DNA polymerase fidelity and the polymerase chain reaction.

K A Eckert and T A Kunkel

Genome Res. 1991 1: 17-24
Access the most recent version at doi:10.1101/gr.1.1.17
DNA Polymerase Fidelity and the Polymerase Chain Reaction

Kristin A. Eckert and Thomas A. Kunkel

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

The distinctive ability of the polymerase chain reaction (PCR) to produce a substantial quantity of DNA from an initially small amount of genetic material is revolutionizing molecular biology. The variety of highly successful applications of PCR technology and the ease with which PCR can be performed should not be misinterpreted as evidence that the PCR is a simple process. In fact, quite the opposite is true. A single cycle of DNA polymerization is a complex process requiring the precise interaction of several components. The repetitive cycles characteristic of PCR provide an additional layer of complexity to the technique. The final PCR product cannot be considered a unique entity, as even a discrete DNA band on an agarose gel may contain a variety of DNA molecules differing from the original genetic information by one or more nucleotides.

How serious a concern is the fidelity of PCR? The answer depends on the precise nature of the application. In many instances, for example, direct characterization of the amplified population by DNA sequence analysis or nucleic acid hybridization, random errors in nucleotide sequence that may be produced during PCR are of little concern. However, polymerase-mediated errors at a frequency of 1 mutation per 10,000 nucleotides can be controlled experimentally.

**ERROR ACCUMULATION DURING PCR**

Error rates in PCR vary according to the precise DNA sequence and the in vitro conditions of DNA synthesis. Several laboratories have estimated the error frequency (mutations per nucleotide per cycle) during PCR catalyzed by the thermostable *Thermus aquaticus* (*Taq*) DNA polymerase by cloning individual DNA molecules from the amplified population and determining the number of DNA sequence changes. As seen from the data in Table 1, observed error frequencies during PCR vary more than 10-fold, from $-2 \times 10^{-4}$ to $<1 \times 10^{-5}$. For applications of PCR products that utilize the population of amplified DNA molecules, random errors at these frequencies will not interfere with the biochemical analyses. However, polymerase-mediated errors at a frequency of 1 mutation per 10,000 nucleotides can be controlled experimentally.

**Table 1** Estimates of Fidelity during PCR Catalyzed by *Taq* Polymerase

<table>
<thead>
<tr>
<th>Target sequence</th>
<th>Number of cycles</th>
<th>Error frequency per nucleotide</th>
<th>Error rate per nucleotide per cycle$^a$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apolipoprotein B</td>
<td>30</td>
<td>22/8,000</td>
<td>1/5,600</td>
<td>6</td>
</tr>
<tr>
<td>HLA-DPB</td>
<td>30</td>
<td>17/6,692</td>
<td>1/5,900</td>
<td>7</td>
</tr>
<tr>
<td>α-Antitrypsin</td>
<td>33</td>
<td>7/4,700</td>
<td>1/11,000</td>
<td>8</td>
</tr>
<tr>
<td>HPRT</td>
<td>32</td>
<td>16/15,000</td>
<td>1/15,000</td>
<td>9</td>
</tr>
<tr>
<td>HLA-A,B</td>
<td>30</td>
<td>20/21,870</td>
<td>1/16,000</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>17/30,710</td>
<td>1/18,000</td>
<td>10</td>
</tr>
<tr>
<td>TCR Vb</td>
<td>25</td>
<td>1/1,500</td>
<td>1/19,000</td>
<td>11</td>
</tr>
<tr>
<td>HIV gag/env</td>
<td>30</td>
<td>$&lt;1/5,411$</td>
<td>$&lt;1/83,000$</td>
<td>12</td>
</tr>
</tbody>
</table>

$^a$Error rate = $1/[(observed error frequency/number of cycles) \times 2]$, assuming a constant efficiency per cycle.$^{(7)}$
per cycle are an important consideration for any PCR application that begins with a small amount of starting material (i.e., less than a total of 10,000 nucleotides of target DNA) or that focuses on individual DNA molecules present in the final PCR population.

The proportion of DNA molecules that contain sequence changes is a function of the error rate per nucleotide per cycle, the number of amplification cycles, and the starting population size. The population of altered DNA molecules arises during PCR from two sources: (1) new errors at each cycle and (2) the amplification of DNA molecules containing errors from previous cycles. On average, the same number of DNA sequence changes in the final population results from polymerase errors during the last amplification cycle as result from the amplification of errors that occurred during previous cycles. The formula \( f = np/2 \) describes the average mutation frequency \( f \) for PCR amplification as a function of the polymerase error rate \( p \) per nucleotide per cycle and the number of cycles \( n \).

The frequency of DNA sequence changes can be controlled by altering the number of cycles \( n \) and/or the polymerase error rate per nucleotide \( p \). The error frequency is expected to increase. Keeping \( p \) constant at one error per 10,000 nucleotides per cycle, the error frequency when \( n = 20 \) cycles is one change per 1000 nucleotides and increases to one change per 400 nucleotides after 50 amplification cycles. Alternatively, the final error frequency can be varied by changing \( p \), the polymerase error rate. Keeping a constant cycle number of \( n = 50 \), the final mutation frequency is only 1/2000 nucleotides if \( p = 1/50,000 \), but increases to 1/400 nucleotides if \( p = 1/10,000 \). Due to the exponential nature of PCR, the occurrence of an early error can increase the final error frequency above the average described by \( f = np/2 \), because the mutant DNA molecules will be amplified with each cycle, resulting in populations with a larger than average number of mutants. The remainder of this review explores the ways in which the polymerase error rate, \( p \), can be manipulated experimentally as a means of controlling the PCR error frequency.

**MOLECULAR NATURE OF MUTATIONS**

A variety of changes in DNA sequence can occur during enzymatic DNA amplification. These mutations vary from single base-substitution mutations to deletions and insertions. The most familiar type of mutation is the base substitution that results from the misincorporation of an incorrect dNTP during DNA synthesis. Twelve distinct base mispairs are possible, varying in composition and in the template base versus the incoming dNTP (e.g., T'dGTP versus G'dCTP). Together with the effects of neighboring template sequence, this creates a variety of possible molecular structures among which any polymerase must discriminate. All DNA polymerases examined to date also generate mutations during DNA polymerization in vitro wherein one or more nucleotides is lost or added. One factor that will influence the types of DNA sequence changes (base substitution versus deletion mutation) is sequence-dependent and is influenced by the stability of the template sequence. In this assay, single-stranded M13mp2 DNA containing the \( \alpha \)-complementation region of the Escherichia coli lacZ gene is used as the template for a single cycle of DNA synthesis. Upon transfection of appropriate E. coli strain with the products of the reaction, accurate DNA synthesis can be detected as dark blue M13mp2 plaques. However, polymer-
ase errors during synthesis of the lacZ gene result in lighter blue or colorless plaques, due to decreased α-complementation of β-galactosidase activity in the infected host cell. The error rates for specific classes of polymerase errors during a single synthesis cycle can be calculated from the frequency of light blue and colorless (mutant) plaques among the total plaques scored and DNA sequence analysis of a collection of the mutants. A more sensitive assay for specific types of mutations can be performed by using an M13mp2 template DNA which contains a base substitution or deletion mutation, resulting in a colorless phenotype. In these assays, polymerase errors are scored as DNA sequence changes that revert the mutant back to a wild-type or pseudo-wild-type phenotype. This approach is especially useful for highly accurate polymerases. Unlike the forward mutation assay, however, M13mp2 reversion assays are focused on a limited subset of errors occurring at only a few sites.

Polymerase-mediated discrimination against errors during synthesis occurs at several steps in the biochemical reaction (Fig. 1) (for review, see ref. 25). The first control point is the binding of the dNTP substrate to the polymerase (step A). The rate of binding of both the correct and incorrect incoming deoxynucleoside triphosphate is assumed to be the same. However, the correct dNTP substrate may remain preferentially bound to the enzyme-DNA complex due to hydrogen bonding and base-stacking interactions. Alternatively, an incorrect dNTP may dissociate from the E-DNA•dNTP complex more readily than a correct dNTP. At this stage, polymerase error rates can be varied experimentally by adjusting the ratio of correct and incorrect dNTP substrates, reflecting the relative probability that a polymerase will bind an incorrect versus a correct dNTP. An example of such a reaction catalyzed by Taq polymerase is shown in Table 2 (part I), wherein base substitution fidelity is decreased eightfold by decreasing the concentration of a single correct nucleotide. Imbalances in the dNTP pools can be either mutagenic or antimutagenic, depending on the error being considered. A second point for error discrimination is the more rapid rate of phosphodiester bond formation for correct rather than for incorrect E-DNA•dNTP complexes (step B). Observations from several laboratories suggest that DNA polymerases use selectivity at both steps A and B to improve base substitution fidelity, but that the relative contribution of the two steps varies from one polymerase to the next. (26-28)

Once the phosphodiester bond is formed between the incoming dNTP and the DNA primer, the rate of pyrophosphate release (step C) may occur more slowly in the case of incorrect incorporations. This third level of discrimination, together with the observation that some DNA polymerases catalyze pyrophosphorolysis and pyrophosphate exchange reactions, provides the possibility that the presence of pyrophosphate in a polymerization reaction may influence fidelity. Both mutagenic (30) and antimutagenic (31) pyrophosphate effects on the fidelity of E. coli DNA polymerase I have been observed.

A fourth level of error discrimination is the selective ability of a polymerase to continue DNA synthesis on a correct primer-template structure rather than on a mutational intermediate (step D). The probability of preserving any mispaired intermediate as a mutation is increased by reaction conditions that increase the rate of extension synthesis, such as high dNTP concentrations, high enzyme concentrations, or long incubation times. The rate of mispair extension varies for both the type of mispair and the polymerase. (32,33) In general, purine-purine mispairs are the most difficult to extend (10^{-5} to 10^{-6}), while purine-pyrimidine and pyrimidine-pyrimidine mispairs are extended more efficiently (10^{-2} to 10^{-4}) as compared to the corresponding correct base pairs (M. Goodman, pers. comm.). For the Taq polymerase, the failure to extend mispairs is not due to differential binding of the polymerase to correct versus incorrect terminal base pairs, but rather is due to a kinetic block to the addition of the next nucleotide onto the mispair (M. Goodman, pers. comm.). As for other DNA polymerases, the intrinsic extension efficiency for Taq polymerase measured under steady-state conditions differs for each type of mispair (M. Goodman, pers. comm.).

Finally, one of the most important control points for the production of mutations is the 3'→5' exonucleolytic removal (proofreading) of 3' ter-

---

**FIGURE 1** Steps in polymerase error discrimination. This reaction scheme for error discrimination by DNA polymerases in vitro has been generalized from data for a variety of enzymes, but relies heavily on data generated using the Klenow fragment of E. coli DNA polymerase I. (21) (E) Polymerase; (DNA_{n}) starting DNA primer-template of length n; (DNA_{n+1}) DNA_{n} primer-template after incorporation of one nucleotide; (DNA_{n+2}) DNA_{n+1} primer-template after extension synthesis; (dNTP) deoxynucleoside triphosphate; (PP_i) pyrophosphate; (dNMP) deoxynucleoside monophosphate. Letters A–E denote control points for error discrimination (see text for details).
minal mispairs to regenerate the start-
tive T7 polymerases can have an ex-
quently high fidelity for particular base-substitution errors, the fidelities are significantly lower when a larger number of sites and classes of errors are monitored in a forward mutation assay.

Given that thermostable DNA polymerases are preferred for PCR, it is noteworthy that preparations of the thermostable DNA polymerase from Thermococcus litoralis (Vent polymerase) contain a 3'→5' exonuclease activity (P. Mattila and J. Ronka, pers. comm.). These investigators have demonstrated that this enzyme has a base-substitution error rate of 1/31,000 in reactions containing 1 mM dNTPs, similar to that observed for the Klenow polymerase (Table 2, part III). Consistent with the possibility that exonuclease-removal of base-
substitution errors is occurring during polymerization, the fidelity of this polymerase is fivefold higher than the fidelity of the exonuclease-deficient Taq polymerase under identical reaction conditions (Table 2, parts II and IV), and the 3'→5' exonuclease activity has a slight preference for exci-
sion of mismatched rather than matched base pairs (P. Mattila and J. Ronka, pers. comm.).

The majority of commercially avail-
able thermostable DNA polymerases do not contain 3'→5' exonuclease activities. When the fidelities of various polymerases are compared using similar reaction conditions, the Taq, Thermus thermophilus (Tth), and Thermus flavus (Replinase) DNA poly-
merases are approximately 10-fold less accurate than are the T4 or native T7 polymerases, consistent with these en-
zymes being proofreading-deficient (Table 2) (P. Mattila and J. Ronka, pers. comm.). Even among exonuc-
lease-deficient DNA polymerases, error rates are not constant. The fidelity of the AMV reverse transcriptase is signifi-
cantly higher than the fidelities of the Taq polymerase or the exonuclease-deficient form of T7 polymerase (Se-
quenase 2.0) (Table 2, part IV). Fluctua-
tions in base-substitution error rates of more than 10-fold have been seen with several DNA polymerases, reflecting differences in polymerase error dis-

### Table 2: Comparative DNA Polymerase Fidelity in the M13mp2 Mutation Assays

<table>
<thead>
<tr>
<th>Polymerase</th>
<th>3'→5' Exonuclease</th>
<th>Base substitution error rate per nucleotide&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Taq</td>
<td>-</td>
<td>1/2,400&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13</td>
</tr>
<tr>
<td>Taq (100 μM dNTP)</td>
<td>-</td>
<td>1/290&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13</td>
</tr>
<tr>
<td>II. Klenow (1 μM dATP)</td>
<td>+</td>
<td>1/110,000</td>
<td>35</td>
</tr>
<tr>
<td>Klenow (1 μM dNTP)</td>
<td>-</td>
<td>1/15,000</td>
<td>35</td>
</tr>
<tr>
<td>III. T4</td>
<td>+</td>
<td>1/300,000</td>
<td>T.A.K. (unpubl.)</td>
</tr>
<tr>
<td>Native T7</td>
<td>+</td>
<td>1/290,000</td>
<td>T.A.K. (unpubl.)</td>
</tr>
<tr>
<td>Klenow</td>
<td>+</td>
<td>1/53,000</td>
<td>13</td>
</tr>
<tr>
<td>T. litoralis</td>
<td>+</td>
<td>1/31,000</td>
<td>P. Mattila (pers. comm.)</td>
</tr>
<tr>
<td>IV. AMV (500 μM dNTPs)</td>
<td>-</td>
<td>1/24,000</td>
<td>51</td>
</tr>
<tr>
<td>Taq</td>
<td>-</td>
<td>1/5,600</td>
<td>13</td>
</tr>
<tr>
<td>Sequenase 2.0</td>
<td>-</td>
<td>1/5,000&lt;sup&gt;c&lt;/sup&gt;</td>
<td>T.A.K. (unpubl.)</td>
</tr>
<tr>
<td>V. Taq (4 mM Mg&lt;sup&gt;2+&lt;/sup&gt;)</td>
<td>-</td>
<td>1/300,000</td>
<td>37</td>
</tr>
<tr>
<td>Taq (pH 5.7)</td>
<td>-</td>
<td>1/180,000</td>
<td>37</td>
</tr>
</tbody>
</table>

<sup>a</sup>Except where noted, error rates per detectable nucleotide were determined from the mutant frequency observed in an opal codon reversion assay as previously described.<sup>52</sup>

<sup>b</sup>The base substitution error rate was calculated for changes at the first position (T→G) of the TGA codon only. As determined by DNA sequence analysis, first position changes account for 85% of the total errors produced by Taq polymerase at equal dNTP concentrations.

<sup>c</sup>The base substitution error rate was calculated as described<sup>[50]</sup> from DNA sequence analysis of mutants produced in the M13mp2 forward mutation assay.

All fidelity measurements were done using 1 mM each of four dNTPs and 10 mM MgCl<sub>2</sub>, unless otherwise indicated by parenthetical notations. The T4, native T7, Sequenase 2.0, AMV, and Klenow polymerase reactions were incubated at 37°C; the Taq polymerase reactions at 70°C, and the T. litoralis polymerase reactions at 72°C.

The efficiency of proofreading as an error discrimination mechanism reflects a balance between the compet-
ing processes of polymerization of the next correct nucleotide (step D) and excision (step E). This balance is not the same for all polymerases, as demonstrated by the observation that the T4 and native T7 DNA polymerases are more accurate than is the Klenow polymerase under identical reaction conditions (Table 2, part III). Furthermore, this balance can be modulated by changing the conditions of in vitro polymerization. For example, proof-
reading by the Klenow polymerase is less effective at 1 mM dNTPs than at 1 μM dNTPs (Table 2, parts II and III); the high substrate concentration increases the rate of extension from mispaired termini (step D), lessening the time available for excision of the error (step E). Table 2 (part III) also illustrates that the overall fidelity of a polymerase reaction is controlled by the least ac-
curate reaction that occurs. Although the proofreading-proficient T4 and na-
surrounding DNA sequence (for a more extensive discussion, see ref. 36).

Not unexpectedly, the fidelity of a polymerization reaction in vitro depends upon the synthesis conditions chosen. The influence of variations in reaction conditions on the fidelity of Taq polymerase has revealed that even this exonuclease-deficient DNA polymerase can be highly accurate. Significant improvements in fidelity resulted from decreasing the MgCl₂ concentration or the pH of the reaction. (37) Under these conditions, the error rate per nucleotide polymerized at 70°C can be as low as 10⁻³ for base-substitution errors, similar to the observed error frequencies of proofreading-proficient enzymes (Table 2, part V). In these studies, fidelity was highest at a MgCl₂ concentration that was equimolar to the total concentration of dNTP substrates present in the reaction, and fidelity decreased as the concentration of free divalent cation increased. (37) Low rates for both base-substitution and minus-one base deletion errors are also obtained when Taq polymerase reactions were carried out using Tris, 4-morpholineethanesulfonic acid (MES), and 1,4-piperazinie (ethane-sulfonic acid) (Pipes) buffers at pH 5-6 (70°C). The base-substitution error rate increased 60-fold as the pH was raised approximately three units. (37) The fidelity of the Taq polymerase is moderately temperature dependent. (13) Two- to threefold increases in both base-substitution and minus-one base deletion error rates were observed as the reaction temperature was increased from 23-30°C to 70-80°C. Finally, the overall error rate of Taq polymerase was observed to increase as the dNTP concentration was increased from 1 μM to 1 mM. (13,37) This suggests that the fidelity of PCR amplification can be improved using low dNTP concentrations, since unextended mispairs will be lost during PCR because they do not yield full-length DNA products for further amplification.

DNA DAMAGE IN PCR
Incubation of DNA at high temperatures during PCR produces DNA damage, thus increasing the potential for mutations. One of the most frequent forms of DNA damage is deamination of cytosine to produce uracil. (13,38) Since uracil has the same coding potential as thymine, faithful replication of a template uracil by a DNA polymerase will result in a C→G→T transition mutation. In PCR, the greatest risk of cytosine deamination is during the denaturation step. The rate constant for deamination of cytosine in single-stranded DNA at 80°C is ≈1 x 10⁻⁸ events/sec, increasing to ≈2 x 10⁻⁷ events/sec at 95°C. A second common type of DNA damage is spontaneous base release resulting from hydrolysis of the N-glycosyl bond. (39) The rate of this hydrolytic reaction is significantly increased at high temperatures and low pH. Deamidation of native DNA at pH 7.4 and 70°C occurs at a rate of ≈4 x 10⁻⁹ events per sec. The release of pyrimidines, while approximately 20 times slower than purines, increases significantly as the temperature is increased to 95°C. The rate of spontaneous purine loss in native DNA is markedly increased at low pH. Following spontaneous base hydrolysis, the resulting apurinic/apyrimidinic (AP) sites in the DNA can inhibit synthesis by DNA polymerases. (39) In some cases, DNA polymerases are capable of replicating past AP sites. As AP sites are noncoding lesions, such bypass replication is error-prone and results in base-substitution mutations, most frequently transversions.

The presence of DNA damage also increases the occurrence of jumping PCR. The amount of intermolecular DNA rearrangements relative to intramolecular amplification during PCR was found to be greater when the starting DNA contained double-stranded breaks, AP sites, or ultraviolet light photoproducts. (20) Strand-switching becomes an important consideration to fidelity when ancient DNA or DNA samples from forensic analyses are used as the original DNA for amplification, because these samples often contain various forms of DNA damage. (40)

ERROR DISCRIMINATION DURING PCR
Generally, the rate of polymerase extension from mispaired or misaligned primer-termini (Fig. 1, step D) is dependent upon the concentration of the next correct dNTP substrate. (32) In the absence of proofreading, high dNTP concentrations decrease polymerase fidelity by decreasing the amount of error discrimination at the extension step, while low dNTP concentrations increase fidelity by reducing the rate of mispair extension. Such error discrimination at the extension step can be utilized for primer-specific PCR amplification. (41,42) Thus, template-primers can be differentially extended during PCR, depending upon the precise nature of the 3’ terminal mispair. (41,43) While primers creating a T/G (primer-template), G/T, C/A, or A/C mismatch provide efficient amplification, (41) the presence of an A/G mismatch reduces synthesis approximately 20-fold, and the presence of A/G, G/A, or C/C mismatches reduces overall PCR product yield about 100-fold relative to amplification from correctly base-paired primers. (43) In addition, when the dNTP concentration is lowered from 800 μM to 6 μM, only T/G mismatches or perfect base pairs are extended. (43)

Recently, Keohavong and Thilly have described a method to quantitate errors produced during PCR using denaturing gradient gel electrophoresis (DGGE). (44) In this technique, a 224-bp DNA fragment containing the HPRT exon 3 sequence is amplified by PCR. After purification of the desired DNA band by polyacrylamide gel electrophoresis, the DNA strands are denatured and allowed to reanneal in solution, generating heteroduplex molecules between DNA strands containing errors (mutant) and wild-type DNA strands. A linear gradient of denaturant in a polyacrylamide gel is then used to separate the mutant/mutant or mutant/wild-type heteroduplexes from the wild-type homoduplex molecules. The proportion of mutant DNA molecules in the population is estimated from either the heteroduplex fraction (44,46) or from the wild-type homoduplex fraction. (45) Finally, the polymerase error frequency per base pair per cycle is calculated from the mutant fraction, the number of duplications, and the DNA target size. If desired, individual heteroduplex bands can be excised to determine the exact nature of the mutation. Unlike the M13mp2 fidelity assay, which measures errors produced during a single round of DNA synthesis, (24) the DGGE/PCR...
technique detects the polymerase errors that accumulate during repeated cycles of amplification. In this system, only those errors that allow full-length synthesis are recovered in the final population, whereas errors that inhibit extension synthesis will be lost. DGGE has been used successfully to quantify the error frequency of PCR catalyzed by various polymerases under differing conditions at the HPRT exon 3 target locus (Table 3). (44–46)

The fidelity of PCR catalyzed by the Taq polymerase as measured by DGGE is \(-1-2 \times 10^{-4}\) (Table 3), in close agreement with the estimates from cloning PCR products (Table 1) and with the fidelity of Taq polymerase in M13mp2 in vitro assays (Table 2). Higher-fidelity PCR was obtained using the proofreading-proficient T4 DNA polymerase (-1/100,000) or the thermostable Thermococcus litoralis DNA polymerase (Table 3). PCR catalyzed by the chemically modified form of T7 DNA polymerase (Sequenase) yielded an error frequency of approximately one error per 25,000 nucleotides per cycle (Table 3). These observations do not conflict with low-fidelity DNA synthesis obtained with Sequenase 2.0 in the M13mp2 assay (Table 2). The difference presumably reflects the fact that the chemically modified T7 polymerase retains a low but detectable 3'→5' exonuclease activity, while Sequenase 2.0, a 28-amino-acid deletion mutant, is completely devoid of exonuclease activity. (47) The disparity may also reflect the selective loss during PCR amplification and/or DGGE analysis of some types of mutations that can be scored in the M13mp2 system. Unexpectedly, the high fidelity observed during PCR by Sequenase requires high concentrations of deoxynucleoside triphosphates, and the error rate increases up to 1/5000 when <0.5 mM dNTPs are used. (46) The reason for this dependence is as yet unknown.

The influence of PCR conditions on the fidelity of Taq polymerase was also examined using the DGGE/PCR technique. (46) In these studies, less extreme synthesis conditions could be used as compared to the M13mp2 fidelity assays, (37) since suboptimal reaction conditions resulted in a decreased efficiency or absence of amplification. Nevertheless, Taq polymerase fidelity was observed to decrease as the reaction pH was increased from pH 8 (25°C) to pH 9 (25°C), and to decrease as the dNTP concentration was increased from 50 μM to 200 μM. (46) Surprisingly, the highest fidelity was observed at still higher dNTP concentrations. At 500 μM dNTPs and 5 mM Mg²⁺, the observed error rate was 1/14,000; however, amplification under these conditions was inefficient and generated several products other than the desired 224-bp HPRT fragment. (46) The fidelity of the Thermococcus litoralis polymerase is constant over a broad range of dNTP concentrations (10 μM to 200 μM). (45) However, when 1 mM dNTPs were used for synthesis, the error rate of T. litoralis increased to 1/100,000, (46) consistent with possible inhibition of proofreading activity at high dNTP concentrations. (34)

DNA target sequences that are repetitive in nature may pose special problems for PCR amplification. Strand-switching events between homologous DNA segments have been observed during PCR amplification of HLA genes, and become more prevalent during the final cycles of amplification when the concentration of DNA becomes high. (10) An increased frequency of spurious DNA products with cycle number was also observed during amplification of tandemly repeated minisatellite DNA. (48) Attempts to amplify a segment of the human 18S rRNA gene using the Taq polymerase generated a 54-bp deletion artifact consistent with the formation of a stable intrastrand hairpin loop. (49) Interestingly, this deletion was not observed when the PCR amplifications were performed using the modified T7 polymerase, illustrating again that DNA polymerases can differ significantly in their biochemical properties.

**TABLE 3** PCR Error Rates as Measured by DGGE

<table>
<thead>
<tr>
<th>Polymerase</th>
<th>Error rate per nucleotide per cycle</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>T4</td>
<td>-1/100,000</td>
<td>44</td>
</tr>
<tr>
<td>Modified T7</td>
<td>1/29,000</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>1/23,000</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>1/18,000</td>
<td>46</td>
</tr>
<tr>
<td>T. litoralis</td>
<td>1/42,000</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>1/15,000</td>
<td>46</td>
</tr>
<tr>
<td>Klenow</td>
<td>1/7,700</td>
<td>44</td>
</tr>
<tr>
<td>Taq</td>
<td>1/11,000</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>1/5,000</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>1/4,800</td>
<td>44</td>
</tr>
</tbody>
</table>

22 PCR Methods and Applications
not in large excess over the total concentration of dNTP substrates, and be buffered to pH 6 (70°C) using Bis-Tris Propane or PIPES (Table 2). These buffers have a pH of between pH 6 and pH 7 and a small temperature coefficient (ΔpK_a/°C), allowing the pH to be maintained stably throughout the PCR cycle.

For amplifications in which fidelity is the critical issue, one should avoid the concept that conditions generating more DNA product are the better conditions. Reactions that optimize fidelity may be relatively inefficient for amplification. The total number of cycles should be kept to the minimum necessary to produce a feasible amount of product DNA for the desired use. Thus, although using low pH, low Mg^{2+}, and/or a low concentration of dNTPs may not produce an impressive amount of amplification product as visualized on an ethidium bromide stained-agarose gel, these conditions may provide an adequate amount of DNA for cloning purposes. Using a small number of cycles will also minimize the contribution of mutations resulting from DNA damage at high temperature. The DNA sequence of several independent clones from the final PCR population should be determined to verify that the original DNA sequence has been maintained during the amplification process. Ideally, the clones should be derived from separate PCR experiments to minimize the risk of overrepresentation of early polymerase errors.

The fidelity of DNA synthesis in vitro is limited by the least accurate reaction that occurs. For example, the T$G$ mispair is one of the most frequently produced and most easily extended base substitution error.(35) Even under high-fidelity reaction conditions, most of the Taq polymerase errors were the result of T$G$ mispairs that led to A$T$ to G$C$ transitions.(13) Therefore, the upper limit for Taq polymerase fidelity will be the frequency with which the enzyme creates this mispair. In general, this limitation is not restricted to a particular mispair. The sequence of the target DNA will have a profound influence on the generation of mutations. Thus, the least accurate reaction may be the result of an exceptionally high error rate at a particular nucleotide sequence, since the phenomenon of "hot spots" for polymerization errors has been clearly established.(16)

ACKNOWLEDGMENTS
We are grateful to Drs. Theodora Devereux and Neal Cariello for helpful comments and critical evaluation of the manuscript. We would also like to thank Drs. Myron Goodman, William Thilly, Lucy Ling, Pekka Mattila, and Neal Cariello for providing results prior to publication.

REFERENCES