Direct probing: covalent attachment of probe DNA to double-stranded target DNA

Jun Fujiwara and Michio Oishi

Biotechnology and Medical Engineering Division, Aisin Cosmos R&D Co. Ltd, 5-2-11 Sotokanda, Chiyoda-ku, Tokyo 101, Japan and 1Kazusa DNA Research Institute, 1532-3, Yana, Kisarazu, Chiba 292, Japan

Received August 18, 1998; Revised and Accepted October 27, 1998

ABSTRACT

We report a novel procedure, which can be applied to probing of specific DNA, for covalently attaching probe DNA to complementary sequences in double-stranded target DNA. Employing hairpin-like oligonucleotide probes in combination with successive use of recA protein and DNA ligase, probes can be attached directly to target DNA molecules without dissociation of the DNA. The hairpin-like structure of the probes was designed so that the terminus of the probe oligonucleotide can be brought into close stereochemical proximity to the terminus of the complementary strand of target DNA for ligation. Because of the elimination of the DNA dissociation and subsequent hybridization (and washing) steps in the currently employed method, the probing process has become greatly simplified and more efficient and may lead to development of fully automated probing systems.

INTRODUCTION

Since bases are paired inwardly in double-stranded DNA molecules, recognition of the specific base sequences that are required in fundamental cellular processes such as replication, repair, recombination and transcription is mediated by enzymes specific to these processes. This enzymatic recognition is considered to be very accurate and extremely efficient. On the other hand, recognition of specific base sequences in most molecular biological experiments is carried out with single-stranded DNA molecules after dissociation of double-stranded DNA in order to expose base sequences. DNA–DNA hybridization, one of the most reliable techniques for recognition and also quantification of specific base sequences in a complex genome, has been used widely for the analysis of DNA, often complementing PCR, which excels in convenience and sensitivity. The current methods of base recognition with single-stranded DNA as in DNA–DNA hybridization, however, are quite cumbersome and time consuming as they require DNA dissociation, transfer of the dissociated DNA to membranes, lengthy incubation for hybridization and subsequent extensive washing of the membranes to remove non-specific DNA. Direct recognition of specific base sequences in double-stranded DNA molecules without DNA dissociation, once established, would greatly simplify most of the current procedures for recognition of specific base sequences by DNA–DNA hybridization.

One way to approach the direct recognition of base sequences is to make use of triplex DNA formation. At present, the only triplex DNA structures stable enough for most practical purposes are those formed through Hoogsteen base pairing (1). In this case, however, the target base sequences are limited to polypurine and polypyrimidine (2–4), thus making it quite impractical to utilize this method as a general means for direct probing or cloning of double-stranded DNA. On the other hand, it has long been established that recA protein, probably the most extensively investigated recombination enzyme, recognizes base sequences in double-stranded DNA and mediates a reaction in which one of the target double-stranded DNA strands is displaced by incoming homologous single-stranded DNA (5–8). The newly paired DNA in the displaced region, however, is not stable enough to withstand manipulations employed for most gene analysis, particularly when recA protein is removed and strands are displaced by short deoxyoligonucleotides (9).

Employing hairpin-like oligonucleotide probes and with successive use of recA protein and DNA ligase, we have been able to demonstrate that deoxyoligonucelotide probes can be covalently attached to target double-stranded DNA molecules without dissociation of DNA. In this article, we present the principles and basic characteristics of the procedure and the actual results of direct probing of specific DNA fragments.

MATERIALS AND METHODS

Deoxyoligonucleotide probes

The deoxyoligonucleotide probes were custom synthesized by Kurabo Co. (Osaka) and Sawady technology (Tokyo). The sequences are listed below, in which the hairpin sequence (underlined) and mismatched bases (bold) are emphasized. TRI-20-63, 5'-GTCATGCAATCGCTAGATGCTTTTCTGTGTACCTGTTG- GTAAGTCCGCGCCGCGGTTCGTTCCCGCGCGCCGC-3'; TRI-20-62, 5'-GTATGCAATCGCTAGATGCTTTTCTGTGTACCTGTTG- GTAAGTCCGCGCCGCGGTTCGTTCCCGCGCGCCGC-3'; TRI-20-64, 5'-GTATGCAATCGCTAGATGCTTTTCTGTGTACCTGTTG- GTAAGTCCGCGCCGCGGTTCGTTCCCGCGCGCCGC-3'; TRI-20-50, 5'-GTCATGCAATCGCTAGATGCTTTTCTGTGTACCTGTTG- GTAAGTCCGCGCCGCGGTTCGTTCCCGCGCGCCGC-3'; TRI-20-55, 5'-GTCATGCAATCGCTAGATGCTTTTCTGTGTACCTGTTG- GTAAGTCCGCGCCGCGGTTCGTTCCCGCGCGCCGC-3'.
AGTGCGCGCCGGC-3'; TRI-10-63, 5'-ATCCGCTAACTATC-GTCTTAGTACCAACCGTAAGACAGCGGCGCCGGCT-TTCCGGCGGCGC-3'; TRI-20-63M1, 5'-GTCATGCATCCCT-TAAGATAGTTTTCCTGGTACAGTGAGTCCGGGCGCGG-TTCCGGCGGGC-3'; TRI-20-63M2, 5'-GTCATGCCATAC-GTAAGATAGTTTCCTGTAGCTGGTACAGTGAGTCCGGGCGCG- GTTTTCCGGGGCGC-3'; TRI-20-63M4, 5'-GTCATGCATC-ATACGTAAGATAGTTTTCCTGTAGCTGGTACAGTGAGTCCGGGCGG- CCGGTTTCCGGGGCGC-3'; TRI-20-63M5, 5'-GTCATGCAT- CAATGTAAGATAGTTTTCCTGTAGCTGGTACAGTGAGTCCGGG- CCCGGTTTCCGGGGCGC-3'; TRI-21-63, 5'-GTTATCTA- CAGAATCAATCGGGAATACGGAAGAAGACATGGGCGG-CGCCGTTTCCGGGGC-3'; TRI-22-63, 5'-TTCAACC-CGGTAAAGACAGCTAATGGCACAGCCATGGGCGG-CGCCGTTTCCGGGGC-3'; TRI-PA-20-63, 5'-GCCGC- CGCGGTTCGCCGGCGCGCAGTCAACGGTACAAGAAAGACATCCAT- GGCCGTTTCCGGGGCGG-3'; P53-1-63, 5'-CAAG- ATGTGTTTGCCAAACTGGCAAGACCTTGCCCTGTCAGGGCGG-CGCGGGTTTCCGGGGCGC-3'; P53-3-63, 5'-GGGTT- TGGGTCGGGGTGGTGGCGCGCCCTTCCAAATGGACCG-GCGGCCGCGGTTTCCGGGGC-3'.

DNA purification

pBluescript DNA was extracted by alkaline procedure (10) from *Escherichia coli* (XL1-Blue) which had been cultured in Terrific broth (10) containing ampicillin (50 µg/ml). DNA in the alkaline supernatant was further purified by CsCl centrifugation (85 000 r.p.m. for 16 h), which was followed by Sephacryl S-400 chromatography after dialysis against TE buffer. The DNA was precipitated with ethanol and dissolved in TE buffer at a DNA concentration of 200 µg/ml. *Drosophila* and human DNA were purchased from Promega.

Direct probing

For direct probing of pBluescript DNA, deoxyoligonucleotide probes (5 ng) labeled at the 5'-terminus with $^{32}$P using T4 polynucleotide kinase (Megalab; Takara) were incubated in a reaction mixture (5 µl) which contained recA protein (0.48 µg; Epicentre Technologies or Pharmacia Biotech), ATP-γ-S (4.8 mM; Sigma), magnesium acetate (2.5 mM) and Tris–acetate (30 mM, pH 7.2), for 15 min at 37°C. A ScaI digest of double-stranded DNA (100 ng, pBluescript DNA; Stratagene) in a buffer (5 µl) containing magnesium acetate (22.5 mM), ATP-γ-S (4.8 mM) and Tris–acetate (30 mM) was then added and incubated for another 30 min. Heat-stable DNA ligase (20 µl, Ampligase; Epicentre Technologies) in 10 µl of a buffer (10× RXN buffer diluted 5-fold with redistilled water, provided by the supplier) was then added and the reaction mixture was incubated for 15 min at 65°C. The reaction was terminated by adding 1.8 µl of EDTA (110 mM)/SDS (5.6%) solution, which was followed by incubation for 15 min at 37°C after addition of protease K (1 µl, 20 mg/ml). DNA was precipitated with ethanol and dissolved in TE buffer (10 µl). The samples were then subjected to neutral or alkaline (0.05 M NaOH, 1 mM EDTA) agarose gel (0.8%) electrophoresis (120 V, 2 h or 20 V, 5 h, respectively), the gels stained with ethidium bromide and autoradiographed (5 h at –80°C) using Fuji X-ray film (RX).

For probing of specific DNA fragments in the fruit fly (*Drosophila melanogaster*) or human genomes, higher concentrations of probes, target DNA and recA (25 ng, 20 µg and 7.84 µg, respectively) were employed and excess unreacted labeled oligonucleotide probes were removed before electrophoresis through a Sephacyr S-400 spin column to reduce background noise (legend to Fig. 7).

RESULTS

Principle of the procedure

The procedure is outlined diagrammatically in Figure 1. The oligonucleotide probe specifically designed for this purpose consists of two parts; a probing sequence complementary to the terminal sequence of target DNA and a universal hairpin structure which makes it possible to covalently link the probe to the target DNA. In the first step, the homologous sequence present at the terminus of target DNA is recognized and one of the strands is displaced by the complementary sequence of the oligonucleotide probe through a recA protein-mediated reaction. Second, the probe is covalently attached to the target DNA by DNA ligase. The hairpin-like structure of the probe brings the 5'-terminus of the probe oligonucleotide into close stereochemical proximity to the 5'-terminus of the complementary strand of target DNA for ligation. In the actual process, the labeled oligonucleotide probe was preincubated with recA protein, then incubated with target DNA to form a probe–target DNA–recA protein complex with a partial triplex structure. DNA ligase is then added to covalently attach the probe to the target DNA. After this reaction, the proteins are removed and the products can then be subjected to further analyses such as gel electrophoresis to detect hybridized DNA fragments. The hairpin probes can also be designed to probe the opposite complementary sequence at the same terminus, other terminal sequences of the same fragment, or both (below).

Direct probing

In Figure 2, we show the results of a model experiment in which a double-stranded DNA [pBluescript SK(−)], a derivative of pBR322] was probed by a 5'-labeled hairpin deoxyoligonucleotide (TRI-20-63), which consists of a 40mer base sequence complementary to the 3'-terminal sequence of a ScaI digest of Bluescript (a derivative of pBR322; Stratagene) DNA attached to a 23mer hairpin–like structure which in turn consists of a 20mer base paired (10 bp) region plus a 3mer hinged region (Fig. 1). The probe was first incubated with recA protein (15 min at 37°C, mixed with target DNA and incubated further (30 min at 37°C). A heat-stable ligase (Ampligase) was added and incubation was continued for 15 min at 65°C. The recA protein and ligase in the reaction mixture were then removed by proteinase K (or phenol) treatment and the samples were autoradiographed after electrophoresis. As seen in Figure 2A, only when both recA protein and DNA ligase were present in the reaction was a signal detected at the position of 3 kb where double-stranded Bluescript DNA is expected to migrate. The signal was still present even after the sample was subjected to alkaline electrophoresis (Fig. 2B), indicating that the probe was covalently attached to the target DNA molecule. Due to elimination of the DNA dissociation steps
and use of recA protein as well as heat-resistant DNA ligase, the entire reaction can be completed in just 1 h, only a fraction of the time required for similar procedures currently in use. Because we employed recA protein (and DNA ligase), complex formation between probe and target DNA was quite efficient as we estimate that >60% of target DNA was associated with probe (measured with a Fuji BAS 2000 image analyzer) under the conditions we employed (data not shown). As expected, signal intensities increased proportionally as the quantities of target DNA were increased (data not shown), suggesting that simple quantitation of target DNA is possible as it is in Southern hybridization.

**Probing by various hairpin probes**

Figure 3 shows the results of probing the same target DNA with various hairpin probes. First, no signals were detected when the target DNA was pretreated before the reaction with bacterial alkaline phosphatase to eliminate the terminal (5′) phosphate which is necessary for ligation of the probe to target DNA (lane 2). Likewise, probes that were one base short (TRI-20-62) or had one extra base (TRI-20-64) at the 3′-terminus did not produce a signal (lanes 3 and 4, respectively). A DNA probe (TRI-20-50) similar to the standard probe (TRI-20-63) but with a DNA sequence which does not form a hairpin structure also failed to produce the signal (lane 5). Another probe (TRI-10-63), which carries a 40mer sequence complementary to those located within (nt 2487–2526) the ScaI fragment rather than to the terminal sequence as in TRI-20-63, also did not produce a signal (lane 6). Furthermore, no signals were detected when EcoRV-digested Bluescript DNA, in which the target sequence is located inside, not at the terminus, was employed as a target DNA with the standard probe (TRI-20-63) (lane 7). From these results, it is clear that the signal is produced only under conditions in which ligation can be achieved between the 3′-terminus of the hairpin probe and 5′-terminus of the target DNA.

In a separate series of experiments, fidelity of the recA protein-mediated reaction in recognizing homologous sequences in target DNA was examined using hairpin probes with varying numbers of mismatched bases. As seen in Figure 4, hairpin probes with one or two mismatched bases in the 40mer complementary

\( \text{Figure 1. Diagrammatic outline of direct probing of specific base sequences in double-stranded DNA by hairpin-like oligonucleotide probes. Target DNA and deoxyoligonucleotide probe are shown in blue and red, respectively. Conversion between the two structures (III and IV) is highly speculative and therefore IV is shown in parentheses.} \)

\( \text{Figure 2. Direct probing of pBluescript DNA: effects of recA protein and DNA ligase. A deoxyoligonucleotide probe (TRI-20-63, 5 ng) labeled at the 5′-terminus with {\textsuperscript{32}}P by T4 polynucleotide kinase (Megalabel; Takara) was incubated in a reaction mixture (5 µl) which contained recA protein (0.48 µg; Pharmacia Biotech), ATP-γ-S (4.8 mM, Sigma), magnesium acetate (2.5 mM) and Tris-acetate (30 mM, pH 7.2), for 15 min at 37°C. A ScaI digest of double-stranded DNA (100 ng, pBluescript DNA; Stratagene) in a buffer (5 µl) containing ATP-γ-S (4.8 mM), magnesium acetate (22.5 mM) and Tris-acetate (30 mM) was then added and further incubated for another 30 min. Heat-stable DNA ligase (20 U, Ampligase; Epicentre Technologies) in 10 µl of a buffer (10× RXN buffer diluted 5-fold with redistilled water) provided by the supplier was then added and the reaction mixture was incubated for 15 min at 65°C. The reaction was terminated by adding 1.8 µl of EDTA (110 mM)/SDS (5.6%) solution which was followed by incubation for 15 min at 37°C after addition of proteinase K (1 µl, 20 mg/ml). DNA was precipitated with ethanol and dissolved in TE buffer (10 µl). The samples were then subjected to neutral or alkaline (0.05 M NaOH, 1 mM EDTA) agarose gel electrophoresis (120 V, 2 h or 20 V, 5 h, respectively), the gels stained with ethidium bromide and autoradiographed (5 h at –80°C) using Fuji X-ray film (RX). (A) Autoradiographic patterns after neutral agarose gel electrophoresis. (B) Autoradiographic patterns after alkaline agarose gel electrophoresis. (C) Ethidium bromide staining patterns after neutral agarose gel electrophoresis. (A–C) Lane 1, without recA protein and DNA ligase; lane 2, without DNA ligase; lane 3, without recA protein; lane 4; control (complete); lane M, DNA size markers (λ HindIII digests) with their approximate molecular sizes in kb. Details in Materials and Methods.} \)
sequences could be ligated to target DNA without any appreciable loss of efficiency (lanes 2 and 3) compared with control probe (lane 1). The efficiency, however, was significantly reduced when three or four mismatched bases were introduced into the probe (lanes 4 and 5) and no signal was detected when five mismatched bases were introduced (lane 6). Thus, recA protein recognizes homologous base sequences with at least two mismatched bases in the 40mer terminal sequence with seemingly equal efficiency as it does completely matched sequences.

In the model experiments described above, we employed target DNA with blunt ends (ScaI-treated Bluescript DNA). As seen in Figure 5, target DNA molecules with other forms of termini such as 5' protruding (AflIII-treated DNA) (lane 1) or 3' protruding (AlwNI-treated DNA) (lane 3) can also serve as targets for ligation with almost equal efficiencies as observed for DNA with blunt ends (lane 2). This indicates that DNA fragments produced by various restriction enzymes can be subjected to direct probing.

Next, probing the complementary sequence of the opposite strand at the same terminus was examined. As seen in Figure 6, ligation of a hairpin probe (TRI-PA20-63) to the 3'-terminus of the opposite strand was equally effective (lane 2) as that to the 5'-terminus of the target DNA (lane 1) with TRI-20-63 as shown above.

It should be noted that despite many attempts we failed to covalently attach the hairpin probe to the termini of complementary single-stranded DNA after conventional hybridization of probe to target DNA. The reason for the difficulty in ligating the probe to single-stranded DNA and not to double-stranded DNA as described above is not fully understood at the present time but is currently under investigation.
Probing a specific sequence in genomic DNA

We examined whether a specific sequence in genomic DNA can be directly probed by hairpin probes. First, one of the homeobox genes (fushi tarazu, ftz) in D. melanogaster (11), which is present as a single copy in a genome of ~180 Mb, was subjected to this investigation. Whole Drosophila genomic DNA was digested with FspI, which produces a 1753 bp ftz gene fragment, and the digest was subjected to direct probing using a hairpin probe (FTZ-1-63) in which a 40mer sequence is complementary to the 3'-terminal sequence of the FspI fragment. As seen in Figure 7A, a band of ~1.75 kb corresponding to the FspI fragment of the ftz gene in size was observed (lane 1, indicated by an arrow). No signals were detected when EcoRV-digested Drosophila DNA in which the sequence corresponding to the probe is not located at the 3'-terminus was used as target DNA (lane 2).

Similar experiments were performed to probe a single copy DNA fragment (p53 gene) in the human genome (~3000 Mb), which is 17 times more complex than the Drosophila genome. Human genomic DNA was digested with BamHI and PvuII, which produces a 1312 bp p53 gene fragment (12), and the digest was subjected to direct simultaneous probing with two hairpin probes in which the respective 40mer sequences are complementary to the 3'-terminal (P53-1-63) and 5'-terminal sequences (P53-3-63) of the p53 fragment. As seen in Figure 7B, after eliminating most of the unreacted labeled oligonucleotides before electrophoresis (legend to Fig. 7) to reduce background noise and thus to detect a single gene segment in a complex genome by an oligonucleotide probe, we were able to detect a band corresponding to the targeted fragment (lane 1, indicated by an arrow). No signals were detected with a control DNA sample (EcoRV-digested DNA), in which the sequences corresponding to the probes are not located at either terminus (lane 2). These results strongly suggest that direct probing of specific base sequences can be achieved even in highly complex genomes such as the Drosophila and human genomes.

DISCUSSION

In this paper we have described a novel procedure which can be applied to probing of specific DNA sequences. By employing a hairpin-like probe, it has become possible to covalently attach...
probe DNA directly to double-stranded target DNA. Elimination of DNA dissociation and subsequent hybridization (and washing) have made the entire process of probing streamlined and efficient, achievable within a fraction of the time required for similar procedures currently in use such as Southern hybridization, while maintaining high probing efficiency. In addition, direct probing has proven to have other more specific advantages over conventional procedures. For example, when one probes a single sample with multiple probes, repeated hybridization (rehybridization) is no longer necessary as probing can be carried out simultaneously in separate tubes for separate probes and samples can be electrophoresed in parallel. This may be particularly advantageous in non-PCR-based polymorphism analysis using multiple probes. Our novel procedure makes it easier to detect a single copy DNA fragment in highly complex genomes using oligonucleotide probes than with current procedures, as shown above in probing the human p53 gene fragment. We attributed this to (i) the high efficiency of recA protein in finding complementary sequences and (ii) removal of unreacted labeled probes before electrophoresis, both of which are unique to this procedure.

In the procedure described, base sequences located at the terminus of the target DNA are recognized and only sequences in that region can be probed. Because of this feature, ambiguous hybridization products derived from illegitimate hybridization with similar but not identical sequences at other locations, which are often encountered in current procedures and require re-examination by other means, should essentially be eliminated. On the other hand, since only sequences at the termini of restriction fragments can be recognized, one may encounter problems when sequences to be recognized cannot be placed at the termini for some reason or when the exact location of the sequence is not known.

The unique characteristics of the reported procedure seem to have virtually eliminated the anticipated obstacles for construction of fully automated probing equipment, which has yet to be realized despite a great potential need. Needless to say, this procedure could be exploited even more effectively by combining it with other molecular biological techniques such as PCR, depending upon the purposes of the experiments involved.

ACKNOWLEDGEMENTS

We thank Drs Y. Shigemori, K. Okumura and T. Yamamoto for their helpful discussions. This work was supported by the R & D Project of the Innovative Technology for the Earth Program which is sponsored by NEDO (New Energy and Industrial Technology Development Organization).

REFERENCES