Nucleic Acids Research, Vol. 20, No. 3 595–600

Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA

Douglas H.Jones and Stanley C.Winistorfer
Department of Pediatrics, University of Iowa College of Medicine, Iowa City, IA 52242, USA

Received June 4, 1991; Revised and Accepted December 20, 1991

ABSTRACT

We present a novel method for the PCR amplification of unknown DNA that flanks a known segment directly from human genomic DNA. PCR requires that primer annealing sites be present on each end of the DNA segment that is to be amplified. In this method, known DNA is placed on the uncharacterized side of the sequence of interest via DNA polymerase mediated generation of a PCR template that is shaped like a pan with a handle. Generation of this template permits specific amplification of the unknown sequence. *Taq* (DNA) polymerase was used to form the original template and to generate the PCR product. 2.2 kb of the β-globin gene, and 657 bp of the 5′ flanking region of the cystic fibrosis transmembrane conductance regulator gene, were amplified directly from human genomic DNA using primers that initially flank only one side of the region amplified. This method will provide a powerful tool for acquiring DNA sequence information.

INTRODUCTION

The polymerase chain reaction (PCR) is a method which permits the amplification and sequencing of specific DNA sequences *in vitro* without cloning (1–3). PCR requires primer annealing sites in each end of the target sequence in order for amplification to occur, and until recently PCR has required knowledge of the sequences initially flanking the target sequence. Consequently, PCR could not be used to amplify and sequence unknown DNA flanking a known sequence. Recently, several strategies have been developed to overcome this limitation of PCR, but none has found widespread application in the direct amplification of unknown flanking DNA from very complex mixtures, such as the genomic DNA of higher eukaryotes (4–14).

This paper describes a method that permits the rapid retrieval and sequencing of unknown DNA that flanks a known site, such that one can 'walk' into an uncharacterized region of DNA without cloning. This method eliminates laborious cloning steps and associated cloning artifacts in obtaining DNA sequence information from complex mixtures, such as human genomic DNA. Thus, this method overcomes major obstacles in the human genome project and has a myriad of applications in molecular biology, including: determination of viral integration sites, determination of transposon integration sites, amplification of fragments adjacent to cDNA such as regulatory regions and intron-exon junctions, generation of yeast artificial chromosome (YAC) (15) endpoints, and chromosome jumping (16,17).

This method requires the generation of a template shaped like a pan with a handle, so the method is termed panhandle PCR. The template is generated by restriction enzyme digestion of genomic DNA followed by ligation to a single-stranded oligonucleotide. This ligated-oligonucleotide has a free 3′ end. This 3′ end is complementary to the known region of DNA. In this method, single-strands of DNA which contain the complement of the ligated-oligonucleotide undergo self-annealing, forming a stem-loop structure. The ligated oligonucleotide can then prime template-directed DNA polymerization. PCR amplification of the unknown DNA can subsequently be carried out because known sequence now flanks both ends of the unknown DNA.

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotides were synthesized on an Applied Biosystems 391 DNA synthesizer (Foster City, CA), desalted over a Sephadex G25 column, dried, and suspended in H_2O. Aliquots were kept at −20°C.

Oligonucleotide phosphorylation

Two μg of the oligonucleotide were incubated with 10U T_4 polynucleotide kinase (New England BioLabs, Beverly, MA) in kinase buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 5 mM Dithiothreitol) at 37°C×30 min. T_4 polynucleotide kinase was subsequently inactivated by heating at 68°C×10 min, and the oligonucleotide was then aliquoted and stored at −20°C until use.

Restriction enzyme digestion followed by Calf Intestinal Alkaline Phosphatase treatment or partial fill-in by Klenow fragment (Step 1)

Five μg of human genomic DNA (Clontech, Palo Alto, CA) was digested with 40U *Bam* HI (New England BioLabs), 20U *Avr* II (New England BioLabs), or 30U *Hind* III (Boehringer Mannheim Biochemicals, Indianapolis, IN) in 100 μl for two hours. The digested genomic DNA was either treated by the addition of 0.05U of calf intestinal alkaline phosphatase in 5 μl (Boehringer Mannheim Biochemicals, Indianapolis, IN) with incubation at 37°C×30 minutes, or underwent a partial fill-in reaction by the addition of 1 μl each of 10 mM dCTP, dATP,

dTTP and 5 U Klenow fragment in 2.5 μl (Boehringer Mannheim Biochemicals) with incubation at 23°C×30 minutes (18). Following either treatment, the DNA underwent glass bead extraction using Geneclean (BIO 101, La Jolla, CA), and was suspended in 50 μl TE (10 mM Tris-HCl pH 8, 1 mM EDTA). Five 5 μl aliquots were frozen for later use as non-oligonucleotide-ligated template controls.

Ligation of phosphorylated oligonucleotide (Step 2)

The remaining 25 μl of genomic DNA was ligated to a previously 5′ phosphorylated 30−35 nucleotide long single-stranded oligonucleotide whose 5′ end is complementary to the single-stranded ends of restriction enzyme digested genomic DNA. The genomic DNA was ligated with fifty-fold molar excess of the phosphorylated oligonucleotide in T_4 DNA ligase buffer (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 0.5 mM ATP, 10 mM Dithiothreitol) using one Weiss U T_4 DNA ligase (Boehringer Mannheim Biochemicals) at 23°C×4 hours. The 30−33 nucleotide long 3′ region of this oligonucleotide was complementary to the known region of genomic DNA. The ligation mixture underwent Geneclean purification and was suspended in 25 μl TE.

For amplification of β-globin sequences, GATCTTCTCTGTCTCCACATGCCCAGTTTCTATTG was ligated to *Bam* HI digested and calf intestinal alkaline phosphatase treated genomic DNA or to *Avr* II digested and partially filled in genomic DNA, or CTAGTTCTCTGTCTCCACATGCCCAGTTTCTATTG was ligated to *Avr* II digested and calf intestinal alkaline phosphatase treated genomic DNA. For amplification of the cystic fibrosis transmembrane conductance regulator (CFTR) sequence, AGCTTGAGCCCAGACGGCCCTAGCAGGGAC was ligated to *Hind* III digested and calf intestinal alkaline phosphatase treated genomic DNA.

Panhandle formation (Step 3)

A 25 μl aliquot of 2×PCR mix (1.25U *Taq* (DNA) polymerase (AmpliTaq, Perkin-Elmer Cetus, Norwalk CT), 100 mM KCl, 20 mM Tris-HCl pH 8.3, 3 mM $MgCl_2$, 0.02% w/v gelatin, 400 μM each dNTP), pre-aliquoted and stored at −20°C, was thawed, 10 μl H_2O was added, and 50 μl mineral oil was layered on top. The tube was pre-heated to 80°C prior to addition of 2 μl of template (diluted in total 10 μl for ease of handling), in order to prevent nonspecific annealing and polymerization (19). The mixture underwent one thermal cycle of denaturation, annealing and polymerization on a thermal cycler (Perkin Elmer Cetus) followed by a transition to 80°C soak. Since the genomic DNA concentration was < 4 ng/μl, the denaturation and reannealing steps result in intra-strand annealing of the ligated synthetic oligonucleotide to its complementary sequence in the genomic DNA (5). This is followed by polymerase extension of the recessed 3′ end. Two separate tubes, one containing the control template (genomic DNA that had been digested but not ligated to the phosphorylated oligonucleotide), and the other containing a reagent control (no DNA), were processed concurrently in all experiments.

For self-annealing to the β-globin gene the following parameters were used: 94°C×1 min, a two minute transition to 60°C×30 sec, then a rapid transition to 80°C. For self annealing to the CFTR gene the following parameters were used: 94°C×1 min, a two minute transition to 72°C×30 sec, then a rapid transition to 80°C.

Set up of initial PCR amplification (Step 4)

To each of the three tubes, 12.5 pmoles of each primer (1 and 2), in a total volume of 5 μl H_2O, was added under the mineral oil while the tubes remained in the heat block at 80°C, resulting in a final concentration of 0.25 μM each primer with 200 μM each dNTP.

β-globin primer 1, AGGCCCTGGGCAGGTTGGTATC; primer 2, TCCTCTTGGGTTTCTGATAGGCACTGAC. (5′ nucleotides not complementarity to β-globin underlined). CFTR primer 1, TTTGGAGACACCGCTGGCCTTTTC; primer 2, GTAATGCCAAAGACCTACTACTCTGGGTGC. (5′ nucleotide not complementarity to CFTR underlined).

Initial PCR amplification (Step 5)

All three tubes underwent 30 PCR amplification cycles followed by a final extension at 72°C×7 min and then came to an 80°C soak.

Thermal cycling parameters used in individual experiments are as follows: β-globin DNA amplification following *Bam* HI digestion, calf intestinal alkaline phosphatase treatment, oligonucleotide ligation: 94°C×30 sec, 56°C×30 sec, 72°C×30 sec. β-globin DNA amplification following *Avr* II digestion, calf intestinal alkaline phosphatase treatment, oligonucleotide ligation: 94°C×30 sec, 56°C×30 sec, 72°C×2 min. β-globin DNA amplification following *Avr* II digestion, partial fill-in reaction, oligonucleotide ligation: 94°C×30 sec, 56°C×30 sec, 72°C×2 min. CFTR DNA amplification following *Hind* III digestion, calf intestinal alkaline phosphatase treatment, oligonucleotide ligation: 94°C×30 sec, 60°C×30 sec, 72°C×1 min.

Set up of Nested PCR amplification (Step 6)

One μl of the unpurified amplified PCR product was removed from each tube by inserting a long thin pipette tip through the mineral oil layer, and placed in a corresponding second set of PCR tubes containing nested primers (3 and 4) preheated to 80°C with the same enzyme, reagents and primer concentrations as in the first PCR amplification.

β-globin primer 3, CAAGGTTACAAGACAGGTTTAAGGAGAC; primer 4, CGTCTCTCTGCCTATTGGTCTATTTTCC. (5′ nucleotides not complementarity to β-globin underlined). CFTR primer 3, AGGCGACCTCTGCATGGTCTCTC; primer 4, GCTGCCGCTCAACCCTTTTTCTCTG. (5′ nucleotide not complementarity to CFTR underlined).

Nested PCR amplification (Step 7)

All three tubes with the nested primers underwent 35 PCR amplification cycles followed by a final extension at 72°C×7 min.

The thermal cycling parameters were identical to those used in the initial PCR amplification (Step 5) except for the β-globin DNA that had undergone *Avr* II digestion, calf intestinal alkaline phosphatase treatment, oligonucleotide ligation: 94°C×30 sec, 60°C×30 sec, 72°C×2 min.

Detection of PCR products

10 μl of each PCR product was analyzed by agarose gel electrophoresis (2% agarose for products < 1 kb and 1% agarose for products > 1 kb) followed by ethidium bromide staining.

Direct sequencing of entire PCR product

Unpurified PCR products, re-amplified with primers 3 and 4, were purified over a Centricon 30 microconcentrator (Amicon,

Danvers MA) and the PCR product was directly sequenced with an Applied Biosystems 373 automated DNA sequencer, using primers 3 and 4 as sequencing primers.

Cloning of individual PCR products

PCR products were cloned into *E. coli* using recombination PCR (20). Briefly, pUC19 was digested with *Hind* III, and 2 ng of the linearized plasmid underwent 20 cycles of PCR amplification and modification with *Taq* (DNA) polymerase using primers whose 5′ ends have 24 nucleotides of complementarity to the 5′ ends of the primers used in the nested primer PCR amplification of step 7. 2.5 μl of each of two crude PCR products: the PCR amplified and modified linear plasmid and the PCR product derived from the nested primers of step 7, were co-transfected into MAX efficiency DH5α competent *E. coli* (GIBCO BRL/Life Technologies, Gaithersburg, MD). Transformed clones were screened for the recombinant using PCR as described (20).

Sequencing of plasmid inserts

Plasmids were purified using Qiagen midi-columns (Qiagen Inc., Studio City CA) following the manufacturer's instructions. Plasmid inserts were sequenced using Sequenase version 2.0 (United States Biochemical, Cleveland, OH) following the manufacturer's protocol, using standard M13 and RM13 primers (Pharmacia, Pleasant Hill, CA)

RESULTS

General strategy

The method is illustrated in Figure 1. The PCR template is generated following digestion of genomic DNA with a restriction enzyme that leaves a 5′ overhang, and ligation of a single-stranded oligonucleotide to the digested genomic DNA, resulting in modification of the 3′ end of each strand. This oligonucleotide is designed to be complementary to the known region of DNA upstream of the unknown region of interest. Denaturation and annealing occurs under dilute conditions so that intra-strand annealing is promoted. Strands of genomic DNA which contain the complement of the ligated-oligonucleotide form a stem-loop structure. The sequence-specific annealing that constitutes the stem contains a recessed 3′ end, so that this oligonucleotide can prime DNA polymerization using the known sequence as the template. This polymerization results in known DNA being

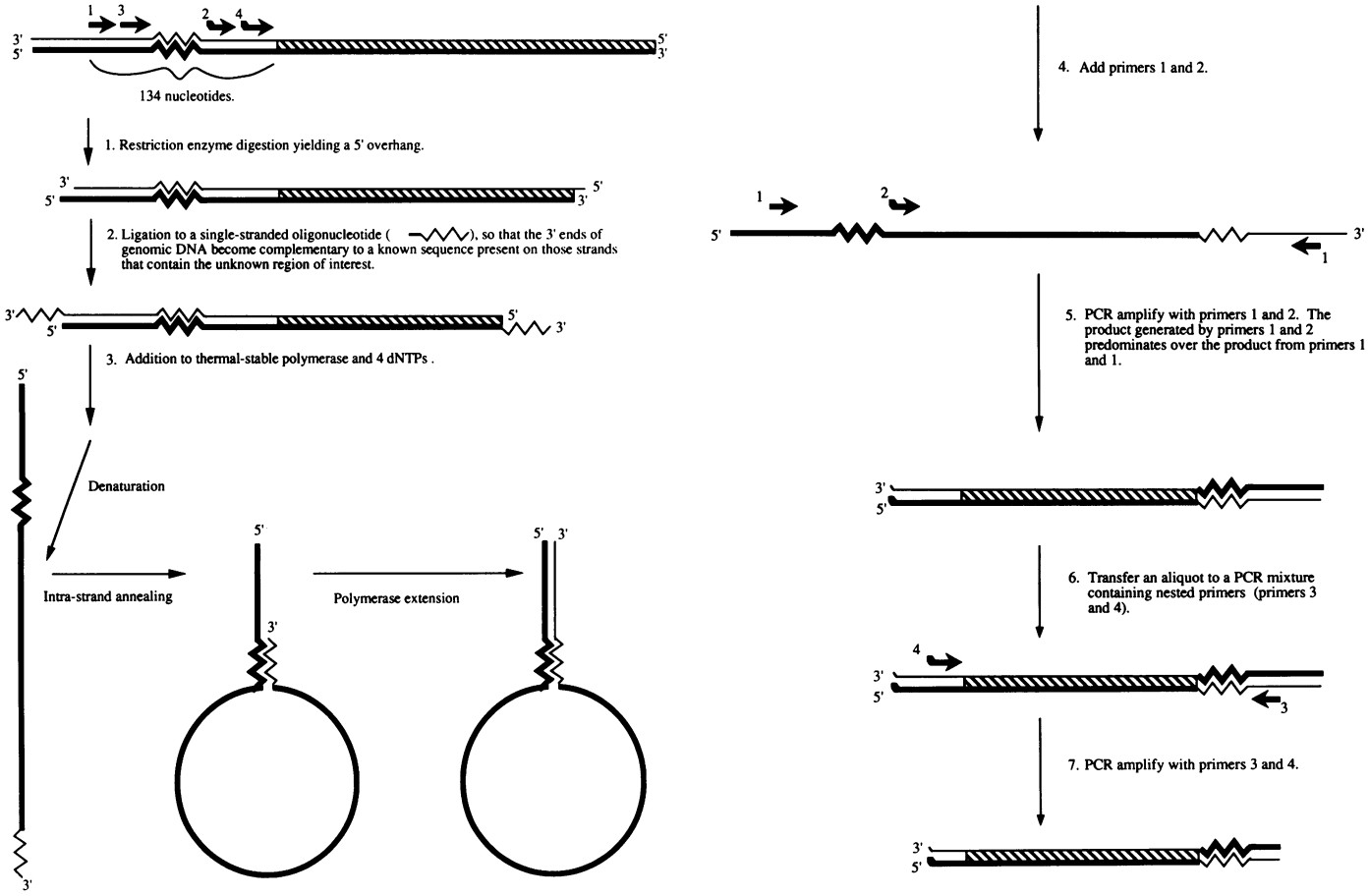

Figure 1. Panhandle PCR: A method to amplify genomic DNA flanking a known sequence without cloning, using PCR. The two complementary strands of genomic DNA are represented by thin and thick lines. Double-stranded unknown DNA that flanks the known region of genomic DNA is striped between the two strands. The jagged portion of the thick line represents the annealing region for the ligated-oligonucleotide. The PCR primers are numbered arrows. Their location in relation to the relevant strands of genomic DNA are shown on top of the diagram for step 1, and the primers are not used until step 4. 134 nucleotides spans the primer annealing sites used to amplify β-globin. Nucleotides added to the 5′ ends of primers 2 and 4 that are not homologous to the thick strand of genomic DNA are represented by upended 5′ ends. This modification was done in order to decrease the possibility of a short circuit in the PCR amplification, resulting from PCR primer annealing to the thick strand of unknown flanking DNA yielding a short product. The numbered steps in this Figure directly correspond to the numbered steps detailed in the Methods section.

appended to the uncharacterized end of the unknown DNA contained in the loop, generating the panhandle template used for subsequent PCR amplification. Since the generation of this template results in known DNA being positioned on both sides of the unknown flanking DNA, PCR can be used to amplify the unknown DNA. In practice, two primers are used in a PCR amplification, one primer which is homologous to the region upstream from the annealing site for the ligated-oligonucleotide, and one primer which is homologous to a region located between the ligated-oligonucleotide annealing site and the unknown flanking DNA. PCR amplification with only one primer, homologous to the region upstream from the ligated-oligonucleotide annealing site (corresponding to primer 1 or 3 in Figure 1) amplified less efficiently (data not shown). Implementation of the method illustrated in Figure 1, using two primers for an initial PCR amplification and nested primers (1) in a subsequent PCR amplification, permitted consistent amplification and sequencing of DNA initially flanking one side of the primer annealing sites directly from the human genome.

Amplification of a 305 bp DNA segment from human genomic DNA using panhandle PCR

The efficacy of this method was tested by the PCR amplification of a region of β-globin DNA from total human genomic DNA, using four amplification primers that initially flank only one side of this sequence. This region consists of the last 10 bp of the first intron and the first 208 bp of the second exon of the β-globin gene (21). Genomic DNA was digested with *Bam* HI, treated with calf intestinal alkaline phosphatase, and ligated to a phosphorylated oligonucleotide prior to generation of the panhandle template and PCR amplification. The product obtained migrated at the predicted 305 bp size (see Figure 2, lane 1), and permitted the identification of the 218 bp of 'unknown' flanking genomic DNA. Sequence analysis of the entire PCR product confirmed its identity. Individual products were analyzed using recombination PCR (20). Eleven of 12 clones tested contained the recombinant. Sequencing of the 305 bp product in 6 clones revealed no errors in 4 clones and 2 errors each in the remaining 2 clones (Both clones had a C for T substitution within the 218 bp of 'unknown' flanking DNA; one clone had an T to C substitution within the ligated-oligonucleotide, and the other clone had a C deleted from a sequence of 3 Cs, two of which lie within primer 4; deletion of a base in a primer sequence is a type of error previously reported in recombination PCR (20)).

Amplification of a >2 kb DNA segment from human genomic DNA using panhandle PCR

In order to explore the ability of this method to amplify a large fragment directly from human genomic DNA, we used our strategy to amplify a 2307 bp product containing 2221 bp of 'unknown' β-globin DNA. This 2221 bp of 'unknown' flanking DNA consists of the last 10 bp of the first intron, the remaining 1333 bp of the β-globin gene and 878 bp flanking the 3′ end of the β-globin gene (21). *Avr* II digested human genomic DNA was treated with calf intestinal alkaline phosphatase and ligated to a phosphorylated oligonucleotide. The phosphorylated oligonucleotide annealed to the same region as in the previous experiment and the same amplifying primers were used. Products were obtained that migrated in the range of the predicted 2307 bp product (see Figure 3a, lane 2). Individual products were cloned using recombination PCR (20). 2 of 6 clones tested contained inserts; one 1.7 kb and the other 2.3 kb. The ends of each insert were sequenced. Sequencing of 195 nucleotides of the 1.7 kb product from the primer 4 end revealed the predicted sequence with no errors. Sequencing of 171 nucleotides of the 1.7 kb product from the primer 3 end revealed that 12 nucleotides

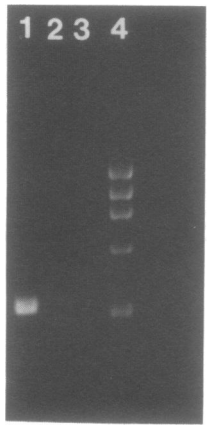

Figure 2. Amplification of a portion of the human β-globin gene directly from human genomic DNA using four primers that initially flank only one side of the amplified region, run on an agarose gel. Lane 1, human genomic DNA after *Bam* H1 digestion, calf intestinal alkaline phosphatase treatment and ligation to phosphorylated oligonucleotide; lane 2, human genomic DNA after *Bam* H1 digestion and calf intestinal alkaline phosphatase treatment without ligation to phosphorylated oligonucleotide; lane 3, no template control; lane 4, φX174 *Hae* III MW markers.

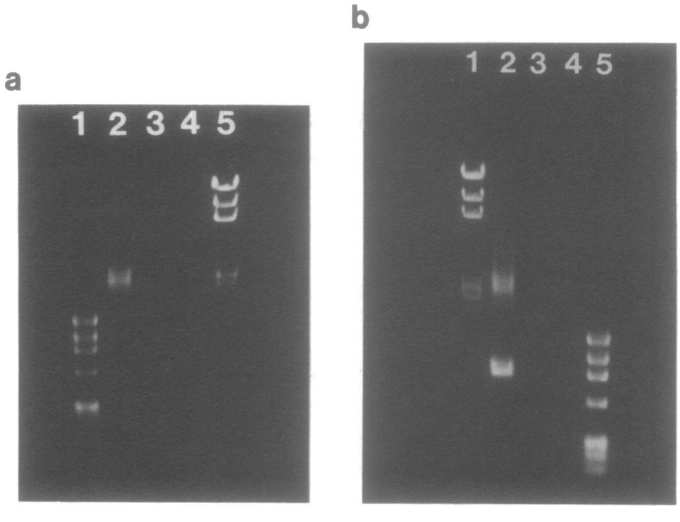

Figure 3a. Amplification of a portion of the human β-globin gene directly from human genomic DNA using four primers that initially flank only one side of the amplified region, run on an agarose gel. Lane 1, φX174 *Hae* III MW markers; lane 2, human genomic DNA after *Avr* II digestion, calf intestinal alkaline phosphatase treatment and ligation to phosphorylated oligonucleotide; lane 3, human genomic DNA after *Avr* II digestion and calf intestinal alkaline phosphatase treatment without ligation to phosphorylated oligonucleotide; lane 4, no template control; lane 5, λ *Hind* III MW markers. **3b.** Amplification of a portion of the human β-globin gene directly from human genomic DNA using four primers that initially flank only one side of the amplified region, run on an agarose gel. Lane 1, λ Hind III MW markers; lane 2, human genomic DNA after *Avr* II digestion, partial fill-in reaction with Klenow fragment and ligation to phosphorylated oligonucleotide; lane 3, human genomic DNA after *Avr* II digestion and partial fill-in reaction with Klenow fragment without ligation to phosphorylated oligonucleotide; lane 4, no template control; lane 5, φX174 *Hae* III MW markers.

were deleted from the 5′ end of the ligated-oligonucleotide, and 569 nucleotides of the β-globin sequence were missing, with no errors following the sequencing of 120 nucleotides of β-globin. Sequencing of the ends of the 2.3 kb insert revealed the predicted sequence with only one error (From the primer 4 end, 195 nucleotides were without error. From the primer 3 end, of 199 nucleotides sequenced, there was a one nucleotide substitution in the ligated oligonucleotide). Thus, this method permitted the amplification of a large piece of flanking DNA directly from human genomic DNA.

To increase the number of restriction sites that can be used to digest genomic DNA prior to ligation to a given phosphorylated oligonucleotide, we used a partial fill-in strategy (18). Following restriction enzyme digestion of genomic DNA with an enzyme that yields a 5′ single-strand overhang sequence, a partial fill-in of the 5′ single-stranded termini of genomic DNA using a DNA polymerase and < 4 dNTPs will render these ends non-complementary to each other, preventing ligation with similarly cleaved fragments. These DNA ends retain a single-stranded 5′ end that can be ligated to a phosphorylated single-stranded oligonucleotide by T_4 DNA ligase. To test this strategy, genomic DNA was cut with *Avr* II and underwent partial fill-in with dATP, dCTP and dTTP. This partial fill-in reaction permitted ligation of the digested genomic DNA to any phosphorylated oligonucleotide that has a 5′ terminal G. The PCR product contained a product close to the predicted 2307 bp size (see Figure 3b, lane 2), and a 850 bp product. These products were cloned using recombination PCR (20). Eleven of 12 clones tested contained the 850 bp insert and one contained a 2.3 kb insert. The ends of two of the 850 bp inserts and the ends of the single 2.3 bp insert were sequenced. Each of the clones sequenced contained the β-globin region of interest. In each of the three clones, the inserts were shorter than the predicted length. In addition, each insert contained short extraneous sequences between the phosphorylated oligonucleotide and the unknown DNA of interest. Sequencing of 230 nucleotides of the 850 bp products from the primer 4 end revealed the targeted β-globin sequence without errors except for a C deleted from the primer 4 sequence in one clone. Sequencing of 190 nucleotides of the 850 bp products from the primer 3 end revealed that one nucleotide was deleted from the ligated-oligonucleotide, followed by 6 nucleotides of unknown origin (ATAGAG); this was followed by the absence of 1462 nucleotides of β-globin sequence, then 122 nucleotides of β-globin with no sequence errors. Sequence analysis of the 2.3 kb insert revealed a product just short of the full length of targeted β-globin sequence (Sequencing of 230 nucleotides from the primer 4 end revealed no errors. Sequencing of 224 nucleotides from the primer 3 end revealed that 5 nucleotides were deleted from the 5′ end of the ligated-oligonucleotide, followed by 8 nucleotides of unclear origin (C-AATATGT); this was followed by the absence of 53 nucleotides of β-globin DNA, then 158 nucleotides of β-globin DNA with one error, a T to C substitution). In each of the three clones, the DNA that lies adjacent to the 'known' region, sequenced from the primer 4 containing end, was intact. Extraneous sequences were appended to the far end of the 'unknown' DNA, adjacent to the ligated oligonucleotide.

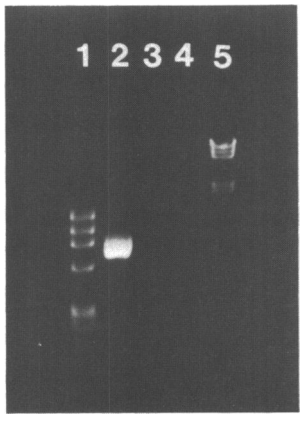

Figure 4. Amplification of 5′ flanking region of the CFTR gene directly from human genomic DNA using four primers derived from cDNA data, run on an agarose gel. Lane 1, φX174 *Hae* III MW markers; Lane 2, human genomic DNA after *Hind* III digestion, calf intestinal alkaline phosphatase treatment and ligation to phosphorylated oligonucleotide; lane 3, human genomic DNA after *Hind* III digestion and calf intestinal alkaline phosphatase treatment without ligation to phosphorylated oligonucleotide; lane 4, no template control; lane 5, λ *Hind* III MW markers.

Amplification of the cystic fibrosis transmembrane conductance regulator (CFTR) gene promoter sequence using primers that anneal to cDNA sequence using panhandle PCR

In order to assess the ability of this method to amplify promoter sequences using known cDNA sequence data, and to evaluate the efficacy of this method on a relatively GC rich region, the CFTR gene promoter was amplified directly from human genomic DNA using cDNA sequence data (22). The predicted 769 bp product was obtained (See Figure 4, lane 2). This fragment contains 686 bp that flanks the primer annealing sites, including 657 bp of genomic sequence that lies 5′ to the cDNA sequence. The unpurified PCR product was directly sequenced, establishing its identity. The two previously published sequences of this region contain three discrepancies, all single nucleotide substitutions (23,24). These differences, placing the first nucleotide of the cDNA sequence at position +1, are: an A vs G in position −500, a C vs A in position −266, and a T vs G in position −258, respectively. Our sequence differed from each of the two published sequences by containing an A in position −500, an A in position −266, and a G in position −258. Sequencing of this region using a conventional PCR approach confirmed our sequence.

DISCUSSION

The experiments in this paper illustrate that panhandle PCR permits the amplification and sequencing of unknown flanking DNA in a complex mixture. In the first experiment, we were able to amplify a 305 bp product that contained 218 bp of β-globin DNA that flanked the primer annealing sites. The second set of experiments were designed to amplify 2307 bp of β-globin DNA containing 2221 bp of DNA that flanked the primer annealing sites. The full length product was obtained. In the third experiment 657 bp of DNA flanking the 5′ end of the CFTR cDNA was amplified using primers derived from cDNA sequence data. The presence of sequence deletions on the ends of several PCR products, seen when attempting to amplify the 2.3 kb product, may result from the secondary structure of the panhandle template, aberrant restriction enzyme cutting, exonuclease digestion of the DNA end prior to oligonucleotide ligation, or

PCR itself. When a partial fill-in approach was used prior to ligation of the phosphorylated oligonucleotide, clones contained small added sequences between the ligated oligonucleotide and the unknown DNA of interest. The cause of this is not clear. Inefficient ligation of the phosphorylated oligonucleotide to a single-stranded end only one nucleotide long may have contributed to the suboptimal result using this partial fill-in approach. Nonetheless, the 'unknown' flanking region was not scrambled in any of the clones. The sequence additions were adjacent to but not within the flanking DNA. Sequence analysis of a total of 2873 bp of flanking DNA in single clones contained 3 errors, indicating a misincorporation rate of 0.1% following formation of the panhandle and 65 cycles of PCR amplification. Sequence analysis of entire PCR products, as was done with the 305 bp β-globin sequence, and the CFTR sequence, is not effected by such misincorporations (3). The sequence differences between our DNA flanking the 5′ region of the CFTR cDNA and each of the two previously reported sequences may constitute sequence polymorphisms.

In the context of cloning technology, panhandle PCR will facilitate the generation of contiguous YAC clones by providing a rapid and reliable method for the cloning of YAC endpoints, permitting the ordering of large fragments of individual chromosomes cloned in YAC vectors (13,14,25,26). This method will also provide a method for amplifying and sequencing successive, overlapping fragments of YAC inserts. Sequence artifacts have been associated with cloning in prokaryotic vectors (27), and with the cloning of chromosome fragments into YAC vectors (28). A method that permits the amplification of DNA segments from bulk human genomic DNA has the potential of supplanting current cloning-based strategies for sequencing the human genome, and will permit genome walking into unclonable regions of DNA (29). This method will also provide an alternative method for chromosome jumping procedures. In the case of chromosome jumping, a large restriction fragment, only one side of which is known, is circularized in order to bring the region with known sequence at one end of the fragment adjacent to the region of unknown sequence at the other end of the fragment. These large circles of DNA could be restriction enzyme digested with a frequent cutter that renders a single-stranded 5′ overhang sequence. A single-stranded oligonucleotide would be ligated to the ends of the resulting fragments. This oligonucleotide would be designed to be complementary to the known region at one end of the original fragment, which is now adjacent to the unknown region of interest. Denaturation of these fragments, followed by self-annealing under high stringency and template-directed polymerization, will result in a DNA fragment in which known DNA flanks both ends of the sequence of the unknown end of the original, large restriction fragment, permitting amplification and sequencing of this unknown end. Panhandle PCR could also be used to generate specific linking fragments (17,30). Linking fragments connect large genomic DNA fragments that have been produced following digestion by a rare cutting restriction enzyme, such as *Not* I. These large fragments are separated by pulsed-gel-field electrophoresis and can be cloned into YAC vectors. To obtain a specific linking fragment, this method would be used to amplify a short stretch of DNA adjacent to an end of one large fragment directly from genomic DNA. Numerous additional applications of this method can also be anticipated, including: the amplification of fragments adjacent to cDNA, such as regulatory regions and intron-exon junctions, and the determination of viral and transposon integration sites. The panhandle PCR method has the potential of sidestepping traditional cloning approaches for obtaining new sequence information.

ACKNOWLEDGEMENTS

This work was supported by the Roy J. Carver Charitable Trust. We thank Charles Grose, Raji Padmanabhan and Jeffrey Murray for helpful discussions.

REFERENCES

1. Mullis,K., Faloona,F., Scharf, S., Saiki, R., Horn G, and Erlich, H. (1986) Cold Spring Harbor Symposia on Quantitative Biology, Cold Spring Harbor Laboratory, Vol.LI, pp. 263–73.
2. Saiki,R.K., Scharf,S., Faloona,F., Mullis,K.B., Horn,G.T., Erlich,H.A. and Arnheim,N. (1985) *Science*, **230,** 1350–1354.
3. Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239,** 487–491.
4. Ochman,H., Gerber,A.S. and Hartl,D.L. (1988) *Genetics*, **120,** 621–23.
5. Triglia,T., Peterson,M.G. and Kemp,D.J. (1988) *Nucleic Acids Res.*, **16,** 8186.
6. Silver,J. and Keerikatte,V. (1989) *J Virol.*, **63,** 1924–1928.
7. Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) *Proc Natl Acad Sci USA.*, **85,** 8998–9002.
8. Loh,E.Y., Elliott,J.F., Cwirla,S., Lanier,L.L. and Davis,M.M. (1989) *Science*, **243,** 217–220.
9. Ohara,O., Dorit,R.L. and Gilbert,W. (1989) *Proc Natl Acad Sci USA.*, **86,** 5673–5677.
10. Shyamala,V. and Ames,GF-L. (1990) *Gene* **84,** 1–8.
11. Pfeifer,G.P., Steigerwald,S.D., Mueller,P.R., Wold,B. and Riggs,A.D. (1989) *Science* **246,** 810–813.
12. Mueller,P.R. and Wold,B. (1989) *Science* **246,** 780–786.
13. Riley,J., Butler,R., Ogilvie,D., Finniear,R., Jenner,D., Powell,S., Anand,R., Smith,J.C. and Markham,A.F. (1990) *Nucleic Acids Res.*, **18,** 2887–2890.
14. Rosenthal,A. and Jones,D.S.C. (1990) *Nucleic Acids Res.*, **18,** 3095–3096.
15. Burke,D.T., Carle,G.F. and Olson,M.V. (1987) *Science* **236,** 806–812.
16. Collins,F.S., and Weissman, S.M. (1984) *Proc. Natl. Acad. Sci. USA.*, **81,** 6812–6816.
17. Poustka,A. and Lehrach,H. (1986) *Trends in Genet.*, **2,** 174–179.
18. Hung,M-C. and Wensink,P.C. (1984) *Nucleic Acids Res.*, **12,** 1863–1874.
19. Mullis,K.B. (1991) *PCR: Methods and Applications* **1,** 1–4.
20. Jones,D.H. and Howard,B.H. (1991) *BioTechniques* **10,** 62–66.
21. Lawn,R.M., Efstratiadis,A., O'Connell,C. and Maniatis,T. (1980) *Cell* **21,** 647–651.
22. Riordan,J.R., Rommens,J.M., Kerem,B-S., Alon,N., Rozmahel,R., Grzelczak,Z., Zielenski,J., Lok,S., Plavsic,N., Chou,J-L., Drumm,M.L., Iannuzzi,M.C., Collins,F.S. and Tsui,L-C. (1989) *Science*, **245,** 1066–1072.
23. Yoshimura,K., Nakamura,H., Trapnell, B.C., Dalemans,W., Pavirani,A., Lecocq,J-P. and Crystal,R.C. (1991) *J Biol Chem*, **266,** 9140–9144.
24. Zielenski,J., Rozmahel,R., Bozon,D., Kerem,B-S., Grzelczak,Z., Riordan,J.R., Rommens,J., and Tsui,L-C. (1991) *Genomics* **10,** 214–228.
25. Breukel,C., Wijnen,J., Tops,C., Klift,H., Dauwerse,H. and Khan,P.M. (1990) *Nucleic Acids Res.*, **18,** 3097.
26. Ochman,H., Medhora,M.M., Garza,D. and Hartl, D.L. (1990) In Innis,M.A., Gelfand,D.H., Sninsky,J.J. and White T.J. (ed.) PCR Protocols—A Guide to Methods and Applications. Academic Press, San Diego, CA. pp. 219–227.
27. Ishiura,M., Hazumi,N., Shinagawa,H., Nakata,A., Uchida,T. and Okada,Y. (1990) *J Gen Microb.*, **136,** 69–79.
28. Neil,D.L., Villasante,A., Fisher,R.B., Vetrie,D., Cox,B. and Tyler-Smith, C. (1990) *Nucleic Acids Res.*, **18,** 1421–1428.
29. Wyman,A.R., Wolfe,L.B. and Botstein,D. (1985) *Proc. Natl. Acad. Sci. USA.*, **82,** 2880–2884.
30. Kandpal,R.P., Shukla,H., Ward,D.C. and Weissman,S.M. (1990) *Nucleic Acids Res.*, **18,** 3081.