Theoretical Analysis of the Kinetics of DNA Hybridization with Gel-Immobilized Oligonucleotides

Mikhail A. Livshits and Andrei D. Mirzabekov
Joint Human Genome Program: Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 117984, Russia, and Argonne National Laboratory, Argonne, Illinois 60439 USA

ABSTRACT A new method of DNA sequencing by hybridization using a microchip containing a set of immobilized oligonucleotides is being developed. A theoretical analysis is presented of the kinetics of DNA hybridization with deoxynucleotide molecules chemically tethered in a polyacrylamide gel layer. The analysis has shown that long-term evolution of the spatial distribution and of the amount of DNA bound in a hybridization cell is governed by "retarded diffusion," i.e., diffusion of the DNA interrupted by repeated association and dissociation with immobile oligonucleotide molecules. Retarded diffusion determines the characteristic time of establishing a final equilibrium state in a cell, i.e., the state with the maximum quantity and a uniform distribution of bound DNA. In the case of cells with the most stable, perfect duplexes, the characteristic time of retarded diffusion (which is proportional to the equilibrium binding constant and to the concentration of binding sites) can be longer than the duration of the real hybridization procedure. This conclusion is indirectly confirmed by the observation of nonuniform fluorescence of labeled DNA in perfect-match hybridization cells (brighter at the edges). For optimal discrimination of perfect duplexes from duplexes with mismatches the hybridization process should be brought to equilibrium under low-temperature nonsaturation conditions for all cells. The kinetic differences between perfect and nonperfect duplexes in the gel allow further improvement in the discrimination through additional washing at low temperature after hybridization.

INTRODUCTION

DNA hybridization with immobilized probes is used in many preparatory and analytical techniques in molecular biology. A promising approach to DNA sequencing that is under development in a number of laboratories (Bains and Smith, 1988; Cantor at al., 1992; Drmanac et al., 1989; Fodor et al., 1991; Guo et al., 1994; Khrapko et al., 1989; Lamture et al., 1994; Livache et al., 1994; Lysov et al., 1988; Mirzabekov, 1994; Pease et al., 1994; Southern et al., 1992; Yershov et al., 1996) is based on the hybridization of a DNA fragment with a set of oligonucleotides of definite length (sequencing by hybridization). A perfect duplex formed by DNA with a given oligonucleotide provides evidence of the presence of the corresponding "word" in the DNA "text." Overlapping of "words" allows (under certain conditions) deciphering of the nucleotide sequence of the DNA.

In one of the sequencing-by-hybridization approaches (Mirzabekov, 1994) the oligonucleotides are immobilized in polyacrylamide gel cells. Such a three-dimensional immobilization (relative to two-dimensional (2D) immobilization on a glass or filter surface) has been shown (Khrapko et al., 1991; Livshits et al., 1992, 1994) to be advantageous in several ways: 1) the capacity of the gel support is much higher than that of 2D supports, 2) the wide spacing of immobilized oligonucleotide molecules prevents their interference during hybridization with DNA, and 3) the apparent stability of DNA binding, which is dependent on the concentration of the immobilized oligonucleotides, can be adjusted. The possibility of compensating for significant base-composition-dependent differences in the stability of perfect duplexes is an essential advantage of this approach.

A theoretical description of the kinetics of DNA hybridization with oligonucleotides fixed in a gel layer is needed for understanding and optimizing the process. Such a description should be helpful in choosing the best physicochemical conditions of hybridization for revealing the cells containing perfect duplexes. The broad scope of applications (DNA sequencing, analysis of genetic polymorphism, diagnosis of genetic diseases, gene expression studies, detection and analysis of viral and bacterial species) may require various adjustments of the conditions.

There are rather few experimental studies of DNA hybridization kinetics. The kinetics of perfect and mismatched hybridization of oligonucleotides ~20 bases long with longer DNA targets in solution at 68°C were monitored by Herning et al. (1991), who used fluorescence polarization measurements. The kinetics, on a scale of hours, was shown to be sensitive to mismatches, the influence of the mismatches being the more pronounced the shorter the oligonucleotides.

The kinetics of hybridization of oligonucleotides (in the same range of lengths) with targets immobilized on a solid support were recently measured by surface plasmon resonance (Gotoh et al., 1995). Association and dissociation constants for 20-base-pair duplexes at 37°C were estimated to be $k_{assoc} \sim 3 \times 10^5\ M^{-1}\ s^{-1}$ and $k_{diss} \sim 1.3 \times 10^{-4}\ s^{-1}$, whereas with a single mismatch the k_{diss} value increased 2.5 times.

Received for publication 4 June 1996 and in final form 12 August 1996.
Address reprint requests to Dr. Mikhail A. Livshits, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov Street, Moscow 117984, Russia. Tel.: 7-095-135-2316; Fax: 7-095-135-1405; E-mail: livshits@imb.ac.ru.
© 1996 by the Biophysical Society
0006-3495/96/11/2795/07 $2.00

A mathematical model for DNA hybridization on a solid surface was recently presented (Chan et al., 1995). The hybridization was modeled as irreversible binding mediated by 2D diffusion of adsorbed molecules to surface-immobilized probes. The conclusion was that, because of 2D diffusion, the hybridization kinetics can be optimized at a limited surface density of the probe.

In the case of DNA hybridization with oligonucleotides immobilized within a gel layer one should expect specific features of the process. During the hybridization process both the quantity and the distribution of DNA within a hybridization cell are changing. In some cases there is clear evidence of nonuniform distribution of DNA within the gel, i.e., brighter fluorescence of labeled DNA on the periphery of a hybridization cell (see Fig. 1). The theoretical conclusion conforming to these observations is that such nonuniformity should be characteristic of the cells with the most-stable duplexes. It is in these cells that the time needed to approach the equilibrium, uniform state is the longest, possibly longer than the time of the actual hybridization procedure. During the time that is characteristic for the binding of DNA to oligonucleotides an abrupt distribution of DNA near the gel surface is formed. Up to the time characteristic for dissociation, the DNA distribution becomes smoother but still remains nonuniform. A homogeneous distribution requires a much longer time, the time characteristic of "retarded diffusion." We introduced the concept of retarded diffusion (Livshits et al., 1992, 1994) when analyzing the kinetics of DNA washing out of the hybridization gel to provide a way to express special features of the molecular motion in a medium containing immobile binding centers. The diffusion of a molecule in such a medium is interrupted by each act of association with a binding site and is renewed only after dissociation. Such motion on average can be characterized as diffusion with a correspondingly lowered diffusion coefficient.

During the hybridization process DNA penetrates the hybridization cell by a mechanism of retarded diffusion, and the tighter the DNA binding in a cell, the longer the process of filling the cell. The duration of the hybridization procedure can be unfavorable for the discrimination of perfect duplexes from imperfect ones if there is enough time for maximum filling of imperfect cells and insufficient time for filling of perfect ones.

The discrimination of perfect duplexes from imperfect ones by comparison of the quantities of DNA bound in different cells can be subject to additional difficulties. It is possible that, within a DNA sequence, several sites are able to form imperfect duplexes with one and the same oligonucleotide. The cell containing such an oligonucletide can "imperfectly bind" DNA in a quantity in no way inferior to that bound in perfect-duplex cells.

It should be noted that such a pseudoconcentration effect does not create the same difficulties for some other methods of perfect/nonperfect discrimination. For example, if DNA is washed out at elevated temperatures, then the discrimination parameter (be it "washout temperature" (Livshits et al., 1994) or the characteristic time of DNA washing out at a given temperature), which is defined by the relative decrease of DNA content, does not depend on the initial DNA concentration.

THEORY

Mathematical formulation

We start the hybridization process by placing a microchip containing an array of oligonucleotides into a hybridization

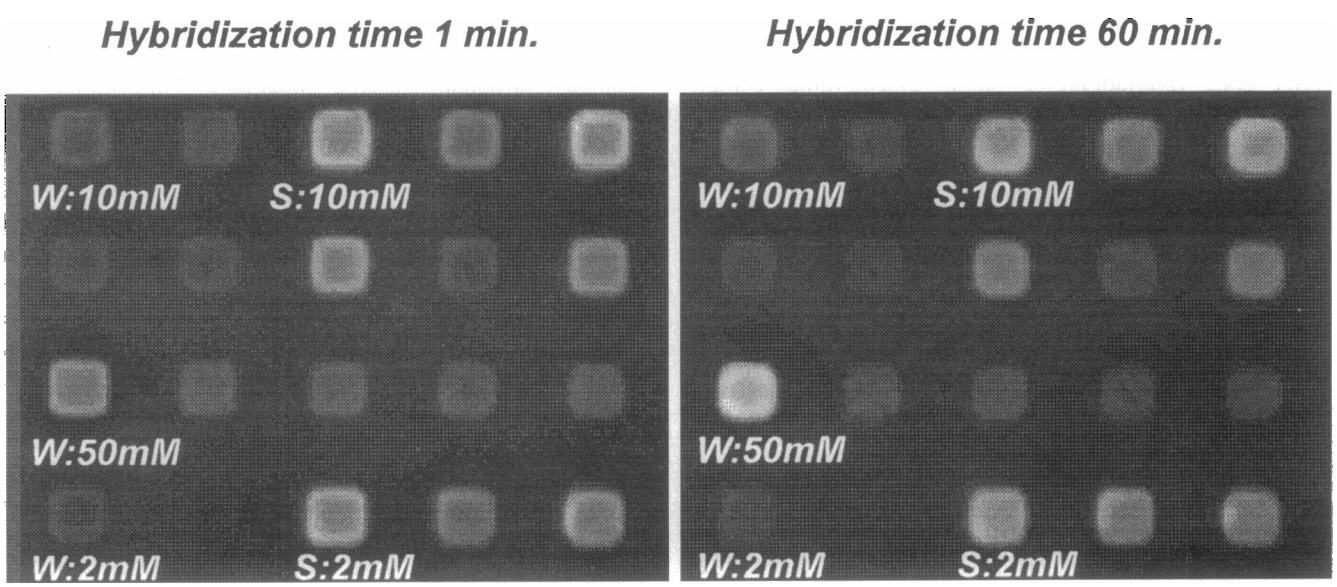

FIGURE 1 Hybridization of fluoresceine-labeled deoxyoligonucleotide 5′-CCGCGCCGTATAATTT-(FL) on a microchip for 1 min and for 60 min. The marked cells of the microchip contained gel-immobilized hexanucleotides, either strongly binding (*S*) gel-CGCGGC-5′ or weakly binding (*W*) gel-GGCATA-5′, at the specified concentrations. The hybridization was performed at 0°C, 4.5 M NaCl, and a sample oligonucleotide concentration of 0.5 μM. The preparation procedures and registration system were the same as described earlier (Yershov et al., 1996).

chamber containing a solution of fluorescently labeled DNA fragments. The hybridization chamber volume V can be some tens to hundreds of microliters. The microchip consists of many ($N \gg 1$) tiny polyacrylamide gel cells, with specific oligonucleotides chemically attached to the gel uniformly throughout the volume of each cell. Typical dimensions of the cells are from $100 \times 100 \times 30$ μm to $40 \times 40 \times 20$ μm, so the cell volume v is hundreds to tens of picoliters. The typical concentration of an immobilized oligonucleotide in a cell for most of these experiments was $m \cong 10^{-4}$ M (the theoretically estimated maximum is $\sim 5 \times 10^{-2}$ M). The number of cells in a sequencing chip containing a complete set of octanucleotides is $N_8 = 4^8 = 65{,}536$; for a chip containing all hexanucleotides, $N_6 = 4^6 = 4{,}096$.

The local concentrations of free DNA $f(x)$ and bound DNA $b(x)$ in a cell change owing to the diffusion of free DNA into the gel layer and to association/dissociation reactions with immobilized oligonucleotide molecules:

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} - k_{assoc}(m - b)f + k_{diss}b,$$
$$\frac{\partial b}{\partial t} = k_{assoc}(m - b)f - k_{diss}b. \quad (1)$$

Here t is time, D is the diffusion coefficient for the DNA fragment in the gel, m is the concentration of the immobilized oligonucleotide, and x is the coordinate across the gel layer counted from the glass surface ($x = 0$). (As the gel thickness, $h \sim 20$–30 μm, is rather small in comparison with the in-plane dimensions of a cell, we consider the one-dimensional model calculation. The influence of three-dimensional features can be assessed by comparison with another model calculation (see the Appendix) in which a hybridization cell is taken to be semispherical.)

At the initial moment, the cells contain no labeled DNA fragments:

$$f(x, 0) = b(x, 0) = 0$$

The boundary conditions for the concentration of free DNA in a cell are zero flux through the gel/glass border (i.e., zero gradient at $x = 0$) and the continuity of the concentration values at the interface with the buffer solution (i.e., the free DNA concentration in the gel at this boundary must be equal to the external bulk DNA concentration):

$$\left(\frac{\partial f}{\partial x}\right)_0 = 0, \qquad f(h) = c_{ext}. \quad (2)$$

It should be noted that here we are making an important assumption that the external DNA concentration is the same in all parts of the hybridization chamber. This means that we consider the transport of the solute over the whole volume of the chamber outside the gel cells to be much faster than in-cell diffusion. Inasmuch as DNA diffusion in free solution, although it is faster than diffusion in a gel, $D_0 \sim 10^{-6}$–10^{-5} cm^2 s^{-1}, can take tens or hundreds of hours for distances of the dimensions of the chip (~1 cm), a kind of forced steering must be used.

The DNA concentration in the buffer, c_{ext}, decreases slowly owing to diffusion of DNA into the hybridization cells:

$$\frac{\partial(c_{ext}V)}{\partial t} = -\sum_{cells}\frac{v}{h}J_c, \qquad J_c = D\left(\frac{\partial f}{\partial x}\right)_h. \quad (3)$$

Owing to the small parameter $v/V < 10^{-5}$ in this equation the evolution of external DNA stock, $c_{ext}V$, is much slower than the change of in-cell DNA concentrations $f(x, t)$ and $b(x, t)$. Within the time scale of the evolution of the f and b distributions the external DNA concentration can be considered constant.

As we are interested in the conditions appropriate for effective discrimination of cells containing perfect duplexes from those containing only mismatched ones, we will be concerned mainly with the case of low saturation levels, $b/m \ll 1$. In such a case, a simplified version of Eq. 1. can be used:

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x^2} - kf + k'b,$$
$$\frac{\partial b}{\partial t} = kf - k'b. \quad (1')$$

(The parameter notation is also simplified: $k_{assoc}m = k$, $k_{diss} = k'$, $k_{assoc}m/k_{diss} = K_{assoc}m = K$.)

Key parameters

For the following discussion estimates of the orders of magnitude of key parameters are needed. The association rate constant for complementary octanucleotides is $k_{assoc} \cong 10^6$ M^{-1} s^{-1} (see Braunlin and Bloomfield, 1991; Cantor and Schimmel, 1980). It is almost independent of nucleotide sequence and shows only a slight dependence on temperature (usually characterized by a small negative effective activation energy). So the characteristic time of the binding is quite short: $\tau_{assoc} = k^{-1} = (k_{assoc}m)^{-1} \cong 10^{-2}$ s.

The dissociation rate constant k_{diss} and the equilibrium binding constant K_{assoc}, on the contrary, depend strongly on both temperature and sequence. For example, the activation energy for dissociation of the duplex of the self-complementary octanucleotide GGAATTCC is $E_d^+ = 43 \pm 5$ kcal/mol, and the enthalpy of the duplex, $\Delta H° = -(52 \pm 2)$ kcal/mol (Braunlin and Bloomfield, 1991). This means an order-of-magnitude decrease in the dissociation rate constant k_{diss} when the temperature is lowered by ~9.5°C and an order-of-magnitude increase in the equilibrium binding constant K_{assoc} when the temperature is lowered by ~8°C.

The dissociation rate constant extrapolated to 0°C (from the value 2000 s^{-1} measured by Braunlin and Bloomfield (1991) at 40°C, 0.2 M Na^+) gives $k'(0°C) \cong 10^{-1}$ s^{-1}. Thus the characteristic time of dissociation is estimated to be

$\tau_{diss}(0°C) = (k'(0°C)^{-1} \cong 10$ s, and the equilibrium constant $K(0°C)$ is $\sim 10^4$. Note also that the stability of hybrid DNA/RNA duplexes may be both higher and lower than that of DNA/DNA duplexes, depending on the GC content and the desoxypyrimidine content of the hybrids (Lesnik and Freier, 1995).

The range of binding constant values for duplexes of different compositions can be assessed from the melting data (Doktycz et al., 1995) obtained for 128 octanucleotide duplexes (at 0.2 μM concentration of chains and 1 M NaCl). Although sequences extremely rich or poor in G-C (expected to correspond to most stable and the most unstable duplexes) were not present in the set, the range of binding constant values was ~4 orders of magnitude.

The characteristic time for diffusion, τ_D, of rather short DNA fragments in the gel on the scale of the depth of a cell (20–30 μm) can be estimated to be of the order of 10 s ($\tau_D \sim h^2/D$, $D \sim 10^{-6}$ cm^2 s^{-1}). Anticipating further analysis, note that the characteristic time of retarded diffusion, $\tau_{RD} \sim Kh^2/D$, at low temperature can be tens of hours and more. The characteristic diffusion length d, estimated to be ~1 μm, is much less than the thickness of the gel layer h (20–30 μm).

Solutions to kinetic equations

Introducing into Eq. 1′ the new variables $z(x, t)$ and $w(x, t)$:

$$f = c_{ext}(1 - z), \qquad b = Kc_{ext}(1 - w), \tag{4}$$

one obtains

$$\frac{\partial z}{\partial t} = k\left(d^2 \frac{\partial^2 z}{\partial x^2} - z + w\right), \qquad \left(\frac{\partial z}{\partial x}\right)_0 = 0, \qquad z(h) = 0;$$
$$\frac{\partial w}{\partial t} = k'(z - w) \qquad z(t = 0) = w(t = 0) = 1 \tag{5}$$

Here $d = (D/k)^{1/2}$ is the characteristic diffusion length.

Let the solution to Eq. 5 be represented by the Fourier expansion

$$z = \sum Z_n(t)\cos\left(\frac{\pi x}{2h}(2n + 1)\right),$$
$$w = \sum W_n(t)\cos\left(\frac{\pi x}{2h}(2n + 1)\right). \tag{6}$$

Then, by virtue of Eq. 5, the time evolution of the amplitudes Z_n and W_n obeys the following equations:

$$\frac{dZ_n}{dt} = k[-(\mu_n + 1)Z_n + W_n],$$
$$\mu_n = \left(\frac{\pi d}{2h}\right)^2 (2n + 1)^2, \qquad \frac{dW_n}{dt} = k'(Z_n - W_n). \tag{7}$$

Because $k \gg k'$ and $K \gg 1$ (hybridization is performed at a temperature low enough for the prevalence of binding over dissociation), the amplitudes Z_n in a very short time $(k(1 + \mu_n))^{-1} < 10^{-2}$ s attain their "stationary values":

$$Z_n = \frac{1}{1 + \mu_n} W_n, \tag{8}$$

which slowly evolve together with W_n:

$$\frac{dW_n}{dt} = -\beta_n W_n. \qquad \beta_n = k' \frac{\mu_n}{1 + \mu_n} \tag{9}$$

The most slowly evolving values are the lowest harmonics, $n = 0, 1, \ldots$, because $\mu_0 \sim 3 \times 10^{-3}$, $\mu_1 \sim 2.5 \times 10^{-2}$, $\mu_2 \sim 7.5 \times 10^{-2}$, $\mu_3 \sim 1.5 \times 10^{-1}$, ... "Senior" harmonics, for which $\mu_n \gg 1$, are changing at the rate $\beta_n \cong k'$.

The initial values for W_n are coefficients of the Fourier expansion of the unit on the interval $(0, h)$:

$$w_n(0) = u_n = \frac{4}{\pi} \frac{(-1)^n}{(2n + 1)}; \quad 1 = \sum u_n \cos\left(\frac{\pi x}{2h}(2n + 1)\right). \tag{10}$$

Finally, solutions to Eq. 9,

$$W_n = u_n \exp(-\beta_n t), \tag{11}$$

with Eqs. 4, 6, and 8 taken into account, describe the evolution of distributions of free $f(x, t)$ and the bound $b(x, t)$ DNA within a hybridization cell on the time scale $t \gg \tau_{assoc} = k^{-1} \sim 10$ ms, in particular for $t \sim \tau_{diss} = (k')^{-1} \sim 10$ s and even for $t \sim \beta_0^{-1} \sim \tau_{RD} \sim 10$ h:

$$f(x, t) = c_{ext}\left(1 - \sum u_n \frac{1}{1 + \mu_n} \exp(-\beta_n t) \cdot \cos\left(\frac{\pi x}{2h}(2n + 1)\right)\right),$$
$$b(x, t) = c_{ext}K\left(\sum (1 - \exp(-\beta_n t))u_n \cdot \cos\left(\frac{\pi x}{2h}(2n + 1)\right)\right). \tag{12}$$

For rather short times at which dissociation is negligible ("times of irreversible binding"), $\tau_{assoc} \ll t \ll \tau_{diss}$, the expressions for the DNA distribution take the following explicit forms:

$$f(x) \cong c_{ext} \frac{\cosh(x/d)}{\cosh(h/d)}, \qquad b(x, t) \cong c_{ext} \frac{\cosh(x/d)}{\cosh(h/d)} kt. \tag{13}$$

(The identity $\Sigma \frac{1}{(2n + 1)^2} = \frac{\pi^2}{8}$, the Fourier expansions (Eq. 10), and

$$\frac{\cosh(x/d)}{\cosh(h/d)} = \sum u_n \frac{\mu_n}{1 + \mu_n} \cos\left(\frac{\pi x}{2h}(2n + 1)\right)$$

were used.)

So, at "times of irreversible binding," $\tau_{assoc} \ll t \ll \tau_{diss}$, the DNA in the gel is concentrated in a narrow ($d \ll h$) surface layer; then, at "dissociation times," $t \sim \tau_{diss}$, the DNA distribution becomes smoother, though it remains nonuniform and tends to become uniform only at "times of retarded diffusion" (Livshits et al., 1992, 1994):

$$t \sim (\beta_0)^{-1} \cong \left(\frac{2h}{\pi}\right)^2 \frac{K}{D} = \tau_{RD}. \qquad (14)$$

The long-lasting gradient of DNA concentration (Eq. 14) is characterized by the whole depth of the cell (h), i.e., it is much smoother than the initial gradient, characterized by the diffusion length (d) under irreversible binding. The total quantity $B(t)$ of DNA bound in a cell for the time t is obtained by integration of the right-hand side of Eq. 12 over the cell's volume:

$$B(t) = c_{ext} K v \left(1 - \frac{8}{\pi^2} \sum \frac{1}{(2n+1)^2} \exp(-\beta_n t)\right)$$
$$= B(\infty)(1 - a_0 \exp(-\beta_0 t) - a_1 \exp(-\beta_1 t) \ldots) \qquad (15)$$

$$a_0 = \frac{8}{\pi^2} \cong 0.81, \quad \beta_0 = k' \frac{\mu_0}{1+\mu_0} \cong \frac{D}{K}\left(\frac{\pi}{2h}\right)^2 = 1/\tau_{RD},$$

$$a_1 = \frac{8}{9\pi^2} \cong 0.09, \quad \beta_1 = k' \frac{\mu_1}{1+\mu_1} \cong 9/\tau_{RD} \ldots.$$

It can be seen that binding of the major part of DNA (81% of the total that will finally be bound in the cell) proceeds at the rate of retarded diffusion.

DISCUSSION AND CONCLUSIONS

Our analysis has led to the conclusion that the kinetics of DNA hybridization to oligonucleotides immobilized within a gel layer is determined by the rate of a specific type of motion of DNA molecules in the medium containing the binding centers. This is "retarded diffusion" (Livshits et al., 1992, 1994), i.e., diffusion interrupted by repeated association/dissociation events with immobile binding centers. The higher the affinity of the binding sites and the larger their number, the slower is the process of filling up the hybridization cell. Note that such dependence can hardly be considered trivial. It may even seem to contradict the expectations. It seems evident that the DNA binding should become faster when the number of binding sites is increased; however, this is correct only for the very initial stages of binding.

At the initial stages, binding is almost nonspecific and is poorly dependent on the nucleotide sequence. Such are the binding rate constants. Imperfect duplexes should be formed in these stages at approximately the same rates as perfect ones. However, the principal DNA consumption takes place at later stages, when penetration of DNA into the cells is governed by the mechanism of retarded diffusion and so proceeds in different cells at different rates. Filling up of the "most perfect" cells can prove to be very slow. The same slow speed should characterize the spreading of DNA (bound primarily only at a surface layer) over the volume of cells. And actually a "frame" of brighter fluorescence at cell edges is usually observed for the duplexes of highest stability, with the contrast decreasing with time. These frames are explicit (though indirect) indications of incomplete hybridization in these cells.

The time of retarded diffusion estimated with the equilibrium constant for a GGAATTCC duplex (Braunlin and Bloomfield, 1991) extrapolated to 0°C is ~10 h. For more-G-C-rich octamer duplexes, this figure can be increased by at least an order of magnitude. Fig. 1 shows that the characteristic time of hybridization with gel-immobilized hexamers is on the scale of tens of minutes or longer.

The hybridization rate should also depend strongly on the length of the DNA fragment. For the diffusion coefficient we have taken the value $D \sim 10^{-6}$ cm^2 s^{-1}, which corresponds to rather small DNA fragments. For longer fragments (and for denser gels), because of the slower diffusion the hybridization time can increase by 1–2 orders of magnitude. The association constant K_{assoc}, which is determined by the duplex energetics, is expected to be independent of the length of the DNA fragment. However, it should be noted that there are rather complicated "dangling-ends" effects on the duplex stability (Senior et al., 1988; Williams et al., 1994). These effects are due to specific stacking interactions between overhanging bases and those involved in the duplex.

In the case of incomplete hybridization the quantities of perfect duplexes are preferentially lowered, making perfect/imperfect discrimination difficult. Additional errors can be due to nonhomogeneous distribution of labeled DNA within the cells.

Let us summarize the main factors that could help to strengthen the contrasts in a hybridization pattern and so favor reliable discrimination of perfect duplexes from nonperfect ones. (An embarrassing influence of these factors is that they also emphasize base-composition-dependent differences between perfect duplexes. Fortunately, there is a possibility of compensating for these undesirable differences by adjusting the concentrations of immobilized oligonucleotides. A complete compensation would be achieved if the concentrations of oligonucleotides multiplied by their association constants (with strictly complementary strands) were equal to each other:

1) It is desirable that the hybridization process be completed (i.e., that equilibrium be attained) in all hybridization cells. This would prevent underestimation of the binding ability of perfect-match cells, which is fraught with mistakes in their identification. However, the estimates show that the time needed for hybridization in the cells with the most stable duplexes can be too long.

2) Nonsaturation conditions are necessary if perfect/imperfect discrimination is assumed to be based on a comparison of the quantities of DNA bound in the corresponding

hybridization cells. The saturation of binding sites by DNA means that the quantity of bound DNA tends to the total number of binding sites in a cell, independently of relative stability of the duplexes in different cells. A simple estimate demonstrates the negative effect of saturation on the discriminating ability of hybridization pattern. The "cell signals" ratio $R = B_{perf}/B_{imperf}$ (the quantity of DNA in a perfect-match cell divided by that of an imperfect one), depending on the perfect-cell saturation level $\theta = B_{perf}/mv$, varies from its maximum value $M = K_{perf}/K_{imperf}$ (the association constants ratio) at the lowest saturation levels ($\theta \to 0$) to unity (no discrimination) at the highest saturation level ($\theta \to 1$):

$$R = M\left(1 - \theta\frac{M-1}{M}\right).$$

(The saturation level of the imperfect cell is assumed to be negligibly low). Two types of circumstances can ensure nonsaturation. Under intensive binding with absorption of almost all DNA, $1 < k_{ass}c^0 < (v/V)\Sigma K_{ass}m$ the total number of binding sites for the DNA fragment must exceed the amount of the fragment subject to hybridization. In this respect, the gel support with its very high capacity for oligonucleotide immobilization is highly advantageous compared with 2D glass or filter supports. However, a DNA shortage is not necessary in the opposite case, that of weak binding, $K_{ass}c^0 < 1$, with absorption of only a minor part of the DNA, $(v/V)\Sigma K_{ass}m < 1$. In such a case, nonsaturation is possible, even with the DNA quantity (c_0V) exceeding the total number of binding sites.

3) The lowering of the temperature is advantageous in two ways. It increases both the association constants and their differences, making the signal reading and perfect/imperfect discrimination easier.

4) The kinetic differences between hybridization cells with perfect and imperfect duplexes provide a great opportunity to improve the discriminating ability of a hybridization pattern. An additional washing procedure of appropriate duration performed after hybridization can diminish the quantity of DNA bound imperfectly, while leaving virtually unchanged the DNA content of cells containing perfect duplexes. To maintain larger kinetic differences and to allow a convenient duration, such additional washing should be carried out at a low temperature.

APPENDIX: KINETICS OF HYBRIDIZATION IN A SEMISPHERICAL GEL CELL

Using a spherically symmetrical Laplacian,

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r,$$

one obtains for the new variables $\varphi(r, t) = rz(r, t)$ and $\chi(r, t) = rw(r, t)$ equations similar to Eq. 5 but with other boundary and initial conditions:

$$\frac{\varphi}{\partial t} = k\left(d^2\frac{\partial^2\varphi}{\partial r^2} - \varphi + \chi\right), \qquad \varphi(R, t) = 0; \tag{5'}$$

$$\frac{\partial\chi}{\partial_t} = k'(\varphi - \chi), \qquad \varphi(r, t = 0) = r, \quad \chi(r, t = 0) = r$$

and hence other Fourier expansions:

$$\varphi = \sum \Phi_n(t)\sin\left(\frac{\pi nr}{R}\right), \qquad \chi = \sum X_n(t)\sin\left(\frac{\pi nr}{R}\right). \tag{6'}$$

Finally, the main differences are new expressions for time scale parameters:

$$\mu_n = \left(\frac{\pi nd}{R}\right)^2, \qquad \beta_n = k'\frac{\mu_n}{1+\mu_n}.$$

In particular, the characteristic "retarded diffusion" time, instead of Eq. 14, becomes

$$\tau_{RD} = (\beta_1)^{-1} \cong \frac{K}{D}\left(\frac{R}{\pi}\right)^2.$$

We are grateful to the members of the joint EIMB-ANL research group for helpful discussions and in particular to A. Drobyshev and D. Prudnikov for providing experimental findings before publication. This research was supported by grants from the Russian Human Genome program and the Russian Foundation for Basic Research (95-04-11241) and by grant 943352401 from the U.S. Department of Energy. We thank D. E. Nadziejka for editorial assistance.

REFERENCES

Bains, W., and G. C. Smith. 1988. A novel method for nucleic acid sequence determination. *J. Theor. Biol.* 135:303–307.

Braunlin, W. H., and V. A. Bloomfield. 1991. ^{1}H NMR study of the base-pairing reactions of d(GGAATTCC): salt effects on the equilibria and kinetics of strand association. *Biochemistry.* 30:754–758.

Cantor, C. R., A. Mirzabekov, and E. Southern. 1992. Report on the sequencing by hybridization workshop. *Genomics.* 13:1378–1383.

Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry. W. H. Freeman & Co., San Francisco, CA.

Chan, V., D. J. Graves, and S. E. McKenzie. 1995. The biophysics of DNA hybridization with immobilized oligonucleotide probes. *Biophys. J.* 69: 2243–2255.

Doktycz, M. J., M. D. Morris, S. J. Dormady, K. L. Beattie, and K. B. Jacobson. 1995. Optical melting of 128 octamer DNA duplexes. Effects of base pair location and nearest neighbors on thermal stability. *J. Biol. Chem.* 270:8439–8445.

Drmanac, R., I. Labat, I. Brukner, and R. Crkvenjakov. 1989. Sequencing of megabase plus DNA by hybridization: theory of the method. *Genomics.* 4:114–128.

Fodor, S. P., J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, and D. Solas. 1991. Light-directed, spatially addressable parallel chemical synthesis. *Science.* 251:767–773.

Gotoh, M., Y. Hasegawa, Y. Shinohara, M. Shimizu, and M. Tosu. 1995. A new approach to determine the effect of mismatches on kinetic parameters in DNA hybridization using an optical biosensor. *DNA Res.* 2:285–293.

Guo, Z., R. A. Guilfoyle, A. J. Thiel, R. Wang, and L. M. Smith. 1994. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* 22: 5456–5465.

Herning, T., E. Tamiya, and I. Karube. 1991. Specific liquid DNA hybridization kinetics measured by fluorescence polarization. *Anal. Chim. Acta.* 244:207–213.

Khrapko, K. R., Yu. P. Lysov, A. A., Khorlyn, V. V. Shick, V. L Florentiev, and A. D. Mirzabekov. 1989. An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett.* 256:118–122.

Khrapko, K. R., Yu. P. Lysov, A. A. Khorlin, I. B. Ivanov, G. M. Yershov, S. K. Vasilenko, V. L. Florentiev, and A. D. Mirzabekov. 1991. A method for DNA sequencing by hybridization with oligonucleotide matrix. *DNA Sequence.* 1:375–388.

Lamture, J. B., K. L. Beattie, B. E. Burke, M. D. Eggers, D. J. Ehrlich, R. Fowler, M. A. Hollis, B. B. Kosicki, R. K. Reich, S. R. Smith, R. S. Varma, and M. E. Hogan. 1994. Direct detection of nucleic acid hybridization on the surface of a charge-coupled device. *Nucleic Acids Res.* 22:2121–2125.

Lesnik, E. A., and S. M. Freier. 1995. Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry.* 34:10,807–10,815.

Livache, T., A. Roget, E. Dejean, C. Barthet, G. Bidan, and R. Teoule. 1994. Preparation of a DNA matrix via an electrochemically directed copolymerization of pyrrole and oligonucleotides bearing a pyrrole group. *Nucleic Acids Res.* 22:2915–2921.

Livshits, M. A., I. B. Ivanov, V. L. Florentiev, and A. D. Mirzabekov. 1992. DNA sequencing by hybridization with oligonucleotide matrix (SHOM). Theory of washing out the DNA after hybridization. *Mol. Biol.* 6:856–865.

Livshits, M. A., V. L. Florentiev, and A. D. Mirzabekov. 1994. Dissociation of duplexes formed by hybridization with gel-immobilized oligonucleotides. *J. Biomolec. Struct. Dynam.* 11:783–795.

Lysov, Yu. P., V. L. Florentiev, A. A. Khorlin, K. R. Khrapko, V. V. Shick, and A. D. Mirzabekov. 1988. Determination of the nucleotide sequence of DNA using hybridization with oligonucleotides: a new method. *Proc. USSR Acad. Sci.* 303:1508–1511.

Mirzabekov, A. D. 1994. DNA sequencing by hybridization—a megasequencing method and a diagnostic tool? *Trends Biotechnol.* 12:27–32.

Pease, A. C., D. Solas, E. J. Sullivan, M. T. Cronin, C. P. Holmes, and S. P. Fodor. 1994. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. USA.* 91:5022–5026.

Senior, M., R. A. Jones, and K. J. Breslauer. 1988. Influence of dangling thymidine residues on the stability and structure of two DNA duplexes. *Biochemistry.* 27:3879–3885.

Southern, E. M., U. Maskos, and J. K. Elder. 1992. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics.* 13: 1008–1017.

Williams, J. C., S. C. Case-Green, K. U. Mir, and E. M. Southern. 1994. Studies of oligonucleotide interactions by hybridisation to arrays: the influence of dangling ends on duplex yield. *Nucleic Acids Res.* 22: 1365–1367.

Yershov, G., V. Barsky, A. Belgovsky, E. Kirillov, E. Kreindlin, I. Ivanov, S. Parinov, D. Guschin, A. Drobyshev, S. Dubiley, and A. Mirzabekov. 1996. DNA analysis and diagnostics on oligonucleotide microchips. *Proc. Natl. Acad. Sci. USA.* 93:4913–4918.