temperatures stated in the tables and captions to figures take the dilution into account and are the conditions reached by the sample in the well of the gel.

Electrophoresis was carried out at 1500 V using a standard DNA sequencing apparatus. Initially, electrophoresis experiments were conducted using gels in the temperature range, 35°–40°C. Since electrophoresis patterns at high temperatures indicated that monomer and dimer were rapidly interconverting in the gel, subsequent electrophoresis experiments were carried out using a gel temperature of 5°C and an equilibration time of at least 2 h at 5°C.

Two types of electrophoresis experiments were carried out, electrophoresis at a constant time interval and electrophoresis at different time intervals, 'load-and-run'. For the former, samples having the same amount of labeled but different amounts of unlabeled DNA were electrophoresed for the same time, 5–6 h. For the load-and-run

gels, a 50-μl sample was made with DNA concentration chosen to give reasonable amounts of both monomer and dimer. At different time intervals, 6-μl aliquots were diluted with glycerol as before and loaded into the wells of the gel. In order to minimize the effects of current disruption on band shape, sample loading for the load-and-run gels was carried out with the current on. The advantage of a load-and-run gel for a system which is slowly equilibrating in the gel matrix is that extrapolation of band intensities to $t = 0$ yields the relative amounts of monomer and dimer present at the time the forms entered the gel from the loading well. These quantities most accurately reflect the equilibrium concentrations of the two forms in the well of the gel.

Autoradiography was done at -20°C using Kodak X-Omat film. Typically, more than one exposure of a gel was required to produce images in the linear response range of the film, optical densities (O.D.) less than approximately 1.1.

5.4. Quantitation of autoradiographic data

The autoradiogram was scanned to a JPEG or TIF file with a Hewlett Packard Scanjet 5P scanner. The scanned images were analyzed with Sigma Scan software (Jandel Scientific) to yield average O.D. versus position for each lane of the autoradiogram using an 11-pixel wide line scan. The lanes themselves were approximately 33 pixels wide. The O.D. profile was fitted to a sum of Gaussians and a linear background. The best fit was characterized by $r^2 > 0.98$ in all cases, and usually $r^2 > 0.995$. When peaks exhibited saturation at their maxima (O.D. > 1.1), data points in the region of saturation were eliminated prior to curve fitting and analysis. The resulting integrated intensities were comparable to those obtained using short autoradiographic exposures of the same gel. In load-and-run gels, the number of Gaussians used for fitting the pattern for all electrophoresis times was determined by inspection of the patterns at the longest electrophoresis time, where the greatest separation between bands occurred.

Acknowledgements

We thank Professor P. N. Borer for helpful discussions concerning this work.

References

[1] N. Panayotatos, R.D. Wells, Cruciform structures in supercoiled DNA. Nature (London) 289 (1981) 466–470.
[2] D.M. Lilley, The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc. Natl. Acad. Sci. USA 77 (1980) 6468–6472.
[3] C.E. Pearson, M. Zannis-Hadjopoulos, G.B. Price, H. Zorbas, A novel type of interaction between cruciform DNA and a cruciform binding protein from HeLa cells. EMBO J. 14 (1995) 1571–1580.
[4] S.Y. Stevens, P.C. Swanson, E.W. Voss, G. Glick, Application of the gel shift assay to study the affinity and specificity of anti-DNA autoantibodies. J. Am. Chem. Soc. 115 (1993) 1585–1586.
[5] R. Berkowitz, J. Fisher, S.P. Goff, RNA Packaging. Curr. Top. Microbial. Immunol. 214 (1996) 177–218.
[6] S. Hoglund, A. Ohagen, J. Goncalves, A.T. Panganiban, D. Gabuzda, Ultrastructure of HIV-1 genomic RNA. Virology 233 (1997) 271–279.
[7] J.L. Darlix, C. Gabus, M.T. Nugeyre, F. Clavel, F. Barre-Sinoussi, Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J. Mol. Biol., 216, 689–699.
[8] S. Roy, S. Weinstein, B. Borah, et al., Mechanism of oligonucleotide loop formation in solution. Biochemistry 25 (1986) 7417–7423.
[9] D. Muriaux, P.-M. Girard, B. Bonnet-Mathonière, J. Paoletti, Ncp7 activates HIV-1 Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. J. Biol. Chem. 270 (1995) 8209–8216.
[10] M. Laughrea, L. Jetté, HIV-1 genome dimerization: formation kinetics and thermal stability of dimeric HIV-1 lai RNAs are not improved by the 1-232 and 296-790 regions flanking the kissing-loop domain. Biochemistry 35 (1996) 9366–9374.
[11] M. Laughrea, L. Jetté, Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation. Biochemistry 35 (1996) 1589–1598.
[12] J.L. Clever, M.L. Wong, T.G. Parslow, Requirements for kissing-loop mediated dimerization of human immunodeficiency virus RNA. J. Virol. 70 (1996) 5902–5908.
[13] J.L. Clever, T.G. Parslow, Mutant human immunodeficiency virus Type 1 genomes with defects in RNA dimerization or encapsidation. J. Virol. 71 (1997) 3407–3414.

[14] R. Marquet, J.-C. Paillart, E. Skripkin, C. Ehresmann, B. Ehresmann, Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. Nucleic Acids Res. 22 (1994) 145–151.

[15] E. Skripkin, J.-C. Paillart, R. Marquet, B. Ehresmann, C. Ehresmann, Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc. Natl. Acad. Sci. USA 91 (1994) 4945–4949.

[16] J.-C. Paillart, R. Marquet, E. Skripkin, B. Ehresmann, C. Ehresmann, Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus Type 1 genomic RNA. J. Biol. Chem. 269 (1994) 27486–27493.

[17] E. Skripkin, J.C. Pillart, R. Marquet, M. Blumenfeld, B. Ehresmann, C. Ehresmann, Mechanisms of inhibition of in vitro dimerization of HIV Type 1 RNA by sense and antisense oligonueleotides. J. Biol. Chem. 271 (1996) 28812–28817.

[18] P.D. Ross, F.B. Howard, M.S. Lewis, Thermodynamics of antiparallel hairpin double helix equilibria in DNA oligonucleotides from equilibrium ultracentrifugation. Biochemistry 30 (1991) 6269–6275.

[19] L.E. Xodo, G. Manzini, F. Quadrifoglio, N. Yathindra, G.A. van der Marel, J.H. van Boom, A facile duplex-hairpin interconversion through a cruciform intermediate in a linear DNA fragment. J. Mol. Biol. 205 (1989) 777–781.

[20] L.E. Xodo, G. Manzini, F. Quadrifoglio, G.A. van der Marel, J.H. van Boom, The duplex-hairpin conformational transition of d(CGCGCGATCGCGCG) and d(CGCGTACWGCG): a thermodynamic and kinetic study. J. Biomol. Struct. Dyn. 6 (1988) 139–152.

[21] L.E. Xodo, G. Manzini, F. Quadrifoglio, G.A. van der Marel, J.H. van Boom, DNA hairpin loops in solution. Correlation between primary structure, thermostability and reactivity with single-strand-specific nuclease from mung bean. Nucleic Acids Res. 19 (1991) 1505–1511.

[22] L.E. Xodo, G. Manzini, F. Quadrifoglio, G.A. van der Marel, J.H. van Boom, Oligodeoxynucleotide folding in solution. Loop Size and Stability of B-Hairpins. Biochemistry 27 (1988) 6321–6326.

[23] D.Z. Avizonis, D.R. Kearns, Kinetic and thermodynamic characterization of DNA duplex-hairpin interconversion for two DNA decamers: d(CAACGGGTTG) and d(CAACCCGTTG). Biopolymers 35 (1995) 187–200.

[24] E.T. Zuo, F.A. Tanious, D. Wilson, G. Zon, G. Tan, R.M. Wartell, Effect of base-pair sequence on the conformations and thermally-induced transitions in oligodeoxyribonucleotides containing only at base pairs. Biochemistry 29 (1990) 4446–4456.

[25] C. Roy, N. Tounekti, M. Mougel, et al., An analytical study of in vitro generated RNA of Moloney murine leukemia virus MoMuLV. Nucleic Acids Res. 18 (1990) 7287–7292.

[26] D.E. Wemmer, S.H. Chou, D.R. Hare, B.R. Reid, Duplex-hairpin transitions in DNA: NMR studies in CGCGTATACGCG. Nucleic Acids Res. 13 (1985) 3755–3772.

[27] M. Hald, J.B. Pedersen, P.C. Stein, F. Kirpekar, J.P. Jacobsen, A comparison of the hairpin stability of the palindromic d(CGCG(A/T)4CGCG) oligonucleotides. Nucleic Acids Res. 23 (1995) 4576–4582.

[28] S. Yang, M.M. Wolska-Klis, J.R. Cann, Gel electrophoresis of reacting macromolecules. Rate-limited self-association. Anal. Biochem. 196 (1991) 192–198.

[29] C. Kleinschmidt, K. Tovar, W. Hillen, Computer simulations and experimental studies of gel mobility patterns for protein binding to DNA. Nucleic Acids Res. 19 (1991) 1021–1028.

[30] J.R. Cann, Theory and practice of gel electrophoresis of interacting macromolecules. Electrophoresis 237 (1996) 1–16.

[31] J. Matthews, R.L. Walker, Mathematical Methods of Physics, Benjamin, New York, 1965.

[32] C.R. Calladine, C.M. Collis, H.R. Drew, M.R. Mott, A study of electrophoretic mobility of DNA in agarose and polyacrylamide gels. J. Mol. Biol. 221 (1991) 981–1005.

[33] T. Maniatis, E.F. Fritsch, J. Sambrook, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

[34] F.B. Howard, C.Q. Chen, P.D. Ross, H.T. Miles, Hairpin formation in the self-complementary dodecarner d-GGTACGCGTACC and derivatives containing GA and IA mispairs. Biochemistry 30 (1991) 779–782.