



Marking of specific sequences in double-stranded DNA molecules —SNP detection and direct observation

Yasushi Shigemori, Hiroataka Haruta, Takao Okada, et al.

Genome Res. 2004 14: 2478-2485

Access the most recent version at doi:[10.1101/gr.2789604](https://doi.org/10.1101/gr.2789604)

References This article cites 14 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/14/12/2478.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Methods

Marking of specific sequences in double-stranded DNA molecules—SNP detection and direct observation

Yasushi Shigemori,^{1,3} Hirotaka Haruta,² Takao Okada,² and Michio Oishi^{1,4}

¹Laboratory of Human Gene Research II, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0812, Japan; ²Research Institute of Biomolecule Metrology, Co. Ltd., Tsukuba, Ibaraki 305-0853, Japan

In this study, we describe a simple method to mark specific sequences in double-stranded DNA molecules. For the marking, we used two specifically designed oligonucleotides, one of which is complementary to the sequence to be marked and the other, serving as a splint, to make the marking stable and detectable by subsequent various analytical means. In the presence of the two deoxyoligonucleotides, whereas RecA protein-mediated reaction converts the sequence to be marked to a regional triple-stranded structure with the complementary (probing) oligonucleotide, DNA ligase transforms it to a stable multi- (possibly quintuple) stranded structure with the splint oligonucleotide. The whole marking process is simple and completed in a single reaction mixture. Because RecA protein makes the marking to proceed with high fidelity, we were able to mark (detect) SNPs in complex genomes like human's. Furthermore, the structure of the marked sequence is stable and quite distinct enough to be readily detectable by biochemical means or direct observation by scanning probe microscopy.

[Supplemental material is available online at www.genome.org.]

At present, the most commonly used method to search for a specific base sequence in DNA molecules is to hybridize a labeled complementary oligonucleotide with a target sequence in DNA molecules and to detect signals derived from the oligonucleotide through various means. However, target DNA molecules must first be converted to the single-stranded form to be recognized by complementary oligonucleotides.

RecA protein is capable of directly recognizing specific base sequences in double-stranded DNA molecules and is known to form a triple-stranded structure with a complementary oligonucleotide, but the triple-stranded structure is quite unstable, particularly when RecA protein is removed (West et al. 1981; Rao et al. 1991, 1993; Rao and Radding 1993). Attempts have been made to make the structure stable enough to withstand various subsequent manipulations (Fujiwara et al. 1998a,b; Potaman et al. 2002; Roulon et al. 2002). Recently, Rice et al. (2004) reported that a D-loop formed with RecA protein is stabilized by converting it to a stable double D-loop.

Using a different approach, we have devised a method to stably mark specific sequences in double-stranded DNA molecules by converting a RecA-mediated triple-stranded structure to a stable one by use of a second deoxyoligonucleotide (a splint oligonucleotide), to which the first probing oligonucleotide is aligned to covalently ligate itself. The resultant regional multi- (possibly quintuple) strand structure is stable and easily detectable by various means, including observation by scanning probe microscopy such as atomic force microscopy (AFM). The procedure is quite simple, and more importantly, marking is very spe-

cific, as the recognition of base sequences by a RecA protein proceeds with extreme high fidelity.

In this study, we report the principle and details of the method (direct marking) that allows us to stably mark almost any specific sequences in double-stranded DNA molecules, even in a complex genomic DNA, essentially by a single process. As practical applications of the method, we also present experimental results in which SNPs in human genomes are detected, as well as the determination of the marked sequence in DNA samples after visualization by AFM.

Results

The principle of the method

The whole procedure is carried out in a single reaction mixture, and events that presumably occur in the mixture are shown diagrammatically in Figure 1. As shown in the figure (see step 1), the probing oligonucleotide consists of three parts; a sequence (a; 20 bp) complementary to the sequence to be marked, two arbitrary arm sequences (b1 and b2; each 10 bp), which bridge the probing sequences to sequences (c1 and c2), each complementary, but with opposite directions to the proximal and distal parts of the splint oligonucleotide sequence (17 bp). The splint oligonucleotide, which could be universal for probing any sequences, consists of sequences complementary to proximal and distal sequences (c1 and c2) of the probing oligonucleotide (see steps 2 and 3).

Linear double-stranded DNA with a sequence to be marked is incubated with probing and a splint oligonucleotide in the presence of RecA protein and DNA ligase. A triple-stranded DNA structure is first formed in the DNA molecules with the probing oligonucleotide through a RecA-mediated reaction (steps 1 and 2). The terminal sequences of the probing oligonucleotide, which are designed to be aligned with the splint oligonucleotide, are

³Present address: Biotechnology Division, Aisin Cosmos R&D Co. Ltd., Kisarazu, Chiba 292-0812, Japan.

⁴Corresponding author.

E-mail oishi@kazusa.or.jp; fax 81-438-52-3911.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.2789604>.

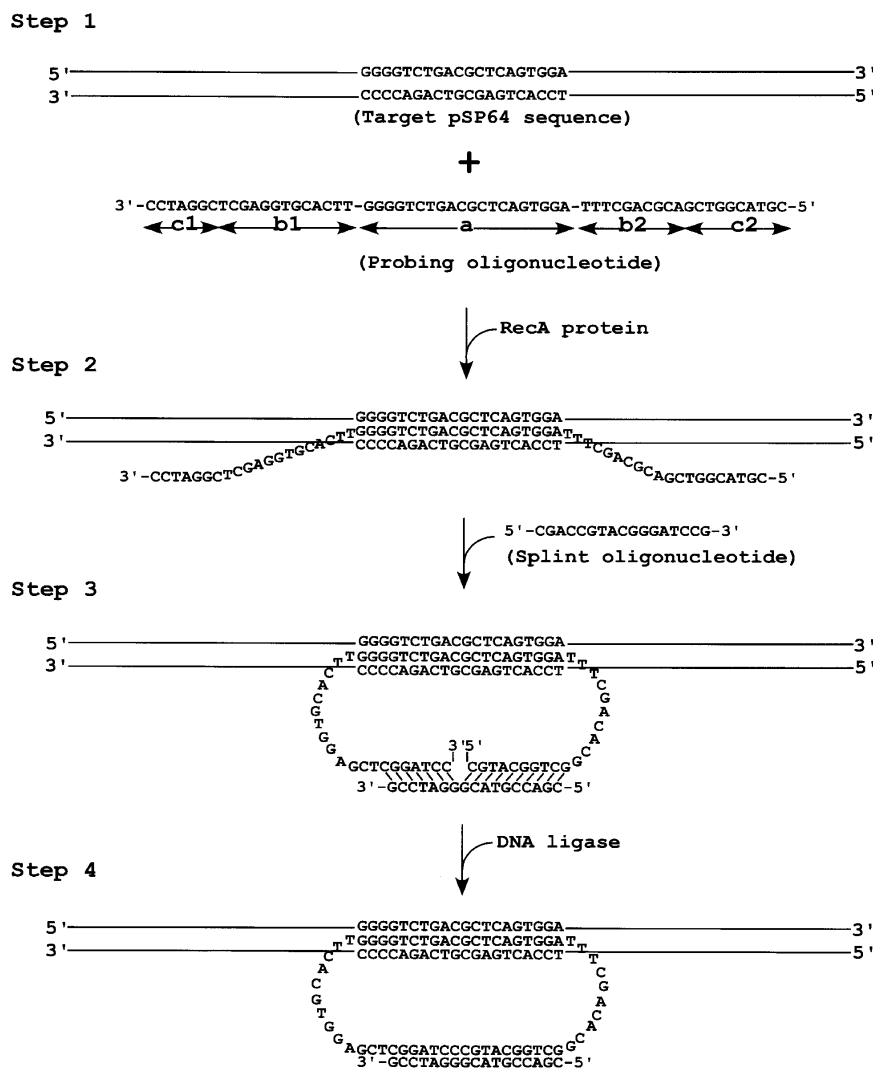


Figure 1. Diagrammatic representation of direct marking of specific base sequences in double-stranded DNA through formation of a multistranded stable structure.

then converted to a quintuple-stranded structure with the splint oligonucleotide (steps 2 and 3). The conversion is completed by covalently closing a nick between the two terminal bases of the probing oligonucleotide with DNA ligase (steps 3 and 4). After removal of RecA protein and other components present in the reaction mixture, the marked DNA can be subjected to various subsequent analyses, as the resultant multi- (quintuple) stranded structure is stable and quite distinct in its structure.

In Supplemental data 1, we show the results of an experiment in which we intended to mark a specific sequence (1892 bp) in pSP64 DNA (3030 bp). The DNA was incubated with a labeled probing oligonucleotide and a splint oligonucleotide in the presence of RecA protein and DNA ligase, and after removal of RecA protein, the products were analyzed by electrophoresis through an agarose, and autoradiographed. It is quite clear that a distinct band corresponding to the pSP64 DNA fragment was seen when all of the components required for the formation of the stable multistranded structure were present (lane 1). Note that when either DNA ligase or a splint oligonucleotide was omitted from the mixture, no bands corresponding to the position were ob-

served (lanes 4 and 5), probably due to dissolution of the postulated unstable triple-stranded structures (see steps 2 and 3, Fig. 1).

The efficiency of the formation of a stable multistranded structure is quite high (see below), but is dependent upon the size of the complementary sequence present in the probing oligonucleotide. As the number of bases complementary to the target sequence increased (from 10 to 40), efficiency of the formation of the stable structure increased accordingly (Fig. 2). We routinely used probing oligonucleotides carrying at least 20 complementary bases.

We compared the results obtained by direct marking with those by conventional Southern hybridization for probing specific sequences in pSP64 DNA fragments with different molecular size (300–1800 bp). As shown in Supplemental data 2A (direct marking) and C (Southern hybridization), essentially the same autoradiographic patterns were obtained by the two procedures. On the other hand, bands revealed by ethidium bromide staining after direct marking (B) were broader than those for Southern hybridization (D), suggesting that DNA fragments carrying the multistranded structure migrated slightly more slowly in electrophoresis as expected.

We also performed experiments similar to the ones described above in which we used four different probing oligonucleotides (Fig. 3A) for the same target DNA fragment. As seen in Figure 3B–E, bands (four, three, two, and one) corresponding to each probing oligonucleotide appeared, confirming that the oligonucleotides are specifically associated with the target sequences in the DNA molecules.

We compared positions of marked products after electrophoresis with those of the original target DNA fragments. As seen in Figure 4, with a 60-bp probing oligonucleotide and a 17-bp splint oligonucleotide, slower mobility of the marked products than the original target DNA fragments (indicated by arrowheads in the figure) was evident, particularly when the molecular size of the target DNA fragments was <900 bp. Although the differences in the mobility do not seem to pose any serious problems in most of the experiments, should one want to know the exact molecular size of the probed DNA fragments with <1 kb or so, it may be necessary to run the gel with control DNA fragments alongside. From the broadness of the bands and other estimations, we have estimated that ~50% of the original DNA sample is converted to the multistranded structure under the conditions described.

Marking (probing) of specific sequences in complex genomic DNA

We examined whether a single specific sequence in complex genomic DNA can be marked and thus probed by the procedure

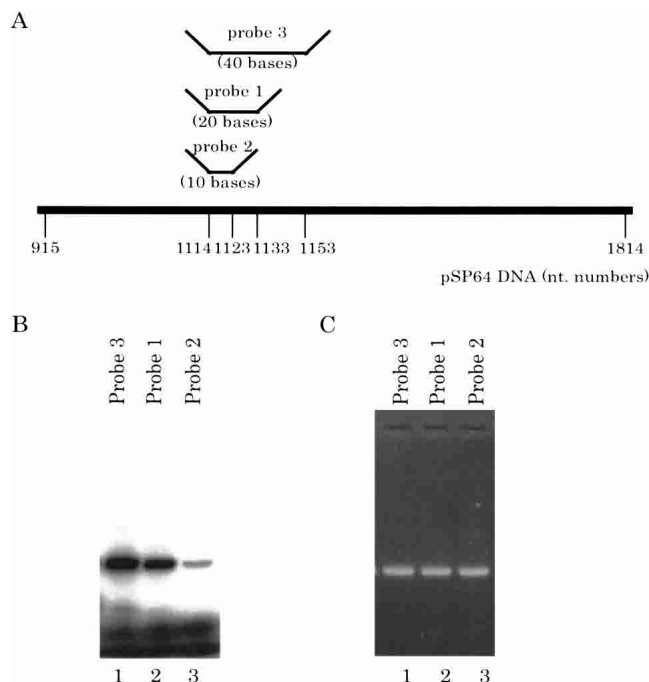


Figure 2. Effect of the size of the complementary sequence in probing oligonucleotides. A reaction mixture (20 μ L) consisting of 100 ng of target DNA (900 bp), a probing oligonucleotide that had been labeled at the 5'-terminus with 32 P using T4 polynucleotide kinase, was subjected to direct marking. (A) Diagrammatic representation of the relative position of the complementary sequences of probe 3, probe 1, and probe 2. (B) Autoradiographic patterns. (C) Ethidium bromide patterns. In both B and C, lane 1, with probe 3 (40 complementary bases); lane 2, with probe 1 (20 complementary bases); and lane 3, with probe 2 (10 complementary bases). For details, see Methods.

described above. Human genomic DNA was digested with restriction enzymes (HindIII, PvuII, and EcoRI) and the digests were subjected to the marking by use of two labeled probing oligonucleotides (probes 7a and 8a, each 80 bp) which correspond to genomic sequences present in two adjacent p53 gene DNA fragments (1909 bp EcoRI–HindIII and 3159 bp HindIII–PvuII fragments, see Fig. 5A), respectively, and a splint oligonucleotide complementary to the terminal sequences of probing oligonucleotides 7a and 8a, respectively. As seen in Figure 5B, when the labeled probing oligonucleotide 7a was used, a band corresponding to 1909 bp EcoRI–HindIII fragment appeared (lane 1). Likewise, when the labeled probing oligonucleotide 8a was used, a band appeared at the position where the 3159-bp HindIII–PvuII fragment should migrate (lane 2). These bands correspond well to the major bands that appeared by Southern hybridization (Fig. 5C). No bands were seen when the splint oligonucleotide was omitted from the reaction mixture (data not shown). Because of the high efficiency of the multistranded structure, as well as direct application of the products to electrophoresis, this procedure makes it extremely easy to probe DNA sequences of interest, even in a complex mammalian genomic DNA.

Fidelity of the marking and SNP detection

One of the advantages of using RecA-assisted marking of specific sequences is its high fidelity in finding sequences of interest in target DNA. In the experiment described above (Fig. 5), we saw

only a single band in the autoradiographic patterns in each of the human genomic DNA samples (Fig. 5B, lanes 1 and 2), whereas extra bands were seen in the patterns obtained by Southern hybridization (Fig. 5C, lanes 1 and 2) even after extensive washing. To examine whether this procedure can be applied to the detec-

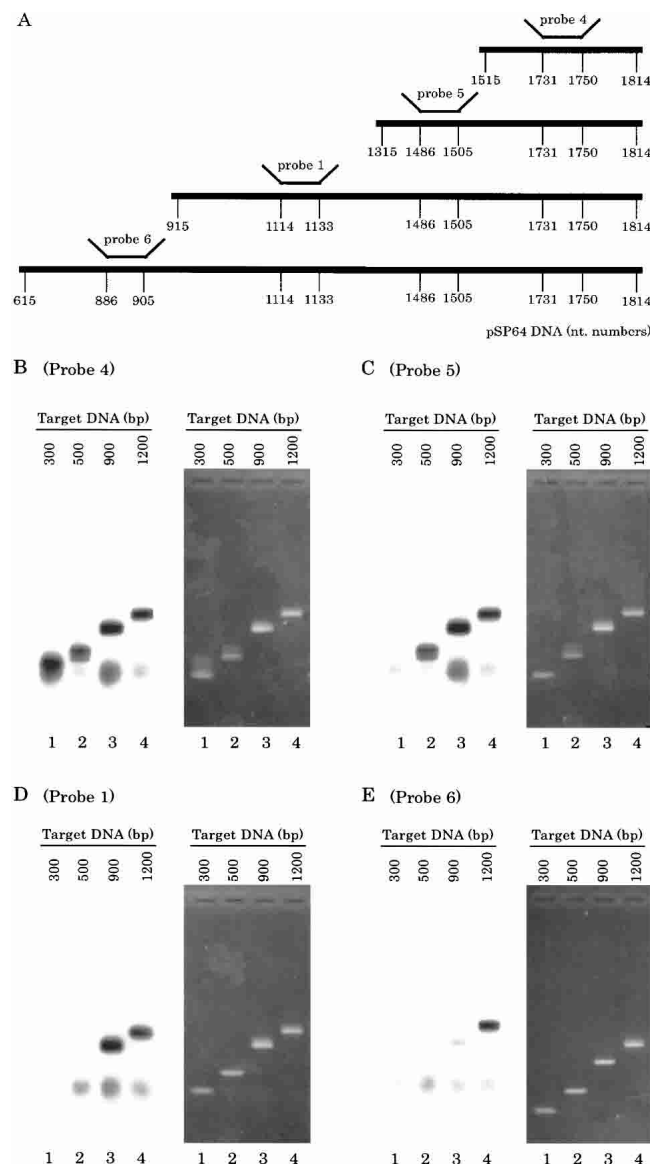


Figure 3. Direct marking with different probing oligonucleotides. DNA fragments with various sizes (300, 500, 900, and 1200 bp) of pSP64 DNA were subjected to direct marking using four different probing oligonucleotides. (A) A diagrammatic representation of the position of probing oligonucleotides used (probes 4, 5, 1, and 6) on the four different target DNA molecules. (B) Autoradiographic (left) and ethidium bromide staining (right) patterns after direct marking using probe 4. (C) Autoradiographic (left) and ethidium bromide staining (right) patterns after direct marking using probe 5. (D) Autoradiographic (left) and ethidium bromide staining (right) patterns after direct marking using probe 1. (E) Autoradiographic (left) and ethidium bromide staining (right) patterns after direct marking using probe 6. Throughout experiments, lane 1, 300-bp DNA fragment; lane 2, 500-bp DNA fragment; lane 3, 900-bp DNA fragment; and lane 4, 1200-bp DNA fragment. Nucleotide numbers (nt. numbers) are those registered in GenBank (accession no. X65328). For details, see Methods.

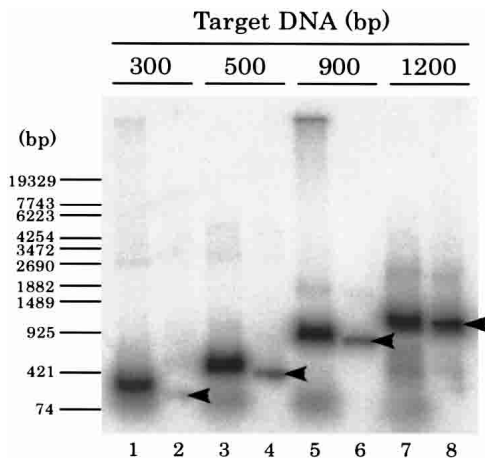


Figure 4. Effects of the formation of the multistranded structure on electrophoretic mobility. DNA fragments (pSP64 DNA) with various sizes (300, 500, 900, and 1200 bp), were subjected to direct marking using a labeled probing oligonucleotide (probe 4; 60 bp) and a splint oligonucleotide (17 bp). As controls, the target DNA fragments were ^{32}P -labeled at the 5'-termini by T4 polynucleotide kinase and electrophoresed in parallel. (Lanes 1,3,5,7) Results of autoradiographic patterns after direct marking; (lanes 2,4,6,8) ^{32}P -labeled target DNA fragments (indicated by arrowheads). (Lanes 1,2) 300-bp DNA fragment; (lanes 3,4) 500-bp DNA fragment; (lanes 5,6) 900-bp DNA fragment; (lanes 7,8) 1200-bp DNA fragment. For details, see Methods.

tion of SNPs, target DNA fragments, in which either one of the four different bases was introduced at the same position, were subjected to direct marking using the corresponding four probing oligonucleotides (Fig. 6A). As seen in Figure 6B–E, signals were detected only when completely matched oligonucleotides (in 20-bp complementary bases) were present, suggesting that the procedure can be applied to SNP detection regardless of the kind of bases at SNP sites.

We performed an experiment with human genomic DNA samples (Samples A, B, and C) from three individuals in which SNPs (G/C and C/G) are present in the exon 5 of p53 gene. (Fig. 7A). Four labeled oligonucleotides covering the SNP site, each with one altered base for the SNP among the total of 20 complementary bases (Fig. 7A), were used for direct marking. As seen in Figure 7B, for sample A, in which the SNP site is homozygous (C/G, C/G), only probe 12 gave a signal (indicated by an arrow). For sample B, in which the SNP site is also homozygous but with a different base pair (G/C, G/C), only probe 13 gave a signal (Fig. 7C). When sample C, in which the pair is heterozygous (C/G, G/C), was used, probe 12 as well as probe 13 gave signals with almost equal, but weaker intensities than those with sample A and B (Fig. 7D). These results demonstrate that this method is applicable to directly detect and identify SNPs in a complex genome.

Taken together, it seems quite clear that the direct-marking procedure presented here not only finds sequences of interest in complex genomes, but also discriminates mismatched marking oligonucleotides almost completely. Thus, this procedure should offer an easy and practical way to detect SNP in complex genomic DNA molecules.

Direct observation of marked DNA molecules

We subjected the marked products to atomic force microscopy (AFM) to see whether the marked region can be observed. Figure

8B shows the AFM pattern of the 1892-bp DNA molecules marked at nucleotide number 1114–1153 (Fig. 8A). Because of a distinct multistranded structure, the marked region in DNA molecules is clearly seen at the position of the molecules where we expect to see the structure without any modifications of the DNA molecules with proteins or other markers (indicated by arrows in the figure). Approximately 50% of the DNA molecules examined under AFM carry the marked structure, confirming high efficiency of formation of the multistranded structure. Figure 8C shows the distribution pattern of the multistranded structure as a function of its relative distance from one end of the double-stranded DNA molecules.

Discussion

Nilsson et al. (1994) first reported to stabilize probing oligonucleotides (padlock probe), which was hybridized with their target sequences in dissociated single-stranded DNA molecules. The procedure we have presented in this study is to directly mark (or probe) target sequences in double-stranded DNA molecules without DNA dissociation. The procedure is quite simple and the marked sequence is stable enough to withstand subsequent manipulations.

Because RecA protein makes the marking proceed with high fidelity as well as with high efficiency, this procedure has enabled us to unambiguously detect DNA fragments with a specific sequence in a complex genome. As a natural extension of the observation, we performed experiments and were able to identify a SNP directly from genomic DNA samples without any prior DNA amplification. The direct marking reported here may be applied to the established procedures that require DNA dissociation such as FISH and haplotype analysis for SNP detection.

Whereas we use RecA protein to search for target sequences in double-stranded DNA molecules, there have been several reports to mark or probe specific sequences without using RecA protein. Although limited to A-T rich sequences in closed circular DNA molecules, specific sequences in double-stranded DNA can be converted to a stable multistranded form with oligonucleotides, with a complementary sequence to those sequences after a brief heat treatment (Potaman et al. 2002). In another approach, partially heat-denatured regions in double-stranded DNA molecules are converted to a triple-stranded structure with a complementary oligonucleotide in the presence of a specific agent (BQQ, 6-[3-(dimethylamino)propyl]amino-11-methoxy-benzo[f]quino-[3,4-b]quinoxaline) (Zain et al. 1999; Geron-Landre et al. 2003). Although these procedures seem to be quite promising, it is yet unknown whether the stably marked structure can be formed specifically at any sequences in double-stranded DNA, particularly in complex genomic DNA. By using peptide nucleic acids (PNAs), Demidov et al. (2001) reported sequence-specific targeting, although limited to polypurine/polypyrimidine stretches, via triple-helice invasion, which forms a stable structure (Demidov and Frank-Kamenetskii 2001).

The marked structure reported in this study is stable and quite distinct because of a multistranded form. As a result, the structure could be detected through various means, which includes direct observation by scanning probe microscopy like AFM. Because the structure is associated with double-stranded DNA, the position is determined more accurately than that associated with single-stranded DNA, in which intramolecular interaction of single strands often poses a problem in quantitative determination of the marked position.

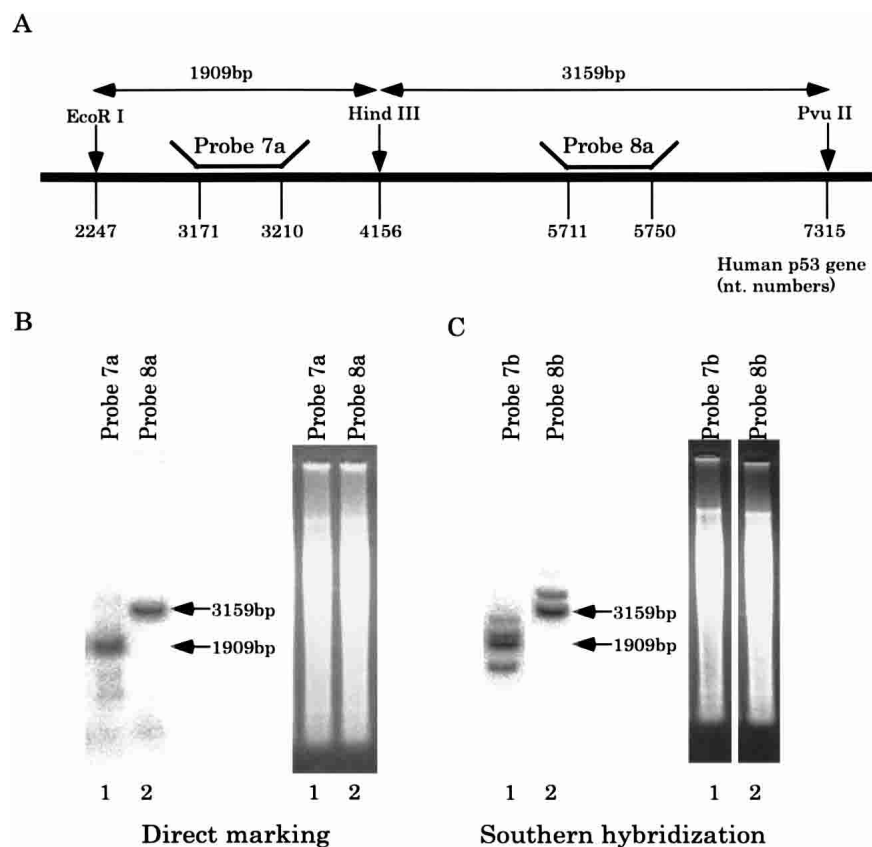


Figure 5. Direct marking of specific sequences in human genomic DNA. Human genomic DNA (digested with EcoRI, HindIII, and PvuII) were subjected to direct marking using a labeled probing oligonucleotide and a splint oligonucleotide (17 bp). (A) A diagrammatic representation of 1909- and 3159-bp DNA fragments in p53 gene. Probe 7a and probe 8a cover 40-bp sequences, respectively. (B) Autoradiographic (left) and ethidium bromide staining (right) patterns after direct marking. (Lane 1) Probed by probe 7a; (lane 2) probed by probe 8a. (C) Autoradiographic (left) and ethidium bromide staining (right) patterns after Southern hybridization. (Lane 1) Probed by probe 7b; (lane 2) probed by probe 8b. For probe 7b and 8b, we used a ^{32}P -labeled DNA fragment produced by PCR with forward primer 6 and backward primer 5 (for probe 7b), forward primer 7 and backward primer 6 (for probe 8b) as described in the Methods. Nucleotide numbers (nt. numbers) are those registered in GenBank (accession no. X54156). Distances in the diagram do not reflect actual distances. For details, see Methods.

Because marked DNA molecules remain intact, subsequent manipulation of the DNA molecules is expected to be quite easy to proceed. For example, not only probing becomes very simple, but cloning of DNA fragments with a sequence of interest can be completed much more easily. Furthermore, the easy marking and probing procedure reported here without DNA dissociation seems to have essentially eliminated anticipated obstacles for construction of fully automated marking and probing equipment, which has not been realized yet, despite a great potential need.

Methods

Deoxyoligonucleotides

Deoxyoligonucleotides were purchased from Sawady Technology, Inc.; Restriction enzymes, T4 DNA ligase, and T4 polynucleotide kinase were obtained from New England Biolabs. Deoxyoligonucleotides were labeled and purified with ^{32}P at the 5'-terminus with T4 polynucleotide kinase in the presence of

$[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (Amersham Biosciences; 6000 Ci/mmol) according to a DNA-labeling kit (Takara-Bio, Inc.; MEGALABEL). The sequences of probing oligonucleotides and splint oligonucleotides used here are listed below. The specific sequences used for probes are underlined. Probe 1; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGACGCTCAGTGGATTACAGTGGAGCTCGGATCC-3', Probe 2; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGACTTCACGTGGAGCTCGGATCC-3', Probe 3; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGACGCTCAGTGGAAACGAAAATCACCCTAAGGGTTCAGCTCGGATCC-3', Probe 4; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGACTTCACGTGGAGCTCGGATCC-3', Probe 5; 5'-CGTACGGTCGACGCAGCTTTATCCGCTCCATCCAGTCTATTCACTGAGCTCGGATCC-3', Probe 6; 5'-CGTACGGTCGACGCAGCTTTGTAGGGCTGCTACAGAGTTTACGTGGAGCTCGGATCC-3', Probe 7; 5'-CGTACGGTCGACGCAGCTTTCTGTATTTGAACTTTTCAGAAGCTCTGAGTTTCCAGATTACAGTGGAGCTCGGATCC-3', Probe 8; 5'-CGTACGGTCGACGCAGCTTTCAA CCTCCTAGCAAAGTAGGGAAAGGAACTTAAGTATTTTACGCTGGAGCTCGGATCC-3'. Probe 9; 5'-CGTACGGTCGACGCAGCTTTGAGGTTGTGAGGCGCTGCCACCATTGAGGCTCGGATCC-3', Probe 10; 5'-CGTACGGTCGACGCAGCTTTGAGGTTGTGAGGCGCTGCCACCATTGAGGCTCGGATCC-3', Probe 11; 5'-CGTACGGTCGACGCAGCTTTGAGGTTGTGAGGCGCTGCCACCATTGAGGCTCGGATCC-3', Probe 12; 5'-CGTACGGTCGACGCAGCTTTGAGGTTGTGAGGCGCTGCCACCATTGAGGCTCGGATCC-3', Probe 13; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGAGGCTCAGTGGATTACAGTGGAGCTCGGATCC-3', Probe 14; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGAGGCTCAGTGGATTACAGTGGAGCTCGGATCC-3', Probe 15; 5'-CGTACGGTCGACGCAGCTTTGGGGTCTGATGCTCAGTGGATTACAGTGGAGCTCGGATCC-3', Splint oligonucleotides; 5'-CGACCGTACGGGATCCG-3'.

Preparation of target DNA

pSP64 DNA (Promega) and human genomic DNA (Promega) were digested with restriction enzymes, treated with phenol:chloroform, precipitated with 2 vol of ethanol, and dissolved in 50 μL of TE-buffer. The solution was mixed with Alkaline Phosphatase (Calf intestine; CIP) dephosphorylation buffer (supplemented), and 0.25 U of CIP (Takara Bio, Inc.), incubated for 30 min at 37°C, and the reaction terminated by adding EDTA (pH 8.0) and SDS to final concentrations of 5 mM and 0.5%, respectively. After incubation for 30 min at 65°C, samples were treated with phenol:chloroform DNA precipitated with 2 vol of ethanol and dis-

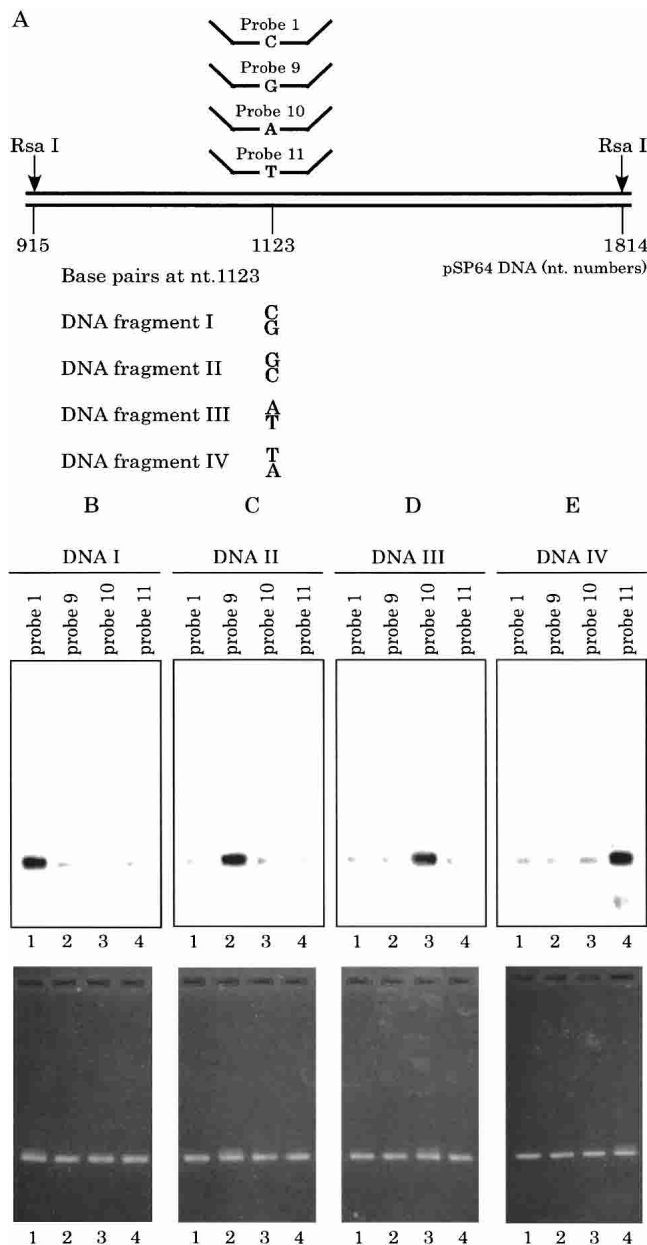


Figure 6. Fidelity of the marking (probing). A DNA fragment (900 bp) of pSP64 DNA with a different bases at position nt. number 1123 was subjected to direct marking using four different probing oligonucleotides (20-bp complementary bases) and a splint oligonucleotide (17 bp). (A) A diagrammatic representation of probing oligonucleotides and target DNA fragments used. One mismatched base is present at position 10 of the 20 complementary base sequences in the probing oligonucleotides. DNA fragment I to IV represented pSP64 DNA with an altered base at position nt. number 1123, respectively. (B) Direct marking for a DNA fragment I. (Lanes 1–4) Probes 1, 9, 10, and 11 were used, respectively. (C) Direct marking for a DNA fragment II. (Lanes 1,–4) Probes 1, 9, 10, and 11 were used, respectively. (D) Direct marking for a DNA fragment III. (Lanes 1–4) Probes 1, 9, 10, and 11 were used, respectively. (E) Direct marking for a DNA fragment IV. (Lanes 1–4) probes 1, 9, 10, and 11 were used, respectively. The construction of the target DNA fragments is described in the Methods. (B–E, top) autoradiographic pattern; (bottom) ethidium bromide staining pattern. Nucleotide numbers (nt. numbers) are those registered in GenBank (accession no. X65328). For details, see Methods.

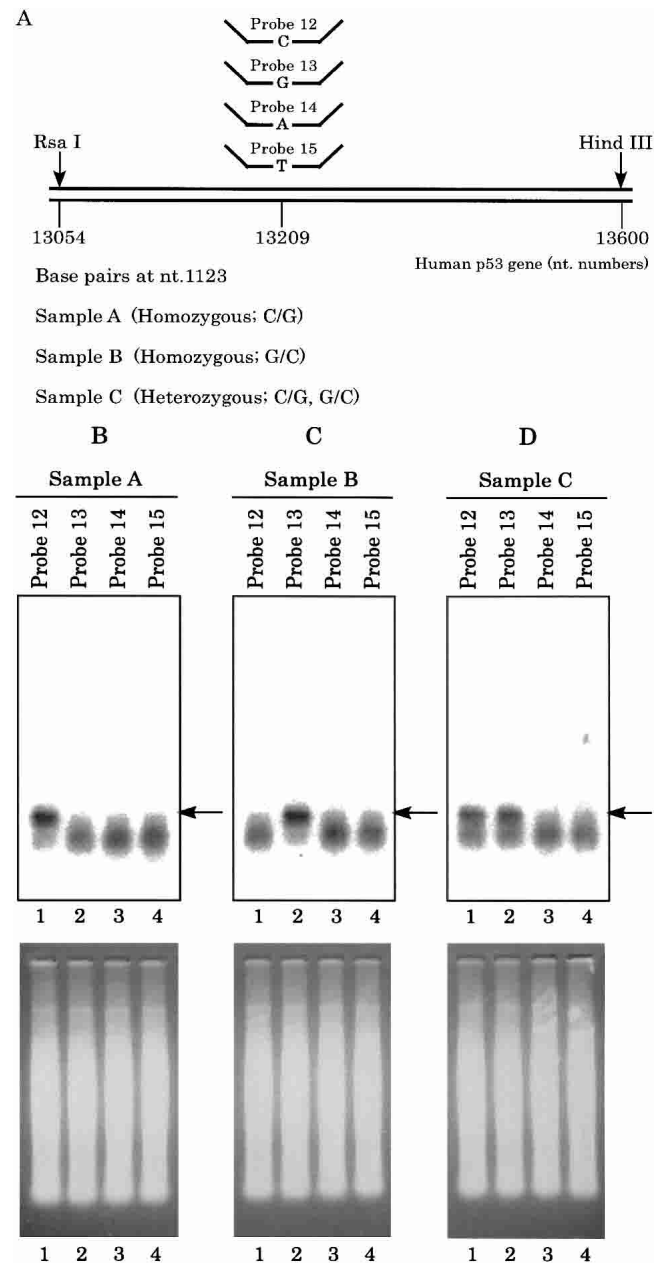


Figure 7. SNP detection. Human genomic DNA samples, with SNP in the exon 5 of p53 gene, were digested with RsaI and HindIII and subjected to direct marking using labeled probing oligonucleotides (40-bp complementary bases) and a splint oligonucleotide (17 bp). (A) A diagrammatic representation of the SNP site in the exon 5 of p53 gene and probing oligonucleotides used. (B) SNP detection for sample A homozygous for the SNP site (C/G). (Lanes 1–4) Probes 12, 13, 14, and 15 were used, respectively. (C) SNP detection for sample B homozygous for the SNP site (G/C). (Lanes 1–4) Probes 12, 13, 14, and 15 were used, respectively. (D) SNP detection for sample C heterozygous for the SNP site (C/G, G/C). (Lanes 1–4) Probes 12, 13, 14, and 15 were used, respectively. (B–D, top) Autoradiographic pattern; (bottom) ethidium bromide-staining pattern. Nucleotide numbers (nt. numbers) are those registered in GenBank (accession no. X54156). The signals are indicated by arrows. For details, see Methods.

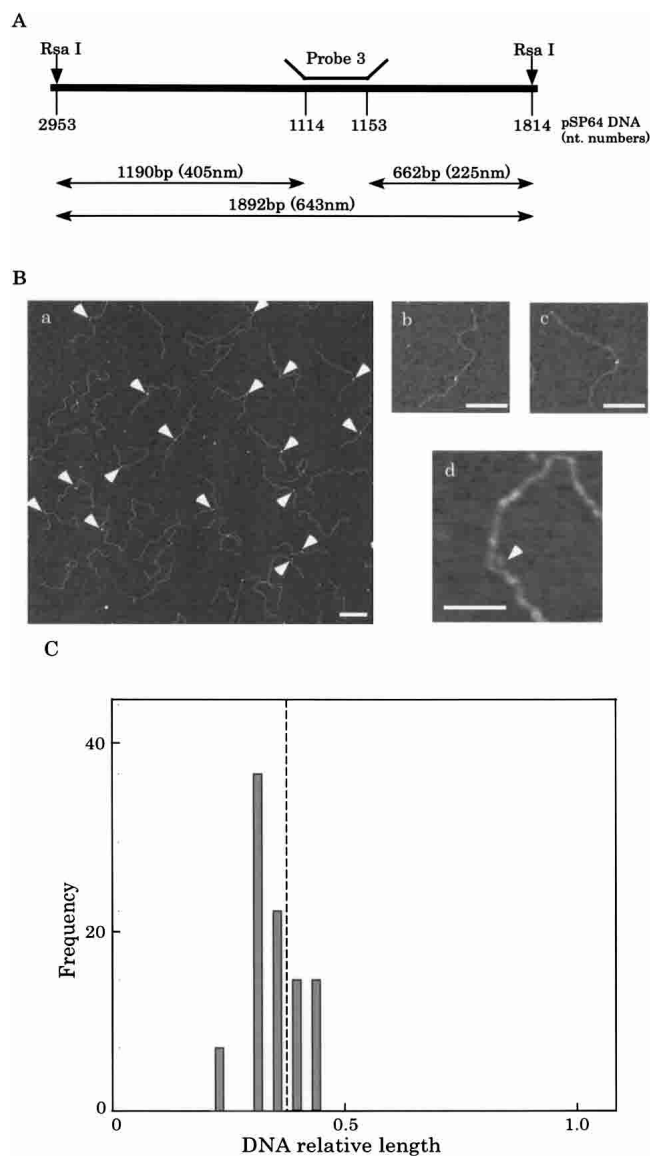


Figure 8. AFM of marked DNA molecules. (A) A diagrammatic representation of the DNA molecules (RsaI-digested pSP64 DNA fragment) that were marked for AFM observation. (B) A typical AFM pattern of the marked DNA molecules. Arrowheads indicate the position of the marked (multistranded) sequence in the DNA molecules. Scale markers indicate 200 nm (a,b,c) and 100 nm (d), respectively. (C) The distribution pattern of the marked sequence as a function of its relative distance from the nearest end of the DNA molecules. The broken line represents the position of the marked region estimated from the location of the sequence in the DNA molecules (see A). A total of 150 molecules were subjected to the analysis. For details, see Methods.

solved in 50 μ L of TE-buffer. Human genomic DNA samples, with SNP in the exon 5 of the p53 gene, were obtained commercially.

Preparation of the target DNA with various molecular sizes

pSP64 DNA fragments with 300, 500, 900, 1200, and 1800 bp were produced by PCR with forward primer 1 and backward primer 1 (for the 300 bp), forward primer 2 and backward primer 1 (for the 500 bp), forward primer 3 and backward primer 1 (for the 900 bp), forward primer 4 and backward primer 1 (for the 1200 bp), and forward primer 5 and backward primer 1 (for the

1800 bp) in a 50 μ L reaction, by use of the following cycling conditions: 25 amplification cycles (98°C for 10 sec, 68°C for 1 min), followed by a final extension at 68°C for 7 min. PCR products were purified by a kit (QIAGEN; QIAquick PCR purification kit). The sequences of the primers used for PCR are listed below. Forward primer 1; 5'-GCCGGGAAGCTAGAGTAAGTAG-3', forward primer 2; 5'-ATCCATAGTTGCCTGACTCCCC-3', forward primer 3; 5'-GGTGGCCTAACTACGGCTACAC-3', forward primer 4; 5'-TCGTGCGCTCTCTGTTCCGACCC-3', forward primer 5; 5'-GCTGCAGGTGACTCTAGAGGA-3' and backward primer 1; 5'-ACTCACCAGTCACAGAAAAGCATC-3'.

Preparation of target DNA (Psp64 DNA) for the detection of mismatched bases

pSP64 DNA fragments with a mismatched base were prepared by PCR using primers with different bases. For DNA fragments II, III, and IV (Fig. 6A), sets of forward primer 3 and backward primer 2, forward primer 3 and backward primer 3, and forward primer 3 and backward primer 4, were used, respectively. The PCR products were ligated with a pSP64 DNA fragment (nucleotides 1125–1814). The ligated products were further amplified by PCR. PCR was performed in a 50- μ L reaction mixture: 25 amplification cycles (98°C for 10 sec, 68°C for 1 min), followed by final extension at 68°C for 7 min. PCR products were purified by a kit (QIAGEN; QIAquick PCR purification kit). The sequences of the primers used for PCR are listed below. Forward primer 3; 5'-GGTGGCCTAACTACGGCTACAC-3', backward primer 2; 5'-TCCACTGAGCCTCAGACCCC-3', backward primer 3; 5'-TCCACTGAGCTTCAGACCCC-3', backward primer 4; 5'-TCCACTGAGCATCAGACCCC-3'.

Marking of specific sequences in double-stranded DNA molecules

In typical experiments marking sequences in pSP64 DNA, a reaction mixture (20 μ L), which contained a restriction enzyme-digested target DNA (200 ng), a probing oligonucleotide (1 pmol, labeled at the 5'-terminus with 32 P by T4 polynucleotide kinase), 3.0 μ g *Escherichia coli* RecA protein or 3.0 μ g *T.th.* RecA protein, 1 mM ATP, 20 mM magnesium acetate, and 30 mM Tris-HCl (pH 7.2), was combined with an equal volume (20 μ L) of a second mixture, which contained a splint oligonucleotide (10 pmol), T4 DNA ligase (50 U), 1 mM ATP, 10 mM MgCl₂, 10 mM dithiothreitol, 25 μ g/mL bovine serum albumin, and 50 mM Tris-HCl (pH 7.5). The mixture was incubated for 60 min at 37°C. After the reactions, the mixtures were subjected to deproteinization by incubating with 20 mM EDTA, 0.5% SDS, and 100 μ g/mL proteinase K (Roche diagnostics) for 30 min at 42°C. The samples were then electrophoresed in 1.2% agarose gel in 1 \times TAE buffer. DNA bands were visualized by staining with ethidium bromide. The gel was dried, autoradiographed, and signals were detected on an imaging plate (Fuji photo film; BAS2000 Image analyzer).

Marking and probing specific sequences in complex genomic DNA

Human genomic DNA was digested with EcoRI, HindIII, and PvuII (Fig. 5) or RsaI and HindIII (Fig. 7). A reaction mixture consisting of the digested DNA (20 μ g), a probing oligonucleotide (10 pmol), which had been labeled at the 5'-terminus with 32 P, 30 μ g *E. coli* RecA protein or 30 μ g *T.th.* RecA protein, 1 mM ATP, and 20 mM magnesium acetate in 200 μ L 30 mM Tris-acetate buffer (pH 7.2) was combined with another mixture consisting of a splint oligonucleotide (100 pmol), 500 U of T4 DNA ligase, 1 mM ATP, 10 mM MgCl₂, 10 mM dithiothreitol, and 25 μ g/mL BSA in 200 μ L 50 mM Tris-HCl buffer (pH 7.5), and incu-

bated for 120 min at 37°C. Originally, we used two successive reactions, one for RecA reaction, the other for DNA ligation. We found, however, the condition described here, successive addition of two reaction mixtures without consecutive incubations, is satisfactory.

The reaction mixture was subjected to deproteinization with 20 mM EDTA, 0.5% SDS, and 100 µg/mL proteinase K for 30 min at 37°C. After removal of unreacted probes by 2× at spin-column chromatography through Sephacryl S-400, the samples were subjected to agarose gel (1.2%) electrophoresis, the gels stained with ethidium bromide, dried, and analyzed on an Image analyzer to detect signals.

Southern hybridization

After electrophoresis, DNA was transferred onto nylon membranes (Pall; Biodyne transfer membrane). Probe DNA (100 ng), produced by PCR, was labeled with [α -³²P]dCTP (Amersham Biosciences; 6000 Ci/mmol) using a DNA labeling kit (Takara-Bio; Random primer DNA labeling kit), and the unincorporated dNTPs removed by spin-column chromatography through Sephadex G-50 (Amersham Biosciences; G-50 spin-column) before use. Hybridization was performed in a hybridization buffer (GIBCO BRL; 2× prehybridization/hybridization solution) for 18 h at 65°C. Filters were washed once with 2× SSC, 0.2% SDS, and twice with 0.2× SSC, 0.2% SDS at 65°C for 30 min. The filters were exposed to an imaging plate to detect hybridization signals. For hybridization probes, we used a ³²P-labeled DNA fragment produced by PCR. The sequences of primers using PCR are listed below. Forward primer 6; 5'-CTGTATTTGAACTTTTCAGA-3', backward primer 5; 5'-CAGCAGTTTGAGACCAGCCTGG-3', forward primer 7; 5'-CAACCTCCTAGCAAAGTAGGGA-3', backward primer 6; 5'-CACCATCTTTGTTCAGGCTGGTC-3'.

Atomic force microscopy (AFM)

Double-stranded DNA molecules (1892-bp fragment of pSP64), in which a 40-bp specific sequence was marked were purified by using a kit (Amersham Biosciences; GFX PCR and Gel Band Purification Kit) according to the manufacturer's instruction. The DNA samples (50 µL) were suspended in a solution of 1 mM NiCl₂ and 10 mM HEPES-NaOH (pH 8.0), dropped onto a freshly cleaved mica, and incubated for 5 min at room temperature to electrostatically immobilize them onto the surface. After rinsing with 5 mL of Milli-Q water, the samples were dried and observed under AFM. AFM imaging was performed using a Nanoscope IIIa (Veeco Instruments) in tapping mode in air. A silicon cantilever NCH (NanoWorld) was used for observation. A series of 150 DNA molecules were scored for AFM. Lengths of DNA fragments were

measured using a computer program specially developed by the Research Institute of Biomolecule Metrology.

Acknowledgments

We thank Drs. O. Ohara, K. Okumura, and K. Kondo for their helpful discussions.

References

- Demidov, V.V. and Frank-Kamenetskii, M.D. 2001. Sequence-specific targeting of duplex DNA by peptide nucleic acids via triplex strand invasion. *Methods* **23**: 108–122.
- Demidov, V.V., Kuhn, H., Lavrentieva-Smolina, I.V., and Frank-Kamenetskii, M.D. 2001. Peptide nucleic acid-assisted topological labeling of duplex dna. *Methods* **23**: 123–131.
- Fujiwara, J. and Oishi, M. 1998a. Direct probing: Covalent attachment of probe DNA to double-stranded target DNA. *Nucleic Acids Res.* **26**: 5728–5733.
- . 1998b. Direct cloning by covalent attachment of probe DNA to target DNA. *Nucleic Acids Res.* **26**: 5734–5737.
- Geron-Landre, B., Roulon, T., Desbiolles, P., and Escude, C. 2003. Sequence-specific fluorescent labeling of double-stranded DNA observed at the single molecule level. *Nucleic Acids Res.* **31**: e125.
- Nilsson, M., Malmgren, H., Samiotaki, M., Kwiatkowski, M., Chowdhary, B.P., and Landegren, U. 1994. Padlock probes: Circularizing oligonucleotides for localized DNA detection. *Science* **265**: 2085–2088.
- Potaman, V.N., Lushnikov, A.Y., Sinden, R.R., and Lyubchenko, Y.L. 2002. Site-specific labeling of supercoiled DNA at the A+T rich sequences. *Biochemistry* **41**: 13198–13206.
- Rao, B.J. and Radding, C.M. 1993. Homologous recognition promoted by RecA protein via non-Watson-Crick bonds between identical DNA strands. *Proc. Natl. Acad. Sci.* **90**: 6646–6650.
- Rao, B.J., Dutreix, M., and Radding, C.M. 1991. Stable three-stranded DNA made by RecA protein. *Proc. Natl. Acad. Sci.* **88**: 2984–2988.
- Rao, B.J., Chiu, S.K., and Radding, C.M. 1993. Homologous recognition and triplex formation promoted by RecA protein between duplex oligonucleotides and single-stranded DNA. *J. Mol. Biol.* **229**: 328–343.
- Rice, M.C., Heckman, B.M., Liu, Y., and Kmiec, E.B. 2004. Fluorescent detection and isolation of DNA variants using stabilized RecA-coated oligonucleotides. *Genome Res.* **14**: 116–125.
- Roulon, T., Coulaud, D., Delain, E., Le Cam, E., Helene, C., and Escude, C. 2002. Padlock oligonucleotides as a tool for labeling superhelical DNA. *Nucleic Acids Res.* **30**: e12.
- West, S.C., Cassuto, E., and Howard-Flanders, P. 1981. recA protein promotes homologous-pairing and strand-exchange reactions between duplex DNA molecules. *Proc. Natl. Acad. Sci.* **78**: 2100–2104.
- Zain, R., Marchand, C., Sun, J., Nguyen, C.H., Bisagni, E., Garestier, T., and Helene, C. 1999. Design of a triple-helix-specific cleaving reagent. *Chem. Biol.* **6**: 771–777.

Received May 13, 2004; accepted in revised form September 16, 2004.