



# GENOME RESEARCH

## Long-distance PCR.

O S Foord and E A Rose

*Genome Res.* 1994 3: S149-S161

---

**References** This article cites 31 articles, 5 of which can be accessed free at:  
<http://genome.cshlp.org/content/3/6/S149.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](#).

---

CRISPR and RNAi Genetic Screening.  
Your new superpower.

LEARN  
MORE



---

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

---

Copyright © Cold Spring Harbor Laboratory Press

## Long-distance PCR

Orit S. Foord  
and Elise A. Rose

Advanced Center for Genetic  
Technology, The Perkin-Elmer  
Corporation, Foster City,  
California 94404

The versatility and power of PCR<sup>(1)</sup> have encouraged its involvement in almost every aspect of molecular biology research. This phenomenon is perhaps best exemplified by the impact the technology has had in genome mapping and sequencing.<sup>(2)</sup> The ability to routinely and specifically amplify and detect PCR products ranging in size from <1 kb to >50 kb, regardless of target template sequence or structure, would revolutionize a broad spectrum of biological research applications.<sup>(2-4)</sup> For example, the ability to amplify fragments up to 20–50 kb would potentially enable the isolation of an entire gene from a cDNA, thereby obviating the time-consuming task of screening a genomic library for the target gene. Additionally, the generation of a 20-kb fragment would span approximately half of the DNA cloned into a cosmid, thus making it possible to access the entire insert from as few as two amplifications initiating from the left and right vector cloning sites. Amplification of 50-kb targets would provide the same benefits to the analysis of bacterial artificial chromosome (BAC) and P1 derived artificial chromosome (PAC) clones, as well as small yeast artificial chromosomes (YACs).<sup>(5-7)</sup> Microbiological manipulations and vector purification efforts therefore would be minimized and subsequent clone analysis facilitated.

Long-distance PCR would make it possible to isolate large genomic fragments from complex genomes, as well as from hybrid cell lines or from microdissected or flow-sorted chromosomal regions.<sup>(8)</sup> Generation of large cDNAs from an uncloned pool of mRNA extension products would also be feasible. These latter applications would be especially useful when studying tissues or species where appropriate libraries are unavailable. Long-distance PCR also enables the amplification of eukaryotic genomic DNA segments containing introns of varying number and lengths, thus facilitating the delineation of intron/exon boundaries.

Applications with particular relevance to genome research include the use of long-distance PCR to analyze the higher-molecular-weight restriction fragment length polymorphisms within a population. It should also be possible to amplify large regions containing genes with expandable triplet repeats, such as the Huntington's Disease gene, making presymptomatic diagnosis by PCR technically feasible. In genome mapping, the use of chromosomal walking to isolate uncharacterized DNA located outside of adjacent known sequences would be greatly enhanced as the reach of this methodology is extended. Expansion of such contiguous clones, including YACs, BACs, PACs, and cosmids, is currently one of the most time-consuming tasks in the construction of physical genomic maps. The ability to generate a wide size range of amplification products would potentially expedite isolation of uncloned DNA (gaps) represented in physical maps and provide a means for maintaining order and orientation of closely linked loci during analysis. The utility of methodologies involving amplification of regions between specific interspersed repetitive elements (or between combinations of types of repeat units) would be enhanced with long-distance PCR, thus facilitating fingerprinting of DNA fragments, as well as ordering and orienting contiguous DNA clones.<sup>(8)</sup> Finally, perhaps one of the most exciting uses of long-distance PCR would be its application in directed transposon-based mapping and sequencing, a potentially powerful strategy for rapidly and efficiently analyzing templates of 100 kb or greater.<sup>(3,4,9)</sup>

Current limitations in the size of PCR products that are reliably generated are on the order of 3–4 kb.<sup>(10)</sup> Although several reports have appeared describing amplification of larger PCR fragments,<sup>(3,4,11-17)</sup> relatively little effort has been made to optimize conditions for the general usage of long-distance amplification. Moreover, research involving long-distance PCR has utilized templates of known size and, in many cases, known sequence. This has resulted in the retrofitting of experimental design to obtain the expected re-

sults. Consequently, attempts to use these methods for long-distance PCR in other primer–template systems have been generally unsuccessful. Few studies have addressed reproducible and specific amplification of large templates of unknown length or sequence.<sup>(4)</sup> For long-distance PCR to be of general use to the scientific community, a robust methodology not only must be available for the routine amplification of a range of specific high-molecular-weight DNA products, but simultaneously, for efficient extension through those genome regions that have been notoriously recalcitrant to PCR. Such regions include those that are especially rich in guanine (G) and cytosine (C) or contain complex secondary structure.<sup>(4)</sup> Fidelity of replication becomes yet another important issue when long-distance PCR fragments are used in cloning or direct sequence analysis.

Implementation of such enhanced PCR capacities would provide significant savings in terms of time, materials, and cost. In designing a general methodology for long-distance PCR, three topics need to be addressed: the components of the reaction mixture, thermocycling conditions, and methods for detecting the amplified product.

### **REACTION MIXTURE COMPONENTS**

The components that are present in the PCR mixture do not singularly ensure success of DNA amplification. Consideration of thermocycling parameters and DNA detection methodologies are equally important in achieving reproducible and specific PCR. This is especially relevant when addressing the ability to generate high-molecular-weight PCR products. A detailed discussion of the individual components included in long-distance amplification reaction mixtures is important, however, as several generally accepted “standard” PCR practices must be altered when attempting to extend the range of PCR product size above 3–4 kb.

#### **Template**

Degraded, damaged, nicked, and unpurified DNA from a multitude of sources has been used successfully in routine PCR. Amplification of higher-molecular-weight PCR products, however, requires that longer stretches of target template be intact. It is therefore important that the majority of DNA present in a long-distance PCR mixture serves as high-quality template. The protocols and precautions of genome researchers involved in cloning and analysis of high-molecular-weight DNA become relevant when attempting to amplify targets approaching 10–50 kb. Isolation procedures that yield the highest quality vector and genomic DNA are therefore encouraged. For example, a purified double-stranded circular vector, such as BAC, PAC, and cosmid clones, containing large cloned inserts (40 kb or greater) should consist primarily of supercoiled DNA. More complex genomic DNA should average, at minimum, 50 kb in size. The analysis of genomic DNA on a low-percentage (e.g., 0.4%) agarose gel or pulsed-field gel electrophoresis is recommended prior to PCR. We have found that whole chromosomal DNA prepared in “pulsed-field quality” low-melting-point agarose,<sup>(18)</sup> followed by agarose digestion,<sup>(19)</sup> generates superior yield of large PCR product as compared with DNA isolated in solution.<sup>(20)</sup> Size analysis of these DNAs using pulsed-field electrophoresis demonstrates that the DNA prepared in low-melting-point agarose ranges several megabases in size, whereas the average size of DNA prepared in solution is 75–500 kb.

In summary, high-integrity DNA provides a greater concentration of intact initial template available for PCR. We note, however, that amplification of larger DNA target templates from more complex DNA sources may need extended denaturation time or use of a higher initial denaturation temperature,

as well as an increase in both temperature and time for efficient primer annealing (see Thermocycling Conditions, below). This later point is of particular relevance when PCR templates of unknown length, sequence nature (e.g., G + C content), or presence of secondary structure are being amplified.

Template concentration directly affects reproducibility, specificity, and yield of amplification product. Excess template increases mispriming and, consequently, the extent of nonspecificity. If too little initial template is used, the amplification of the higher-molecular-weight, lower-copy-number target is less reproducible, with the PCR product often not detectable. Initial template concentration should thus be optimized to achieve a balance between specificity and yield given the expected size range of the PCR products being generated. In long-distance PCR studies that have involved use of plasmid or  $\lambda$  templates, various investigators have amplified their targets from 0.1 to 100 ng of starting template.<sup>(3,4,12–17)</sup> Studies involving the use of genomic DNA in long PCR have used as little as 100 ng of starting template.<sup>(11)</sup>

### **Primers**

Optimal primer design provides a balance between specificity and efficiency (or yield) of amplification.<sup>(21)</sup> Specificity reflects the frequency with which mispriming occurs, resulting in unrelated amplification products.<sup>(21)</sup> Many of the generally accepted rules for primer design apply to long-distance PCR. For example, we and others have found that primer length is optimal at 18–25 nucleotides.<sup>(3,4,11–15)</sup> When designing primers for amplification of regions from complex genomes, or regions known to be particularly G + C rich, increased length provides additional stability. The addition of each incremental nucleotide to the primer will confer approximately four times more specificity.<sup>(21)</sup> As in standard PCR, complementarity at the 3' ends of the primers should be avoided.<sup>(22)</sup> This is important in long-distance PCR, as any competition for reagents (such as occurs in the formation of primer-dimer) will diminish yield of the desired product. It is also often beneficial to incorporate dinucleotides of G's and/or C's at the 3' ends of primers.<sup>(23)</sup> As a general rule, factors that increase the stability of the 3' end of an oligonucleotide (where polymerase extension is initiated) will enhance both specificity and efficiency of the PCR.

The concentration of primers used in long-distance PCR does not differ from standard PCR conditions. Primer concentration should be nonlimiting but not sufficiently in excess to promote mispriming and primer oligomerization. Successful long-distance PCR products have been generated using 10 pmoles to 200 nmoles of each primer per 50- to 100- $\mu$ l reaction.<sup>(3,4,11–15,17)</sup> Adherence to absolute concentrations of oligonucleotides therefore does not appear to be crucial to the success of long-distance PCR. For reasons of cost and convenience, for example, we routinely use 30 pmoles of each primer for a 50- $\mu$ l PCR.

There are a number of software programs available for primer design. For long-distance PCR, Rychlik and Rhodes<sup>(24)</sup> suggest the use of the nearest neighbor algorithm for selection of melting temperature. In our experience, use of this calculation for primer design yields superior long-distance PCR results. The melting temperature of primers should be as high as possible to promote specificity without significantly compromising yield of product. In addition, the nucleotide composition of primers should reflect the nucleotide composition of the region from which the desired target is to be amplified (e.g., if the target region is 50% G + C, then the primers should also be 50% G + C).

Although it is generally recommended that the melting temperatures of primer pairs be balanced,<sup>(25)</sup> differences in the melting temperatures of

primer pairs are not of particular consequence in routine short-distance PCR. For the generation of long-distance PCR products, however, where extension times are increased, we have found that the primer design should be optimized so that both primers have melting temperatures within 1°C of each other.<sup>(3,4)</sup> Even very small divergence in primer pair melting temperatures promotes mispriming.<sup>(3,24)</sup>

### Enzymes

More than a dozen different thermostable enzymes have been used for PCR amplification. The utility of many of these enzymes in long-distance PCR has been evaluated in several reports.<sup>(3,4,11–17)</sup> DNA polymerases with 5' → 3' exonuclease ("nick translation") activity have yielded consistently superior performance in long-distance PCR as compared with those enzymes exhibiting 3' → 5' exonuclease ("proofreading") activity, both 3' → 5' and 5' → 3' exonuclease activities, or neither exonuclease activity. In our early studies we examined enzymes with various combinations of 5' → 3' and 3' → 5' exonuclease activities, in the buffers recommended by their manufacturers (see Table 1).<sup>(3)</sup> Although some of the characteristics of these enzymes have yet to be defined, they clearly differ in a number of additional activities that are potentially important for long-distance PCR, including extension rate, processivity, fidelity, thermostability, and thermal activity profile.

In our hands, under the conditions utilized, *rTth* DNA polymerase (the recombinant form of the DNA polymerase from *Thermus thermophilus*; Perkin-Elmer) at 0.5–2.5 U/50 μl reaction provided the most consistently successful performance in long-distance PCR.<sup>(3)</sup> We and others have also found that Hot Tub DNA polymerase (isolated from *Thermus ubiquitius*; Amersham) can be used to reliably generate PCR products as large as 6–15.6 kb.<sup>(3,13,16)</sup> AmpliTaq DNA polymerase (the recombinant form of the DNA polymerase from *Thermus aquaticus*; Perkin-Elmer) and native *Taq* DNA polymerase (Perkin-Elmer) have also been used to generate long-distance PCR products, but results have not been as specific or reproducible as required.<sup>(3,16,17)</sup> In contrast, enzymes with 3' → 5' exonuclease activity, such as *Vent* DNA polymerase (isolated from *Thermus litoralis*; New England Biolabs),<sup>(3,15)</sup> or *Pfu* DNA polymerase (isolated from *Pyrococcus furiosus*; Stratagene)<sup>(17)</sup> failed to consistently generate specific PCR products >6 kb in length.<sup>(3,15)</sup> Most of our efforts have therefore focused on optimizing conditions for use of *rTth* DNA polymerase in long-distance PCR.<sup>(3,4,9)</sup>

Recombinant *Tth* DNA polymerase has a half-life of 20 min at 95°C (sufficient to remain active over 30 or more cycles during which the enzyme is transiently exposed to such denaturation temperatures), optimal DNA polymerase activity in the same temperature range at which stringent primer annealing occurs (65–75°C), 5' → 3' exonuclease activity, and no detectable 3' → 5' exonuclease activity.<sup>(28,29)</sup> The fact that *rTth* DNA polymerase lacks

**TABLE 1** Exonuclease Activities of DNA Polymerases

DNA polymerase	5' → 3' exonuclease	3' → 5' exonuclease	References
AmpliTaq	+	–	26, 27
<i>rTth</i>	+	–	28, 29
Hot Tub	+	–	30
<i>Vent</i>	–	+	31
<i>Tma</i>	+	+	32
AmpliTaq Exo-mutain	–	–	33
Stoffel fragment	–	–	27

(+) Enzyme possesses activity; (–) enzyme lacks activity.

3' → 5' exonuclease activity minimizes the likelihood that primers and single-stranded templates are destroyed during PCR. The enzyme's processivity (i.e., the number of nucleotides replicated before the enzyme dissociates from the template) is 30–40 nucleotides, and its extension rate is ~60 nucleotides per second.<sup>(29)</sup>

Following our initial success with *rTth* in moderately long-distance PCR, we suggested that the use of combinations of enzymes with different activities might be beneficial in amplifying long templates.<sup>(3)</sup> For example, it is possible that small amounts of 3' → 5' exonuclease activity, which edits nucleotide mismatches, might enhance the efficiency of enzymes such as *rTth* for DNA amplification over long distances. There is a high probability that an enzyme containing 3' → 5' exonuclease activity exists in bacteria such as *Thermus thermophilus* in vivo and may play a role in concert with *rTth* in chromosomal repair. In isolating a pure DNA polymerase from this thermostable bacterium, however, other enzymes that may function together with *rTth*, as well as necessary accessory proteins, have been eliminated. It is also important to note that benefits derived from the use of multiple enzymes in long-distance PCR might result from activities other than inherent exonuclease activities, such as processivity and extension rate. Regardless of the mechanism<sup>(3)</sup> Barnes<sup>(17)</sup> has recently reported PCR amplification of up to 35 kb from  $\lambda$  templates using a combination of an amino-terminal deletion mutant of *Taq* DNA polymerase (KlenTaq I) and *Pfu* DNA polymerase.

The thermostable enzymes currently used for PCR amplification are most likely repair enzymes, which characteristically replicate short stretches of nucleotides with low processivity and low extension rates. This is in contrast to replication enzymes,<sup>(34)</sup> which display high processivity and rates of extension. For example, a well-characterized replication enzyme is the nonthermostable polymerase of *Escherichia coli*, DNA polymerase III (Pol III holoenzyme).<sup>(35)</sup> Pol III holoenzyme has an extension rate of 750 bases per second and processivity of 100–500 bases.<sup>(35)</sup> The use of such replication enzymes alone might significantly enhance our ability to obtain high yield of very long PCR products (20–50 kb). The isolation of thermostable replication enzymes for use in PCR would clearly make this approach testable. Because the *E. coli* Pol III holoenzyme is not thermostable, however, an alternative would be to modify reaction conditions in a manner enabling one to exploit the enzyme's useful activity. For example, it may be possible to use creative oligonucleotide design to achieve primer–template annealing and dissociation, so that enzyme driven polymerization can occur repetitively without requiring a high-temperature denaturation step.

Finally, we have speculated that benefits to long-distance PCR might be realized through the use of alternative thermostable DNA polymerases (including mutant enzymes with alterations to specific active sites) and the combination of such thermostable enzymes with thermostable pyrophosphatases and accessory proteins (e.g., thermostable helicases and single-stranded DNA displacement proteins).<sup>(3)</sup> Studies concerning the use of such proteins in long-distance PCR are currently in progress.

### **Buffers**

Standard PCR protocols recommend a buffer comprised of 10 mM Tris-HCl (pH 8.3–8.4 at 20–25°C) for both *Taq* DNA polymerase and *rTth* DNA polymerase. Although often resulting in successful PCR, this common Tris-HCl buffer system was historically intended to represent a starting point from which reaction conditions could be optimized for specific primer–template systems. It does not represent the optimal buffer system for long-distance PCR. Early in our studies we realized that two parameters needed to be ad-

dressed to define more appropriate buffer conditions. First, we obtained superior results when we increased the molarity of Tris 2.5-fold from 10 to 25 mM.<sup>(3)</sup> This suggests that more buffering capacity is necessary when utilizing the longer extension times required for amplification of large templates. Second, we increased the pH of the reaction mixture from the standard 8.3–8.4 to 8.9–9.0 (at 20–25°C),<sup>(3)</sup> owing to the fact that the pKa drops significantly with increasing temperature,<sup>(36)</sup> and we wanted to maintain neutral pH at the chosen extension temperature. In a normal functioning cell, *in vivo*, pH is highly regulated. It thus seems logical that one would want to keep neutral pH tightly controlled at the temperature at which the enzyme is expected to function optimally as a polymerase (i.e., 70–72°C). If a pH 9.0 (at 25°C) Tris buffer is used, the reaction mixture can be maintained at approximately neutral pH throughout the range of temperatures required for primer annealing and extension by either *Taq* DNA polymerase or *rTth* DNA polymerase.

We find that long-distance PCR with *rTth* DNA polymerase works well with a buffer containing 5% glycerol (vol/vol), 25 mM Tris-HCl (pH 8.9 at room temperature), 100 mM KCl, 0.75 mM EGTA, 0.05% Tween 20, 2.5 mM MgCl<sub>2</sub>, and 0.01% gelatin.<sup>(3)</sup> The EGTA is present to chelate manganese and inhibit any reverse transcriptase activity that might otherwise be exhibited by this enzyme. As in standard PCR with *rTth*, a magnesium-dependent DNA polymerase, optimal activity is observed in a concentration range of 1.5 to 2.5 mM MgCl<sub>2</sub>.<sup>(29)</sup> Total dNTPs are added to 1 mM (250 μM each), again in nonlimiting quantities sufficient for generation of the desired long products.<sup>(3)</sup>

Given the large temperature dependence of pKa in Tris buffers, Ponce and Mikol<sup>(14)</sup> proposed the use of a less-temperature-sensitive Tricine buffer. These investigators used a 30 mM Tricine buffer (pH 8.4) at room temperature, to obtain specific 6.2-kb products with *Taq* DNA polymerase. The use of less temperature-sensitive buffers such as Tricine, or a combination of buffers with pKa's falling within the neutral pH range at optimal enzyme extension temperatures (more similar to *in vivo* conditions), thus might enhance the ability to obtain long PCR products.

Additional components that proved useful in amplifying long templates included 5% glycerol and 0.01% gelatin. The presence of gelatin in excess of 0.05% inhibited long-distance PCR. Finally, we siliconize PCR tubes routinely to prevent the enzyme and low-copy templates from sticking to the walls.<sup>(3)</sup>

### Wax-mediated Hot Starts

We speculated that optimization of long-range PCR would benefit from the use of approaches that maximize the specificity of PCR. Such methods should ensure that all available primers, nucleotides, and enzyme molecules would be used to generate the desired product. Hot starts enhance PCR specificity by eliminating production of nonspecific products resulting from pre-PCR mispriming and primer oligomerization (primer-dimer) during the initial steps of PCR.<sup>(37)</sup> AmpliWax PCR Gems (Perkin-Elmer) proved to be an effective and convenient tool for implementing this technique. We find that uniformity and yield of PCR also improve using this approach. The uniformity effect is most likely attributable to providing a vapor barrier of consistent mass in each reaction tube, and the enhanced yield from focusing all reagents on production of the specific desired product. By permitting specific PCR to occur at temperatures that would otherwise result in the formation of non-specific products, wax-mediated hot starts provide the additional benefit of decreasing the time required for PCR optimization.

### Reaction Volume

Historically, most PCR amplifications have been carried out in 100 μl vol-

umes. This also applies to most published studies on long-distance PCR. Recently, however, there has been a trend toward the use of lower reaction volumes. Further, we have found that smaller reaction volumes (e.g., 50  $\mu$ l) yield superior results.<sup>(3)</sup> We speculate that this is attributable to improved heat transfer, especially given that such results are further enhanced by the use of thin-walled MicroAmp reaction tubes in the Perkin-Elmer GeneAmp PCR System 9600, an instrument designed for tight fit of these tubes.<sup>(38)</sup> This point is important as the use of tubes that do not fit properly into thermal cyclers, or that have not been designed for thermal cycling, results in inefficient heat transfer and would be expected to yield inferior results if used in long-distance PCR. One might also expect that rapid cycling in glass capillary tubes would provide excellent heat transfer. Heat transfer could be improved even further if PCR could be performed in submicroliter volumes in microfabricated instruments. Technical advances in this area may provide significant benefits to amplification of long templates.

### **THERMOCYCLING CONDITIONS**

One cycle in a PCR amplification reaction consists of three steps that can be carried out at two or three separate temperatures. A typical cycle consists of a denaturation step at 90–97°C, an annealing step at 40–60°C, and an extension step at 70–75°C. It is often helpful to precede the first denaturation step with an initial predenaturation at 95°C for 3 min. A common standard PCR protocol might involve 30 cycles, each consisting of denaturation for 1 min at 94°C, annealing for 1 min at 55°C, and extension for 1 min at 72°C. For optimal PCR, however, the specific times and temperatures used for each step are chosen depending on the characteristics of the primer–template system being used. Two-temperature PCR can be employed because most thermostable DNA polymerases used in PCR are actively extending off the primers over the entire temperature range between the points commonly chosen for annealing and extension (enabling one to use a combined anneal/extend step).

Although denaturation occurs rapidly, longer denaturation times might be expected to produce superior results when amplifying long templates. At 95°C, *Taq* DNA polymerase still maintains 50% of its original activity after 40 min, and *rTth* has a half-life of 20 min, both of which are more than sufficient to provide adequate activity throughout the 25–30 cycles of most PCRs during which the enzyme is transiently exposed to these high temperatures.<sup>(39)</sup>

Annealing temperatures are chosen on the basis of the primer melting temperatures (see discussion of primers above) and estimated template lengths and G + C content. We routinely use the equations defined by Rychlik et al.<sup>(12,24)</sup> for the calculation of the optimal temperature for primer–template annealing conditions. These calculations have proven to be the most reliable in defining optimal annealing conditions when template length and nucleotide composition are unknown. Under these conditions, we calculate the annealing temperature for a size range of templates. The G + C content of templates is estimated based on the general nucleotide content of the genomic region of study.

*Taq* DNA polymerase extends at a rate of 0.25 nucleotides/sec at 22°C, 1.5 nucleotides/sec at 37°C, 24 nucleotides/sec at 55°C, >60 nucleotides/sec at 70°C, and 150 nucleotides/sec at 75–80°C.<sup>(39)</sup> Thus, at commonly chosen extension temperatures such as 70–72°C, *Taq* DNA polymerase would be expected to be extending at a rate of >3.5 kb/min. As a general rule, extension times of 1 min/kb are more than sufficient to generate the expected PCR product.<sup>(39)</sup> The same rule applies when using *rTth*, as this enzyme has an extension rate that is very similar to that of *Taq* DNA polymerase.

We have used the rationale described above in defining optimal conditions for long-distance PCR using a GeneAmp PCR System 9600 (Perkin-Elmer). Although other types of DNA thermal cyclers may have different performance characteristics, the general guidelines for designing long PCR thermocycling conditions should still apply. It should be noted, however, that any features of an instrument's design that compromise ability to reliably and reproducibly regulate temperature or incubation times or that result in inferior heat transfer would be expected to make the generation of long PCR products more difficult.

For a 12.2-kb template, using *rTth* DNA polymerase and the type of reaction mixture described in the sections above, we have demonstrated success in generating specific long-distance PCR products by running 10 cycles each consisting of a 1-min denaturation at 94°C, a 1-min annealing step at 55°C, and a 3-min extension step at 72°C, followed by 20 cycles in which the extension step is automatically extended 30 sec in each cycle. Automated hot starts are performed using AmpliWax PCR Gems, as described above.<sup>(3,4,9)</sup>

The ability to program the GeneAmp PCR System 9600 to automatically increase extension times in the later cycles of the reaction proved to be particularly useful. DNA polymerases extend primers discontinuously through a succession of reactions, moving on and off of the DNA template. As the PCR product begins to accumulate in the later cycles, the ability to extend a significant proportion of primers over a long distance in a given unit of time might be limited by the relative decrease in enzyme molecules per template. We speculated that increasing the extension time in each of the later PCR cycles would increase the likelihood that we would successfully synthesize long PCR products. Whether or not this rationale is correct, we found that use of autosegment extension resulted in more consistent generation of long PCR products.

### **DETECTION**

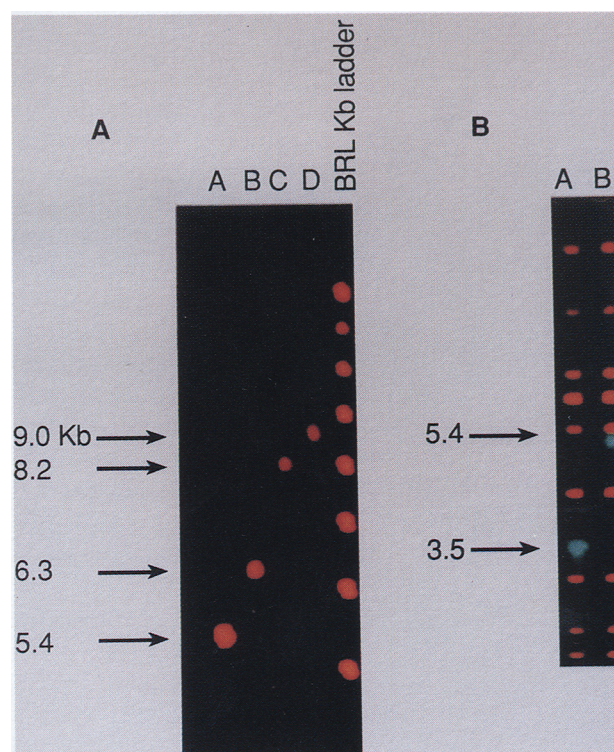
The method used for PCR product detection is equally as important as reaction mixture components and thermocycling conditions. Yield of PCR product decreases significantly with increasing product length, and high-molecular-weight PCR products are often generated at low copy number. For example, starting with 10–50 ng of template and using a single PCR protocol, we have reported that the yields of 3.5–5.4-kb PCR products were on the order of 20–100 ng/ $\mu$ l, whereas only 10 pg/ $\mu$ l of a 12.2-kb fragment was produced.<sup>(3,4)</sup> The need for sensitivity in detecting long-distance PCR products becomes even more important as one considers future trends toward the use of smaller reaction volumes. Although it may be possible to increase yield of long-distance PCR products as an alternative to employing more sensitive detection methodologies, our attempts have been compromised by decreases in specificity of amplification.

To date, the most commonly used method of detection of PCR products has been the use of ethidium bromide fluorescence with a standard UV light box. To achieve sufficient sensitivity to detect low-copy-number PCR products of high molecular weight, it is generally necessary to utilize the entire reaction volume. To achieve the necessary sensitivity, as well as to overcome background problems resulting from nonspecific amplification, probing with radioactively labeled single-copy DNA or oligonucleotide probes has often been utilized.<sup>(11,40)</sup> This approach suffers from the general problems associated with the use of radioactivity, including instability of probes, use of biohazardous materials, length of time required for detection, and difficulty in automation. Southern hybridization with oligonucleotide probes conjugated to enzymes or biotin, followed by either colorimetric or chemilumi-

nescent detection, overcomes many of the problems associated with the use of radioactivity. These methods still suffer from lack of sufficient sensitivity, however, if one needs to detect subnanogram quantities of PCR product. The same problems apply to use of probes labeled with currently available fluorescein dyes.

Use of primers labeled with biotin to incorporate signal-generating moieties directly into the PCR product do not provide adequate sensitivity for detecting high-molecular-weight PCR products present at low abundance. Primers that are end-labeled with fluorescent tags, although providing enhanced sensitivity compared with the other methods described above, are still inadequate (see Fig. 1).<sup>(3)</sup>

Detection of large amplification products is improved significantly if the number of signal-generating units incorporated per DNA molecule is increased.<sup>(3,4,9)</sup> Use of a DNA probe containing multiple biotinylated dUTPs, or direct incorporation of biotinylated dUTPs into the PCR product, provides enhanced signal per molecule but, because of the nature of the signal, still lacks sufficient sensitivity. We have found that optimal sensitivity is obtained either by intercalation of a fluorescent agent into the PCR product<sup>(3,4)</sup> or by direct incorporation of fluorescently labeled nucleotides during the PCR am-



**FIGURE 1** Detection of PCR fragments. (A) Lanes A–D display PCR fragments generated using the  $\lambda$  138::Tn5supF model system described previously.<sup>(3,15)</sup> Utilizing the ABI 362 Fluorescent Fragment Analyzer, a profile of PCR fragments ranging from 5.4 to 9.0 kb is shown. PCR products were detected using the inherent fluorescence of ethidium bromide. One to five microliters of each 50- $\mu$ l PCR mixture and ABI 2 $\times$  gel loading buffer were loaded onto a 0.8% agarose gel. The gel was run for 6 hr in 1 $\times$  TBE buffer, at ambient temperature, with 0.05  $\mu$ mole ethidium bromide included in both buffer and gel. (B) Detection of PCR products using primers conjugated with JOE (Applied Biosystems). GENESCAN-30000 labeled with ROX (Applied Biosystems) is run in the same lane as the sample to serve as a size marker as well as a quantitative marker. Samples were separated on a 0.8% agarose gel, and the gel was run for 6 hr in 1 $\times$  TBE buffer, at ambient temperature. Fluorescent PCR products were detected using the ABI 362 Fragment Analyzer. Lanes A and B correspond to the 3.5-kb and 5.4-kb PCR products generated using the  $\lambda$  138::Tn5supF transposon model system described previously.<sup>(3,15)</sup>

plification.<sup>(9)</sup> Such fluorescent PCR products may be detected at femtomolar concentrations.

In a previous series of experiments, we detected PCR products as large as 9 kb using the inherent fluorescence of the DNA intercalating agent ethidium bromide (Fig. 1).<sup>(3)</sup> The use of an Applied Biosystems model 362A Fluorescent Fragment Analyzer, which employs an argon-based laser system, enabled rapid and sensitive detection and sizing of PCR products. A major asset of this system is the ability to use internal size standards to adjust for slight variations in migration of DNA containing intercalated dyes. Another advantage of using this system is that it allowed us to analyze only 1–5  $\mu$ l of a 50- $\mu$ l PCR sample mixture, leaving the remaining PCR product available for other post-PCR procedures.

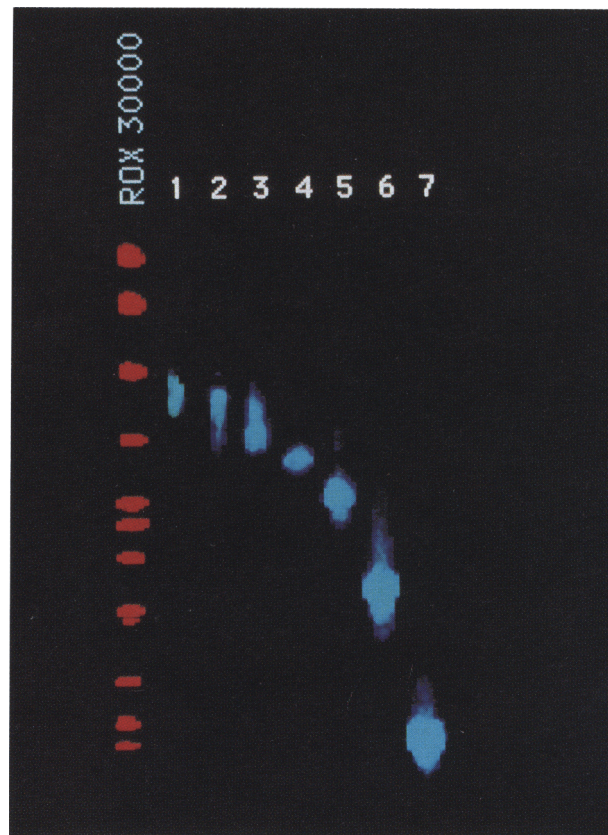
In more recent studies<sup>(4)</sup> we describe a modified protocol, adapted for our specific detection requirements, involving the use of an alternative fluorescent intercalating reagent, the thiazole orange dimer, TOTO-1 [1,1'-(4,4',7,7'-tetramethyl-4,7-diazaundecamethylene)-bis-4-(3-methyl-2,3-dihydro-[benzo-1,3-thiazole]-2-methylidene)-quinolinium tetraiodide].<sup>(41,42)</sup> The sensitivity of DNA detection achievable with TOTO-1, as compared with the use of conventional ethidium bromide staining of large PCR products, is several orders of magnitude greater. Again, using a model 362A Fluorescent Fragment Analyzer, we were able to detect PCR products ranging in size from 1 to 12.2 kb, where the larger PCR products were present at femtomolar concentrations in submicroliter volumes (Fig. 2).<sup>(4)</sup>

We have also used a series of newly developed fluorescent dUTP analogs for detecting long PCR products generated at low copy number.<sup>(9)</sup> Similar to other fluorescent nucleotides, these compounds contain either rhodamine or fluorescein derivatives as fluorophores. The dUTP analogs utilized in these studies differ in structure, however, from previously described fluorescent nucleotides. This modification in synthesis results in an ~500–1000-fold greater incorporation efficiency of these derivatives, and the detection of PCR products is correspondingly enhanced. Analysis of PCR fragments labeled with these novel nucleotides, again using the Fluorescent Fragment Analyzer, enables us to detect attomole amounts of DNA using 1/100 of a 50- $\mu$ l PCR mixture. This ability to analyze such small volumes with adequate sensitivity of detection will be critical in the future as further reductions in the scale of PCR are achieved. Use of these newly developed dUTP compounds has proven to be the most sensitive detection system available to us for analysis of DNA templates containing ~50–80% A+T. For genome regions containing a higher G+C content, we are investigating the use of an analogous dCTP fluorescent derivative.

Although all of the methodologies discussed above have utilized standard agarose gel electrophoresis to separate PCR products of different sizes, it should be noted that DNA fragments do not separate in a logarithmic fashion as their length approaches 50 kb. Even in the 20- to 50-kb range, longer and lower percentage agarose gels may be required to obtain adequate separation and accurate sizing. Pulsed-field gel electrophoresis is recommended for resolution of DNA fragments >50 kb in length. Enhanced detection of such very large DNA fragments can be achieved in an automated fashion, using an Applied Biosystems model 373A Fluorescent Fragment Analyzer together with field inversion gel electrophoresis (E. Rose, unpubl.).

## CONCLUSION

Great strides have been made over the past few years in extending the length over which one can reliably amplify specific DNA templates. In the near future it will be critical to gain more experience applying this technology to



**FIGURE 2** Lanes 1–7 contain PCR products of 12.2, 10.3, 8.2, 6.3, 5.4, 3.5, and 1.6 kb, respectively. Fragments were generated using the  $\lambda$  138::Tn5*supF* transposon model system described previously.<sup>(3,15)</sup> Samples were separated on a 0.8% agarose gel and the gel was run for 6 hr in 1 $\times$  TBE buffer, at ambient temperature. Fluorescent PCR products were detected using the ABI 362 Fragment Analyzer. Samples 3–7 were diluted 1:150 before reacting with the TOTO-1 dye and being subjected to electrophoretic analysis. Note that despite the dilution of 1:150, the 1.6- to 8.2-kb fragments were overloaded. The amount of PCR product decreases with increasing fragment size, however; therefore, it was necessary to use 1  $\mu$ l of the samples containing 10.3-kb and 12.2-kb PCR fragments. The ROX 30000 size standards (Applied Biosystems) were used in this experiment, as intercalation of DNA fragments (within this size range) with TOTO-1 does not significantly alter the mobility during electrophoretic separation.

unknown templates, as opposed to well-defined model systems. We have no doubt that further improvements will be developed involving enzyme formulations, reaction components, thermocycling conditions, and detection methodologies for longer PCR products. We believe that the ability to routinely and specifically amplify targets in the 20- to 50-kb range and beyond will have a revolutionary impact on many aspects of molecular biology research. Attention will thus shift from development of the methods themselves to the truly exciting work involving applications of long-distance PCR to address a vast array of basic questions in the study of biology.

#### ACKNOWLEDGMENTS

We gratefully thank Mary-grace Brubacher for review of this article. E.A.R. is funded by National Institutes of Health grant RO1-HG00565-01.

#### REFERENCES

1. Mullis, K.B. and F. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods Enzymol.* **1155**: 335–350.

2. Rose, E.A. 1991. Applications of the polymerase chain reaction to genome analysis. *FASEB J.* **5**: 46–51.
3. Ohler, L. and E.A. Rose. 1992. Optimization of long distance PCR using a transposon-based model system. *PCR Methods Applic.* **2**: 51–59.
4. Ohler, L., M. Zollo, E. Mansfield, and E.A. Rose. 1993. Use of a sensitive fluorescent intercalating dye to detect PCR products of low copy number and high molecular weight. *PCR Methods Applic.* **3**:85–140.
5. Shizuya, H., B. Birren, U.J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M.I. Simon. 1992. *Proc. Natl. Acad. Sci.* **89**:8794–8797.
6. Ioannou, P., C. Amemiya, A. Garnes, P. Kroisel, C. Chen., M. Batzer, A. Carrano, H. Shizuya, and P.J. DeJong. 1994. A new bacteriophage P<sub>1</sub>-derived vector for the propagation of large human DNA fragments. *Nature Genet.* **6**: 84–89.
7. Green, E.D. and M.V. Olson. 1990. Systematic screening of yeast artificial chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci.* **87**: 1213–1217.
8. Nelson, D.L., S.A. Ledbetter, L. Corbo, M.F. Victoria, R. Ramirez-Solis, T. Webster, D.H. Ledbetter, and C.T. Caskey. 1989. Alu polymerase chain reaction: A method for rapid isolation of human specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci.* **86**:6686–6690.
9. (a) Ohler, L., A. Blasband, M. Zollo, and E.A. Rose. 1993. The use of newly developed fluorescent nucleotide analogues for detection of high molecular weight, low copy number PCR products. *Genome mapping and sequencing*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, p.185; (b) Blasband, A. and E.A. Rose, in prep.
10. Erlich, H., D. Gelfand, and J. Sininsky. 1991. Recent advances in the polymerase chain reaction. *Science* **252**:1643–1650.
11. Jeffreys, A.L., V. Wilson, R. Neumann, and J. Keyte. 1988. Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **16**: 10953–10971.
12. Rychlik, W., W.J. Spencer, and R.E. Rhoads. 1990. Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Res.* **18**: 6409–6412.
13. Maga, E.A. and T. Richardson. 1991. Amplification of a 9.0-kb fragment using PCR. *BioTechniques* **11**: 185–186.
14. Ponce, M.R. and J. Micol. 1991. PCR amplification of long DNA fragments. *Nucleic Acids Res.* **20**: 623.
15. Krishnan, B.R., D. Kersulyte, I. Brikun, C.M. Berg, and D.E. Berg. 1991. Direct and crossover PCR amplification to facilitate Tn5<sup>supF</sup>-based sequencing of lambda clones. *Nucleic Acids Res.* **19**: 6177–6182.
16. Kainz, P., A. Schmiedlechner, and B. Strack. 1992. In vitro amplification of DNA >10 kb. *Anal. Biochem.* **202**: 46–49.
17. Barnes, W.M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from  $\lambda$  bacteriophage templates. *Proc. Natl. Acad. Sci.* **91**: 2216–2220.
18. Rose, E.A., T. Glaser, C. Jones, C.L. Smith, W.H. Lewis, K.M. Call, M. Minden, E. Chamagne, L. Bonetta, H. Yeager, and D. Housman. 1990. Complete physical map of the WAGR region of 11p13 localizes a candidate Wilms' tumor gene. *Cell* **60**: 495–508.
19. Agarase package insert, Sigma Chemical Co. 1993.
20. Blin, N. and D.W. Stafford. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.* **3**: 2303.
21. Dieffenbach, C.W., J. Lowe, and G.S. Dveksler. 1993. General concepts for PCR design. *PCR methods and applications.* **3**: 530–537.
22. Kwok, S., C.E. Kellog, N. McKinney, D. Spasic, L. Goda, C. Levenson, and J.J. Sninsky. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus 1 model studies. *Nucleic Acids Res.* **18**: 999–1005.
23. Lowe, T., J. Shareifkin, S.Q. Yang, and C.W. Dieffenbach. 1990. A computer program for selection of oligonucleotide primers for the polymerase chain reaction. *Nucleic Acids Res.* **18**: 1757–1761.
24. Rychlik, W. and R.E. Rhodes. 1989. A computer program for the choosing of oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucleic Acids Res.* **17**: 8543–8551.
25. Innis, M.A. and D.H. Gelfand. 1990. Optimization of PCRs. In: *PCR protocols: A guide to methods and applications* (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 3–12. Academic Press, San Diego, CA.
26. AmpliTaq DNA Polymerase package insert, Perkin-Elmer Cetus. 1991.
27. Lawyer, F.C., S. Stoffel, R.K. Saiki, S. Chang, P.A. Landre, R. Abramson, and D. Gelfand. 1993. High level expression, purification and enzymatic characterization of full length *Thermus aquaticus* DNA polymerase and a truncated form deficient in 5'–3' exonuclease activity. *PCR Methods and Applic.* **2**: 275–287.
28. rTth DNA polymerase package insert, Perkin-Elmer Cetus. 1991.
29. Myers, T.W. and D.H. Gelfand. 1991. Reverse transcription and DNA amplification by a *Thermus Thermophilus* DNA polymerase. *Biochemistry* **30**: 7661–7666.

30. Hot Tub DNA Polymerase package insert, Amersham. 1991.
31. VENT DNA Polymerase package insert, New England Biolabs. 1991.
32. UITma DNA Polymerase package insert, Perkin-Elmer. 1993.
33. Abramson, R., Roche Molecular Systems. (pers. comm.)
34. Kornberg, A. and T. Baker. 1991. *DNA Replication*. W.H. Freeman, New York.
35. Kong, X., G. Onrust, M. O'Donnell, and J. Kuriyan. 1992. Three dimensional structure of the B subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* **69**: 425–437
36. Good, N.E. and S. Izawa. 1972. *Methods Enzymol* **24B**: 53.
37. Chou, Q., M. Russell, D. Birch, J. Raymond, and W. Bloch. 1991. Prevention of pre-PCR mis-priming and primer dimerization improves low copy number amplification. *Nucleic Acids Res.* **20**: 7.
38. Haff, L., J.G. Atwood, J. Dicesare, E. Katz, E. Picozza, J. Williams, and T. Woudenberg. 1991. A high performance system for automation of the polymerase chain reaction. *BioTechniques* **10**: 102–112.
39. Gelfand, D.H. 1989. Taq DNA Polymerase. In: *PCR Technology—Principles and applications for DNA amplification*. (ed. H.E. Erlich), pp. 17–22. Stockton Press, New York.
40. Strausbaugh, L.D., M.T. Bourke, M.T. Sommer, M.E. Coon, and C.M. Berg. 1990. Probe mapping to facilitate transposon based DNA sequencing. *Proc. Natl. Acad. Sci.* **86**: 5908–5912.
41. Rye, H.S., S. Yue, D.E. Wemmer, M.A. Quesada, R.P. Haugland, R.A. Mathies, and A.N. Glazer. 1992. Stable fluorescent complexes of double stranded DNA with bis-intercalating asymmetric cyanine dyes: Properties and applications. *Nucleic Acids Res.* **20**: 2803–2812.
42. Glazer, A.N. and H.S. Rye. 1992. Stable dye-DNA intercalation complexes as reagents for high-sensitivity fluorescent detection. *Nature* **359**: 859–861.