2998–3004 *Nucleic Acids Research, 1996, Vol. 24, No. 15*

DNA sequencing by hybridization to microchip octa- and decanucleotides extended by stacked pentanucleotides

Sergei Parinov[1], Victor Barsky[1], Gennady Yershov[1,+], Eugene Kirillov[1], Edward Timofeev[1], Alexander Belgovskiy[1] and Andrei Mirzabekov[1,2,+,*]

Joint Human Genome Program: [1]Engelhardt Institute of Molecular Biology, 32 Vavilov Str., B-334, Moscow 117984, Russia and [2]Center for Mechanistic Biology and Biotechnology, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

Received March 28, 1996; Revised and Accepted June 22, 1996

ABSTRACT

The efficiency of sequencing by hybridization to an oligonucleotide microchip grows with an increase in the number and in the length of the oligonucleotides; however, such increases raise enormously the complexity of the microchip and decrease the accuracy of hybridization. We have been developing the technique of contiguous stacking hybridization (CSH) to circumvent these shortcomings. Stacking interactions between adjacent bases of two oligonucleotides stabilize their contiguous duplex with DNA. The use of such stacking increases the effective length of microchip oligonucleotides, enhances sequencing accuracy and allows the sequencing of longer DNA. The effects of mismatches, base composition, length and other factors on the stacking are evaluated. Contiguous stacking hybridization of DNA with immobilized 8mers and one or two 5mers labeled with two different fluorescent dyes increases the effective length of sequencing oligonucleotides from 8 to 13 and 18 bases, respectively. The incorporation of all four bases or 5-nitroindole as a universal base into different positions of the 5mers permitted a decrease in the number of additional rounds of hybridization. Contiguous stacking hybridization appears to be a promising approach to significantly increasing the efficiency of sequencing by hybridization.

INTRODUCTION

Hybridization of DNA with oligonucleotides has been developed for DNA sequencing [sequencing by hybridization (SbH)] and sequence analysis (1–9). A simple example demonstrates this approach. A 10 base-long single-stranded DNA (ssDNA), AGTCGGAATA, when hybridized with all possible 8mers, gives rise to perfect duplexes upon binding with only three 8mers: 1, TTCCGACT; 2, ATTCCGAC; and 3, TATTCCGA. Overlapping of two 8mers of known sequence, 1 and 2 or 2 and 3, by seven bases allows one to consider them as adjacent in DNA, to arrange them in the proper order, and to reconstruct the sequence of the DNA. Such a hybridization can be carried out with an array (4) or microarray (sequencing microchip) (6–9) of immobilized oligonucleotides; however, a full sequencing microchip for SbH is too complex. An array of 65 536 8mers is needed to sequence DNA up to 200 bases long, and an array of 67 108 864 13mers is needed for sequencing DNA a few thousand bases long. In both cases, only a negligible fraction of the oligonucleotides in the array, 200 or a few thousand respectively, are hybridized with DNA. To overcome this drawback of SbH, we have suggested (1,5) using of an array of short oligonucleotides (8mers) and extending the length of only those that are hybridized to the DNA.

This extension can be carried out by contiguous stacking hybridization without any additional oligonucleotide synthesis. A major factor in DNA duplex stability is stacking interactions between adjacent bases in a DNA strand. The stacking interactions stabilize the DNA duplex even in the absence of a phosphodiester bond or a phosphate group (5–7,10–12). Thus, rather weak duplexes of 5mers with DNA are significantly reinforced by stacking with more stable, adjacent, 8mer duplexes to allow the assembly of a 13 bp (base pairs) duplex lacking phosphodiester bonds between the 8mer and the 5mer (Fig. 1A) (6). The stacked 5mers can also be ligated to longer oligonucleotides or DNA (13).

Sequence analysis and identification of mutations in genomic DNA were performed on sequencing microchips in pilot experiments (7). The use of oligonucleotides immobilized in a gel, with its high immobilization capacity, is crucial to carrying out contiguous stacking hybridization with the quite unstable 5mers. Computer simulations demonstrated that hybridization with 5mers makes an 8mer chip as efficient as a 13mer chip and capable of sequencing DNA a few thousand bases long, rather than 200 bases (14). Stacking interactions have been also used to compose assembled modular primers for DNA polymerase reaction from several shorter oligonucleotides (11,12).

* To whom correspondence should be addressed

+Present address: Center for Mechanistic Biology and Biotechnology, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

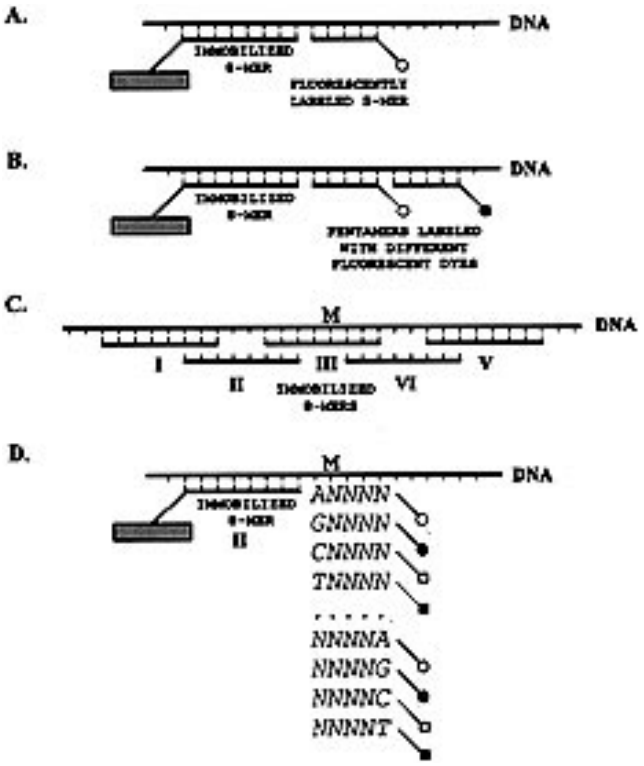

Figure 1. A scheme of CSH of DNA with an immobilized 8mer and one (**A**) or two (**B**) 5mers labeled with different fluorescent dyes. The 5mers do not hybridize to DNA under the chosen conditions; however, 5mer hybridizations were made possible by adjacent 8mer duplexes as a result of stacking interactions between their terminal bases. (**C**) Scheme of a simplified microchip for diagnostics that allows one to localize mismatch M. Immobilized 8mers are overlapped with each other by three bases. Therefore 1000 bases can be covered by only 200 such overlapped 8mers. (**D**) Scheme with 20 5mers, which allows 1024 rounds of hybridization with all possible 5mers to be replaced by only five hybridizations. The 5mers have a meaningful base *(A, G, C* or *T)* in only one discriminating position and have a universal base (or all four bases, *'N'*) in all other positions. Each 5mer is labeled with one of four different dyes specifying each discriminating base. For details, see text.

Contiguous stacking hybridization can also simplify a microchip for the purpose of diagnosis of genetic diseases. The position of a mutation can be located by hybridization of DNA with a chip containing, for example, 8mers having a three-base overlap (Fig. 1C). In the following hybridization rounds with properly stacked 5mers, the mutated base can be reliably identified. To identify any base change within, for example, 1000 bases by using such a procedure, a microchip with only 200 immobilized oligonucleotides is enough (7). Using conventional means, such identification can be carried out with 4000 oligonucleotides covering all possible base changes (15). However, the use of contiguous stacking hybridizations with all possible 1024 5mers to analyze DNA of unknown structure seems cumbersome. An algorithm was developed (14), that minimizes the number of additional hybridization steps by assembling sets of pentamers to be added in each round of hybridization. Another hybridization strategy uses a set of 20 5mers, each containing a 'discriminating' base at one position and a universal base or a mixture of all four bases at all four other positions—or, equivalently, a set of 5mers in which all four bases are represented in all four other positions—and labeled with a different fluorescent compound for each of the four bases on the 'discriminating' position (7). This strategy enables one to decrease the number of CSH rounds to five (Fig. 1D).

In the present investigation, we have tested the effect of different factors, such as presence of mismatches in pentamers and immobilized oligonucleotides and stacking between different bases, on the method. The usefulness of introducing 5-nitroindole (16,17) and all four bases into pentamers to simplify the procedure has been demonstrated. Two rounds of stacking hybridization (14), with two 5mers carrying different labels contiguously stacked to each other and to an immobilized 8mer to form 18-bp contiguous duplex (Fig. 1B), also demonstrated.

MATERIALS AND METHODS

Synthesis of oligonucleotides

Oligonucleotides were synthesized on a 394 DNA/RNA synthesizer (Applied Biosystems Inc.) by using reagents from Applied Biosystems Inc. Oligonucleotides for immobilization were synthesized with 3-methyluridine located on the 3′ end. The 5′-fluorescent pentamers were synthesized by using HEX and FAM amidites (Applied Biosystems Inc.). The phosphoramidite of 5-nitroindole was prepared as described (16).

Microchip manufacturing

A microchip with oligonucleotides immobilized within polyacrylamide gel elements (8% acrylamide; 0.28% bis-acrylamide) of 100 μm × 100 μm × 20 μm (in the wet condition), spaced by 200 μm and attached to a glass surface, was prepared as described (6,7). The polyacrylamide gel was activated by replacing some amide groups with hydrazide groups. Oligodeoxynucleotides for immobilization containing 3′-terminal 3-methyluridine were activated by oxidation with $NaIO_4$ to produce dialdehyde groups for coupling with the hydrazide groups of the gel (6). The solution of activated oligonucleotide was transferred onto the microchip element with a specially devised one-pin robot (7). Each gel element contained ~100 fmol of immobilized oligonucleotide.

Monitoring hybridization on a microchip

Oligonucleotides were fluorescently labeled with HEX or FAM and hybridized with a microchip. The fluorescence of every microchip element was monitored in real time (0.1–1 s) (7) with a specially devised multicolor epifluorescence microscope equipped with a charge-coupled device (CCD) camera (TE/CCD512SF, Princeton Instruments) and two sets of fast exchangeable FAM- and HEX-specific (OMEGA Optical) filters. The image of the microchip on the CCD camera was displayed and analyzed on a personal computer using specially developed software; the results were printed out (7).

Hybridization

Microchips were hybridized with 1 μM 21 base-long ssDNA in 2–5 μl of the hybridization buffer (1 M NaCl, 1 mM EDTA, 1% Tween 20 and 5 mM sodium phosphate, pH 7.0) at 0°C for 15 min on a Peltier thermotable (working range from –5.0 to +60.0°C). Unbound DNA was washed out at 0°C for 10–20 s in 50–100 μl of the hybridization buffer. Then microchips were hybridized with mixtures of fluorescent 5mers (1–3 μM each) at 0°C for

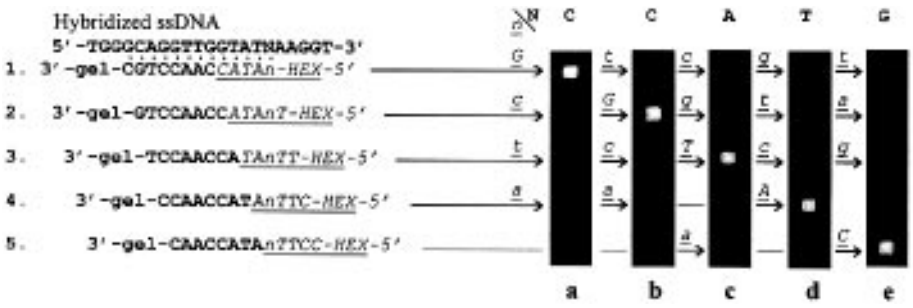

Figure 2. The effect of base stacking, base composition, mismatches, gaps or overlaps. The microchip contained five 8mers, (shown in bold) immobilized within 100 μm × 100 μm × 20 μm microchip gel elements. The 8mers were complementary to a 21 base-long ssDNA and were displaced by one base from each other. One of the four DNAs containing C (**a** and **b**), A (**c**), T (**d**) or G (**e**) at the 'N' position in the region complementary to stacked 5mers was hybridized first to the microchip and then exposed to a mixture of four 5mers (shown in underlined italic) displaced from each other by one base which stacked to the immobilized 8mers. The 5mers were labeled at the 5′-end with HEX. One of the 5mers in the mixture contained a base at the '*n*' position that was complementary to the 'N' in the hybridized DNA; the three other 5mers contained a mismatched base at the '*n*' position; these bases '*n*' are shown for each experiment (a–e) over the arrows in capital or lower case letters when they form perfect or mismatched base pairs respectively. After each hybridization, either only 5mers or all mobile oligonucleotides were washed off, and the microchip was hybridized again.

5 min in 2–5 μl of the hybridization buffer. The hybridization solution was covered by a cover-glass slide (Corning) with attached 0.1 mm-thick spacers. Following hybridization, pentamers were washed off with 50–100 μl of the hybridization buffer at 20°C for 10 min (duplex DNA hybridized to the immobilized oligonucleotides remains stable under these conditions). Then other 5mers were hybridized with the microchip under the same conditions. The hybridized DNA was washed off under more stringent conditions: 100 μl of 100 times-diluted hybridization buffer at 20°C for 10 min.

RESULTS

The validity of contiguous stacking hybridization for DNA sequence analysis and identification of the β-thalassemia mutations was demonstrated for the first exon and the first intron of β-globin gene (7). A synthetic 21 base-long ssDNA from the same region of β-globin gene was used to test the effect of different factors on hybridization of DNA with immobilized 8mers and 10mers and fluorescently labeled 5mers in solution (see Figs 2–4).

The effect of different mismatches in different positions of 5mers

Four 21 base-long ssDNA fragments containing four base changes at position 'N' and 20 HEX-labeled pentamers containing all four base changes in all five positions were used for hybridization. The microchip contained five gel-immobilized 8mers complementary to the DNA and displaced by one base. The experimental procedure consisted of two steps. The microchip was first hybridized with one of the four DNAs and then with a mixture of four 5mers. The 5mers in the mixture were displaced from each other by one base in order to stack to different 8mers of the microchip. One of the 5mers in the mixture was fully complementary to the DNA; the other three formed duplexes having different single mismatches. Such an experimental scheme allows one to estimate in a single experiment the relative hybridization efficiency by comparing the hybridization signals from four 5mers that form one perfect and three mismatched duplexes stacked with different 8mers. Several consecutive rounds of hybridization were carried out on the same microchip after washing off only the 5mers with the hybridization buffer (duplexes of DNA with the immobilized oligonucleotides remain stable under these conditions) or washing off the 5mers and the complimentary DNAs with the 0.01 dilution of the hybridization buffer.

Some results are shown in Figure 2. In experiment (a), the microchip was hybridized first with ssDNA containing C at the 'N' position. Then the duplexes formed on the microchip were hybridized with a mixture of HEX-labeled 5mers: no. 1, *GATAC*; no. 2, *TcATA*; no. 3, *TTtAT*; and no. 4, *CTTaA*. Only no. 1 produced a strong fluorescent signal; it was fully complementary to the DNA and was stabilized in the duplex by *C***C** base stacking with the immobilized 8mer. The other 5mers contained different mismatches (*g*-C in no. 4; *t*-C in no. 3; and *c*-C in no. 2 at the second, third and fourth positions from the 3′-end respectively) and did not hybridize. In the following experiment (b), the 5mers were washed out, and the microchip with the same prehybridized DNA was hybridized again with another mixture of four 5mers. In this mixture, only the 5mer *TGATA* formed a perfect duplex, stabilized by *A***C** stacking. Before the next experiment (c), both the 5mers and the DNA were completely washed off, and the microchip was hybridized with DNA containing A in the position 'N', followed by hybridization with a new 5mer mixture. Only the pentamer *TTTAT*, which formed a perfect duplex stabilized by *T***A** stacking, gave a fluorescence signal. A full set of hybridizations with four DNA fragments and 20 5mers were carried out, and the results are summarized in Table 1. In most experiments, perfectly stacked 5mer duplexes showed hybridization intensities >10 times higher than those from mismatched duplexes. Even the least destabilizing 5′-terminal, a *g*-T mismatch, was reliably discriminated: the hybridization intensity from the mismatched pentamer *gATAC* was at least five times weaker than that from the perfect pentamer *CTTAA* (Fig. 2d).

The effect of mismatches in immobilized oligonucleotides

The microchip in Figure 3 contained 10 immobilized 10mers. Five 10mers were displaced by one base from each other and were fully complementary to the DNA; the other five formed similar duplexes with the DNA but with weakly destabilizing G-t mismatches in the 5′-terminal and internal positions. Figure 3

Table 1. Contiguous stacking hybridization of four ssDNAs with five immobilized 8mers and 20 fluorescently labeled 5mers (for details see Fig. 2)

	Hybridized ssDNA	N in ssDNA				Stacked bases
		A	G	C	T	5′—3′
	5′-TGGGCAGGTTGGTATNAAGGT-3′					
1.	3′-gel-CGTCCAACCATAa-HEX-5′	–	–	–	+	CC
2.	3′-CATAg-HEX-5′	–	–	+	±	– –
3.	3′-CATAc-HEX-5′	–	+	–	–	– –
4.	3′-CATAt-HEX-5′	+	±	–	–	– –
5.	3′-gel-GTCCAACCATAaT-HEX-5′	–	–	–	+	AC
6.	3′-ATAgT-HEX-5′	–	–	+	–	– –
7.	3′-ATAcT-HEX-5′	–	+	–	–	– –
8.	3′-ATAtT-HEX-5′	+	–	–	–	– –
9.	3′-gel-TCCAACCATAaTT-HEX-5′	–	–	–	+	TA
10.	3′-TAgTT-HEX-5′	–	–	+	–	– –
11.	3′-TAcTT-HEX-5′	–	+	–	–	– –
12.	3′-TAtTT-HEX-5′	+	–	–	–	– –
13.	3′-gel-CCAACCATAaTTC-HEX-5′	–	–	–	+	AT
14.	3′-AgTTC-HEX-5′	–	–	+	–	– –
15.	3′-AcTTC-HEX-5′	–	+	–	–	– –
16.	3′-AtTTC-HEX-5′	+	–	–	–	– –
17.	3′-gel-CAACCATAaTTCC-HEX-5′	–	–	–	+	AA
18.	3′-gTTCC-HEX-5′	–	–	+	–	GA
19.	3′-cTTCC-HEX-5′	–	+	–	–	CA
20.	3′-tTTCC-HEX-5′	+	–	–	–	TA

+, Strong fluorescence signal from hybridized 5mer; ±, noticeable signal (~10–20% of the strong signal); –, no hybridization signal.

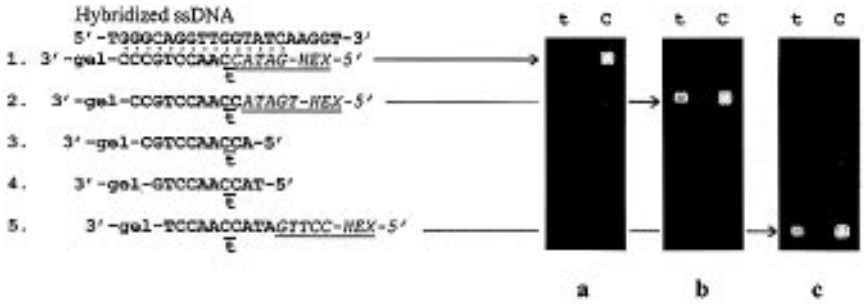

Figure 3 . The influence of mismatches in immobilized oligonucleotides. A microchip with two sets of five decamers either fully complementary to ssDNA or forming a G-t mismatch at a different position in the duplex was first hybridized with DNA and then successively with three fluorescently labeled 5mers *GATAC* (**a**), *TGATA* (**b**) and *CCTTG* (**c**).

shows the results of hybridization of the microchip, first with DNA and then with 5mers. No hybridization of the complementary 5mer is observed when the immobilized 10mer DNA duplex containing the mismatch is at the 5′-terminal position (Fig. 3a). The terminal mismatch appears to interfere with proper stacking interactions. On the other hand, the internal mismatches decrease the 10mer hybridization and, as a result, diminish the hybridization (Fig. 3b and c). The destabilizing effect of internal mismatches on DNA-microchip hybridization and therefore on stacking hybridization is, as expected, greater in 8mers than in 10mers (not shown).

Contiguous stacking hybridization with two stacked pentamers

Figure 4 shows hybridization of DNA with immobilized 8mers and two contiguously stacked 5mers (double stacking). The microchip consisted of seven immobilized 8mers overlapping by one base. The microchip was first hybridized with the unlabeled ssDNA and then with a mixture of two pentamers, *FAM-GATAC* and *HEX-ACCTT*, each labeled with a fluorescent dye. The *FAM-GATAC* contained 50% unlabeled 5mer because fluorescent label at the 5′-end considerably weakens stacking interactions (not shown). Therefore, the experiment demonstrates indeed the double stacking of DNA with the immobilized 8mer, unlabeled *GATAC* and *HEX-ACCTT*. Successive monitoring of FAM and HEX fluorescence was performed by changing the microscope filter sets for these dyes. Both 5mers form stable, single-stacked, 13-bp duplexes with the corresponding 8mer on the microchip when hybridized to DNA (Fig. 4, oligo 2-FAM and oligo 7-Hex). Neither forms the duplexes if there is ≥1 nt gaps or overlaps with immobilized 8mers. The effective formation of an 18 bp duplex (lacking two phosphodiester groups) was observed when DNA was hybridized with the immobilized 8mer and the 5mers stacked to the 8mer and to each other (Fig. 4, oligo 2-HEX). It is essential to stress here that stacking between two 5mers is not enough to produce a stable duplex with DNA

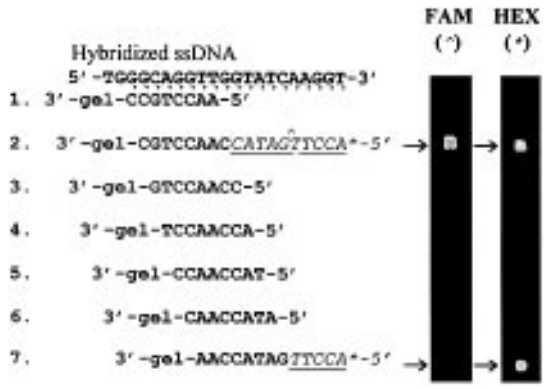

Figure 4. Multiple continuously stacked hybridization. A microchip containing seven immobilized overlapped 8mers was hybridized first with a complementary ssDNA and then with a mixture of two differently labeled 5mers: *FAM-GATAC* and *HEX-ACCTT*. The 5mer *FAM-GATAC* contained 50% of the same unlabeled oligonucleotide. HEX and FAM fluorescence was monitored by changing the microscope filter sets.

(Fig. 4, oligo 1). Stacking with longer and more stable duplexes apparently is essential for keeping 5mer duplexes together.

Contiguous stacking hybridization with pentamers containing all four bases and a universal base in different positions

The incorporation of all four bases (*N*) or a universal base into one or several positions of stacked 5mers (Fig. 1D) was suggested for increasing the efficiency of stacking hybridization (7). This incorporation can decrease the number of additional hybridizations, simplify microchips for DNA sequence analysis, and stabilize the 5mer duplexes. However, the introduction of four bases into the one, two, three and four positions within a 5mer decreases the portion of the complementary 5mers in the hybridization mixture by factors of 4, 16, 64 and 254, respectively, and may result in a much higher background. Such experiments were tested with a complementary 5mer (60% A/T bases) in a mixture of similar oligomers containing all four bases in the one, two and three positions (Table 2, pentamer mixtures 2, 3 and 4). A contiguous stacked duplex with this 5mer survives a mild, but not a harsh, washing procedure. Hybridization with *NATAC* was easily detected during hybridization. For detecting *NNTAC*, a mild washing procedure was applied (hybridization solution was removed and 10 µl of fresh hybridization buffer was added). We were unable to detect hybridization of the 5mer over a high background when *N* was present in three positions: *NNNAC*. We have tested 5-nitroindole as a universal base (*X*) which can match all four bases in the complementary DNA strand (16,17). Incorporation of *X* into the 5′-terminal position provides a stabilizing effect on the duplexes of the pentamer *XATAC* relative to the tetramer *ATAC* (Table 2, oligonucleotides 5 and 6), and of the 6mer *XTTATA* relative to a 100% A/T 5mer *TTATA* (Table 2, oligonucleotides 10 and 11). On the other hand, incorporation of *X* into an internal position destabilizes stacked 5mer duplexes (Table 2, 5mer 8). Even an additional 5-terminal *X* in the corresponding 6mer (Table 2, oligonucleotide 9) does not make it stable enough to use 5-nitroindole at an internal position in CSH under the hybridization conditions used.

Table 2. The effect of the presence of 5-nitroindole or all four bases in the 5mers on CSH

	ssDNA # 5'-TGGGCAGGTTGGTATCAAGGT-3'	
1.	3'-gel-CCCGTCCAACCATAG-HEX-5'	+
2.	3'-CATAN-HEX-5'	+
3.	3'-CATNN-HEX-5'	±
4.	3'-CANNN-HEX-5'	–
5.	3'-CATAt-HEX-5'	–
6.	3'-CATA-HEX-5'	–
7.	3'-CATAX-HEX-5'	+
8.	3'-CAXAG-HEX-5'	–
9.	3'-CAXAGX-HEX-5'	–
	ssDNA #2 5'-TGGGCAGGTTGGTATAAAGGT-3'	
10.	3'-gel-CCGTCCAACCATATT-HEX-5'	*
11.	3'-ATATTX-HEX-5'	+

A microchip containing two immobilized 10mers was hybridized with the first ssDNA fragments and then subsequently with fluorescently labeled 5mers 1–9. After washing off the DNA and the 5mers, the microchip was hybridized with the second DNA and subsequently with the two 5mers 10 and 11 subsequently. *X*, 5-nitroindole as a universal base; *N*, all four bases; +, strong hybridization signal detected during hybridization; ±, weak hybridization signal that can be observed only after washing; –, no hybridization; *, the duplex with 100% A/T pentamer 10 does not survive the washing procedure.

The effect of stacked bases, base composition, gaps and overlaps on the contiguous stacking hybridization

Among the 16 possible neighboring pairs of stacked bases we have tested the following eight: *C***C**, *A***C**, *T***A**, *A***T** *A***A**, *C***A**, *G***A**, (Figs 2 and 3) and *T***G** (Fig. 4). In all cases, reliable strong signals were observed upon hybridization to DNA of 5mers stacked with 8mers. The stability of the stacked 5mer duplexes is much effected by A/T content. The duplexes of 80–100% A/T 5mers are not stable enough. They dissociate and are washed off the microchip when their concentration is decreased by dilution with several volumes of fresh hybridization buffer. On the other hand, more stable 5mers, having at least 40% G/C content, survive washing at 0°C with 100-fold excess of the buffer. Figures 2 and 4 show that no hybridization signals were observed when one to four base-long gaps or overlaps interfered with hybridization of 5mers. Apparently, single 5mers alone, or even 5mers stacked with each other (Fig. 4), do not form stable duplexes with DNA. However, weak signals can be identified upon hybridization with one-base overlap of the 5mers *GATAC* or *TGATA* and immobilized 10mers 2 or 3, respectively (Fig. 3). The synthetic 10mers were not purified before immobilization, which could account for the presence on the microchips of some admixture of immobilized 9mers. These 9mers may form stable duplexes with DNA and hybridize with contiguously stacked 5mers, thus simulating hybridization with a one base overlap. In the case of the 8mer microchip containing 7mers as admixture, the duplexes of DNA with these 7mers were too unstable to participate in CSH and to simulate overlapped hybridization. It appears that the method is more sensitive to the presence of admixtures in immobilized oligonucleotides than direct sequencing by hybridization.

DISCUSSION

There are several approaches to preparing an array of immobilized oligonucleotides: solid-phase synthesis of oligonucleotides by conventional methods on a two-dimensional surface, (e.g., a glass) (4); a highly efficient, photo-activated parallel synthesis of a large oligonucleotide microarray (8,9); and immobilization of presynthesized oligonucleotides (5,6,18,19).

In the sequencing by hybridization to oligonucleotide microchips approach, presynthesized oligonucleotides are immobilized within three-dimensional gel microchip elements, which provides several essential advantages. The large capacity of the gel for immobilization, together with the high hybridization capacity of gel-immobilized oligonucleotides, increases the sensitivity of the hybridization monitoring. The apparent stability of DNA duplexes formed with gel-immobilized oligonucleotides depends on their concentration (6). High concentrations of immobilized 8mers allow their efficient hybridization with DNA and the monitoring of the formation of unstable duplexes with stacked 5mers. The use of the gel-immobilized oligonucleotides appears to be crucial for contiguous stacking hybridization with 5mers.

There is a difference in the stacking interactions among all 16 possible pairs of adjacent bases. Among them, eight were tested in the 5mer hybridizations with immobilized 8mers and 10mers. Even one of the weakest, TA stacking (20–22), was strong enough (Table 1, oligonucleotides 9 and 20) to be used in CSH. As expected, an essentially lower stability was observed for A/T-rich stacked duplexes with 5mers. Nevertheless, monitoring of hybridization with 80–100% A/T 5mers can be carried out without including a washing procedure. The 5mer duplex stability can be significantly increased by adding 5-nitroindole as a universal sixth base into the 5′-terminal position (Table 2). This addition can be recommended as a general approach for stabilizing the derived hybrids.

The duplexes for 5mers, 8mers and 10mers vary in stability due to different numbers of base pairs. The duplex stabilities can be calculated approximately by the use of empirical rules (20–22). As the result of these differences, the hybridization of DNA on a microchip containing immobilized 8mers and 10mers was carried out separately from the second round of hybridization with the 5mers. The conditions for these hybridizations were found from preliminary experiments. The fluorescence measurements can be carried out simulteneously for all microchip elements in real time with our specially devised fluorescence microscope equipped with a CCD camera (7). This enables us to measure in parallel the melting curves for all duplexes formed on microchips and identify an optimal temperature for the detection of each duplex (unpublished information).

The accuracy of SbH depends on how reliably perfect duplexes can be discriminated from those containing mismatches (in particular, terminal mismatches). The presence of a mismatch in the duplex apparently has a higher destabilizing effect for 5mers than for 8mers and longer oligonucleotides. A mismatch in the stacked 5mer usually blocks hybridization completely, whereas in an immobilized 10mer, the mismatch just lowers the extent of hybridization (Figs 2 and 3). Even the least-destabilizing (G-T) terminal mismatch in the 5′-terminal position significantly (5-fold) decreases the hybridization signal from the stacked 5mer and can be reliably discriminated (Fig. 2d and e). Thus, hybridization with 5mers can significantly increase the accuracy of sequencing and of identifying base changes, compared with direct SbH methology.

In the presence of short repeats in DNA, sequence reconstitution based on hybridization data can be ambiguous. Repeats as short as 7 and 12 bases lead to ambiguity upon hybridization with the 8mer microchip alone or together with a 5mer stacked with the immobilized 8mer respectively. ‘Double’ hybridization on the 8mer microchip with two differently labeled 5mers, stacked to each other and to immobilized 8mer (Fig. 4), provides the essential information needed to reconstitute a DNA sequence a few thousand bases in length containing up to 16 base repeats (10).

The identification of known base changes in DNA by the hybridization for genetic determination of known mutations, for example, requires hybridization with a few 5mers; however, gene polymorphism studies or *de novo* sequencing may require hybridization with up to 1024 5mers. To simplify such a cumbersome procedure to one involving only five rounds of hybridization, a modification using 20 5mers was envisioned (Fig. 1D). With these 5mers, the discriminating bases are present in only one position of the 5mers and the 5mers are labeled with four different fluorescent dyes for the four varied bases. All other positions are occupied with all four bases or with a base with multiple specificity that matches all four bases. We have demonstrated here that this procedure works with a four-base substitution at two positions and with the 5′-terminal, 5-nitroindole (Table 2). 5-Nitroindole, as expected (17), stabilized the duplexes only at their terminal positions and destabilized them at internal ones. Inosine and other universal bases (23) could also be tested for their ability to increase the efficiency of the hybridization.

These experiments suggest that contiguous stacking hybridization may be a promising approach for sequence analysis, possibly including *de novo* DNA sequencing, and studies of genomic diversity.

ACKNOWLEDGEMENTS

We are indebted to Dr Yu. Lysov, Aleksei Drobyshev and Natalia Mologina for their contribution to this work and crucial comments. We thank L. Novikova and D. E. Nadziejka for editorial assistance. This work was supported by grants 558 and 562 of the Russian Human Genome Program, and by contract 943352403 with Argonne National Laboratory (USA).

REFERENCES

1 Lysov,Y.P., Florentiev,V.L., Khorlyn,A.A., Khrapko,K.R., Shick,V.V. and Mirzabekov,A.D. (1988) *Dokl. Akad. Nauk SSSR* (Russian), **303**, 1508–1511.
2 Bains,W. and Smith,G.C. (1988) *J. Theor. Biol.*, **135**, 303–307.
3 Drmanac,R., Labat,L., Brukner,I. and Crkvenjakov,R. (1989) *Genomics*, **4**, 114–128.
4 Southern,E., Maskos,U. and Elder,J.K. (1992) *Genomics*, **13**, 1008–1017.
5 Khrapko,K.R., Lysov,Y.P., Khorlin,A.A., Shick,V.V., Florentiev,V.L. and Mirzabekov,A.D. (1989) *FEBS Lett.*, **256**, 118–122.
6 Khrapko,K.R., Lysov,Y.P., Khorlin,A.A., Ivanov,I.B., Yershov,G.M., Vasilenko,S.K., Florentiev,V.L. and Mirzabekov,A.D. (1991) *J. DNA Seq. Mapp.*, **1**, 375–388.
7 Yershov,G., Barsky,V., Belgovskiy,A., Kirillov,E., Kreindlin,E., Ivanov,I., Parinov,S., Guschin,D., Drobyshev,A., Dubiley,S. and Mirzabekov,A. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 4913–4918.
8 Fodor,S.P.A., Read,J.L., Pirrung,M.C., Stryer,L.T., Lu,A. and Solas,D. (1991) *Science*, **251**, 767–773.

9 Pease,A.C., Solas,D., Sullivan,E.J., Cronin,M.T., Holmes,C.P. and Fodor,S. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 5022–5026.
10 Pieters,J.M.L., Mans,R.M.W., van den Elst,H., van der Marel,G.A., van Boom,J.H. and Altona,C. (1989) *Nucleic Acids Res*, **17**, 4551–4565.
11 Kieleczawa,J., Dunn,J.J. and Studier,F.W. (1992) *Science*, **258**, 1787–1792.
12 Kotler,L.E., Zevin-Sonkin,D., Sobolev,I.A., Beskin,A.D. and Ulanovsky,L.E. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4241–4245.
13 Broude,N.E., Sano,T., Smith,C.L. and Cantor,C.R. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3072–3076.
14 Lysov,Y.P., Chernyi,A.A., Balaeff,A.A., Beattie,K.L. and Mirzabekov,A.D. (1994) *J. Biomol. Struct. Dyn*, **11**, 797–812.
15 Lipshutz,R.J, Morris,D., Chee,M., Hubbell,E., Kozal,M., Shah,N., Shen,N., Yang,R. and Fodor,S.P.A. (1995) *BioTechniques*, **19**, 442–447.
16 Loakes,D. and Brown,D.M. (1994) *Nucleic Acids Res.*, **22**, 4039–4043.
17 Loakes,D., Brown,D.M., Linde,S. and Hill,F. (1995) *Nucleic Acids Res.*, **23**, 2361–2366.
18 Lamture,J.B., Beattie,K.L., Burke,B.E., Eggers,M.D., Ehrlich,D.J., Fowle,R., Hollis,M.A., Kosicki,B.B., Reich,R.K., Smith,S.R., Varma,R.S. and Hogan,M.E. (1994) *Nucleic Acids Res.*, **22**, 2121–2125.
19 Ghu,Z., Guilfoyle,R.A., Thiel,A.J., Wang,R., and Smith,L.M. (1994) *Nucleic Acids Res.*, **22**, 5456–5465.
20 Brestlauer,K.J., Frank,R., Blocker,H. and Marky,L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3746–3750.
21 Wetmur,J.G. (1991) *Crit. Rev. Biochem. Mol. Biol.*, **26**, 227–259.
22 Doktycz,M.J., Morris,M.D., Dormady,S.J., Beattie,K.L. and Jacobson,K.B. (1995) *J. Biol. Chem.*, **270**, 8439–8445.
23 Kong Thoo Lin,P. and Brown,D.M., (1992) *Nucleic Acids Res.*, **20**, 5149–5152.