Volume 16 Number 15 1988 Nucleic Acids Research

A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions

Russell Higuchi, Barbara Krummel[1] and Randall K.Saiki

Department of Human Genetics, Cetus Corporation, Emeryville, CA 94608 and [1]Department of Biochemistry, U.C. Berkeley, Berkeley, CA 94720, USA

Received May 19, 1988; Revised and Accepted July 8, 1988

ABSTRACT

Specific, end-labeled DNA fragments can be simply and rapidly prepared using the polymerase chain reaction (PCR). Such fragments are suitable for use in DNase I protection footprint assays, chemical sequencing reactions, and for the production and analysis of paused RNA polymerase transcription complexes. Moreover, a general means of introducing a specific mutation at any position along the length of such PCR-generated fragments is described. These procedures, which can circumvent the need for large-scale phage or plasmid growths, preparative gel-electrophoresis and the screening of molecular clones, can facilitate the rapid study of sequence-specific interactions of proteins and DNA. A rapid means of removing excess oligonucleotide primers from completed PCRs is also described.

INTRODUCTION

The preparation of large amounts of specific DNA fragments is now an almost universal prerequisite to experiments in molecular biology. This is particularly true for the study of interactions between proteins and specific DNA sequences. In some cases, the chemical synthesis of DNA fragments is the most direct preparative method. However, it is difficult and expensive to chemically produce fragments much larger than 110 bp (1). The growth and restriction enzyme digestion of recombinant DNA plasmids is the usual method of preparing larger fragments.

PCR, the primer directed, enzymatic amplification of specific DNA sequences (2,3,4), can also be used to prepare significant quantities of larger DNA fragments (up to at least 4 kb; R. Griffith and R. Higuchi, unpublished results) and offers many advantages over the preparation of restriction fragments. Some of these advantages are apparent in Fig. 1, in which the preparation of an end-labeled DNA fragment is outlined for both the restriction enzyme mediated and the PCR methods. The production of microgram amounts of restriction fragments may take days and involves the growth of bacterial cultures of at least several

hundred ml, extractions with organic solvents, precipitations, and preparative gel-electrophoresis. The production of such fragments by PCR takes hours and involves fewer steps that, given the automation of PCR (4), are not labor intensive.

Since as little as a single molecule of the sequence to be amplified can suffice (4), a "miniprep" of a bacterial plasmid DNA, if not a single bacteria or bacterial colony, provides more than enough staring material for PCR amplification. PCR is capable of specifically amplifying portions of genes that occur only once in genomes as complex as the human genome (4). The relative lack of sequence complexity of bacterial DNA implies that PCR will have more than adequate specificity in this context.

In this paper, we first show that PCR fragments can substitute for restriction fragments in a study of the sequence specific binding of *E. coli* RNA polymerase to a promoter. Next, since such studies frequently examine the effect of specific base substitutions on the interaction, we demonstrate that PCR can also substitute for bacterial cloning methods to produce, *in vitro,* a DNA fragment containing such a mutation. The method is general in that it should be able to put almost any deletion, insertion, or substitution anywhere along the length of the DNA fragment.

MATERIALS AND METHODS

Enzymes: *E. coli* RNA polymerase holoenzyme was purified from *E. coli* DG156 as previously described (5). Bovine pancreas DNase I was purchased from Worthington Biochemicals and Bentonite (VWR) treated to remove RNase (6). Calf intestinal phosphatase (8U/μl) was from Boehringer Mannhein. Restriction enzymes *Rsa* I and *Sma* I were from NEB and IBI respectively. T4 polynucleotide kinase (10U/μl) was from NEB, and *Thermus aquaticus* (Taq) DNA polymerase was from Perkin-Elmer Cetus. Nucleotides: The dinucleotide ApU was from Sigma, dNTPs and HPLC purified rNTPs were from Pharmacia, other rNTPs were from PL Biochemicals. (γ-^{32}P)-ATP and (α-^{32}P)-GTP (both 3000 Ci/mmole)were from NEN.

Oligonucleotide PCR primers (all sequences are written 5' to 3'):

RH 18: ACTGCCGGGCCTCTTGCG
RH 20: CCCGGACCCCGGTGTCGATT
RH 28: CCCACCCGAAGGTGAGCCAG
RH29: GTTAGCCATAAAGTGATAACC
RH 30: ACACGGCGAAAAGCCATCCC
RH31: GGGATGGCTTTTCGCCGTGT

Preparation of DNA restriction fragments that contain the phage T7A1 promoter.

A 529 base pair (bp) fragment containing the phage T7A1 promoter (7) was produced by a digest of 100 μg of plasmid pAR1707 (provided by Dr. A. Rosenberg; described in reference 8) with restriction enzymes *Sma* I and *Rsa* I (one unit each per μg DNA). The restriction fragments were separated by electrophoresis on a 5% nondenaturing polyacrylamide gel and the 529 bp fragment was identified after ethidium bromide staining. The fragment was electroeluted from the gel slice and concentated by ethanol precipitation.(6).

To obtain the same fragment ^{32}P-labelled at the 5' end of the template strand, 25-50 μg of plasmid pAR1707 was linearized with *Sma* I, then dephosporylated using calf-intestinal phosphatase and labeled at the 5'-OH ends using [γ-^{32}P] ATP and polynucleotide kinase. Uniquely end-labeled fragments were generated by digestion with *Rsa* I. These fragments were separated as above and the fragment of appropriate size identified after autoradiography. This fragment was then prepared as described above. (A typical preparation yielded 1-2 μg of fragment with a specific activity of 4 x 10^5 dpm/pmole.)

Preparation of PCR fragments that contain the phage T7A1 promoter.

Reactions were performed in a 0.5 ml microfuge tube in a final volume of 0.25 ml. In this volume, 1.2 nM plasmid (approx. 4 μg/ml) pAR1707 , 175 μM all four deoxynucleoside triphosphates, 0.2 μM each oligonucleotide primer and 0.04 U/μl Taq DNA polymerase were mixed in 50 mM KCl, 10mM tris-HCl (pH8.3), 2.5 mM $MgCl_2$ and 0.01% gelatin (w/v) (4). Cycling reactions were performed in a Perkin-Elmer Cetus DNA Thermal Cycler. Each cycle consisted of denaturation at 93 °C for 60 sec., annealing at 37 °C for 40 sec., and extention at 72 °C for 1 min. and 40 sec. Sixteen cycles were performed. Extension time of the last cycle was 10 minutes to ensure that all single stranded DNA had been copied or reannealed to form double strands. The completed reactions were extracted once with an equal volume of a 1:1 mixture of phenol and chloroform and once with a two-fold volume of butanol. The DNA was concetrated on an Amicon centricon 30 as recommended by the manufacturer after dilution to 2 ml with TE (10 mM tris-HCl (pH 8.0), 0.1 mM EDTA). A typical yield was 2-5 pmoles of product per 0.1 ml reaction volume.

To obtain a uniquely end-labeled fragment, the reaction was peformed as above except that one of the primers (50 pmoles) used had first been 5' ^{32}P-end labeled using [γ-^{32}P] ATP and polynucleotide kinase (6).

DNase I Footprinting

Purified *E. coli* RNA polymerase holoenzyme (17 μg/ml; 36 nM) was incubated with promoter DNA fragment labeled at the 5' end of the template strand (7 μg/ml; 18 nM) in "transcription buffer" (44 mM tris-HCl pH 8.0,14 mM

2-mercaptoethanol, 14 mM $MgCl_2$, 20 mM NaCl, 2% glycerol, 0.04 mM EDTA, 0.04 mg/ml acetylated BSA) in a final volume of 10 µl for 10 minutes at 30 °C to form "open" promoter complexes. A total of approximately 50,000 to 100,000 cpm end-labeled DNA was used in each 10 µl reaction. A one-fourth volume of "carrier mix" (transcription buffer with 8 mM $CaCl_2$ and 0.34 mg/ml sonicated salmon sperm DNA) was added followed by the addition of DNase I to a final concentration of 0.01 µg/ml. Incubation was for 1' 30". Reactions were quenched by the additon of a 30-fold excess of sonicated salmon sperm DNA. A one-tenth volume of 60% sucrose, xylene cyanol, bromophenol blue was added and the reactions were loaded immediately onto a 4% acrylamide gel (38:1 acrylamide: bisacrylamide). Nondenaturing PAGE was used to separate enzyme/DNA complexes from free DNA (9). The protein/ DNA complexes or free DNA were identified by autoradiography, eluted from gel slices and prepared for denaturing PAGE (8% acrylamide, 7 M urea) as described (9).

Chemical Sequencing

Chemical sequencing of DNA was performed as described by Maxam and Gilbert (10).

Transcription Reactions

In vitro transcription experiments were performed as described (8). RNA polymerase and DNA fragment concentrations were both 18 nM. For ternary complex formation conditions the concentration of the dinucleotide, ApU, was 50 µM while HPLC purified ATP, CTP and [α-^{32}P]-GTP (35,000 cpm/pmole) were present at 5 µM each. Elongation conditions included 500 µM ATP, CTP, GTP and UTP (not HPLC purified).

Directed mutagenisis by PCR

PCR was performed as above using mismatched primers (see Fig. 4) to introduce mutations. 10 µl of each of the two completed, primary PCRs (using the primer pairs RH29 and RH31 and RH30 and RH 28 on pAR1707 template - see Fig. 2) were mixed and excess primers removed as described below. 10 µl of the Centricon eluent was added to a second, 0.1 ml PCR using only RH29 and RH28. 12 additional cycles of PCR were performed. The product of this PCR was used to seed more PCR as additional mutagenized fragment was needed.

Removal of excess primers by Centricon 100 ultrafiltration

PCR product was diluted in 2.0 ml of TE and added to the upper compartment of the Centricon device (obtained as protypes, courtesy of M. Bear, Amicon). The manufacturer's instructions for the Centricon 30 were followed, using 20 min. of centrifugation at 1000G. Centrifugal force in excess of 1000G resulted in more PCR product loss. Only a single centrifugation cycle was

generally used, in which 95% of the excess primer is removed.

Parameters affecting PCR

Some of the parameters affecting PCR in general are covered in reference 4. In this study we routinely used primers approximately 20 nt in length that have an average GC content of 50%. Primers with potential secondary structure were avoided. A higher annealing temperature (55 °C instead of 37 °C) can disrupt secondary structure and results in more specific amplification (4). This was not needed here but could be required in other systems. The amount of starting material can be lessened as the number of cycles is increased, although the incidence of misincorporation is proportional to the number of doublings (see below). A decrease in the amount of enzyme can result both in a decrease in specificity (4) and in an increase in yield. 2-5 pmoles of product per 0.1 ml PCR is the usual yield under the conditions described here. The lengthened extension period (10' at 72 °C) of the last cycle is intended to allow complete reannealing and/or extension so that little single stranded DNA remains. The amount of single stranded DNA remaining is probably less than 1% of total product (U. Gyllensten and H. Erlich, P.N.A.S. in press). PCR is a rapidly developing methodology. The "ideal" protocol may change quickly. The authors welcome inquiries as to updates on the methods described.

RESULTS

The production of an end-labelled DNA fragment for DNAse I footprint analysis

To summarize, a small amount of plasmid DNA is used as a template for primer directed, enzymatic DNA amplification. The number of amplification cycles can be varied to accomodate more or less starting DNA. One of the two amplification primers carries a ^{32}P phosphate group at its 5' end (see Fig. 1), which results in a PCR product with only one of the termini labelled. PCR consists of mixing Taq DNA polymerase, buffer, dNTPs and primers followed by temperature cycling. This temperature cycling causes the repeated DNA denaturation (at 94 °C) followed by primer annealing (37 °C) and primer extension (72 °C) that results in the doubling and redoubling of the copy number of DNA sequences flanked by the primers.

The yield of DNA will eventually plateau; that is, little new DNA is synthesized with additional cycles. When this occurs depends on the amount of starting material and the efficiency of the given PCR. The amount of amplified DNA at this plateau is usually 2 - 5 pmoles per 0.1 ml PCR (about 1 microgram of a several hundred bp fragment). From DNA samples of relatively low sequence complexity, *e.g.*, plasmid DNA, there is usually little amplification of DNA

sequences other than those desired (see fig. 5A for examples).

The particular DNA sequences amplified and the primers used in this study are diagrammed in fig. 2. The restriction fragment defined by the indicated *Rsa* I and *Sma* I sites was made and labeled at the end created by *Sma* I cleavage (see

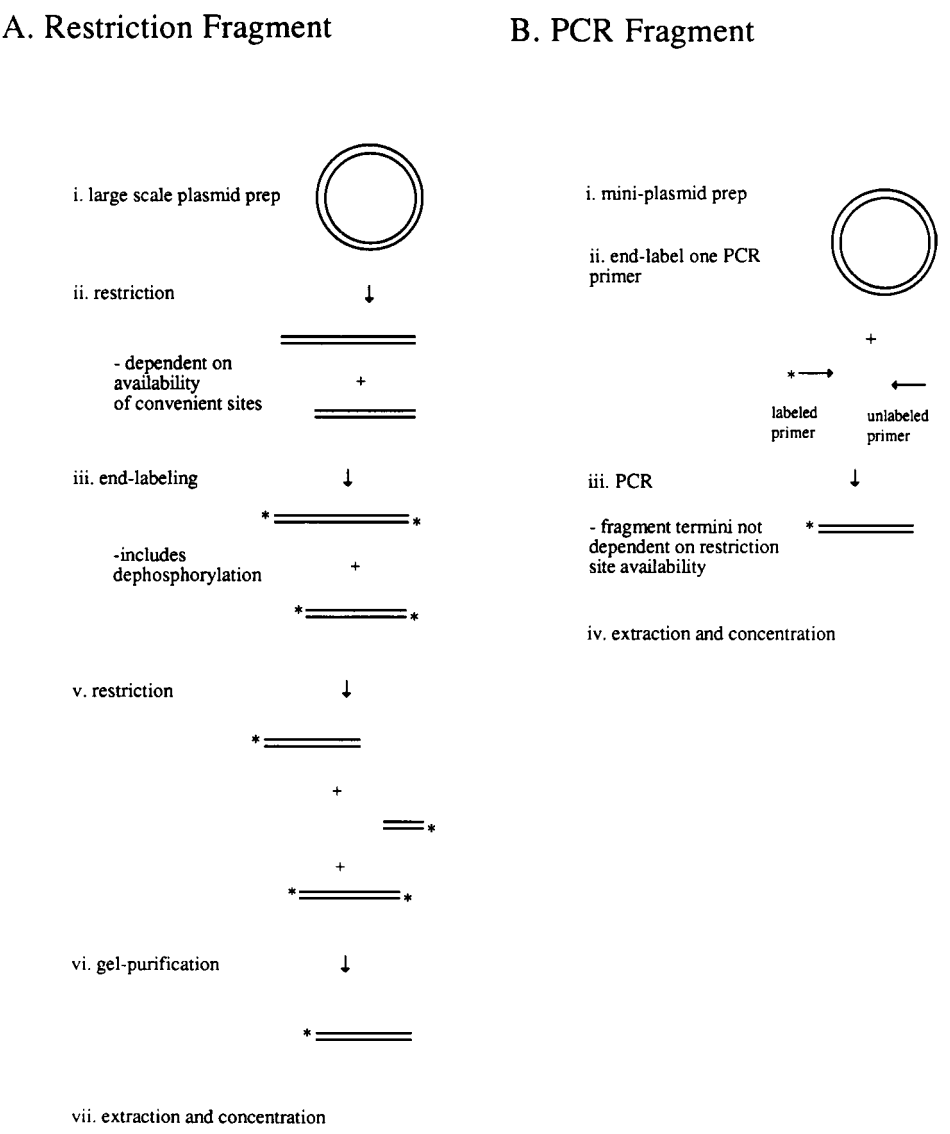

Figure 1. Diagram of steps needed to produce end-labeled restriction fragments compared to steps needed to produce end-labeled fragments via PCR.

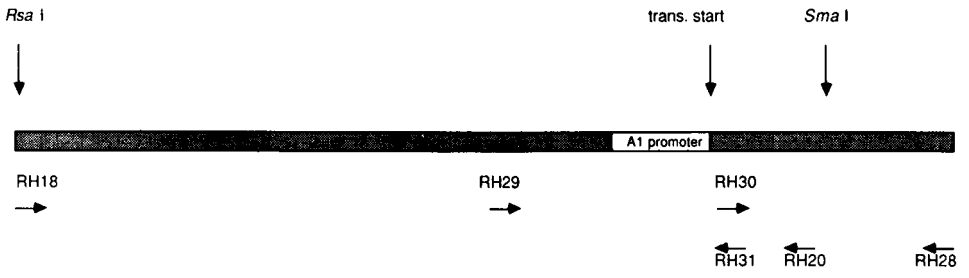

Figure 2. Diagram of T7 A1 promoter region showing amplification primers used. Primer sequences are shown in Materials and Methods. Total length of the region is 614 bp.

Methods). An analagous blunt-ended fragment, shorter by 10 bp at the *Sma* I end, was made and labeled using PCR with the primers RH18 and RH20. Using a kinased primer RH20, 2 - 3 μg of the 519 bp fragment containing approx. 2 x 10^6 dpm/pmole was made in a 0.25 ml PCR. This specific activity was 2 -5 times higher than that of fragments made by restriction enzyme digestion, due probably to the greater efficiency of phosphate transfer by T4 kinase to an oligonucleotide than to a DNA fragment.

After organic extraction to remove DNA polymerase and dialysis by ultrafiltration, this PCR DNA product was used for binding by RNA polymerase from *E. coli* . The product DNA contains the phage T7 A1 promoter which is active in early phage gene synthesis (7). The enzyme is bound specifically at the promoter and forms "open" promoter complexes ($E\text{-}P_o$) in which the DNA template is unwound in preparation for RNA synthesis (11). These $E\text{-}P_o$ complexes, formed with both the PCR and restriction enzyme derived fragments, were treated with DNAse I and resolved from unbound DNA by nondenaturing PAGE (see Materials and Methods). DNA extracted from the nondenaturing gel was resolved by electrophoresis on high resolution polyacylamide gels to show the DNAse I "footprint" generated by the ability of the bound enzyme to block cleavage (12).

Fig. 3A shows the PCR generated and end-labeled fragment used in Maxam-Gilbert, chemical-degradation sequencing (10). All fragments used in Fig. 3 are labelled at the 5' end of the template strand. The appropriate sequence is obtained; this sequence is shown correlated with a specific DNase I digestion pattern obtained from the uncomplexed DNA fragment. Fig 3B shows the footprints obtained with either the PCR or restriction fragments containing the A1 promoter region. There is no apparent difference in the pattern of protection obtained from

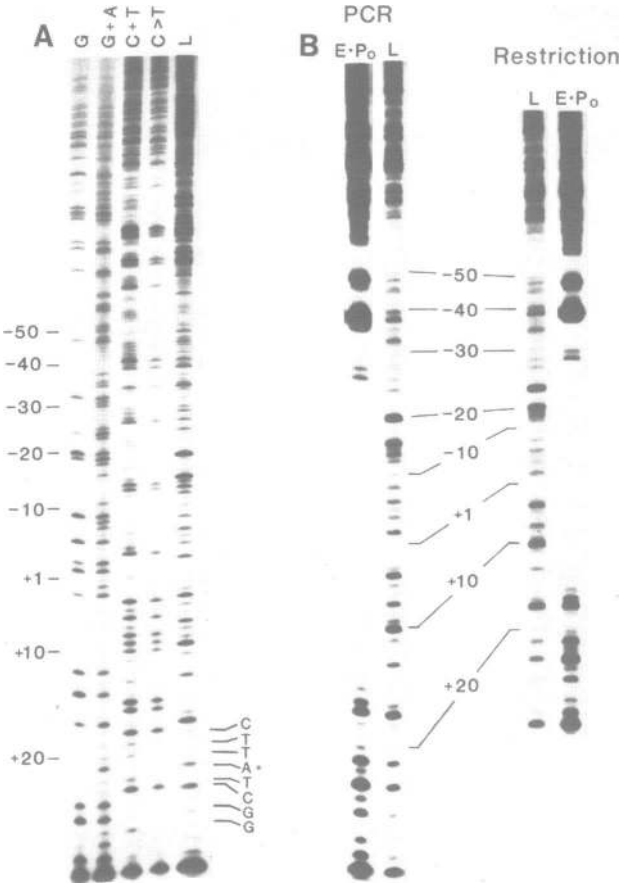

Figure 3. Comparison of E. coli RNA polymerase DNase I footprints on PCR and restriction enzyme generated fragments that contain the phage T7 A1 promoter. Footprinting experiments were performed as described in Methods. (A) The first four lanes are Maxam-Gilbert sequencing reactions of the PCR fragment. "L" is a DNase I ladder of the same fragment in the absence of RNA polymerase. DNase I cleavage products have a 3' hydroxyl while the chemical sequencing products have a 3' PO_4; this accounts for the slight mobility difference. (B) Lanes labeled "E-Po" show footprints of open promoter complexes (complexes with the holoenzyme bound specifically at the promoter) while lanes labeled "L" show cleavage patterns of the naked fragment. The numbering is relative to the transcription start site of +1. The PCR and restriction fragment results were obtained from two separate gels that were run to different extents, hence the restriction fragment footprint is more compact. However, as the numbering of the DNase I ladder indicates, the cleavage patterns and extent of protection is identical on the two fragments.

the PCR prepared and restriction enzyme derived fragment. The footprints shown for comparison are from two gels that were run to slightly different extents. This is indicated by the numbering which corresponds to base pair positions relative to the transcription start site. DNase I shows some sequence specificity that results in a distinctive cleavage pattern of the naked DNA fragment (lanes "L" in Fig. 3). These patterns are identical and reproducible for the two fragment preparations. The presence of the RNA polymerase bound in the E-P_o complex protected a large region of the DNA template (from approx. -55 to +20). Within these boundaries there are specific sites of enhanced cleavage referred to as DNase I hypersensitive sites. These include sites from -35 to -40, -45 to -50 and at position +16. Again these patterns are identical for the two fragments. Additional *E. coli* RNA polymerase E-P_o footprints have been obtained using PCR generated fragments that contain different promoter sequences and have given similar reproducible results (B. Krummel, manuscript in prep.). In more than 10 separate PCR fragment preparations, we have not observed any sequence variations in the regions of interest as a result of the amplification procedure.

Site-directed mutagenesis of PCR fragments

Mutagenesis at the primer region of a PCR product has been demonstrated before. Single bases mismatched between the amplification primer and the template become "fixed" to the primer sequence as a result of amplification (13; see Fig. 4A). Similarly, sequences added on to the 5' end of amplification primers, for example restriction enzyme cleavage sites, become incorporated into the final PCR product even though these "add-on" sequences do not match the target sequence (13). It is difficult, however to efficiently synthesize oligonucleotides much greater than 110 bases in length, which places a limit on how far mutations introduced via PCR can be relative to the termini of the final product.

A general scheme for introducing a mutation in a PCR produced DNA fragment, anywhere along its length, is diagrammed in fig. 4. Two primary PCR reactions produce two overlapping DNA fragments, both bearing the same mutation introduced via primer mismatch, in the region of overlap. This mutation is not restricted to base-substitution. Insertions or deletions can be created as 5' "add-on" sequences in a similar fashion (to produce a deletion, the primer sequence contains the deletion near the 5' end such that the sequence 5' of the deletion is effectively an "add-on" sequence). The overlap in sequence allows the two fragments to recombine in two possible ways after their mixture, denaturation, and renaturation (Fig. 4B). Only one of these combinations produces a structure with recessed 3' OH ends that can be extended by a DNA polymerase to produce a complete duplex fragment. These extended segments can

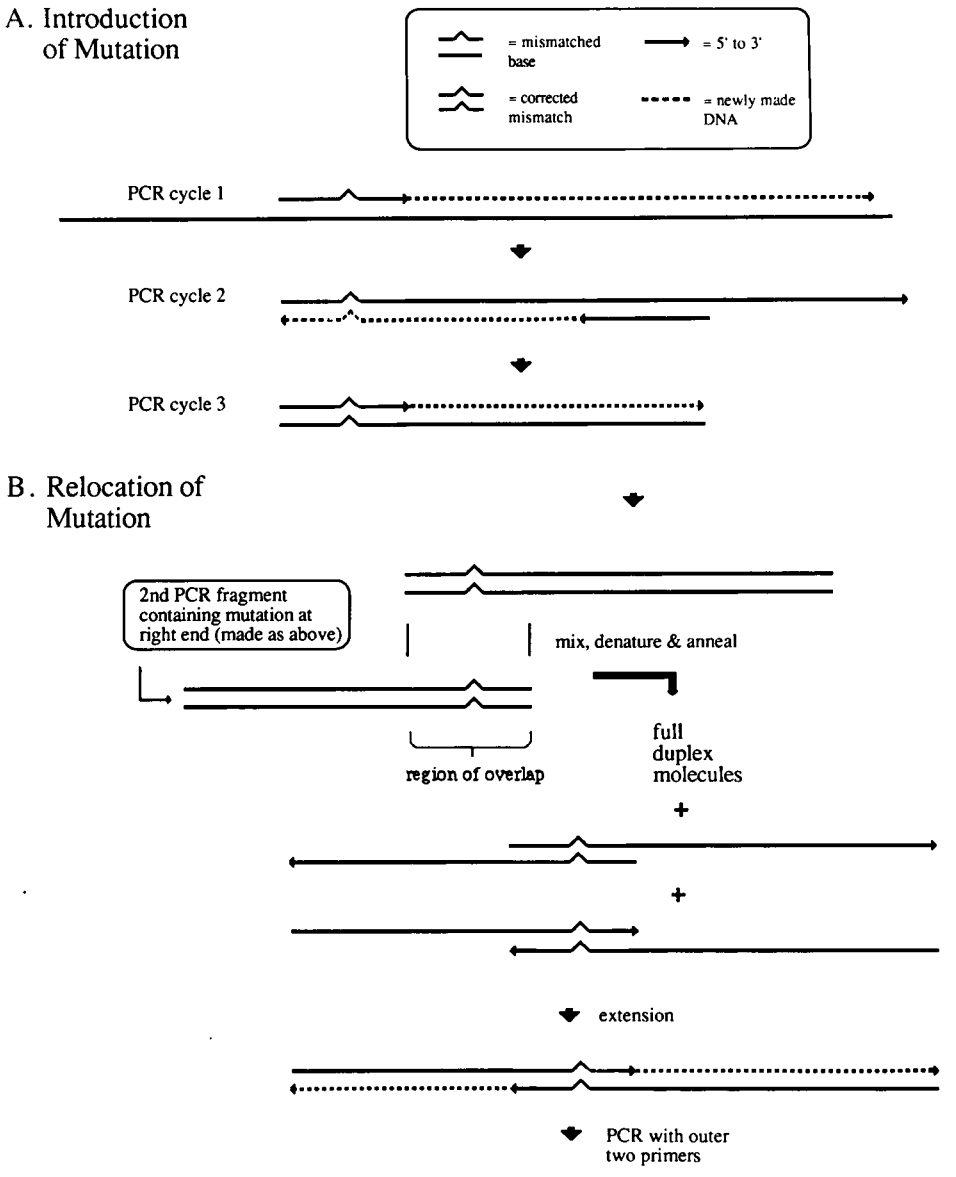

Figure 4. Diagram of method to produce a DNA fragment with a specific mutation at any position. Mutations can be introduced up to 110 nt, the limit of efficient oligonucleotide synthesis (1), from the end of a PCR produced fragment as shown in (A). To create larger fragments altered more than 110 nt from the termini, two primary PCR products can be recombined and reamplified as shown in (B).

then serve as templates for the secondary reamplification of the combined sequences using only the outermost two of the four primers employed to produce the primary fragments.

In this study, the goal was to make a mutated fragment with a specific change at position +21 in the transcribed region. Primary PCRs were performed with primers RH29 & RH31 and with primers RH30 & RH28 on plasmid pAR1707 template as diagrammed in Fig. 2. Excess primers were removed from the completed primary PCRs by centrifuge-driven ultrafiltration on a Centricon 100 (Amicon; see below and Fig. 5C). Small aliquots of purified, primary PCR products were then mixed, denatured, reannealed and subjected to 15 additional cycles of PCR using only primers RH29 and RH28.

These primary and recombined PCR products are shown resolved by electrophoresis on agarose gels in Fig. 5A. The desired products predominate in these reactions. The sequence of the final, recombined PCR product was determined directly, without subcloning, by enzymatic, dideoxy-termination sequencing (14,15). An autoradiograph of the sequencing gel is shown in Fig. 5C and reveals that the desired mutation, a T to an A change in the transcribed strand, has been successfully introduced (compare to Fig 3A). Examination of this autoradiograph shows no evidence of the original T residue over the background in the T lane at that position. That this modified DNA can also interact appropriately with RNA polymerase is demonstrated in Fig. 6 and described below.

In vitro transcription with PCR generated DNA template

To compare the use of PCR generated DNA versus restriction enzyme derived fragments as template by *E. coli* RNA polymerase *in vitro*, we have used a procedure described previously (8). Transcription can be initiated at the T7A1 promoter with the dinucleotide ApU. If ATP, CTP, and GTP are also included, elongation continues to position 20 where the enzyme pauses due to the absence of UTP (Fig. 6B). The α-^{32}P labelled RNA transcripts are analyzed by denaturing PAGE (Fig. 6A) The predominant transcript in the abscence of UTP from either the original restriction fragment or the unmodified PCR fragment is the same 20 nt transcript. There is also some 26 nt transcript as the result of read-through due probably to low levels of UTP in the three nucleoside triphosphates - the next U residue is at position 27. By contrast, using the PCR produced fragment in which the U residue at position 21 is changed to an A (in the transcript sequence) results predominantly in 26 nt transcript and the expected readthrough to position 32.

Removal of excess primers from PCR reactions

A convenient means of removing excess primers from completed PCR reactions in order to facilitate subsequent manipulations, such as above, is the use

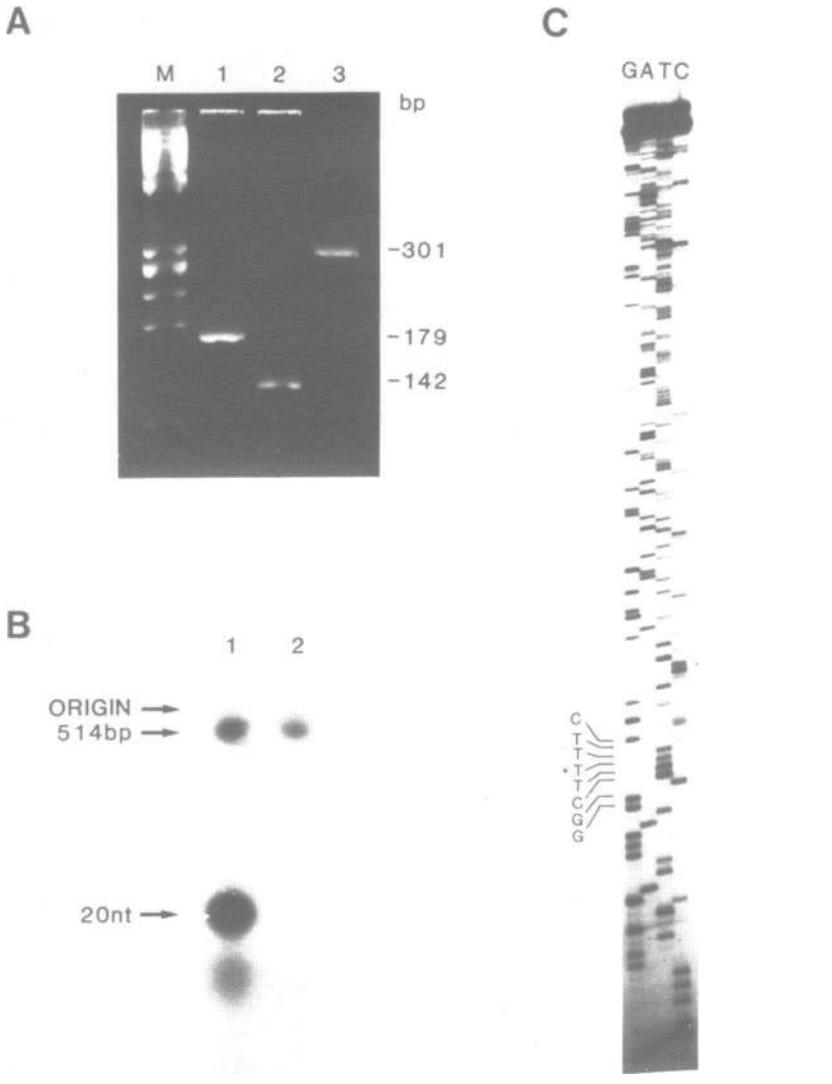

Figure 5. (A) Production of primary (lanes 1 and 2), and recombined, secondary (lane 3), specifically mutagenized PCR products by the method diagrammed in Fig. 4 and described in detail in the text. The PCR products are shown resolved by electrophoresis on a 4%, Nusieve-agarose (FMC) gel, here photographed under UV light after staining with etidium bromide. M is a marker digest of phiX174 DNA cut with Hae III. (B) Removal of excess primers from a PCR reaction by ultrafiltration on a Centricon 100 (Amicon). Shown is an autoradiograph of an electrophoretic fractionation as in (A) of (lane 1) an aliquout of completed PCR reaction containing both excess ^{32}P labeled, 20 nt primer and labeled 514 bp

product, and (lane 2), an equivalent aliquout of the retentate of a single Centricon 100 spin of this PCR reaction. (C) Direct sequence of the recombined, secondary PCR product shown in (A.). Dideoxy-termination sequencing reactions on this product were performed as described (15, 16) using RH28 (Fig. 2) as an end-labeled sequencing primer. The A to T mutation is indicated (compare to Fig. 3A).

of a centrifuge driven, ultrafiltration device. A Centricon 100 contains a filter that allows the passage of oligonucleotides (oligos of up to 48 nt were tested) and does not allow the passage of ds DNA as small as 100 bp. The filter in a Centricon 30, by comparison, has high retention of 20 nt oligos. The use of Centricon 100s is demonstrated in Fig. 5B. By electrophoresis on an agarose gel, the retentate and eluate of a Centricon 100 fractionation of a PCR containing a ^{32}P-end-labeled 20 nt primer are compared to the unfractionated reaction. In the retentate, the product 514 bp fragment is >95% free of the 20 nt primer.

DISCUSSION

The major advantages of PCR over standard restriction enzyme preparation of DNA fragments in our studies have been (1) the ability to obtain a DNA template of a defined sequence regardless of restriction enzyme sites (2) ease and short time of preparation (3) convenient addition of radioactive label or 5' "add on" sequences to either end via the synthetic primers (4) diminished exposure to radioactive material during the preparative procedures and (5) in our hands, higher specific-activity of labeled DNA fragments.

In this paper we have shown that a PCR produced and end-labeled fragment can be used in Maxam-Gilbert sequence analysis, DNAse I protection "footprinting", and the isolation and analysis of specific *E. coli* RNA polymerase transcription complexes. We have also demonstrated a simple, general procedure to introduce, using PCR, specific mutations anywhere along the length of these fragments. The sum of these procedures is a convenient system for the determination of the effect of altering specific DNA sequences on protein/DNA interactions.

One could imagine as well an experimental design where a mutagenized gene coding for a protein of interest could be produced by PCR and expressed *in vitro* using a coupled transcription/translation system (16). This would provide a means of studying the affect of changes introduced at the DNA level on the activity of the protein products without the use of molecular cloning, as long as the study did not require large amounts of the protein in purified form. DNase I protection

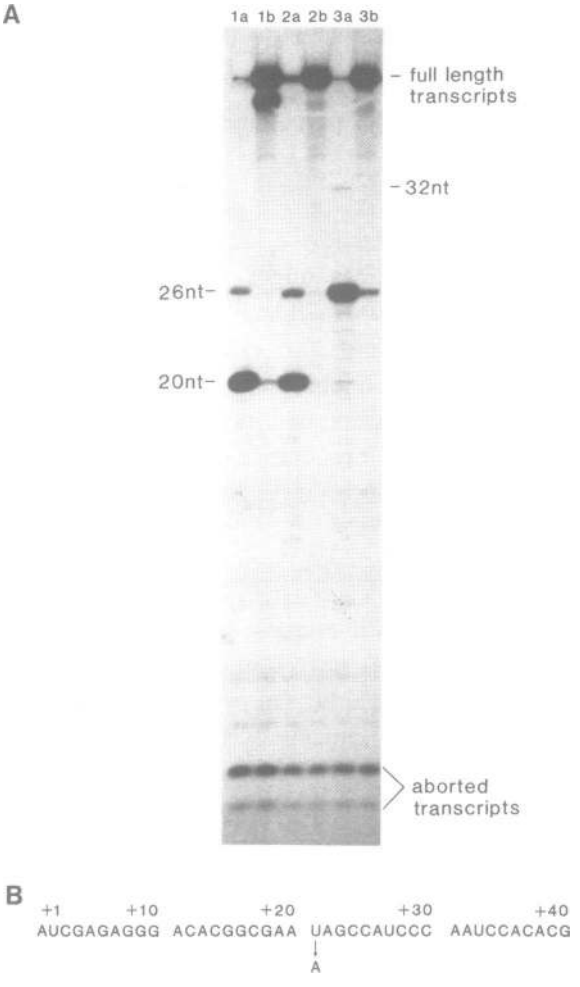

Figure 6. Comparison of transcription from the phage T7A1 promoter on a restriction fragment and PCR original or altered fragments. (A) Transcription conditions were as described in Methods. Lanes 1, 2 and 3 show transcripts from the restriction fragment, PCR original fragment and the PCR altered fragment respectively. Numbers followed by "nt" are transcript lengths in nucleotides. Lanes 1a, 2a and 3a show transcripts synthesized in ternary complex formation conditions where no UTP is present. Lanes 1b, 2b and 3b show the elongation of these transcripts to full length after the addition of 400 μM of all four NTPs.
(B) The RNA sequence of the first 40 nucleotides of the T7A1 transcription unit is shown. +1 indicates the first nucleotide in the transcript. The sequence change in the transcript (U-->A) that results from the PCR generated mutagenesis of the DNA template is indicated.

footprints, for example, can be readily obtained using crude mixtures of proteins in which the protecting protein is at low level (17). Thus a study where both the DNA and a specific binding protein can be mutated and prepared entirely *in vitro* is possible. For such studies, however, the rate at which Taq DNA polymerase misincorporates nucleotides will need to be taken into consideration.

Because it is estimated that Taq DNA polymerase will insert an incorrect nucleotide every 6 x 10^3 nucleotides incorporated (4; this estimate is still preliminary), and because these misincorporations are cumulative over each PCR cycle of replication, there can be in any one DNA molecule produced after, for example, 30 such replication cycles a base-pair change every 400 basepairs (4). When these molecules are examined in the aggregate, as in the experiments described in this paper, there are at any single position presumably 399 molecules with the correct basepair for each with an incorrect one. If this DNA is cloned, however, any 400 basepair length of cloned DNA is likely to have a basepair change. Therefore, applying the above methodology to large fragments that are to be cloned (for expression, for example) carries the caveat that additional, unwanted mutagenesis may occur. There are analogous concerns about the cloning of completely synthetic genes (1) due to the misincorporation rate of organic synthesis. At least until more experience is obtained with cloned, PCR produced DNA, verification of such clones by sequencing will probably be necessary.

Similarly, RNA transcripts of this aggregate PCR product will have errors at the same rate. If such a single, 400 nt transcript, which has a nearly even chance of carrying a nucleotide substitution, is translated, the resultant protein molecule will have about two-thirds of that probability of carrying an amino-acid substitution and a much lower chance of being prematurely terminated. In the scenerio described above, in which proteins made off PCR product are used in a DNase I protection assay, these random substitutions will probably not have a major effect because (1) the majority of these changes will not affect the DNA/protein interaction and (2) the effect of any one random change that does affect the interaction will be blurred among the effect of all the changes that do.

We also note that the mutation rate of the PCR process can be reduced by the use of the large fragment of DNA polymerase I (2,3) or by the use of the antimutator T4 DNA polymerase (18). These enzymes of course lack the convenient property of thermostability and would need replacement after each denaturation cycle of PCR. PCR with these enzymes is also less specific at the lower temperature of primer extension required, although this may not be a problem from templates of low sequence complexity such as plasmids. It may also

be possible to find conditions in which the rate of misincorporation by Taq DNA polymerase is reduced. Since the error frequency is cumulative with the number of replications, it will always be advisable to keep the number of PCR cycles at a minimum whenever possible.

Finally, we note that the method of combining PCR products into larger PCR products shown here can be used to combine sequences from different templates. All that is needed is to add to the sequences from one template to a region of overlap with the other sequence via a 5' "add-on", and it should be possible to combine almost any two sequences at a given junction, independent of available restriction sites. More than two sequences can be specfically joined by the serial application of this principle. Thus complex constructions which would otherwise involve several cycles of restriction enzyme or nuclease digestion, fragment purification, ligation, molecular cloning and the screening of clones that may take weeks to complete may be done much more quickly. Because of the uncertainty noted above about the mutation rate of the PCR process, however, if this construct is to be used for transgenic purposes it would be advisable to first clone the final construct and verify its sequence. Therefore, small constructs which are easily sequenced are suggested until more experience is acquired.

ACKNOWLEDGEMENTS

We thank M. Chamberlin and H. Erlich, in whose laboratories this work was carried out. B.K. was supported by a training grant (GM07232) from the National Institutes of Health. We thank M. Brow, who first did chemical sequencing on end-labeled PCR product. We also thank M. Bear and Amicon for Centricon 100 devices, C. Levenson, D. Spasic and L. Goda for oligonucleotides, D. Gelfand and S. Stoffel for Taq DNA polymerase, S. Nilson for help with graphics, and M. Kriegler for helpful comments on this manuscript.

REFERENCES

1. Wosnick,M.A., Barnett,R.W., Vicentini,A.M., Erfle,H. Elliott,R., Sumner-Smith,M., Mantei,N., and Davies,R.W. (1987) Gene **60**, 115-127.
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. Faloona,F. and Mullis,K. (1987) Methods in Enzymology **155**, 335-350.
4. 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.
5. Gonzalez,N., Wiggs,J., and Chamberlin,M.J. (1977) Arch. Biochem. Biophys. **182**, 404-408.
6. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) Molecular Cloning, a Laboratory Manual, Cold Spring Harbor, NY.
7. Dunn,I.J. and Studier,F.W. (1981) J. Mol. Biol. **148**, 303-330.

8. Levin,J.R., Krummel,B., and Chamberlin,M.J. (1987) J. Mol. Biol. **196**, 85-100.
9. Straney,D.C. and Crothers,D.M. (1985) Cell **43**, 449-459.
10. Maxam,A.M. and Gilbert,W. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 560-564.
11. Siebenlist,U. (1979) Nature **279**, 651-652.
12. Schmitz,A. and Galas,D.J. (1979) Nuc. Acids Res. **6**, 111-137.
13. Scharf,S.J., Horn,G.T., and Erlich,H.A. (1986) Science **233**, 1076-1078.
14. Higuchi,R., von Beroldingen,C., Sensabaugh,G., and Erlich,H.A. (1988) Nature **332**, 543-546.
15. Wrischnik,L.A., Higuchi,R.G., Stoneking,M., Erlich,H.A., Arnheim,N., and Wilson,A.C. (1987) Nuc. Acids Res. **15**, 529-542.
16. Bujard,H., Gentz,R., Lanzer,M., Stueber,D., Mueller,M., Ibrahimi,I., Haeuptle,M-T., and Dobberstein,B. (1987) Methods in Enzymology **155**, 416-433.
17. Briggs,M.R., Kadonaga,J.T., Bell,S.P., and Tjian,R. (1986) Science **234**, 47-52.
18. Keohavong,P., Kat,A.G., Cariello,N.F., and Thilly,W.G. (1988) DNA **7**, 63-70.