

## **Gel-Immobilized Microarrays of Nucleic Acids and Proteins**

*Production and Application for Macromolecular Research*

**Jordanka Zlatanova and Andrei Mirzabekov**

### **1. Introduction**

Biochips are small platforms with spatially arrayed macromolecules (or pieces thereof) that allow the collection and analysis of large amounts of biological information. The principle of the technology is based on specific molecular recognition interactions between the arrayed macromolecules and the test molecule of interest. Classical examples of such recognition reactions are the interactions between the two complementary strands of a double-helical DNA molecule, between a single-stranded DNA stretch and the messenger RNA copied from it during transcription, between an antigen and an antibody, and between small ligands and their nucleic acid or protein partners.

It has become customary to compare biological microchips with electronic microchips with respect to their ability to perform multiple simple reactions in parallel in a high-throughput fashion. Biochips are expected to revolutionize biology in the same way as the electronic chips revolutionized electronics earlier in the twentieth century. Testimony to such a revolutionary role can be found in recent science polls, which ranked the biochip technology among the 10 most important scientific developments in 1998 (*1*). There are many different types of biochips (*2*). This chapter focuses on the biochip developed at the Engelhardt Institute of Molecular Biology in Moscow and the Biochip Technology Center at Argonne National Laboratory, Argonne, IL.

## 2. General Description of the MAGIChip™ Technology

MAGIChips™ (Micro Arrays of Gel-Immobilized Compounds on a Chip) are arrays that we have been developing for the past several years (3–5). This array is based on a glass surface that has small polyacrylamide gel elements affixed to it (**Fig. 1**). The size of the pads can differ from  $10 \times 10 \times 5 \mu\text{m}$  to  $100 \times 100 \times 20 \mu\text{m}$ , with volumes ranging from picoliters to nanoliters. Each individual gel element can function as an individual test tube because it is surrounded by a hydrophobic glass surface that prevents exchange of solution among the elements. This property is crucial to performing pad-specific reactions, e.g., polymerase chain reaction (PCR) amplification of the hybridization signal of specific sequences of interest.

The production of such microchips involves the following consecutive steps: creation of the microarray of gel elements (pads) on the glass surface (micromatrix), and application and chemical immobilization of different compounds (probes) onto the gel pads (**Fig. 1**). Once the blank micromatrix has been converted into a microchip containing the immobilized probes, the test sample is added and the reaction of molecular recognition takes place under specified conditions. To be able to monitor the results of such molecular interactions, the test sample needs to be labeled, usually by attaching various kinds of fluorescent labels to it.

Finally, the results of the molecular recognition reaction need to be monitored and analyzed. The type of monitoring instrumentation used depends on the required level of performance, and the type of label attached to the test molecule. The analysis of the reaction patterns is automated using specially designed software. In the next sections, we describe in more detail the separate steps of the production and use of the biochip. We also describe some specific features of the different types of biochips—oligonucleotide, cDNA, and protein chips—giving specific examples of their application. Because our efforts have been focused so far on nucleic acid biochips, most of what follows applies to those chips. Some developments concerning protein biochips are described at the end of this chapter.

### 2.1. Production of the Micromatrix

The matrix of glass-attached gel elements is prepared by photopolymerization (6). The acrylamide solution to be polymerized is applied to a manually assembled polymerization chamber consisting of a quartz mask, two Teflon spacers, and a microscopic glass slide, clamped together by two metal clamps (**Fig. 2A**). The internal side of the quartz mask has ultraviolet (UV)-transparent windows arranged in a specified spatial manner in nontransparent 1- $\mu\text{m}$ -thick chromium film (**Fig. 2B**). The assembled chamber containing the

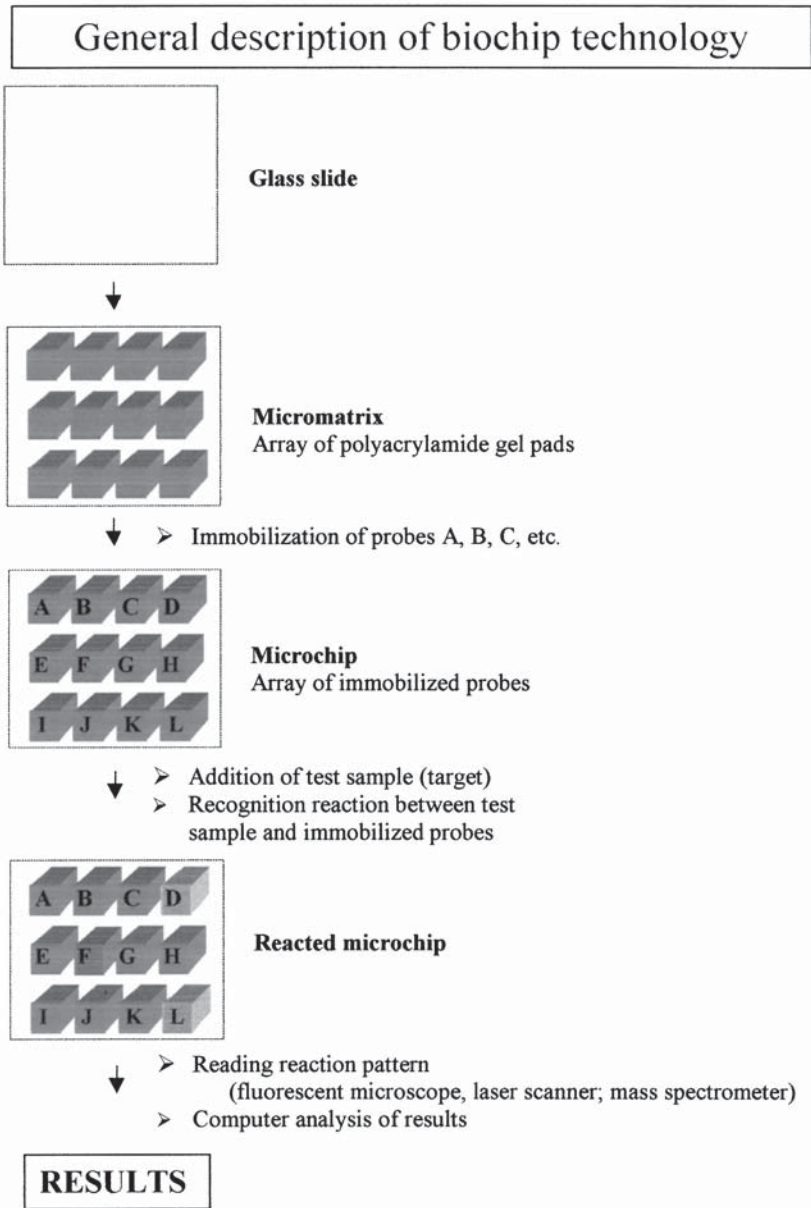


Fig. 1. Overall scheme of the MAGIChip™ technology.

acrylamide solution is exposed to UV light to allow polymerization in only those positions of the chamber that are situated directly under the transparent

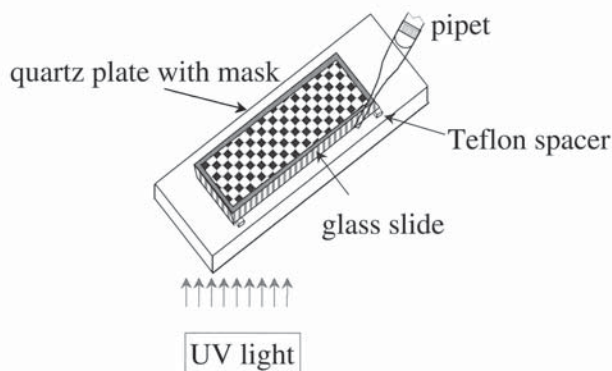
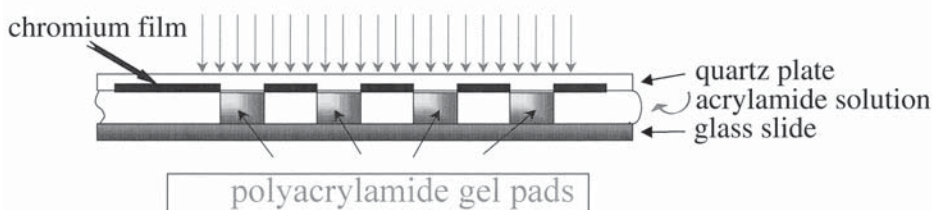
**A****B**

Fig. 2. (A) Scheme of the polymerization chamber. (B) Scheme of the photopolymerization.

windows. Following polymerization, the chamber is disassembled, and the matrix is washed, dried, and kept at room temperature in sealed chambers.

We have recently introduced a method for production of matrices that combines the polymerization step with the step of probe immobilization (7). In this method acrylamide is copolymerized with oligonucleotides or proteins containing unsaturated residues. In the case of oligonucleotides, such unsaturated units are incorporated during standard phosphoramidite synthesis; in the case of proteins, the protein is chemically attached to the acrylamide monomer containing double bond.

## 2.2. Probe Activation

Oligonucleotides or DNA fragments to be immobilized in the gel elements should be activated to contain chemically reactive groups for coupling with the activated gel elements. The chemistry of probe activation is chosen in concert with the chemistry of activation of the polyacrylamide gels. Thus, e.g., immo-

# B

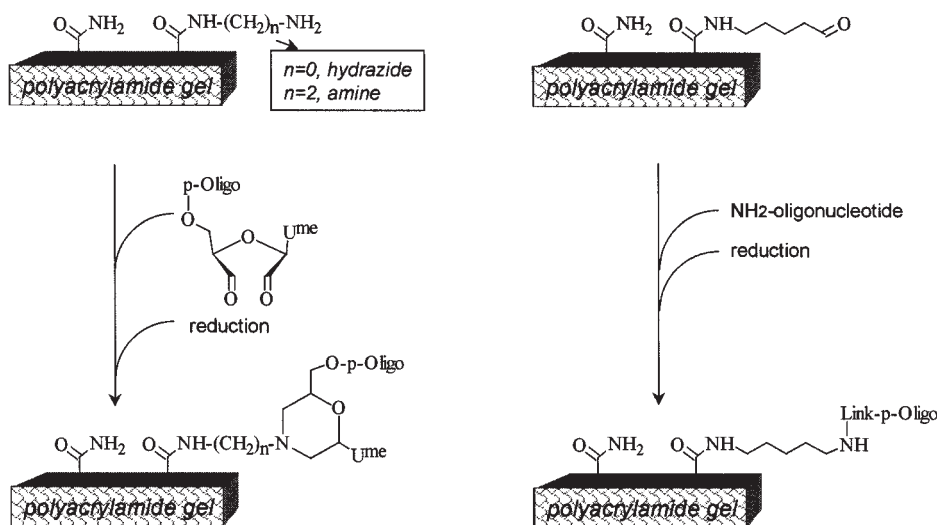


Fig. 3. Chemistry of immobilization of oligonucleotide probes into polyacrylamide gel pads.

bilization in aldehyde-containing gels would require the probe to be functionalized by the introduction of amino groups (8) (Fig. 3). If the gels are activated by the introduction of amino groups, the probes may be oxidized to contain free aldehyde groups (9) (Fig. 3). The probe can be prepared by introduction of chemically active groups in terminal positions of the oligonucleotides during their chemical synthesis; alternatively, active groups can be introduced within the chain of nucleotides (chemically synthesized or naturally occurring) in a number of ways (8,10). The probe activation chemistry is well developed and allows for high-yield, reproducible coupling with the gel matrix.

### 2.3. Application of Probes to Micromatrix and Their Chemical Immobilization in Gel Pads

Routinely, the probes for immobilization are transferred into the gel elements of the micromatrix using a home-designed dispensing robot (*II*). The fiber-optic pin of the robot has a hydrophobic side surface and a hydrophilic tip surface, and operates at dew point temperature to prevent evaporation of sample

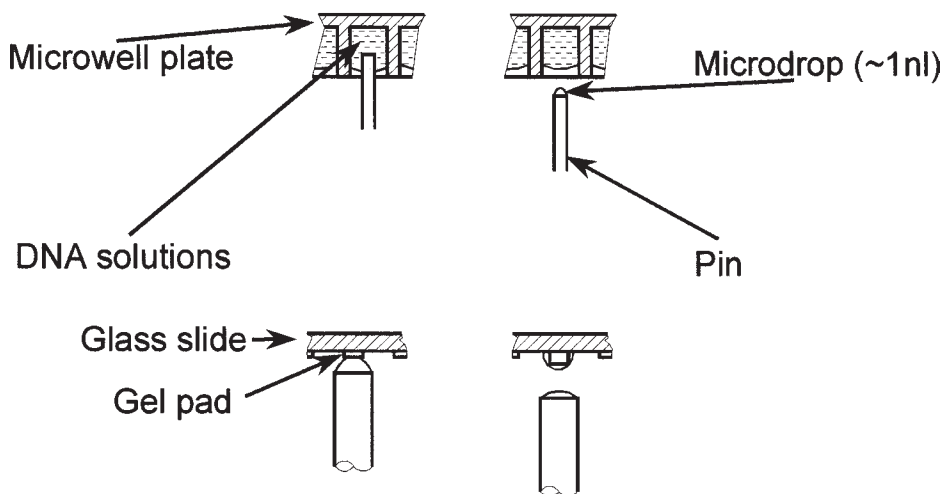


Fig. 4. Sectional view of the loading pin shown at the various phases of the loading process.

during transfer. The top of the pin is introduced into the probe solutions that are kept in microtiter plates, upon which a small 1-nL droplet is formed on the tip; the pin then touches the gel element surface, and the sample is transferred (**Fig. 4**). A manual version of this procedure is also available, in which the application is carried out with a pipet under a regular microscope (**6**).

The chemical immobilization of the activated probes to the gel elements is the next step of microchip production. We have been routinely using two methods for immobilization. In the first method, the gel supports contain amino or aldehyde groups allowing coupling with oligonucleotides bearing aldehyde or amino groups, respectively (**Fig. 3**) (**9**). In the second method, the polyacrylamide gel matrix is activated by introducing hydrazide groups that interact with the 3'-dialdehyde termini of activated oligonucleotides. The disadvantage of this method is that the hydrazide chemistry does not provide sufficient stability of attachment in repeated hybridization experiments.

## 2.4. Preparation of Target

The target molecules need to be labeled to allow monitoring of the interaction reactions. Although we have used radioactive labeling in the past, our present technology is based on labeling with fluorescent dyes. The advantages of using fluorescence detection are many, including the possibility to monitor processes in real time, high spatial resolution, and lack of radiation hazards. Several criteria need to be met by a labeling procedure:

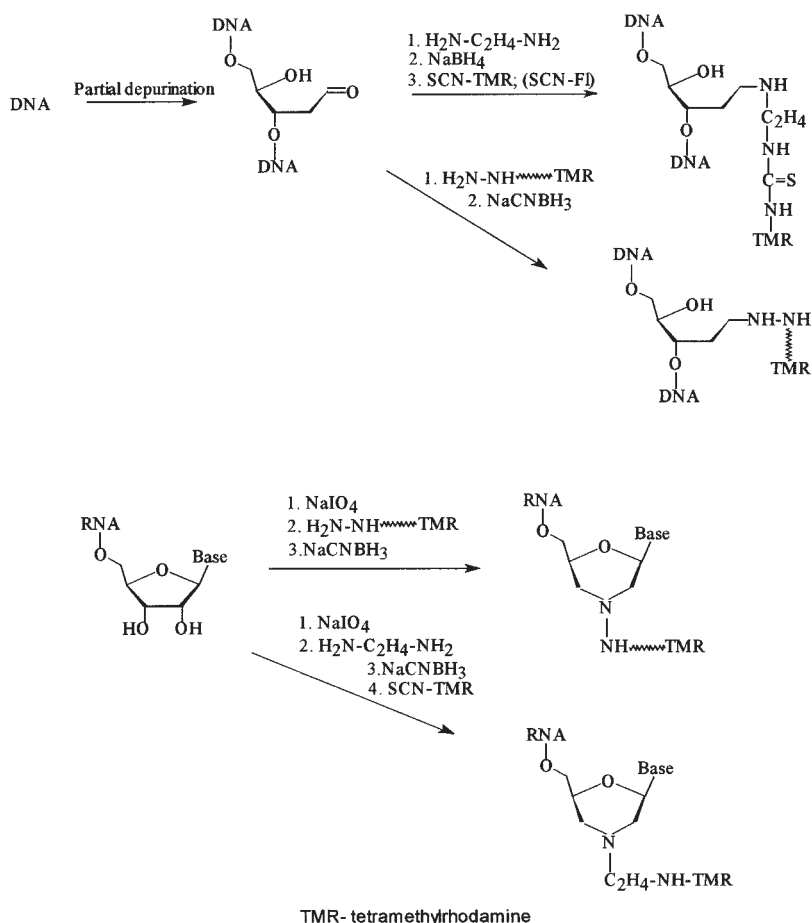


Fig. 5. Chemistry of the procedures used for labeling of DNA or RNA targets. TMR, tetramethylrhodamine.

1. It should be simple, fast, and inexpensive.
2. It should be applicable to both RNA and DNA targets.
3. It should be compatible with the fragmentation often required to decrease secondary structure formation (*see below*).
4. It should allow incorporation of one label into one fragment to ensure proper quantitation of the hybridization intensity.
5. It should allow coupling of multiple dyes.

We have developed a useful procedure (**10**) that is based on the introduction of aldehyde groups by partial depurination of DNA or oxidation of the 3'-terminal ribonucleoside in RNA by sodium periodate (**Fig. 5**). Fluorescent dyes

with attached hydrazine group are efficiently coupled with the aldehyde groups, and the bond is stabilized by reduction. An alternative procedure (10) uses ethylenediamine splitting of the DNA at the depurinated sites, stabilization of the aldimine bond by reduction, and coupling of the introduced primary amine groups with isothiocyanate or succinimide derivatives of the dyes. New methods for efficient simultaneous radical-based fragmentation and labeling are also being developed. Other published procedures based on reaction of abasic sites in DNA with fluorescent labels containing an oxyamino group (12) can also be used in target preparation.

### **2.5. Performing Hybridization: Some Theoretical Considerations**

The basic principle underlying the use of oligonucleotide and DNA biochips is the discrimination between perfect and mismatched duplexes. The efficiency of discrimination depends on a complex set of parameters (13,14), such as the position of the mismatch in the probe, the length of the probe, its AT-content, and the hybridization conditions. Thus, e.g., central mismatches are easier to detect than terminal ones, and shorter probes allow easier match/mismatch discrimination, although the overall duplex stability decreases as the length of the oligomer decreases, which may lead to prohibitory low hybridization signals with shorter probes.

Significant differences may exist in duplex stability depending on the AT content of the analyzed duplexes. This difference stems from the rather large difference in the stability of the AT and CG base pair (two vs three hydrogen bonds). The situation is further complicated because the stability is also sequence dependent: duplexes of the same overall AT content may have different stabilities depending on the mutual disposition of the nucleotides. Several approaches have been used to equalize the thermal stability of duplexes of differing base compositions, including using probes of different lengths, and performing the hybridization in the presence of tetramethylammonium chloride, or betaine (15).

If the technology allows monitoring of melting curves of duplexes formed with individual probe (16), then it is possible to optimize the reaction conditions in order to improve the discrimination of perfect/mismatched duplexes. Note that the melting temperature,  $T_m$ , of duplexes formed with matrix-immobilized oligonucleotides is a function of the concentration of the test sample (and is independent of the concentration of the immobilized species). The higher the concentration of the test sample, the more thermodynamically favorable the binding, and, hence, the higher the  $T_m$ . When melting is carried out in excess of target molecules, i.e., under conditions of saturation of all binding sites in the gel pad at low temperature, then no match/mismatch discrimination is possible at this temperature. Raising the temperature to the  $T_m$



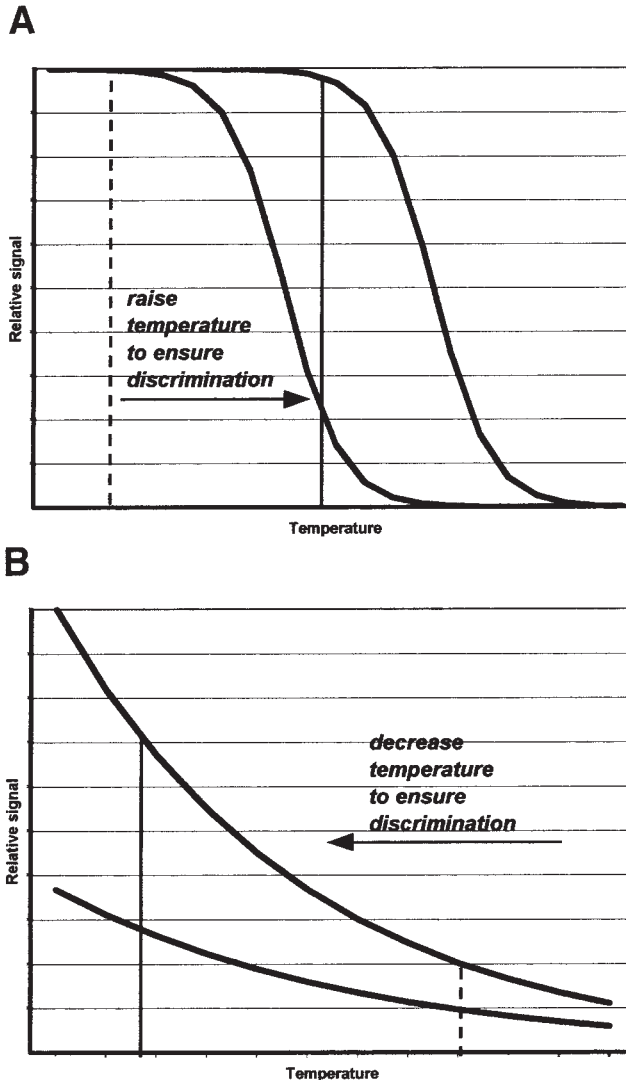


Fig. 6. Melting curves of duplexes formed under conditions of (A) target excess and (B) low concentration of target. The schemes illustrate the importance of correct temperature selection to ensure discrimination between duplexes of different thermodynamic stabilities. The temperatures that will not allow discrimination are represented by dashed lines, whereas the temperatures at which discrimination will be easily achieved are represented by the solid lines.

will ensure the discrimination (Fig. 6). On the other hand, if hybridization is carried out at low concentration of the target, so that even the low temperature

will not lead to saturation of the probe, it will be necessary to decrease the temperature to enhance discrimination. In such a case, both the perfect and the mismatched signals and the difference between them will be increased.

The previous discussion refers to thermodynamic equilibrium differences in the stability of the perfect and mismatched duplexes, and will be valid only under equilibrium hybridization conditions. Discrimination may, however, be achieved through alternative, kinetic differences. For instance, posthybridization washes can drastically reduce the mismatched signals, almost without affecting the perfect duplexes, in view of the faster dissociation of the mismatched ones.

An interesting twist in approaching the AT content problem came from the unexpected experimental observation that if the oligonucleotide probes are immobilized in three-dimensional gel pads, the apparent dissociation temperature,  $T_d$  (defined as the temperature at which the initial hybridization signal decreases 10-fold during step wise heating, posthybridization washing), is actually dependent on the concentration of the immobilized oligonucleotides (5). (For the usual first-order dissociation reaction in solution, the kinetics should be probe concentration-independent.) Our analysis suggests that the diffusion of the dissociated test molecules through the gel pad is retarded by encountering and reversibly binding to other probes immobilized at high density within the gel pad. This retarded diffusion is then probe concentration dependent and creates the apparent probe concentration dependence of the dissociation as a whole. This experimental observation was used to derive an algorithm that allows the design of "normalized" oligonucleotide matrices in which a higher concentration of AT-rich and lower concentration of GC-rich immobilized oligonucleotides can be used to equalize apparent dissociation temperatures of duplexes differing in their AT content, thus facilitating true match/mismatch discrimination.

Finally, we need to note the possibility of using chemically modified nucleotides to improve the discrimination. Examples of such use have been reported (17), and our own unpublished experiments clearly demonstrate the feasibility of such an approach.

Another issue that requires careful consideration is the effect secondary structures in single-stranded nucleic acids may have on the hybridization. The same conditions that favor duplex formation between the immobilized probe and the target will also favor intrastrand duplexing, thus making the target sequence inaccessible for intermolecular complex formation. The use of peptide nucleic acids as probes, rather than standard oligonucleotides, has been described (18) to circumvent this obstacle. We have chosen to prevent the formation of stable secondary structures in the target molecules by performing random fragmentation and fluorescent labeling of the targets under conditions

in which the duplexes are melted, e.g., by high temperature. The use of such fragmented targets for hybridization is efficient and produces signals of high intensity.

In summary, even this brief description makes it clear that the design of the biochip and the hybridization conditions should be carefully selected to give unambiguous and reproducible results.

## **2.6. On-Chip Amplification Reactions**

Use of biochip technology will be greatly broadened if on-chip amplification of the hybridization reaction could be performed. This is a highly desirable feature in cases when the nucleic acid of interest presents only a relatively small portion of the molecular population applied on the chip, e.g., when one is dealing with single-copy genes or with mRNAs of low abundance. With this in mind, we are developing methods for on-chip amplification.

In a single-base extension approach (**19**), a primer is hybridized to DNA and extended with DNA polymerase by a dideoxyribonucleoside triphosphate that matches the nucleotide at a polymorphic site. In our method (**20**), we perform the single-base extension reaction isothermally, at elevated temperatures, in the presence of each of the four fluorescently labeled ddNTPs (**Fig. 7A**). Performing the extension at a temperature above the melting temperature of the duplex between the DNA and the immobilized primer allows rapid association/dissociation of the target DNA. Thus, the same DNA molecule interacts in succession with many individual primers, leading to amplification of the signal in each individual gel pad. In an alternative procedure, the biochip contains four immobilized primers that differ at the 3' end by carrying one of the four possible nucleotides, matching the polymorphic site (**Fig. 7B**). In this case, extension of the primer will occur only in the gel pads where the primer forms a perfect duplex with the target DNA. Both procedures were applied to the identification of  $\beta$ -globin gene mutations in  $\beta$ -thalassemia patients, and to the detection of anthrax toxin gene (**20**).

We are also in the process of performing bona fide PCRs directly on the chip, with high expectations of success. In principle, the capability of the chip to perform individual PCRs in individual gel pads depends on the possibility of isolating each pad from its neighbors, which is trivial with our technology but may present insurmountable obstacles in other available chip platforms.

## **2.7. Readout**

For the analysis of the hybridization results obtained with fluorescently labeled target molecules, we use instrumentation constructed in collaboration with the State Optical Institute in S. Petersburg, Russia (**Fig. 8**). The instruments are based on research-quality fluorescence microscopes employing cus-

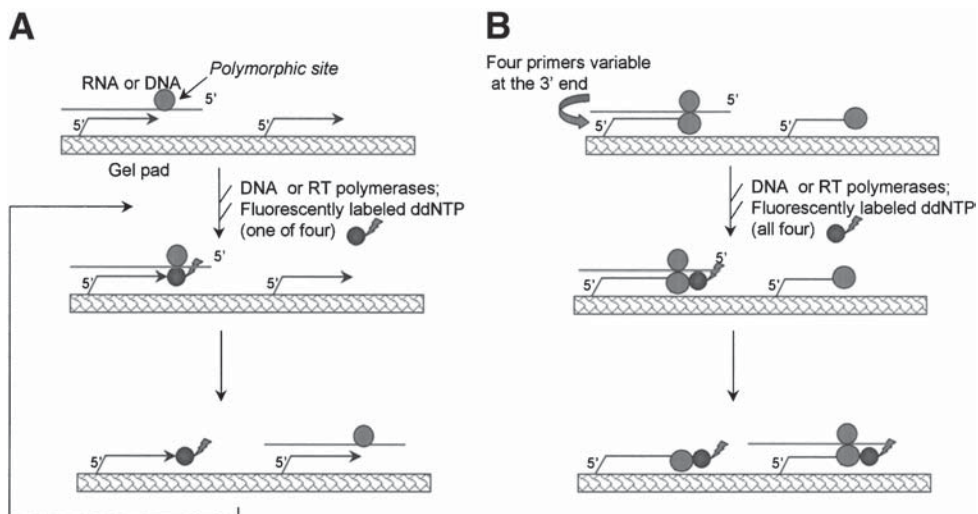
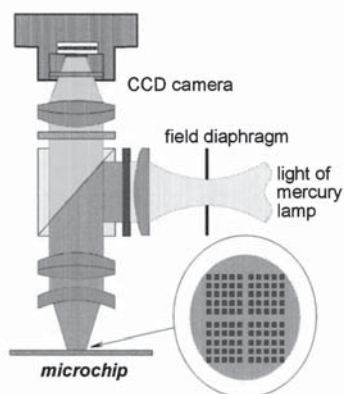
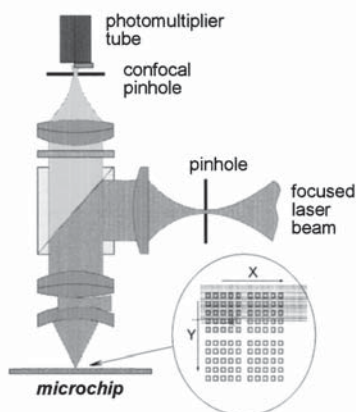


Fig. 7. Schemes illustrating the principle of single-nucleotide extension. (A) Multibase extension. (B) Multiprimer extension. The extension reactions are performed isothermally, at high temperature, to ensure "jumping" of the target molecule from one immobilized primer to another. For further details, *see* text.

tom-designed, wide-field, high-aperture, large-distance optics, and a high-pressure mercury lamp as a light source for epiillumination. Interchangeable filter sets allow the use of fluorescein, Texas Red, and tetramethylrhodamine derivatives as labeling dyes. The instruments are equipped with a controlled-temperature sample table, which allows changing the temperature in the range from  $-10$  to  $+60^{\circ}\text{C}$  in the chip-containing reaction chamber during the course of the experiment. The position of the thermostable can also be changed in a stepwise manner, to allow two-dimensional movement of the sample and analysis of different fields of view. A cooled charge-coupled device (CCD) camera is used to record the light signals from the chip, which are then fed into the analyzing computer program for quantitative evaluation of the hybridization signals over the entire chip.

At present, we are using four different variants of the microscope-based reading instrumentation that differ in their performance level. An important advantage of these devices is that they allow real-time monitoring of the changes in hybridization signal in each individual gel pad under a wide variety of experimental conditions. Most important, they allow monitoring of melting curves, which, in some cases, may be crucial in the proper match/mismatch discrimination. Such instrument capabilities are also important in studies of

FLUORESCENCE MICROSCOPE + CCD IMAGE DETECTORCONFOCAL LASER SCANNER

	<u>MICROSCOPE</u>	<u>SCANNER</u>
Readout mode	parallel, in real time	serial
Readout area	field of view (now - 8 x 8 mm)	several cm <sup>2</sup>
Relative pixel size	0.001 to 0.002 typ.	0.0005 typ.
Background due to out-of-focus fluorescence and light scattering	relatively high	negligible
Effective dynamic range	< 300 : 1	> 1000 : 1
# of wavelengths is # of:	filters (now - 4)	lasers (now - 2)

Fig. 8. Types of reading instrumentation, fluorescence microscope, and laser scanner and their basic characteristics.

the specificity of binding of sequence-specific ligands to single- and double-stranded DNA, because such specific binding raises the melting temperature of the ligand-bound duplexes to a measurable and interpretable degree. The feasibility of such an approach has recently been demonstrated in studies of Hoechst binding to DNA (21).

Although the most widely used in our current experimental practice, the conventional imaging fluorescence microscopy is not the only approach to microchip readout that is under development in our group. In many cases, when parallel measurements of gel pad signals are essential because of possible data loss, a more cost-effective solution of the readout problem can be offered using laser-scanning platforms. Because of inherently low background and excellent uniformity of the fluorescence excitation and detection, microchip scanners are especially well suited for precise quantitative measurements of signals varying over the range of three or even more orders of magnitude. However, all commercially available scanners are closed-architecture instruments optimized

with the surface-immobilization microchips in mind. Typically, they lack such a useful feature as temperature control of the sample table and employ an objective lens with a working distance too small to accommodate microchips packaged in a hybridization cell.

To meet the specific requirements of gene expression studies and cost-sensitive diagnostic applications, we have recently developed a laser scanner of unique, nonimaging design that makes use of the well-defined geometry of the gel-based microchips. The scanner employs a 2-mW HeNe laser as an excitation source and a low-noise PIN photodiode as a detector. The laser wavelength (594 nm) almost perfectly matches the absorption band of Texas Red. The numerical aperture of the miniature objective lens is 0.62. Yet, its working distance (approx 3 mm) is long enough for scanning packaged microchips. A microchip is mounted on a stationary controlled-temperature sample table of a design similar to that used in our fluorescence microscopes. All parameters of the scanning, data visualization, and processing are set up via the host computer. The hybridization pattern can be stored in a file either in the raw-data format or as an array of integral fluorescence intensities calculated on-line per each gel pad. Using a Texas Red dilution series microchip, we determined the detection threshold ( $3\sigma$ ) of the scanner to be approx 2 amol of Texas Red per gel pad, with a linear dynamic range being up to three orders of magnitude in terms of integral signal intensities. These characteristics are close to those of a commercial ScanArray 300 scanner (General Scanning) that we use for routine microchip inspection.

## **2.8. Informatics**

The digitized images of hybridization patterns obtained with the help of the CCD camera are further treated with the help of specialized software. This treatment includes automatic image analysis that determines the localization of the rows and columns of the matrix gel pads and their centers. For each element that contains a large number of pixels, the program calculates the total intensity of the hybridization signal. The program allows, if the need arises, filtering of the image in order to remove any noise coming from fluorescent impurities (e.g., dust particles) in the gel. The computer then performs, based on the calculated intensities of all gel pads on the chip image and stored information on standard image patterns, recognition operations. Such operations help the investigator obtain the final results in a user-friendly format.

## **3. Types of MAGIChips and Examples of Application**

### **3.1. Oligonucleotide Chip**

#### **3.1.1. Customized Oligonucleotide Arrays**

Customized oligonucleotide biochips are designed to interrogate test samples of known nucleotide sequences. Such sequences may be those of

known genes in cases when one is interested in their expression levels under specified conditions, or sequences of genes that are known or expected to contain single-nucleotide mutations, or to be polymorphic in a given population of individuals. In all these cases, the proper choice of oligonucleotide probes to be immobilized is of crucial importance to the success of the assay.

The choice of the oligonucleotide probes depends on the particular application, but there are certain basic considerations common to all cases. A basic requirement is to minimize the number of probes to be immobilized because such minimization may lead to a simplified design of the chip, which, in turn, makes its production cheaper and its use more convenient. A smaller chip will be easier to handle and read and will require a smaller amount of test sample. On the other hand, the set of selected probes must be big enough to allow unambiguous identification of the test samples; ambiguity might arise in view of the inherent to the procedure variations in the intensity of the hybridization signal in individual gel pads.

With these general requirements, a set of potential probes is created for each interrogated sequence that form perfect duplexes with that sequence. Then, some of the potential probes that may create ambiguities in the interpretation of the hybridization pattern are excluded from this list based on the AT vs GC content, and the propensity to form hairpins and other types of stable secondary structures that may drastically affect the intensity of hybridization. At a final stage of selection, all members of the shortened list are checked for their uniqueness: the probes should not form duplexes similar in their stability to the perfect duplexes with any material present in the mixture applied to the chip. This round of selection thus compares the sequence of each selected probe to the sequences in the set of test samples that are supposed to be distinguished from each other on the chip. The comparison of duplex stabilities takes into consideration not only the number of existing mismatches but also their location with respect to the probe because the stabilities of the perfect and mismatched duplexes may differ insignificantly if the mismatched nucleotide is close to the end of the probe, especially in cases of longer probes.

Examples of successful applications of customized oligonucleotide chips include detection of  $\beta$ -thalassemia mutation in patients (11,20,22), gene identification (23), allele-type identification in the human HLA DQA1 locus (24), identification of polymorphic base substitutions in patients with neurological disorders (25), and identification of and discrimination among closely related bacterial species (26). For the diagnostics of  $\beta$ -thalassemia mutations, a simple chip was designed that contained six probes corresponding to different  $\beta$ -thalassemia genotypes and hybridized with PCR-amplified DNA from healthy humans and patients (11). The hybridization results showed the expected significant differences in signal intensity between matched and mismatched duplexes, thus allowing reliable identification of both homozygous and heterozygous mutations.



For the bacterial biochip, oligonucleotides complementary to the small-subunit rRNA sequences were immobilized on the chip and hybridized with either DNA or RNA forms of the target sequences from nitrifying bacteria (26). This biochip successfully identified the tested microorganisms. In addition, the system was used to verify the utility of varying the concentrations of the immobilized oligonucleotides to normalize hybridization signals, and to demonstrate the use of multicolor detection for simultaneous hybridization with multiple targets labeled with different fluorescent dyes. In another application, chips based on sequences from 16S rRNA were successfully used to discriminate among closely related pathogenic and nonpathogenic bacterial species.

### 3.1.2. Generic Oligonucleotide Arrays

A more universal array that can be used to interrogate any target sequence is the so-called generic array that uses complete sets of small oligonucleotides of a given length. Such arrays were originally proposed for *de novo* DNA sequencing (3,27,28), but some intrinsic problems have, for the time being, hampered their implementation to such sequencing. These are mainly connected to sequence reconstruction ambiguities stemming from the presence of repeats along the DNA, and because relatively short stretches of nucleotides (of the length of the immobilized probes) can be found in many positions of a complex DNA sample. These restrictions severely limit the length of the DNA fragment that can be successfully decoded on short oligonucleotide arrays. Although the length and complexity of the readable DNA can be increased by using arrays of longer oligonucleotides, the gain in length is expected to grow linearly with the length of the probes (29), whereas the number of probes in a complete set of  $n$ -mers increases exponentially with probe length ( $4^n$ ). Thus, a complete array of 6-mers contains 4096 members, of 7-mers, 16,384 members; of 8-mers, 65,536; of 9-mers, 262,144; and so on. The production of such arrays still poses a practical challenge.

We have suggested that the use of contiguous stacking hybridization (4) can largely overcome the need of such impractically large arrays. In this approach, the initial hybridization of the target DNA with the array containing the full set of oligonucleotides of fixed length  $L$  is followed by additional multiple rounds of hybridizations with fluorescently labeled oligonucleotides of length  $l$  (Fig. 9). These labeled oligonucleotides will form extended duplexes with the target DNA strand owing to the strong stacking interactions between the terminal bases of the existing DNA duplex with the immobilized probe and of the otherwise unstable duplex with the short labeled probe (Fig. 9). The stacking interactions stabilize the DNA duplex even in the absence of a phosphodiester bond or a phosphate group. The stacked  $l$ -mers can also be ligated to the probe ([30]; for on-chip ligation, see ref. 23). Theoretical calculations have demonstrated



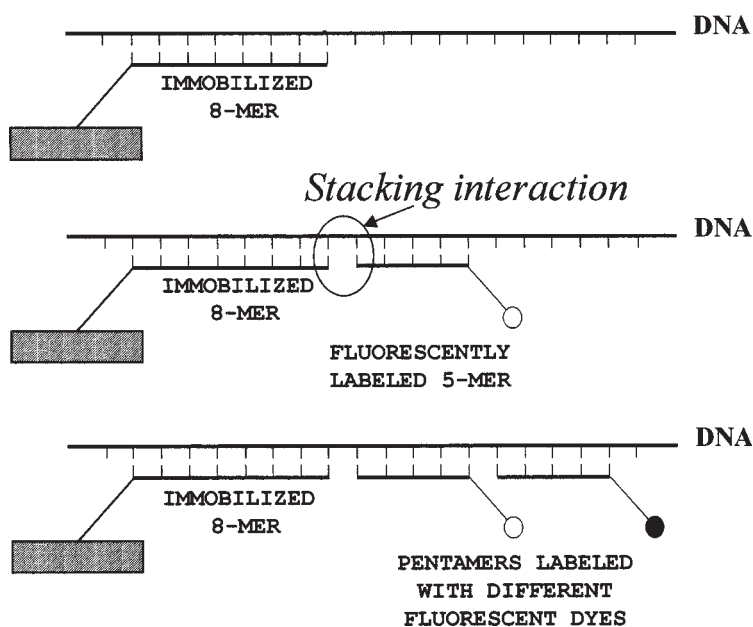


Fig. 9. Principle of contiguous stacking hybridization.

that contiguous stacking hybridization considerably increases the resolution power of the matrix, which approaches the power of  $(L + I)$ -matrix (29). An algorithm has been developed that allows the minimization of the number of additional hybridization steps, by assembling sets of  $I$ -probes to be added together at each hybridization round (29).

Recently, an ingenious approach that uses the ligation of DNA targets to arrays containing duplex probes with 5'-mer overhangs has been reported by Affymetrix (31). Even this approach allowed only for unambiguous sequence verification, not *de novo* sequencing, of relatively long targets (1200 bp). Note that although ligation certainly expands the resolution power of the chip, such chips cannot be used more than once. This limitation will be especially undesirable with more complex, expensive arrays.

In principle, contiguous stacking hybridization may be used with customized oligonucleotide chips too (32). In this work, an alternative hybridization strategy uses sets of labeled "stacking" oligonucleotides, each containing a "discriminating" base at one position and a universal base or a mixture of all four bases at all other positions. Such an approach allows 1024 rounds of hybridization with all possible 5-mers to be replaced by only five rounds of hybridization. This study has demonstrated that the 5-mers are stabilized in duplexes even with weak stacking bases and that mismatches in any of the five

positions of the 5-mers drastically destabilize the stacked duplexes. The destabilizing effect of mismatches in the 5-mers was also convincingly shown in a study of point mutations of the  $\beta$ -globin gene associated with  $\beta$ -thalassemia (11). Because the destabilizing effect of mismatches increases dramatically upon decreasing the length of the duplex (i.e., it is much stronger in 5-mers compared with, e.g., 8-mers) the use of contiguous stacking hybridization with short additional probes will be much more advantageous in mismatch discrimination than the conventional one-probe procedure.

### 3.2. cDNA Microarrays

In cDNA microarrays, the immobilized probes are individual cDNAs obtained by reverse transcription on mRNA populations extracted from cells in different physiological and developmental states. cDNA arrays have been widely used to study gene expression (33–35). A potential problem with cDNA arrays on the MAGICChip could be connected to the difficulty of introducing and evenly distributing long molecules into the gel pads. To facilitate diffusion of longer cDNAs, we are successfully developing polyacrylamide gels of various compositions that contain larger average pore sizes. Such gels will also be used in protein chips, where it may be necessary to immobilize rather bulky protein molecules within the gel pads (see **Subheading 3.3.**). An even more practical solution to the diffusion problem lies in randomly fragmenting the cDNA into relatively small pieces before immobilization.

### 3.3. Protein MAGICChips

It is our goal to expand the capabilities of our technology by producing protein chips. Such chips should contain different proteins immobilized as probes in a way that preserves their biological activity. The feasibility of producing such gels has been demonstrated (6,36).

One potential limitation with protein chips may stem from the difficulty of diffusion within the gel pads of molecules of high molecular masses. A way to circumvent this limitation is being sought in the production of polyacrylamide gels of larger pore sizes. In principle, larger-pore gels can be prepared by increasing the ratio of the crosslinker *N,N'*-methylenebisacrylamide (Bis) to acrylamide (37), or by the use of alternative crosslinkers such as *N,N'*-diallyltartardiamide (6), or a mixture of Bis and *N,N'*-(1,2-dihydroxyethylene)-bisacrylamide (36–38).

We have tested two protein-immobilization procedures (36). The first is based on activation of the polyacrylamide gel with glutaraldehyde (39). In the second procedure, which is applicable to glycoproteins such as antibodies, the gel is activated by partial substitution of the amide groups with hydrazide groups, and the polysaccharide component of the protein is activated by  $\text{NaIO}_4$

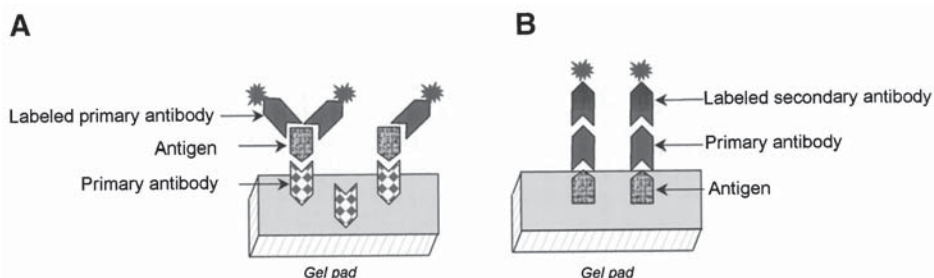


Fig. 10. Indirect approaches for detection of antigen-antibody reactions performed on the chip: the target is unlabeled, and a positive reaction is registered by using labeled molecules that specifically recognize the target. (A) The chip contains immobilized primary antibodies to interrogate the sample for the presence of a specific antigen. Detection is through binding of fluorescently labeled primary antibodies that recognize antigenic determinants distinct from the ones recognized in the primary reaction. (B) The chip contains immobilized antigen that is recognized by primary antibodies present in the solution. The reaction is detected via binding of secondary antibodies (antibodies to the first antibody), as in classical immunochemical techniques.

oxidation. The reaction between the hydrazide and aldehyde groups efficiently crosslinks the protein to the gel. This procedure is similar to one of the methods used for oligonucleotide immobilization.

Protein microchips preserve the high specificity of molecular recognition reactions observed in solution. For instance, the interaction between antigens and their specific antibody partners may occur on-chip in a variety of experimental setups. Either the antigen or the antibody can be immobilized, and both direct and indirect monitoring reactions can be performed. In the direct methods, one uses target molecules labeled with fluorescent dyes or coupled to enzymes catalyzing color precipitate-forming reactions. In the indirect methods, the target is unlabeled, and the reaction is detected by using a labeled molecule that specifically recognizes the target. Examples include the use of secondary antibodies to detect primary antibodies bound to the immobilized antigen, or the use of labeled primary antibodies to detect antigens bound to immobilized specific antibodies (**Fig. 10**).

Finally, the protein biochip allows the study of enzymatic activity of immobilized enzymes, by overlaying the chip with solutions containing the respective substrates. The reaction is monitored by following the formation of color or fluorescent precipitates; more important, this can be done in real-time experiments. We have been successful in demonstrating the feasibility of this approach with enzymes of different molecular masses: horseradish peroxidase (44 kDa), alkaline phosphatase (140 kDa), and  $\beta$ -D-glucuronidase (290 kDa).

The enzyme biochips present an important future application for combinatorial drug discovery, in view of the possibility of detecting the effect of inhibitors on enzymatic activity (36).

## Acknowledgments

We acknowledge the members of our team for help in preparation of the manuscript, in particular Drs. V. Barsky, S. Bavykin, M. Livshitz, Y. Lysov, A. Perov, D. Proudnikov, V. Shick, E. Timofeev, and G. Yershov. We also acknowledge the skillful secretarial help of Felicia King. This work was supported by the U.S. Department of Energy, Office of Health and Environmental Research, under contract no. W-31-109-Eng-38; the Defense Advanced Research Project Agency, under interagency agreement no. AO-E428; CRADA grant no. 9701902 with Motorola and Packard Instrument Company; and by grants no. 558 and 562 of the Russian Human Genome Program.

## References

1. (1998) The runners-up. *Science* **282**, 2156–2157.
2. (1999) *Nat. Genet.* **21**(Suppl.).
3. Lysov, Y., Chernyi, A., Balaeff, A., Beattie, K., and Mirzabekov, A. (1994) DNA sequencing by hybridization to oligonucleotide matrix. Calculation of continuous stacking hybridization efficiency. *J. Biomol. Struct. Dyn.* **11**, 797–812.
4. Khrapko, K., Lysov, Y., Khorlin, A., Ivanov, I., Yershov, G., Vasilenko, S., Florentiev, V., and Mirzabekov, A. (1991) A method for DNA sequencing by hybridization with oligonucleotide matrix. *DNA Sequence* **1**, 375–388.
5. Khrapko, K., Lysov, Y., Khorlin, A., Shick, V., Florentiev, V., and Mirzabekov, A. (1989) An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett.* **256**, 118–122.
6. Guschin, D., Yershov, G., Zaslavsky, A., Gemmell, A., Shick, V., Proudnikov, D., Arenkov, P., and Mirzabekov, A. (1997) Manual manufacturing of oligonucleotide, DNA and protein microchips. *Anal. Biochem.* **250**, 203–211.
7. Vasiliskov, A. V., Timofeev, E. N., Surzhikov, S. A., Drobyshev, A. L., Shick, V. V., and Mirzabekov, A. D. (1999) Fabrication of microarray of gel-immobilized compounds on a chip by copolymerization. *Biotechniques* **27**, 592–606.
8. Proudnikov, D., Timofeev, E., and Mirzabekov, A. (1998) Immobilization of DNA in polyacrylamide gel for manufacturing of DNA and DNA-oligonucleotide microchips. *Anal. Biochem.* **259**, 34–41.
9. Timofeev, E., Kochetkova, S., Mirzabekov, A., and Florentiev, V. (1996) Regioselective immobilization of short oligonucleotides to acrylic copolymer gels. *Nucleic Acids Res.* **24**, 3142–3148.
10. Proudnikov, D. and Mirzabekov, A. (1996) Chemical methods of DNA and RNA fluorescent labeling. *Nucleic Acids Res.* **24**, 4535–4542.
11. Yershov, G., Barsky, V., Belgovskiy, A., Kirillov, E., Kreindlin, E., Ivanov, I., Parinov, S., Guschin, D., Drobyshev, A., Dubiley, S., and Mirzabekov, A. (1996)

- DNA analysis and diagnostics on oligonucleotide microchips. *Proc. Natl. Acad. Sci. USA* **93**, 4913–4918.
12. Boturn, D., Defancq, E., Ducros, V., Fontaine, C., and Lhomme, J. (1997) Quantitative one step derivatization of oligonucleotides by a fluorescent label through abasic site formation. *Nucleosides Nucleotides* **16**, 2069–2077.
  13. Livshits, M., Florentiev, V., and Mirzabekov, A. (1994) Dissociation of duplexes formed by hybridization of DNA with gel-immobilized oligonucleotides. *J. Biomol. Struct. Dyn.* **11**, 783–795.
  14. Livshits, M. and Mirzabekov, A. (1996) Theoretical analysis of the kinetics of DNA hybridization with gel-immobilized oligonucleotides. *Biophys. J.* **71**, 2795–2801.
  15. Mirzabekov, A. (1994) DNA sequencing by hybridization—a megasequencing method and a diagnostic tool? *Trends Biotech.* **12**, 27–32.
  16. Fotin, A. V., Drobyshv, A. L., Proudnikov, D. Y., Perov, A. N., and Mirzabekov, A. D. (1998) Parallel thermodynamic analysis of duplexes on oligodeoxyribonucleotide microchips. *Nucleic Acids Res.* **26**, 1515–1521.
  17. Mirzabekov, A. D. (1999) *Biological Agent Detection and Identification DARPA Conference Proceedings*, April 27–30, Santa Fe, NM, pp. 57–67.
  18. Marshall, A. and Hodgson, J. (1998) DNA chips: An array of possibilities. *Nature Biotech.* **16**, 27–31.
  19. Syvanen, A. C., Aalto-Setälä, K., Harju, L., Kontula, K., and Soderlund, H. (1990) A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **8**, 684–692.
  20. Dubiley, S., Kirillov, E., and Mirzabekov, A. (1999) Polymorphism analysis and gene detection by minisequencing on an array of gel-immobilized primers. *Nucleic Acids Res.* **27**, e19.
  21. Drobyshv, A. L., Zasedatelev, A. S., Yershov, G. M., and Mirzabekov, A. D. (1999) Massive parallel analysis of DNA Hoechst 33258 binding specificity with a generic oligodeoxyribonucleotide microchip. *Nucleic Acids Res.* **27**, 4100–4105.
  22. Drobyshv, A., Mologina, N., Shick, V., Pobedimskaya, D., Yershov, G., and Mirzabekov, A. (1997) Sequence analysis by hybridization with oligonucleotide microchip: identification of beta-thalassemia mutations. *Gene* **188**, 45–52.
  23. Dubiley, S., Kirillov, E., Lysov, Y., and Mirzabekov, A. (1997) Fractionation, phosphorylation, and ligation on oligonucleotide microchips to enhance sequencing by hybridization. *Nucleic Acids Res.* **25**, 2259–2265.
  24. Shik, V. V., Lebed, Y. B., and Kryukov, G. V. (1998) Identification of HLA DQA1 alleles by the oligonucleotide microchip method. *Mol. Biol.* **32**, 679–688.
  25. LaForge, K. S., Shick, V., Sprangler, R., Proudnikov, D., Yuferov, V., Lysov, Y., Mirzabekov, A., and Kreek, M. J. (2000) Detection of single nucleotide polymorphisms of the human mu opioid receptor gene by hybridization or single nucleotide extension on custom oligonucleotide gelpad microchips: Potential in studies of addiction. *Am. J. Med. Genet.*, in press.
  26. Guschin, D., Mobarry, B., Proudnikov, D., Stahl, D., Rittman, B., and Mirzabekov, A. (1997) Oligonucleotide microchips as genosensors for determi-

- native and environmental studies in microbiology. *Appl. Environ. Microbiol.* **63**, 2397–2402.
27. Bains, W. and Smith, G.C. (1988) A novel method for nucleic acid sequence determination. *J. Theor. Biol.* **135**, 303–307.
  28. Drmanac, R., Labat, I., Brukner, I., and Crkvenjakov, R. (1989) Sequencing of megabase plus DNA by hybridization: Theory of the method. *Genomics* **4**, 114–128.
  29. Lysov, Y., Florentiev, V., Khorlin, A., Khrapko, K., Shick, V., and Mirzabekov, A. (1988) A new method to determine the nucleotide sequence by hybridizing DNA with oligonucleotides. *Proc. Natl. Acad. Sci. USSR* **303**, 1508–1511.
  30. Broude, N. E., Sano, T., Smith, C. L., and Cantor, C. R. (1994) Enhanced DNA sequencing by hybridization. *Proc. Natl. Acad. Sci. USA* **91**, 3072–3076.
  31. Gunderson, K. L., Huang, X. C., Morris, M. S., Lipshutz, R. J., Lockhart, D. J., and Chee, M. S. (1998) Mutation detection by ligation to complete *n*-mer DNA arrays. *Genome Res.* **8**, 1142–1153.
  32. Parinov, S., Barsky, V., Yershov, G., Kirillov, E., Timofeev, E., Belgovskiy, A., and Mirzabekov, A. (1996) DNA sequencing by hybridization to microchip octa- and decanucleotides extended by stacked pentanucleotides. *Nucleic Acids Res.* **24**, 2998–3004.
  33. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.
  34. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14**, 457–460.
  35. Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J. M. (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* **21**, 10–14.
  36. Arenkov, P., Kukhtin, A., Gemmell, A., Voloschuk, S., Chupeeva, V., and Mirzabekov, A. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **278**, 123–131.
  37. Righetti, P. G., Brost, B. C. W., and Snyder, R. S. (1981) On the limiting pore size of hydrophilic hydrogels for electrophoresis and isoelectric focusing. *J. Biochem. Biophys. Methods* **4**, 347–363.
  38. O'Connell, P. B. H. and Brady, C. J. (1976) Polyacrylamide gels with modified cross-linkages. *Anal. Biochem.* **76**, 63–73.
  39. Hermanson, G. T. (1996) *Bioconjugate Techniques*, Academic, San Diego, CA.



<http://www.springer.com/978-0-89603-822-6>

DNA Arrays

Methods and Protocols

Rampal, J.B. (Ed.)

2001, XI, 264 p., Hardcover

ISBN: 978-0-89603-822-6

A product of Humana Press