© 2001 Oxford University Press

DNA microarrays with stem–loop DNA probes: preparation and applications

Natalia E. Broude[1,*], Karen Woodward[2], Robert Cavallo[2], Charles R. Cantor[1,3] and David Englert[4]

[1]Center for Advanced Biotechnology and Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA, [2]Packard BioChip Technologies, LLC Meriden, CT 06450, USA, [3]Sequenom Inc., San Diego, CA 92121, USA and [4]Packard BioScience Inc., Meriden, CT 06450, USA

Received May 29, 2001; Revised July 15, 2001; Accepted August 1, 2001

ABSTRACT

We have developed DNA microarrays containing stem–loop DNA probes with short single-stranded overhangs immobilized on a Packard HydroGel chip, a 3-dimensional porous gel substrate. Microarrays were fabricated by immobilizing self-complementary single-stranded oligonucleotides, which adopt a partially duplex structure upon denaturing and re-annealing. Hybridization of single-stranded DNA targets to such arrays is enhanced by contiguous stacking interactions with stem–loop probes and is highly sequence specific. Subsequent enzymatic ligation of the targets to the probes followed by stringent washing further enhances the mismatched base discrimination. We demonstrate here that these microarrays provide excellent specificity with signal-to-background ratios of from 10- to 300-fold. In a comparative study, we demonstrated that HydroGel arrays display 10–30 times higher hybridization signals than some solid surface DNA microarrays. Using Sanger sequencing reactions, we have also developed a method for preparing nested 3′-deletion sets from a target and evaluated the use of stem–loop DNA arrays for detecting *p53* mutations in the deletion set. The stem–loop DNA array format is simple, robust and flexible in design, thus it is potentially useful in various DNA diagnostic tests.

INTRODUCTION

The last several years have witnessed remarkable development of new microarray-based automated techniques allowing parallel analysis of multiple DNA samples. Two major current applications of DNA microarrays are gene expression profiling (1,2) and gene mutation analysis (3,4). Gene expression profiling using DNA microarrays has already provided results that were not achievable a few years ago. For example, it allowed the molecular classification of leukemias (5) and other types of cancers (6) and revelations about genetic network architecture (7).

At the same time, mutation studies with DNA microarrays are still at an early stage. Among the variety of genomic mutations, single nucleotide polymorphisms (SNPs) are the most suitable targets for DNA microarray analysis, while multiple mutations, insertions, deletions and rearrangements are more challenging (4,8). However, even SNPs are detected by microarrays with relatively poor accuracy. Therefore, an independent verification of new sequence variants is necessary (4,9). Different approaches have been suggested to improve the accuracy of high throughput SNP genotyping (for reviews see 4,10). For example, generic microarrays with a pre-selected set of oligonucleotides and two color detection of tag sequences, introduced into targets during single base extension reactions, have been suggested (11). Generic microarrays and zip code hybridization were also used in a method based on ligation discrimination of allele variants (12,13).

Another problem with the use of DNA microarrays for mutation analysis is their cost and limited availability. Flexibly designed and inexpensive DNA microarrays are in great demand for robust and high throughput mutation analysis.

In earlier studies, positional sequencing by hybridization (PSBH) showed excellent discrimination of matched versus mismatched targets (14–16). Two major features distinguish PSBH from conventional hybridization methods. (i) Partially duplex DNA probes with short single-stranded overhangs are used for hybridization instead of traditional totally single-stranded probes. (ii) Two enzymatic reactions, ligation and primer extension, can be used to further enhance discrimination. Recently, the PSBH format has been applied in a large-scale mutation analysis using high density DNA microarrays (17) that were generated by solid surface chemical synthesis. Single-stranded oligonucleotides were made partly double-stranded by hybridization with a generic oligonucleotide complementary to a common portion of each probe immobilized on the surface. By ligation of the DNA targets to a complete 8mer duplex probe array, Gunderson *et al.* (17) analyzed complex targets up to 2.5 kb with >90% accuracy. Re-sequencing using 9mer arrays allowed reading of up to 1.2 kb of target with 99.5–99.9% accuracy.

Partially duplex DNA arrays, however, have not yet seen wide application. One of the reasons hindering their application is the lack of efficient and practical methods of preparation

*To whom correspondence should be addressed at: Center for Advanced Biotechnology, 36 Cummington Street, Boston, MA 02215, USA.
Tel: +1 617 353 8490; Fax: +1 617 353 8501; Email: nebroude@hotmail.com

Table 1. Oligonucleotides used in this study

##	Sequence (5′-3′)	Application
1	GATGAATTCC GAGCCACGC**T** GGCGTGGCTC GGAATTCATC AGCTCG	Self-complementary oligos with a 6 base single-stranded overhangs and an oligo target complementary to the overhangs. Aminated dT for immobilization is in bold. (see Figs. 2, 3)
2	GACGAATTCC ACACGC**T**GGC GTGTGGAATT CGTC AGCTCG	
3	GACGAATTCC ACACGC**T**GCG TGTGGAATTC GTC AGCTCG	
4	GCTGAATTCC GCC**T**GGGCGG AATTCAGC AGCTCG	
5	GCTGAATTCC GCC**T**GGCGGA ATTCAGC AGCTCG	
6	CGCCGTCCTATTACGAGCT	
7	GCTGAATTCC GCC**T**GGCGGA ATTCAGC GGCAGC	Self-complementary DNA probes used in the experiments with *p53* targets. (see Fig. 4 and 5)
8	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TGGGG	
9	GCTGAATTCC GCC**T**GGCGGA ATTCAGC CCTGG	
10	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TCCCT	
11	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TGCTCC	
12	GCTGAATTCC GCC**T**GGCGGA ATTCAGC CTTAGT	
13	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TCGCTT	
14	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TGCTCG	
15	GCTGAATTCC GCC**T**GGCGGA ATTCAGC GCAGT	
16	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TGGGC	
17	GCTGAATTCC GCC**T**GGCGGA ATTCAGC AGCTCG	
18	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TTTTT	
19	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TATAGC	
20	GCTGAATTCC GCC**T**GGCGGA ATTCAGC CAGCAG	
21	GCTGAATTCC GCC**T**GGCGGA ATTCAGC CAGTT	
22	GCTGAATTCC GCC**T**GGCGGA ATTCAGC GTTAA	
23	GCTGAATTCC GCC**T**GGCGGA ATTCAGC TGGGT	
24	CTCACCACGAGCTGCCCCC	PCR primers used to amplify wild and mutant *p53* fragments (mutations are in bold)
25	TTGGGCAGTGCTCGCTTAGTG	
26	CTCACCACGAGCTGC**A**CCC	
27	TTG**T**GCAGTGCTCGCTTAGTG	
28	GGAGCATTGACAGGA CGAGCT	Oligonucleotide targets used in ligation experiments with HydroGel arrays in single-stranded form and as partially duplex targets when annealed with oligo 39 (see Table 2 and Fig. 6)
29	GGAGCATTGACAGGA CCCCA	
30	GGAGCATTGACAGGA GGAGCA	
31	GGAGCATTGACAGGA ACTAAG	
32	GGAGCATTGACAGGA AAGCGA	
33	GGAGCATTGACAGGA CGAGCA	
34	GGAGCATTGACAGGA ACTGC	
35	GGAGCATTGACAGGA GCTATA	
36	GGAGCATTGACAGGA CTGCTG	
37	GGAGCATTGACAGGA AACTG	
38	GGAGCATTGACAGGA TTAAC	
39	TCCTGTCAATGCTCC	

and immobilization of duplex probes. Also lacking are methods for generating targets compatible with the partially duplex DNA arrays. Here we describe a method for preparation of stem–loop DNA microarrays on a hydrogel substrate (HydroGel chips). Microarrays are fabricated by immobilizing pre-synthesized self-complementary single-stranded oligonucleotides which adopt partially duplex stem–loop structures upon denaturing and re-annealing (Fig. 1; 18). This method is rather simple, it provides excellent specificity and allows the preparation of large numbers of identical DNA microarrays that are easily adaptable to different applications. We have also developed a method of generating targets compatible with partially duplex DNA arrays for mutation detection.

MATERIALS AND METHODS

Oligonucleotides

Non-phosphorylated, unlabeled or fluorescently labeled oligonucleotides were purchased as custom synthesis products from Integrated DNA Technologies Inc. (Coralville, IA). Phosphorylated self-complementary oligonucleotides with an internal dT with a primary NH_2 group for immobilization (amino modifier C6) were purchased from Operon Technologies Inc. (Alameda, CA). A list of oligonucleotides used in this study is presented in Table 1.

Preparation of diisothiocyanate glass substrates

Aminosilane-derivatized glass slides (Sigma) were converted to diisothiocyanate derivatives by a modification of the method of Guo *et al.* (19). Aminosilane slides were incubated for 2 h in a 0.2% solution of 4-phenylene diisothiocyanate (Aldrich) in pyridine/dimethyl formamide (1:9 v/v). Each slide was washed twice with 15 ml of methanol, twice with 15 ml of acetone, dried under nitrogen and stored in a dessicator.

Printing arrays using the Packard BioChip Arrayer I

DNA probes were delivered to the microarray substrates in a single droplet (average volume 350 pl) using the Packard BioChip Arrayer I, a non-contact piezoelectric microdispenser. It delivers minute and consistent volumes generating uniform, homogeneous spots, 200 µm in diameter at a 300 µm to 1 mm

pitch, and is especially well suited for printing arrays on either planar glass or porous substrates. Arrays were printed on glass substrates or on Packard HydroGel chips.

Printing partially duplex DNA probe arrays

HydroGel chips (Packard BioChip Technologies, Meriden, CT) were activated for immobilization of amine-derivatized oligonucleotides by soaking the slides in a 2% solution of trifluoroacetic acid (Fisher Scientific) for 10 min on a shaking platform. The slides were rinsed three times in distilled water, soaked in water for 5 min and dried by low speed centrifugation (<500 *g*) in a table top centrifuge with a 45° angle rotor. Aldehyde glass slides were purchased from CEL Associates (Houston, TX) or TeleChem International Inc. (Sunnyvale, CA). 3D Link microarray slides were purchased from Surmodics Inc. (Sun Prairie, MN).

DNA probes were dispensed onto all slides in triplicate or quadruplicate from a 384-well source plate containing 20–200 µM oligonucleotide solutions in 50 mM borate buffer, pH 9.0. After dispensing on HydroGels, the Schiff bases formed between the aldehyde groups of HydroGel and the amino groups of oligonucleotide probes were reduced by a 30 min soak in a 100 mM solution of sodium borohydride (Aldrich). Slides were washed for 30 min in water, then transferred into a solution containing 10 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$ and 100 mM NaCl, heated to 60°C, then slowly cooled to room temperature to allow annealing of the partially duplex oligonucleotides.

Diisothiocyanate slides with the dispensed probes were washed in borate buffer, pH 9.0, containing 100 mM glycine for 30 min, then with three changes of distilled water and dried by centrifugation. TeleChem, Cel Associates and Surmodics slides were processed according to protocols recommended by the manufacturers.

Oligonucleotide probes were covalently immobilized onto each surface at two concentrations (35 and 3.5 fmol per 350 pl droplet). The arrayed probes were then hybridized and ligated with a complementary DNA target, which was either radioactively (^{32}P) or fluorescently (Texas red) labeled. Both detection methods generated consistent data.

Testing stem–loop DNA probes in solution

Each self-complementary oligonucleotide probe was heat denatured and re-annealed by slow cooling in *Eco*RI buffer. A re-annealed aliquot of each probe was either digested with *Eco*RI restriction enzyme or ligated with a complementary 10 nt target. All reactions were analyzed on a 10% non-denaturing polyacrylamide gel. The data were consistent with formation of stem–loop structures predicted for the self-complementary probes (not shown).

Preparation of 3′-nested DNA targets compatible with partially duplex DNA arrays

Sequencing ladders were prepared from the double-stranded plasmid, single-stranded M13mp18 DNA and from an 800 bp *p53* gene PCR fragment (all from an AutoRead sequencing kit; Amersham Pharmacia Biotech). Both isothermal sequencing using an AutoRead sequencing kit and cycle sequencing using a fmole DNA cycle sequencing kit (Promega, Madison, WI) were performed to generate double-stranded and single-stranded DNA targets, respectively. The sequencing products were treated with exonuclease III (exo III) (100 U/µl; New England Biolabs), that was serially diluted in 10-fold increments in the buffer supplied for exo III. An aliquot of 1 µl of each dilution was used for digestion of the sequencing ladders at 37°C for 5 min. The reaction was carried out in the presence of single-strand binding protein (SSB) (2.2 µg/µl; Promega) to minimize the influence of secondary structure. The reaction was terminated by adding stop solution containing 6 mg/ml Dextran blue in deionized formamide, denatured at 80–90°C for 3 min, loaded onto a 6% polyacrylamide gel and analyzed on an ALF sequencing instrument (Amersham Pharmacia Biotech).

Hybridization/ligation experiments

Hybridization and ligation of the targets on HydroGel arrays was conducted in a 50 µl final volume. Each 50 µl solution contained 50 mM Tris–HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 50 µg/ml BSA, 100 mM NaCl, 0.1% Triton X-100, 3–5 U T4 DNA ligase (Promega) and 10 pmol ^{32}P-labeled oligonucleotide target (20 000 c.p.m./pmol). In the experiments with 3′-nested targets, 50 ng 46 bp *p53* fragments were used in cycle sequencing reactions and then treated with 1 µl of exo III (0.1 U/µl) in the presence of 2–4 µg SSB (Promega) as described above. DNA samples were purified using a nucleotide purification kit (Qiagen) and labeled with [γ-^{32}P]ATP. The labeled sequencing ladders were purified from excess radioactivity by passing through Qiagen columns and eluted with 50 µl of deionized water. These DNA samples were supplemented with all components necessary for the ligation reaction and applied to the array. Adhesive-backed frames (MJ Research) were fastened to a glass slide containing two 12 × 12 mm HydroGel arrays with immobilized DNA probes. An aliquot of 50 µl of the hybridization/ligation solution was spread over the entire surface of the array by placing and sealing a flexible plastic coverslip (MJ Research) to the slide. The sealed arrays were placed in 50 ml capped tubes (Falcon) and rotated overnight in a hybridization oven (PersonalHyb; Stratagene) at ambient temperature (3–5 h ligation produced similar results). After ligation, the arrays were washed three times for 30 min each with 50 ml of 1× SSPE buffer at room temperature or 37°C and dried in air. They were wrapped in Saran wrap and exposed to a Packard SR Cyclone screen for from 3 h to 5 days, depending on the experiment. The images were generated with a Storm PhosphorImager (Molecular Dynamics, Sunnyvale, CA) or a Cyclone Imager (Packard Instrument Co., Downers Grove, IL) at 50 µm resolution. The images were analyzed by software provided by Molecular Dynamics.

RESULTS

Stem–loop DNA probes immobilized on HydroGel provide a higher signal than when they are immobilized on some solid surfaces

Figure 1 outlines the array design which we utilized in this study. Self-complementary oligonucleotides were designed to form partially duplex DNA probes with single-stranded overhangs upon denaturing and re-annealing. Stem–loop oligonucleotides contained a stem-located *Eco*RI site and a T residue with an NH_2 group within the loop for immobilization

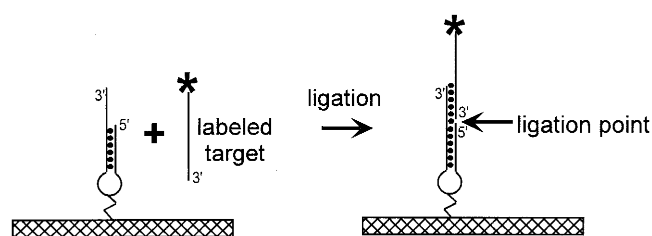

Figure 1. Outline of the experiments with stem–loop DNA probes immobilized on a surface.

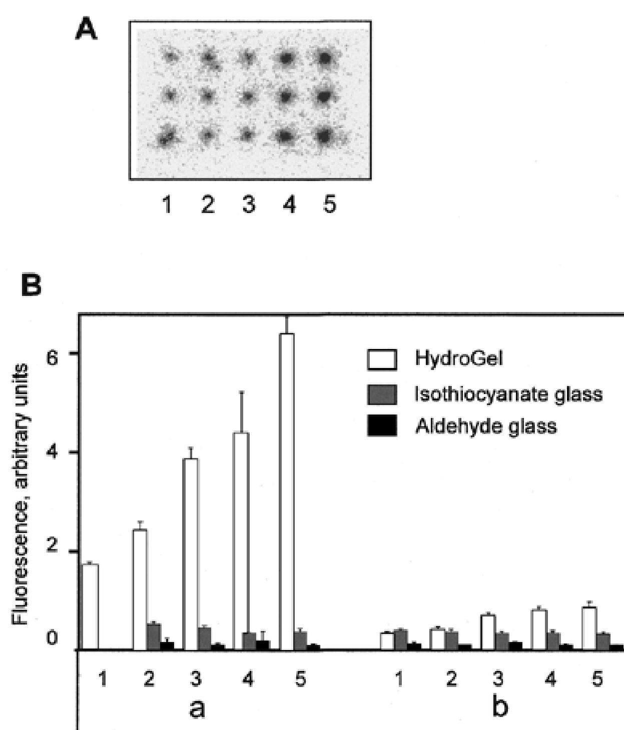

Figure 2. Hybridization/ligation efficiency of a target with stem–loop DNA probes immobilized on different surfaces. Five oligonucleotide probes (1–5, Table 1) were immobilized on different surfaces and ligated with 10 pmol labeled 19 nt long target CGCCGTCCTATTACGAGCT. (**A**) Representative image obtained on a PhosphorImager with 35 fmol each probe immobilized on HydroGel. The target was ^{32}P-labeled (20 000 c.p.m./pmol). (**B**) Oligonucleotide probes were immobilized on HydroGel, isothiocyanate glass or aldehyde glass at 35 (a) and 3.5 (b) fmol concentration. Ligation with 10 pmol of the same19 base long target as in (A), but 5′-labeled with Texas red.

(Table 1). To find the optimal design we varied the length and composition of the stem (Table 1). When tested in solution upon denaturing and annealing, all oligonucleotides displayed a behavior fully consistent with formation of the stem–loop structures predicted for the self-complementary probes (data not shown).

Next, five aminated DNA probes were dispensed and immobilized to three different surfaces: (i) HydroGel chips; (ii) commercially purchased aminosilane glass slides chemically modified with 1,4-phenylene diisothiocyanate; and (iii) commercially available aldehyde glass slides. These probes were denatured, re-annealed and hybridized and ligated with complementary DNA targets. Figure 2A shows a representative experiment with ^{32}P-labeled target and Figure 2B summarizes fluorescence data obtained with different surfaces.

One of the first observations noted was that the enzymatic ligation step provided a substantial increase in the signal level using stem–loop DNA probes with 5–6 nt single-stranded overhangs (data not shown). The second observation was that under the given set of reaction conditions, the net signals achievable on the HydroGel arrays were higher than on either of the glass substrates tested. At the lower concentration of probe, the net signal on the HydroGels was ~2-fold higher (Fig. 2Bb). However, at the higher probe concentrations the HydroGel ligation signal increased, while the signal on the two planar glass substrates did not (Fig. 2Ba). As a result, the signal on Hydrogel arrays was 10–15-fold higher than on the diisothiocyanate surface and 20–30-fold higher than on the aldehyde surface (Fig. 2B). It was also noted that probes 4 and 5 consistently performed better than the other three probes (Fig. 2A and B). Probes 4 and 5 have the shortest double-stranded stem, however, they have a GC-rich cluster adjacent to the loop, which is absent in probes 1–3 (see Table 1). We hypothesize that this GC-rich cluster nucleates formation of the stem and thus contributes to the stability of the stem–loop structure, which in turn enhances hybridization and subsequent target ligation. We used the duplex portion of probe 5 to design new probes in all subsequent studies.

Stem–loop DNA arrays provide excellent reproducibility

To assess the performance of the stem–loop arrays, we estimated the variability of the signal within the arrays in the ligation experiments described above. The coefficient of variation within the arrays was <5%. To test the reproducibility of the entire assay, nested *p53* targets (see below) were prepared and labeled in different experiments and hybridization/ligation experiments were carried out with four different *p53* samples: wild-type and three different *p53* mutants. A high correlation between the hybridization signals of replica arrays was obtained for all DNA targets tested (r^2 = 0.98), which is well within standards needed in clinical diagnostics.

Specificity of target ligation using stem–loop DNA microarrays

To evaluate the specificity of the probe–target interaction on the HydroGel chip arrays we performed experiments with 10 different stem–loop probes immobilized on each array. Arrays were hybridized and ligated with a single oligonucleotide target that was Cy5- or fluorescein-labeled. Each target was complementary to the overhang of one of the probes and had one or several mismatches with others. The signal of the perfectly matched probe was compared with the average signal of the mismatched probes and this ratio was considered as a target-to-noise ratio (T/N). The results show that T/N varies for different probes from 10- to 267-fold (Table 2). As expected, on average the T/N was higher for the GC-rich 6 base overhangs. The data showed very little cross-hybridization between closely related sequences. For example, probes 2 and 5 differ by a single base at the ligation site. The ligation signal for matched target 5 was 10-fold higher than that for mismatched target 2 after hybridization with probe 5. Accordingly, target 5 hybridized with probe 2 20-fold less efficiently than the perfectly matched target 2. These data demonstrate the excellent specificity of the stem–loop DNA microarrays.

Table 2. Specificity of ligation reaction using HydroGel arrays with stem–loop DNA probes[a]

	Probe overhang sequence (5′→3′)	T/N[b]	SD
1	TGGGG	21	8.3
2	TGCTCC	37	15.2
3	CTTAGT	10	1.4
4	TCGCTT	120	20.4
5	TGCTCG	139	37.3
6	GCAGT	27	35.5
7	TATAGC	36	9.2
8	CAGCAG	127	31.4
9	CAGTT	37	17.3
10	AGCTCG	267	125.5

[a]T/N was calculated in the experiments when each oligonucleotide target was hybridized and ligated to an array of immobilized probes containing one perfect match and the rest of mismatched probes. T/N is signal over average mismatch hybridization.
[b]Each data point is the average of four replicates.

Sanger sequencing ladders as DNA targets for stem–loop arrays

Partially duplex DNA probes interrogate only a terminal target sequence (Fig. 1). Therefore, to make the entire length of the target accessible to hybridization, one must create a set of nested deletion targets. In our design, the polarity of the single-stranded overhang in immobilized DNA probes requires nesting of the 3′-ends of the target. We generated 3′-nested DNA targets using Sanger sequencing reactions, which provide a set of fragments with single base resolution. This approach, however, requires removal of the 3′-terminal dideoxynucleotide (ddN) to permit subsequent enzymatic ligation. For this purpose we used exo III treatment. Exo III from *Escherichia coli* is an enzyme that removes mononucleotides stepwise from the 3′-blunt or 3′-recessed termini of duplex DNA (New England Biolabs Catalog, 2000). The enzyme is believed to be inactive with single-stranded DNAs (20; New England Biolabs Catalog, 2000), although there are data indicating that exo III can digest single-stranded templates with low levels of activity (21,22). It was not known whether exo III is capable of removing 3′-terminal ddN residues. Therefore, experiments were performed to answer these questions.

Figure 3A outlines the experimental scheme. After Sanger sequencing reactions, DNA fragments have 3′-terminal ddN residues. If exo III removes ddN residues, the DNA fragments will become shorter and move faster on a denaturing polyacrylamide gel. These experiments were also designed to choose the optimal concentration of exo III, which would lead to minimal shortening of DNA fragments and permit reading of the sequence.

Figure 3B shows a sequencing gel obtained with plasmid DNA that was isothermally sequenced and digested with different concentrations of exo III before PAGE analysis. The results show that decreasing concentrations of exo III correlate with increasing length of the DNA fragments (compare clones 2–4) and that homologous DNA fragments treated with exo III have faster mobility than their untreated counterparts (marked by arrows). This indicates that exo III digests the DNA Sanger ladders. Figure 3C presents the sequencing reads obtained both before and after exo III digestion. Digestion of DNA with two different concentrations of exo III allowed correct reading of the sequence, which means that the set of DNA fragments was digested relatively uniformly by exo III. Similar results were obtained with the products of cycle sequencing, which are mostly single-stranded (23) (data not shown).

Stem–loop DNA arrays detect mutations in *p53* DNA targets

A 46 nt long fragment of the *p53* gene as well as three mutants of the same fragment were generated by PCR. All PCR fragments were subjected to Sanger sequencing reactions followed by exo III digestion and 5′-labeling.

Sixteen stem–loop probes were dispensed and immobilized on HydroGel arrays. These probes consisted of the duplex portion of probe 5 (see Table 1) and contained 5 or 6 base single-stranded overhangs. Eleven overhangs were complementary to different sequences of the 46 nt wild-type *p53* fragment and five overhangs did not contain complements in the *p53* fragment (see Table 1 for overhang sequences and Fig. 4E and F for overhang maps). These arrays were hybridized and ligated with the ^{32}P-labeled sequencing ladders prepared from the wild-type and three mutant *p53* fragments.

Figure 4A shows that seven positive probes displayed signals with the wild-type *p53* target and none of the negative probes showed cross-hybridization. The probes displaying the strongest positive signals correspond to the shortest (probes 1–4) and to the longest (probe 11) DNA fragments present in the sequencing ladder (see Fig. 4F). A lower signal was obtained with probes 8 and 9 and no signal was visible with probes 5–7 and 10. Hybridization patterns of mutant 1 (Fig. 4B), mutant 2 (Fig. 4C) and double mutant *p53* (Fig. 4D) targets were entirely consistent with the mutations, allowing unambiguous mutation detection. Specifically, with mutant 1, signals from probes 2 and 3 were reduced as compared with the wild-type variant. With mutant 2, signal from probe 11 was substantially reduced. With the double mutant, the pattern was a superimposition of the patterns for mutants 1 and 2 (Fig. 4). These data are completely consistent with the expected results.

To address identification of the heterozygous mutations, we performed analysis of 1:1 mixtures of the wild-type and mutant 1 *p53* DNA fragments. In this experiment the arrays contained several marker probes which distinguish the wild-type and mutant sequences. Specifically, probe TGGGG detects only wild-type sequence and probe TGGGT is specific for the mutant sequence. Probe TGGGC is independent of mutation and should detect both wild-type and mutant sequence. Figure 5A shows a representative image, while Figure 5B presents quantitative results. The signals from probes 2 (GGGGT) and 3 (TGGGT) were compared to the signal from probe 1 (TGGGC) with the wild-type and mutant 1 and a 1:1 mixture of wild-type plus mutant 1 *p53* DNA fragments. The results showed that the pattern of the signal from the mixture of wild-type and mutant *p53* fragments presents a superimposition of the patterns for the wild-type and mutant *p53* fragments. These data indicate that the stem–loop arrays are capable of detecting heterozygous mutations.

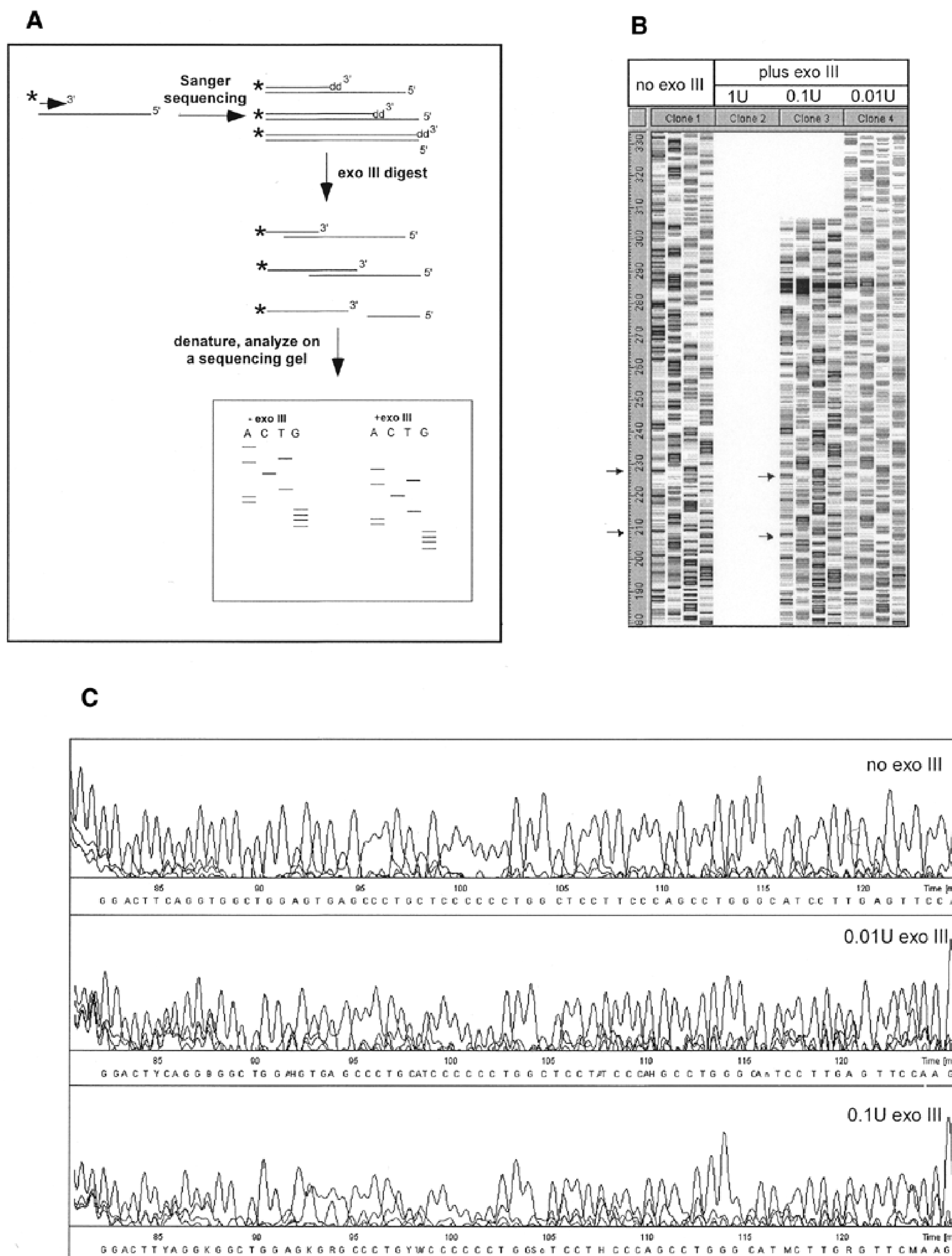

Figure 3. Exonuclease III digests of Sanger sequencing ladders. (**A**) Outline of the experiments. (**B**) Image of the sequencing gel generated by software provided with the ALF sequencing instrument (Pharmacia Biotech, Uppsala, Sweden). Sanger sequencing ladders without and with exo III treatment. Arrows mark matching fragments. (**C**) Sequencing reads obtained before and after exo III digestion.

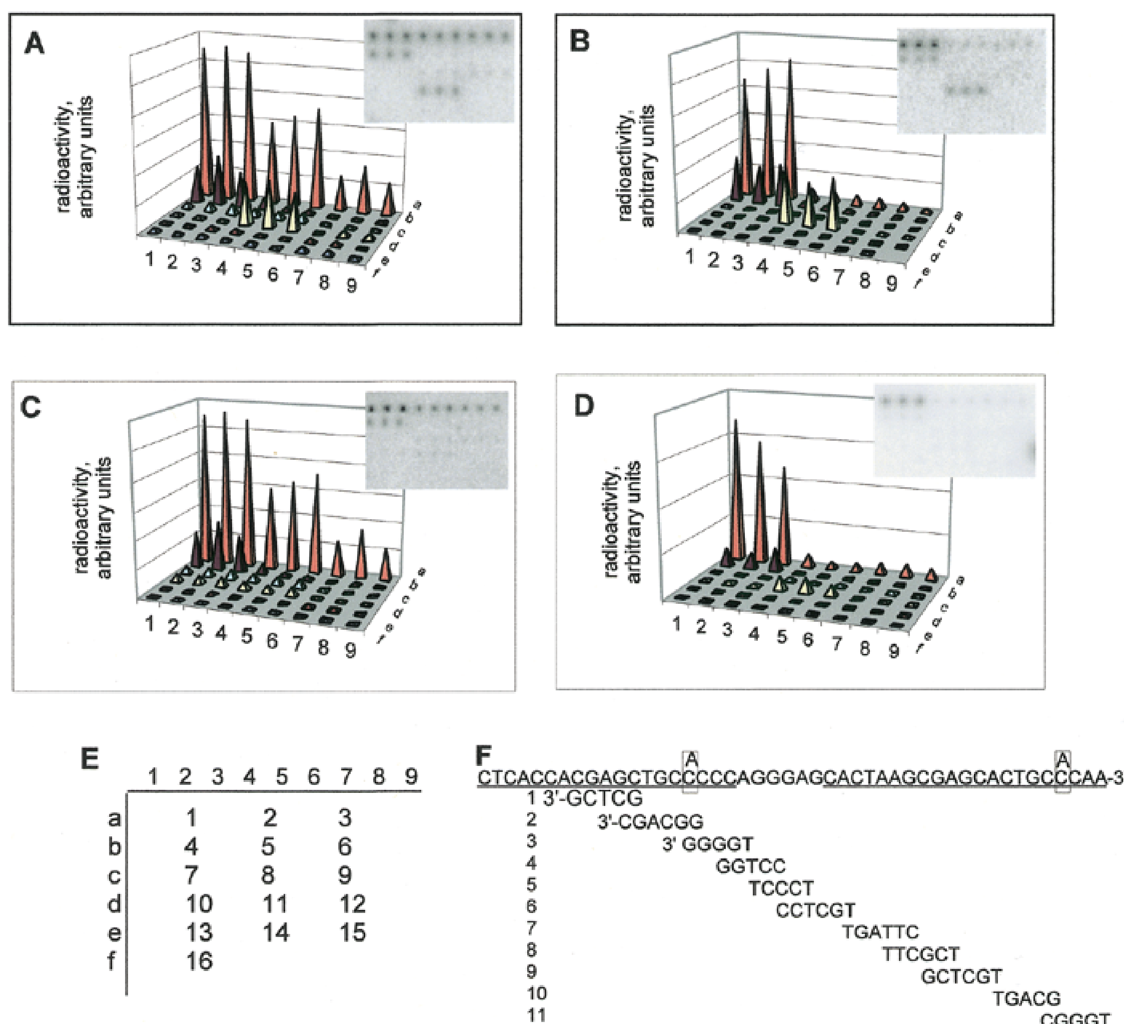

Figure 4. *p53* mutation detection using DNA microarrays with stem–loop DNA probes. A 46 bp fragment of the *p53* wild-type and three mutant genes were used as templates in Sanger sequencing reactions. The products were digested with exo III and labeled with $[\gamma\text{-}^{32}P]$ATP and T4 polynucleotide kinase. The labeled products were hybridized and ligated to stem–loop DNA microarrays containing 11 positive (1–11) and five negative probes (12–16). Each probe was spotted in triplicate. (**A–D**) Representative images (Storm PhosphorImager) obtained with wild-type sequence (CTCACCACGAGCTGCCCCCAAGGGAG-CACTAAGCGAGCACTGCCCAA), mutation 1 (CTCACCACGAGCTGCACCCAAGGGAGCACTAAGCGAGCACTGCCCAA), mutation 2 (CTCACCACG-AGCTGCCCCCAAGGGAGCACTAAGCGAGCACTGCACAA) and double mutation 1 + 2 (CTCACCACGAGCTGCACCCAAGGGAGCACTAAGCGAGCA-CTGCACAA), respectively, (mutations are underlined). Quantitation was done using software provided by Molecular Dynamics. (**E**) Outline of the array design. (**F**) Outline of target sequences and complementary probes. Only sequences of single-stranded overhangs of immobilized DNA probes are shown; the numbers correspond to the numbers of the probes on the array in (E). The negative probes were TTTTT, TATAGC, CAGCAG, CAGTT and GTTAA. The mutated bases are in boxes.

The lack of signal with several probes that were expected to be positive is most likely explained by the lower concentration of the corresponding targets in the sequencing ladder. The gradual decrease in signal intensity from probe 1 to probe 4 (Fig. 4A) correlates with increasing length of the target and supports this explanation. It is known that for short targets it is sometimes difficult to generate Sanger ladders with peaks of uniform intensity (14), though the sequence can be deduced from these data. For longer targets (>80 bp) we had no problems in generating a uniform distribution of nested deletions using our approach (see Fig. 3). We should emphasize, however, that even with several probes being non-informative, it was possible to unambiguously deduce the correct sequences of all targets tested due to the built-in redundancy of the array design.

DISCUSSION

We have developed a new format for DNA microarrays, stem–loop probes with short single-stranded overhangs immobilized on a porous, 3-dimensional substrate, HydroGel, and demonstrated their excellent performance in pilot experiments. Several features of stem–loop DNA probes make them superior to conventional single-stranded probes for target hybridization. First, contiguous stacking interactions between duplex DNA probe and a perfectly matched single-stranded target provide

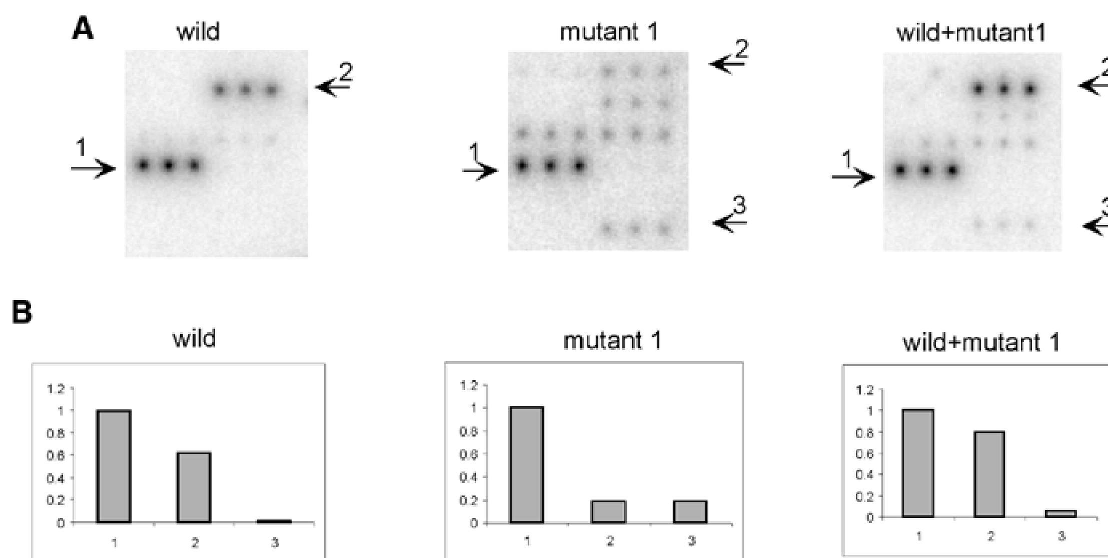

Figure 5. Stem–loop DNA microarrays detect heterozygous mutations. Wild-type and mutant 1 *p53* DNA fragments and a 1:1 mixture were prepared as 3′-nested deletion sets as described above, labeled with ^{32}P at the 5′-end and hybridized with stem–loop DNA arrays containing the following marker probes: 1, TGGGC; 2, TGGGG; 3, TGGGT. (**A**) Images of the arrays. (**B**) Quantitation of signals with probes 1, 2 and 3. Each probe was spotted in triplicate.

additional free energy minimization, increasing the stability of the resulting probe–target complex (18,24–26). This extra stabilization is lacking in mismatched duplexes, therefore match/mismatch discrimination is enhanced. Recently published data on hybridization kinetics of single-stranded targets with stem–loop DNA probes as compared with linear DNA probes clearly demonstrated superiority of the stem–loop probes. Additionally, the stability of the complexes was also higher for the stem–loop probes (27).

Secondly, stem–loop oligonucleotides present an example of structurally constrained DNA probes, for which thermodynamic analysis predicts enhanced specificity of matched target discrimination as a general feature (28). In line with these considerations, Mir and Southern (29) demonstrated that, in addition to linear sequence, higher order structure of nucleic acids is an important determinant of efficient hybridization; duplex formation with oligonucleotide probes was particularly favored at unpaired bases near helical regions of a tRNA target. Recently published data on immobilization on a surface of stem–loop shaped oligonucleotides through multiple anchors also provided experimental evidence of increased efficiency of stem–loop structures in enzymatic reactions as compared with linear oligonucleotides (30).

Thirdly, short single-stranded overhangs of the stem–loop DNA molecules are much more sensitive to mismatches than longer probes. They may also provide better availability of the probe to the target due to decreased probe–surface interaction.

Finally, non-specific binding of targets with duplex DNA is substantially less than with single-stranded DNA. Addition of enzymatic ligation, which covalently links the targets with the stem–loop probes, allows more stringent washing conditions and aids in specificity. Our results indeed showed the exquisite specificity of stem–loop arrays (Table 2), which corroborate our earlier data obtained with partially duplex probes immobilized on magnetic beads (14).

Stem–loop DNA arrays are made from self-complementary oligonucleotides, which fold into duplexes with single-stranded overhangs upon denaturing and re-annealing. Because the probes are formed from a single molecule, this eliminates the need to hybridize a complementary oligonucleotide and covalently stabilize the duplexes (17,31). This method is most compatible with deposition of pre-synthesized oligonucleotides onto a microarray surface. Pre-synthesis of DNA probes allows quality control before the probes are immobilized on the surface, which ensures high reproducibility of the arrays. Therefore, this method is most useful when large numbers of identical arrays are needed. A disadvantage of this method is a practical limitation to the number of the probes that can be printed on the microarray (hundreds or thousands, rather than tens of thousands). However, for many applications, a moderate number of specially designed probes (several hundred) is sufficient to re-sequence rather long sequences (32), and for diagnostic purposes even several probes are needed.

The hybridization signal intensity on DNA microarrays is a crucial factor that limits many applications of this technology. Signal intensity depends on both the concentrations of the target and the immobilized DNA probe and on the availability of both for hybridization. In this study we found that under identical conditions the signal intensities obtained on HydroGel chip arrays were 10–30-fold higher than on some glass surfaces (Fig. 2). Mirzabekov's group has also reported polyacrylamide-immobilized DNA probes that generate higher signals than probes immobilized on a non-porous surface (33). The higher signals observed on the HydroGel chip substrates may be due to a combination of the larger amount of accessible probe and the favorable environment for biomolecular reactions provided by the 3-dimensional hydrophilic matrix. Using oligonucleotides with multiple anchors, which will increase

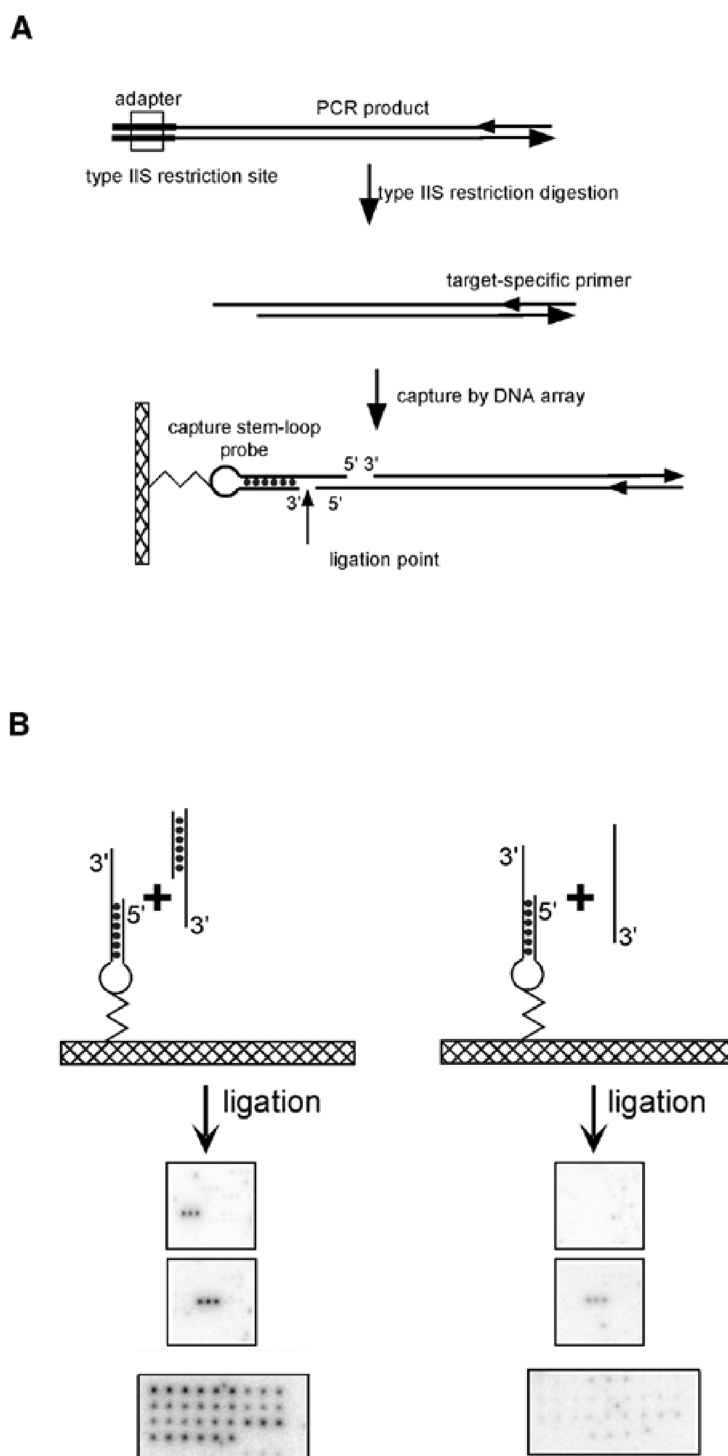

Figure 6. Partially duplex DNA targets ligate to stem–loop arrays with higher efficiency than single-stranded targets. (**A**) Outline of experiments on capturing DNA products from multiplex PCR. (**B**) Comparative hybridization of single-stranded (right) and partially double-stranded (left) targets. In this case each target or target mixture was annealed with oligonucleotide 39, complementary to the common sequence in each target. (Top) Target 33; (middle) target 30; (bottom) equimolar mixture of 11 targets (28–38) (Table 1).

the concentration of the immobilized probe (30), can increase the signal on HydroGel chips further.

We have also developed a method for 3′-nested target preparation compatible with partially duplex DNA arrays, which is based on Sanger sequencing reactions. Our results showed that exo III efficiently removes 3′-ddN residues from Sanger ladders and makes them a convenient substrate for ligation to stem–loop DNA arrays. Although not all DNA probes were informative with the 46 nt long *p53* DNA template, the redundant nature of microarrays allowed correct identification of all

mutations. We believe that this method of target preparation could be a viable method to analyze 100–500 bp long DNA fragments where uniform concentrations of fragments in Sanger sequencing ladders can be generated.

Another possible application of stem–loop DNA probes is sequence-specific capture of DNA fragments (18,27). Both single- and double-stranded DNA targets can be applied for hybridization. For example, such arrays can be used to capture the single-stranded cleavage products from the cleavase invader assay (34,35). An arbitrary tag sequence complementary to the array probe overhangs would be positioned at the 3′-end of the cleavage fragment. Thus the products from a multiplexed invader assay might be captured at different positions on a stem–loop DNA microarray.

Separation and analysis of the double-stranded multiplex PCR products (36–39) could be another application of stem–loop DNA microarrays. By introducing a type IIS restriction site into PCR products, one can create arbitrary protruding single-stranded termini. These gene-specific ends can be used for capture by hybridization with the overhangs of the stem–loop probes (Fig. 6). This idea is completely compatible with the recently developed multiplex PCR method based on PCR suppression (39). This method uses one gene-specific primer per amplicon and one primer common for all targets. It is very convenient to build-in a type IIS restriction site, e.g. for *Hga*I, within this common primer (Fig. 6A). After digestion of the multiplex PCR products with the restriction enzyme each amplicon will have a unique single-stranded overhang, which can be used as an anchor to separate the fragments on the array.

In preliminary experiments we compared signals from single- and double-stranded targets with the same overhang sequence after ligation with the stem–loop DNA arrays (Fig. 6B). The results showed that for all probes tested, double-targets displayed a higher signal and the gain in signal varied from 1.5- to 100-fold. Importantly, match/mismatch discrimination was practically the same for both types of targets.

In conclusion, HydroGel-immobilized stem–loop DNA probes appear to be an attractive alternative to solid surface-immobilized single-stranded DNA microarrays for comparative sequencing. The new technique offers several advantages, such as exquisite specificity, high detection sensitivity, robustness and flexibility in design. We expect that arrays of stem–loop oligonucleotides can be used in a variety of applications.

ACKNOWLEDGEMENTS

The authors are thankful to Vadim Demidov and anonymous reviewers for critical reading of the manuscript and valuable suggestions.

REFERENCES

1. Lockhart,D.J., Dong,H., Byrne,M.C., Follettie,M.T., Gallo,M.V., Chee,M.S., Mittmann,M., Wang,C., Kobayashi,M., Horton,H. and Brown,E.L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.*, **14**, 1675–1680.
2. Wodicka,L., Dong,H., Mittmann,M., Ho,M.H. and Lockhart,D.J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.*, **15**, 1359–1367.
3. Hacia,J.G., Brody,L.C., Chee,M.S., Fodor,S.P. and Collins,F.S. (1996) Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-color fluorescence analysis. *Nature Genet.*, **14**, 441–447.
4. Hacia,J.G. (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nature Genet.*, **21** (suppl. 1), 42–47.
5. Golub,T.R., Slonim,D.K., Tamayo,P., Huard,C., Gaasenbeek,M., Mesirov,J.P., Coller,H., Loh,M.L., Downing,J.R., Caligiuri,M.A., Bloomfield,C.D. and Lander,E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531–537.
6. Brenton,J.D., Aparicio,S.A. and Caldas,C. (2001) Molecular profiling of breast cancer: portraits but not physiognomy. *Breast Cancer Res.*, **3**, 77–80.
7. Tavazoie,S., Hughes,J.D., Campbell,M.J., Cho,R.J. and Church,G.M. (1999) Systematic determination of genetic network architecture. *Nature Genet.*, **22**, 281–285.
8. Chee,M., Young,R., Hubbell,E., Berno,A., Huang,X.C., Stern D., Winkler,J., Lockhart,D.J., Morris,M.S. and Fodor,S.P. (1996) Accessing genetic information with high-density DNA arrays. *Science*, **274**, 610–614.
9. Cargill,M., Altshuler,D., Ireland,J., Sklar,P., Ardlie,K., Patil,N., Lane,C.R., Lim,E.P., Kalayanaraman,N., Nemesh,J., Ziaugra,L., Friedland,L., Rolfe,A., Warrington,J., Lipshutz,R., Daley,G.Q. and Lander,E.S. (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet.*, **22**, 231–238.
10. Whitcombe,D., Newton,C.R. and Little,S.(1998) Advances in approaches to DNA-based diagnostics. *Curr. Opin. Biotechnol.*, **9**, 602–608.
11. Fan,J., Chen,X., Halushka,M.K., Berno,A., Huang,X., Ryder,T., Lipshutz,R.J., Lockhart,D.J. and Chakravarti,A. (2000) Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Res.*, **10**, 853–860.
12. Favis,R., Day,J.P., Gerry,N.P., Phelan,C., Narod,S. and Barany F. (2000) Universal DNA array detection of small insertions and deletions in BRCA1 and BRCA2. *Nat. Biotechnol.*, **18**, 561–564.
13. Favis,R. and Barany,F. (2000) Mutation detection in K-ras, BRCA1, BRCA2 and p53 using PCR/LDR and a universal DNA microarray. *Ann. N. Y. Acad. Sci.*, **906**, 39–43.
14. 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.
15. Fu,D.J., Broude,N.E., Koster,H., Smith,C.L. and Cantor,C.R. (1996) Efficient preparation of short DNA sequence ladders potentially suitable for MALDI-TOF DNA sequencing. *Genet. Anal.*, **12**, 137–142.
16. Fu,D.J., Broude,N.E., Koster,H., Smith,C.L. and Cantor,C.R. (1995) A DNA sequencing strategy that requires only five bases of known terminal sequence for priming. *Proc. Natl Acad. Sci. USA*, **92**, 10162–10166.
17. 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.
18. Lane,M., Benight,A.S., Faldasz,B.D. (1998) US patent no. 5,770,365.
19. Guo,Z., Guilfoyle,R.A., Thiel,A.J., Wang,R. and Smith,L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.*, **22**, 5456–5465.
20. Smith,A.J. (1979) The use of exonuclease III for preparing single stranded DNA for use as a template in the chain terminator sequencing method. *Nucleic Acids Res.*, **6**, 831–848.
21. Okano,K. and Kambara,H. (1995) DNA probe assay based on exonuclease III digestion of probes hybridized on target DNA. *Anal. Biochem.*, **228**, 101–108.
22. Shida,T., Noda,M. and Sekiguchi,J. (1996) Cleavage of single- and double-stranded DNAs containing an abasic residue by *Esherichia coli* exonuclease III (AP endonuclease VI). *Nucleic Acids Res.*, **24**, 4572–4576.
23. Murray,V. (1989) Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nucleic Acids Res.*, **17**, 8889.
24. Lysov,Yu.P., Chernyi,A.A., Balaeff,A.A., Beattie,K.L. and Mirzabekov,A.D. (1994) DNA sequencing by hybridization to oligonucleotide matrix. Calculation of continuous stacking hybridization efficiency. *J. Biomol. Struct. Dyn.*, **11**, 797–812.
25. Lane,M.J., Paner,T., Kashin,I., Faldasz,B.D., Li,B., Gallo,F.J. and Benight,A.S. (1997) The thermodynamic advantage of DNA oligonucleotide ‘stacking hybridization’ reactions: energetics of a DNA nick. *Nucleic Acids Res.*, **25**, 611–616.
26. Vasiliskov,V.A., Prokopenko,D.V. and Mirzabekov,A.D. (2001) Parallel multiplex thermodynamic analysis of coaxial base stacking in DNA

duplexes by oligodeoxyribonucleotide microchips. *Nucleic Acids Res.*, **29**, 2303–2313.
27. Riccelli,P.V., Merante,F., Leung,K.T., Bortolin,S., Zastawny,R.L., Janeczko,R. and Benight,A.S. (2001) Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes. *Nucleic Acids Res.*, **29**, 996–1004.
28. Bonnet,G., Tyagi,S., Libchaber,A. and Kramer,F.R. (1999) Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc. Natl Acad. Sci. USA*, **96**, 6171–6176.
29. Mir,K. and Southern,E. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.*, **17**, 788–792.
30. Zhao,X., Nampalli,S., Serino,A.J. and Kumar,S. (2001) Immobilization of oligodeoxyribonucleotides with multiple anchors to microchips. *Nucleic Acids Res.*, **29**, 955–959.
31. Bulyk,M.L., Gentalen,E., Lockhart,D.J. and Church,G.M (1999) Quantifying DNA-protein interactions by double-stranded DNA arrays. *Nat. Biotechnol.*, **17**, 573–577.
32. Pevzner,P.A., Lysov,Yu.P., Khrapko,K.R., Belyavsky,A.V., Florentiev,V.L. and Mirzabekov,A.D. (1991) Improved chips for sequencing by hybridization. *J. Biomol. Struct. Dyn.*, **9**, 399–410.
33. Fotin,A.V., Drobyshev,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.
34. Kwiatkowski,R.W., Lyamichev,V., de Arruda,M. and Neri,B. (1999) Clinical, genetic and pharmacogenetic applications of the Invader assay. *Mol. Diagn.*, **4**, 353–364.
35. Oldenburg,M.C., Siebert,M. (2000) New Cleavase Fragment Length Polymorphism method improves the mutation detection assay. *Biotechniques*, **28**, 351–357.
36. Chamberlain,J.S., Gibbs,R.A., Ranier,J.E., Nguyen,P.N. and Caskey,C.T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res.*, **16**, 11141–11156.
37. Edwards,M.C. and Gibbs,R.A. (1994) Multiplex PCR: advantages, development and applications. *PCR Methods Appl.*, **3**, S65–S75.
38. Henegariu,O., Heerema,N.A., Dlouhy,S.R., Vance,G.H. and Vogt,P.H. (1997) Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques*, **23**, 504–511.
39. Broude,N.E., Zhang,L., Woodward,K., Englert,D. and Cantor,C.R. (2001) Multiplex allele-specific target amplification based on PCR suppression. *Proc. Natl Acad. Sci. USA*, **98**, 206–211.