PFU DNA Polymerase: A Study of Amplification Error Rate and Subsequent Implications for High Fidelity Mutational Spectrometry

by

Sheila Gay Buzzee

B.S. Biology, Xavier University
December, 1989

Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Civil and Environmental Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 1994

Copyright by the Massachusetts Institute of Technology 1994. All rights reserved.

Author

Department of Civil and Environmental Engineering

January 26, 1994

Certified by

William Thilly, Professor of Toxicology
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by

Professor Joseph Sussman
Chairman, Department Committee on Graduate Students

WITHDRAWN

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
MIL LIBRARIES
MAR 21 1994
PFU DNA Polymerase: A Study of Amplification Error Rate and Subsequent Implications for High Fidelity Mutational Spectrometry

by

Sheila Gay Buzzee

Submitted to the Department of Civil and Environmental Engineering on January 26, 1994, in partial fulfillment of the requirements for the degree of Master of Science in Civil and Environmental Engineering.

Abstract

This thesis tested the *Pyrococcus furiosus* (PFU) thermostable DNA polymerase during PCR of the human mitochondrial segment encoding for tRNA glycine (bp 10,009 to 10,216). This segment is of particular interest for highly precise human mutational spectrometry studies, and the enzyme was tested to determine what amplification error rate may be anticipated in this application, and what the implications are for future human mutational studies. The enzyme was tested under conditions in which cellular DNA was amplified through 20, 40, and 60 sequential doublings so that the enzyme error rate could be determined along a wide curve. The enzyme was tested in tandem with *Tli* (Vent™) DNA polymerase for control purposes and error rate comparison. The enzyme was tested under standard laboratory use conditions, and tested additionally in conditions designed to fully extend incomplete extension products, hypothesized to account for a significant portion of observable background noise during mutational spectrometry using gel electrophoresis. As was previously reported by Lundberg et al [20], PFU DNA polymerase was found under standard laboratory practice to have an error rate lower than *Tli* (Vent™) DNA polymerase. The observed error rate is no more than 1.6 x 10^-6 errors per base pair duplication. Furthermore this error rate was observed to be significantly reduced by additional DNA treatment to fully extend incomplete extension products. This additional treatment consisted of purification of PCR product, resuspension of the DNA, and preparation of a reaction mixture which mimics the PCR recipe but is lacking primers. This new reaction is thermocycled in the presence of PFU DNA polymerase under identical temperature conditions as PCR. This procedure allows for the complete extension of all DNA in the reaction mixture, but some DNA degradation occurs due to the high exonuclease activity of PFU.

Thesis Supervisor: William G. Thilly
Title: Professor of Toxicology, Professor of Civil and Environmental Engineering
Acknowledgments

I offer my sincere thanks and profound gratitude to Dr. John Hanekamp for his consistent support, advice, technical assistance, and unfailing friendship. Thanks also to Riitta and Jia, my wonderfully supportive and sympathetic office mates, to Hilary for keeping me company in lab at even the latest of hours, to Dr. Konstantin Khrapko and Dr. Gengxi Hu for advice, hints when I felt stuck, and sincere encouragement, and to Mark and Paula and Paulo and Aoy and all the other friendly faces for helping me keep my sense of humor no matter how tired I got. Thanks in particular to Professor Thilly for providing me with a unique environment for learning, and a forum of collaborators to discuss and experiment and create the future for ourselves. I wish you and your scientific melting pot many years of fruitful research.

Thanks most importantly to my private fans and supporters- you will never know how much your encouragement has fueled me and fed me as I struggled through the long hours and (seeming) technical impossibilities. Thanks to you all, in no particular order: Mom, Dad, Sarah, Nancy, Stefan, Sarah S., Jamie, The Friday Lunch Bunch, Pat and Brian, and Ken and Jen.
Contents

1 Introduction 9

2 Literature Overview 11
   2.1 Mutational Spectrometry 11
   2.2 Denaturing Gel Electrophoresis 11
      2.2.1 Denaturing Gradient Gel Electrophoresis (DGGE) 11
      2.2.2 Constant Denaturant Capillary Electrophoresis (CDCE) 13
   2.3 Polymerase Chain Reaction (PCR) 13
      2.3.1 Thermostable Enzymes and Enzyme Fidelity 14
      2.3.2 Enzyme Efficiency 15
   2.4 Pyrococcus furiosus (PFU) DNA Polymerase 16

3 Materials and Methods 17
   3.1 DNA Isolation 17
   3.2 Polymerase Chain Reaction 18
      3.2.1 Standard PCR Procedure and Sequential Amplification 19
      3.2.2 Additional Extension Procedure 21
   3.3 DNA Purification 22
      3.3.1 Gel Electrophoresis 22
      3.3.2 DNA Elution 22
List of Figures

4.1a PAGE visualization of first PFU amplification 28
4.1b PAGE visualization of second round of PFU amplification 28
4.1c PAGE visualization of third round of PFU amplification 28
4.1d PAGE visualization of fourth round of PFU amplification 28
4.1e PAGE visualization of fifth round of PFU amplification 28
4.2 DGGE of Tli Vent™ versus PFU, 20 duplications 30
4.3 DGGE of PFU 20, 40, and 60 duplications 30
4.4 DGGE of terminal extension cycling and unaltered product, 20 duplications 31
4.5 CDCE of PFU amplified DNA 33
4.6 CDCE of PFU amplified multiple ss peaks 34
4.7 Phosphor scan of PFU versus Tli Vent™ amplified DNA 36
4.8 Phosphor scan of terminal extension and unaltered product, 20 duplications 37
## List of Tables

| 2.1 | Reported enzyme fidelities | 15 |
| 3.1 | PCR reaction recipe: *Tli* Vent<sub>TM</sub>, PFU, PFU as per Lundberg et al [20] | 20 |
| 4.1a | PCR results, first amplification | 26 |
| 4.1b | PCR sequential amplification, round two results | 27 |
| 4.1c | PCR sequential amplification, round three results | 27 |
| 4.1d | PCR sequential amplification, round four results | 29 |
| 4.1e | PCR sequential amplification, round five results | 29 |
| 4.2 | PFU and *Tli* Vent<sub>TM</sub> fidelity quantification by phosphor scan | 35 |
List of Abbreviations

bp base pair(s)
BSA bovine albumin serum
CDCE constant denaturant capillary electrophoresis
DGGE denaturing gradient gel electrophoresis
dNTP deoxyribonucleoside triphosphate
EtdBr ethidium bromide
EtOH ethyl alcohol
f enzyme fidelity
HeF heteroduplex fraction
hmd high melting domain
lmd low melting domain
MOPS 3(N-morpholino)propanesulfonic acid
nt Nucleotide(s)
PAGE polyacrylamide gel electrophoresis
PCR polymerase chain reaction
PFU DNA polymerase obtained from *Pyrococcus furiosus*
rpm revolutions per minute
ss single stranded
TAE 40mM Tris-acetate, 1mM EDTA (pH = 8.0 at 25°C)
TBE 45mM Tris-borate, 1mM EDTA (pH = 8.0 at 25°C)
TEMED N,N,N',N'-tetramethylethlenediamine
UV ultraviolet
WT wild type
Chapter 1

Introduction

Mutational spectrometry is the science of determining unique and reproducible mutant patterns in DNA, and of characterizing both qualitatively and quantitatively the nature of those mutations. [1, 4, 21, 27, 28, 29] In order to observe rare, low frequency mutants and avoid working with very large numbers of cells, it is necessary to amplify DNA for study using the polymerase chain reaction (PCR). [16, 24] Polymerase enzyme fidelity, $f$, is of obvious importance to mutational spectrometry- an enzyme which has poor fidelity will introduce false signals into the system. Incorrectly amplified sequences, if present in levels comparable to the expected mutant fraction, will be indistinguishable from actual mutants and may mask mutant signals altogether.

This thesis tested the *Pyrococcus furiosus* (PFU) thermostabile DNA polymerase during PCR of the human mitochondrial segment encoding for tRNA glycine (bp 10,009 to 10,216). This multicopy segment is of particular interest for highly precise human mutational spectrometry studies, and the enzyme was tested to determine what amplification error rate may be anticipated in this application, and what the subsequent implications are for future human mutational studies. This enzyme has been reported previously to have a very high fidelity, $1.6 \times 10^{-6}$ errors per base pair duplication, in a study of DNA amplification, cloning, and plating for mutant colonies. [20]
Cloning is less precise than electrophoretic mutational spectroscopy and may contain internal bias in situations of inadequate sample size. Additionally the methods utilized by Lundberg et al include resin binding to remove polymerase activity and DNA elution; both have the potential to induce mutations in DNA [10, 20]. As the enzyme fidelity reported by Lundberg is the highest of any thermostabile enzyme described to date, PFU DNA polymerase was deemed potentially very important to electrophoretic mutational spectrometry, and worthy of further investigation.

PFU DNA polymerase was tested under conditions in which cellular DNA was amplified through 20, 40, and 60 sequential doublings so that the enzyme error rate could be determined along a wide curve. The enzyme was tested in tandem with Tli (Vent™) DNA polymerase for control purposes and error rate comparison. The enzyme was tested under standard laboratory use conditions, and tested additionally in conditions designed to fully extend incomplete extension products, hypothesized to account for a significant portion of observable background noise during mutational spectrometry using gel electrophoresis. This additional treatment consisted of purification of PCR product, resuspension of the DNA, and preparation of a reaction mixture which mimics the PCR recipe but is lacking primers. This new reaction is thermocycled in the presence of PFU DNA polymerase under identical temperature conditions as PCR. This procedure is expected to improve the measured enzyme fidelity. Fidelity with no additional treatment is expected to be approximately $10^{-6}$, as reported previously by Lundberg et al. [20]
Chapter 2

Literature Overview

2.1 Mutational Spectrometry

The science of mutational spectrometry began in 1958, when for the first time, the rII genes of the bacteriophage T4 were treated with 5 bromouracil and the sequence analyzed for both position and kind of mutants. [1] This work was furthered by study of the lac I gene of *E. coli*, leading to a series of techniques capable of detecting base-pair substitutions which lead to a stop codon. [4, 21] Mutant enrichment, purification, and sequencing were combined in a study of the cI gene of bacteriophage λ; the basic techniques therein are still fundamentals of mutational spectrometry research. [27] The most recently developed focus concerns the potential for mutant type and position to shed light on the probable mechanisms involved in mutant induction. [7, 17]

2.2 Denaturing Gel Electrophoresis

2.2.1 Denaturing Gradient Gel Electrophoresis (DGGE)

The concept upon which DGGE technique is based is technically quite basic-
missense mutations confer upon a ds DNA strand an altered stability profile, leading to a change in the denaturing or melting temperature ($T_M$) of the strand in question.

[29] In practice, a DNA sample is run through a polyacrylamide gel matrix which contains increasing concentrations of a urea/formamide chemical denaturing gradient. The combination of elevated temperature and chemical denaturing process allows the separation of sequences which differ by only one base pair over a length of 100 bp. The difference in $T_M$ for a single base pair substitution may be as little as 0.4 °C, but a separation of 2 to 8 centimeters can be achieved during gel electrophoresis. [18, 29] In situations where the deviation from wild type sequence is greater than one base substitution, the separation observed via DGGE analysis should be greater. [29]

DNA segments which separate well on DGGE analysis are those which have adjacent contiguous high and low melting domains. As a ds DNA strand of this type migrates down the gel matrix, its velocity is inversely proportional to its size; but for molecules of identical length, the velocity is inversely proportional to the fraction of time that the lmd is in the denatured or melted state. When the lmd is denatured, the hmd remains in ds conformation, creating a trident or star, effectively arresting migration. As the DNA moves into higher and higher denaturant conditions, the proportion of time that the lmd remains denatured increases, until the DNA becomes relatively immobile. As more thermodynamically stabile strands travel farther before becoming fixed in position, and a banding pattern is obtained in which the stable WT homoduplex focuses below less stable mutant homoduplexes, and below the most unstable WT-mutant heteroduplex bands. Mutant bands may thus be separated cleanly from WT DNA and made available for sequencing to determine the exact nature of the mutation. [2]
2.2.2 Constant Denaturant Capillary Electrophoresis (CDCE)

As it has been demonstrated that DNA handling associated with DGGE conditions induces alterations in DNA [10], and also in light of the length of time required to obtain a spectrum via DGGE analysis (several days to a week or longer), modifications to the process have been investigated. The first modification of DGGE was to employ a gel slab of constant chemical denaturant composition. [11] Building on this concept, but with the advantage of rapid analysis and small sample volume, is the CDCE. [15] This technique employs a linear polyacrylamide gel matrix of low viscosity and constant chemical denaturant to separate DNA strands based on size and melting profile as they pass through a narrow 'hot zone' in the matrix. [15] As in DGGE, ss DNA migrates at a constant velocity, dependant only on the length of the strand.

2.3 Polymerase Chain Reaction (PCR)

Mutational spectrometry depends upon the Polymerase Chain Reaction (PCR) to amplify DNA for study. This technique utilizes a DNA polymerase in appropriate salt and ionic conditions, and at its effective temperature, to exponentially increase the number of ds DNA copies in a given reaction mixture. [16, 22, 24] By specific primer selection, the procedure may be modified to enrich for known mutants [3], but the study of unknown mutants and unselected spectra require high numbers of DNA duplication and high polymerase fidelity.

The fidelity of a DNA polymerase is a measure of the enzyme accuracy while duplicating a template DNA strand. Inaccurately duplicated DNA leads to what is
known as background noise during mutational spectrometry. This background may appear as a dull smear, with some possible visible bands, through the lane of a $^{32}$P-labeled DGGE sample; or as an elevated baseline signal, with some low peaks, in CDCE. The appearance of these phenomena may serve to obscure the true mutant signals sought in electrophoresis.

2.3.1 Thermostabile Enzymes and Enzyme Fidelity

Early PCR, based on the work of Kleppe et al, was developed with the thermolabile Klenow fragment of *E. coli* DNA polymerase. [16, 24] The lability of the enzyme necessitated adding a new aliquot for each cycle of replication. The introduction of Taq, a thermostabile DNA polymerase, made PCR more convenient and minimized opening of the reaction tubes, which can lead to contamination. [25]

Enzyme fidelity, $f$, is reported as errors per base pair doubling. Reported $f$s for current thermostabile DNA polymerases, as measured by techniques of mutational spectroscopy, range from $2.1 \times 10^{-4}$ to $2 \times 10^{-5}$. [14, 19, 20] Those polymerases which possess exonuclease activity, such as Tli and Pfu, tend to have higher fidelity than exonuclease-deficient enzymes. (see table 2.1)

The reported spontaneous rate of mutations in the human genome is $3 \times 10^{-7}$ per gene per year [9]. In order to observe a human DNA sample, a typical 100-bp lmd observed in electrophoretic mutational spectrometry would undergo a typical $10^6$-fold amplification. This amplification, at enzyme fidelity of $2.1 \times 10^{-4}$ to $2 \times 10^{-5}$, would yield .04 to .42 mutations per each ds DNA segment in the reaction; a background too high to observe a true spontaneous mutation spectrum.
# Table 2.1 Reported Enzyme Fidelity

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>fidelity</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taq</td>
<td>$2 \times 10^{-4}$</td>
<td>14</td>
</tr>
<tr>
<td>Taq</td>
<td>$7.2 \times 10^{-3}$</td>
<td>19</td>
</tr>
<tr>
<td>Klenow</td>
<td>$1.3 \times 10^{-4}$</td>
<td>14</td>
</tr>
<tr>
<td>T7</td>
<td>$3.4 \times 10^{-5}$</td>
<td>14</td>
</tr>
<tr>
<td>T7</td>
<td>$4.4 \times 10^{-5}$</td>
<td>19</td>
</tr>
<tr>
<td>T4</td>
<td>$3 \times 10^{-6}$</td>
<td>14</td>
</tr>
<tr>
<td>Vent</td>
<td>$4.5 \times 10^{-5}$</td>
<td>19</td>
</tr>
<tr>
<td>PFU</td>
<td>$1.6 \times 10^{-6}$</td>
<td>20</td>
</tr>
</tbody>
</table>

## 2.3.2 Enzyme Efficiency

During exponential DNA duplication in PCR, the following equation describes the accumulation of product:

\[
N_f = N_o(1 + Y)^n
\]  

(2.1)

where $N_f$ is the final ds DNA copy number, $N_o$ is the initial template DNA copy number, $Y$ is the polymerase efficiency, and $n$ is the number of PCR cycles of exponential duplication. [24] The first two cycles of PCR do not necessarily follow this equation, but when the overall reaction is considered, this relationship serves as a perfectly applicable one. By quantifying the amount of DNA before and after PCR and knowing the number of reaction cycles, it is possible to calculate the effective polymerase efficiency.
2.4 *Pyrococcus furiosus* (PFU) DNA Polymerase

The marine organism *Pyrococcus furiosus* (DSM 3638) was isolated from geothermal marine sediments in Vulcano, Italy. [5] A thermostabile DNA polymerase possessing exonuclease activity has been isolated from this organism, and is reported to have high fidelity during DNA duplication. [20, 26]

Lundberg et al reported PFU DNA polymerase to have a higher fidelity than other thermostabile DNA polymerases currently in use. [20] The test methods included PCR, resin-mediated removal of polymerase activity, DNA eleution, fragment ligation, clone incubation, and plaque scoring. The average fidelity was reported to be $1.6 \times 10^{-6}$, the highest fidelity observed was $1.2 \times 10^{-6}$. [20] For this reason, PFU DNA polymerase is anticipated to have very high fidelity and is herein tested for applicability to electrophoretic mutational spectrometry and for sequential DNA amplification. Additionally, an attempt is made to characterize and eliminate a potential source of amplification error for this enzyme.
Chapter 3

Materials and Methods

3.1 DNA Isolation

Genomic wild type DNA was isolated from human TK6 cells grown in culture. In 1978, TK6 cells were derived from the WI-L2 cell line, which was originally isolated from a spherocytosis patient at the Wistar Institute in Philadelphia. [28] TK6 is a human male lymphoblastoid cell line, and is heterozygous at the thymidine kinase locus.

Using a Quiagen (Studio City, CA) pack 500 anion-exchange column, genomic DNA was isolated from TK6 cells from culture. Cells were spun down, washed twice with 50 mM MOPS (Sigma, St. Louis, MO), and brought up into a solution of 10 ml MOPS with a 0.9% vol/vol SDS content (Bethesda Research Laboratories, Bethesda, MD), 100 μg RNAse A and 1 mg/ml Proteinase K (Sigma, St. Louis, MO). The mixture was incubated for one hour at 37°C followed by thirty minutes at 60°C. Using a 22 gauge needle and a syringe, the mixture was sheared through 10 passes of the needle. The mixture was brought to a concentration of 850 mM KAc using a KAc solution at pH 4.8, and this mixture was then centrifuged at 5000 rpm for 20 minutes to precipitate cell debris and SDS.
A Quiagen column was pre-equilibrated and the DNA-containing supernatant loaded. Twice, the column was washed with 10 ml of the Quiagen QC buffer (1 M NaCl, 50 mM MOPS, 15 % EtOH, pH 7.0); the DNA was subsequently eleuted with 15 ml of the Quiagen QF buffer (1.25 M NaCl, 50 mM MOPS, 15% EtOH, pH 8.2). One volume of isopropyl alcohol was added to precipitate the DNA, and the mixture was centrifuged for 30 minutes at 10,000 rpm. Pelleted DNA was washed with 70% EtOH and allowed to dry. The DNA was then resuspended in TE, pH 7.6, and analyzed via UV spectrometry to determine the quality and quantity of the DNA obtained.

### 3.2 Polymerase Chain Reaction

Polymerase Chain Reaction was used to amplify a 216 bp segment of the mitochondrial tRNA glycine sequence from genomic DNA. Primers J3 and CW7 were utilized:

**J3:** 5' ATG GAG AAA GGG ACG CGG GC 3'

**CW7:** 5' GTA CCG TTA ACT TCC AAT TAA C

For DGGE analysis, J3 was radiolabeled with $^{32}$P; for CDCE analysis, J3 was fluorescein labeled. The primers were synthesized in absence of UV light, without EtBr staining, by Synthetic Genetics (San Diego, CA). Primer J3 was radiolabeled using Polynucleotide kinase (Boehringer Mannheim Biochemica) in a reaction mixture containing 50 mM Tris HCl, 10 mM MgCl$_2$, 0.1 mM ETDA, 5 mM dithiothreitol, 0.1 mM spermidine, 200 μci of $^{32}$P ATP, 5 mM Tricine, 600 pM primer J3, and 1 unit of PNK. The reaction was incubated at 37οC for 45 minutes and then precipitated in 100 mM NaOAc and 67% EtOH at -70οC, spun 30 minutes at 4οC, and the supernatant removed. Precipitation is repeated, followed by two 80% EtOH washes.
3.2.1 Standard PCR Procedure and Sequential Amplification

In order to determine the enzyme error rate per base pair doubling, it is desired to observe DNA samples which have undergone different numbers of duplications. In this way the observed error rate may be compared to ensure that the smear or bands ascribed to enzyme incorporation error actually increase linearly with increased DNA duplication. In order to obtain samples of DNA with different degrees of duplication, Quiaegen purified genomic DNA was diluted to $2 \times 10^5$ copies per µl. This template DNA was amplified from $10^6$ to $10^{12}$ copies using PFU and Vent DNA polymerases in a PCR reaction, and quantified via Polyacrylamide Gel Electrophoresis (PAGE) using EtBr staining and a pBR322 Msp I Digest marker ladder (New England Biolabs, Beverly, MA). The DNA was then diluted in ddH2O back to $10^6$ copies per µl, and then this template was used for another round of DNA amplification using identical PCR conditions. The DNA was then quantified, diluted, and the amplification repeated once more. In this manner DNA was obtained which has undergone 20, 40, and 60 duplications. In a separate experiment, amplified DNA was diluted from $10^{12}$ to $10^9$ copies per µl, and then sequentially amplified another 10 duplications at a time. This resulted in DNA which had undergone 20, 30, 40, 50, and 60 duplications. During all sequential amplifications, enzyme amplification efficiency remained between 65% and 85%, as defined by the PCR equation; lower efficiency warranted repeating the replication cycle as the reaction was then considered not representative. Rearrangement of the PCR equation [24] to calculate efficiency:

$$Y = \exp \left[ \frac{\ln(N_f/N_0)}{n} \right] - 1$$  \hspace{1cm} (3.1)
where $Y = \text{efficiency}$, $N_f = \text{final DNA copy number}$, $N_o = \text{initial DNA copy number}$, $n = \text{number of DNA duplication cycles}$, and $\ln = \text{natural logarithm}$.

<table>
<thead>
<tr>
<th>PCR Reaction component:</th>
<th>Buzzee</th>
<th>Lundberg</th>
<th>$Tli$ Vent</th>
</tr>
</thead>
<tbody>
<tr>
<td>genomic DNA</td>
<td>$10^6$ copies</td>
<td>$10^8$ copies</td>
<td>$10^6$ copies</td>
</tr>
<tr>
<td>Tris</td>
<td>$-\text{Cl, 20 mM}$</td>
<td>$-\text{HCl, 20 mM}$</td>
<td>$20$ mM</td>
</tr>
<tr>
<td>pH</td>
<td>$8.2$</td>
<td>$8.8$</td>
<td>$8.5$</td>
</tr>
<tr>
<td>KCl</td>
<td>$10$ mM</td>
<td>$10$ mM</td>
<td>$10$ mM</td>
</tr>
<tr>
<td>$(\text{NH}_4)_2\text{SO}_4$</td>
<td>$6$ mM</td>
<td>$6$ mM</td>
<td>$10$ mM</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>$2$ mM</td>
<td>$2$ mM</td>
<td>$7.5$ mM ($\text{MgSO}_4$)</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>$0.1%$</td>
<td>$0.1%$</td>
<td>$0.1%$</td>
</tr>
<tr>
<td>BSA</td>
<td>$1$ $\mu$g</td>
<td>none</td>
<td>$10$ mg</td>
</tr>
<tr>
<td>primer J3</td>
<td>$0.3$ $\mu$M</td>
<td>$(250$ ng)$^*$</td>
<td>$1$ $\mu$M</td>
</tr>
<tr>
<td>primer CW7</td>
<td>$0.3$ $\mu$M</td>
<td>$(250$ ng)$^*$</td>
<td>$1$ $\mu$M</td>
</tr>
<tr>
<td>dNTP mixture</td>
<td>$0.2$ $\mu$M</td>
<td>$200$ $\mu$M</td>
<td>$1.5$ $\mu$M</td>
</tr>
<tr>
<td>DNA polymerase</td>
<td>1 unit PFU</td>
<td>1 unit PFU</td>
<td>1 unit $Tli$ Vent</td>
</tr>
<tr>
<td>H$_2$O to balance total volume:</td>
<td>$100$ $\mu$l</td>
<td>$100$ $\mu$l</td>
<td>$100$ $\mu$l</td>
</tr>
</tbody>
</table>

Table 3.1. PCR Reaction Recipe. *Primers and PCR conditions as described by Lundberg et al [20]

All reagents except the DNA polymerase are combined and vortexed together in a $0.5$ ml Eppendorf tube. The tube is placed in a Perkin Elmer/Cetus thermocycler and heated to $94^\circ$C for 5 minutes for initial denaturing of the DNA. The reaction is cooled to $53^\circ$C for 5 minutes for initial primer annealing, after which the reaction is
interrupted and 1 unit of DNA polymerase is added to each tube, which is then heated to 72°C for 5 minutes for the first extension cycle. Subsequent autocycling is of 94°C for 1 minute, 53°C for 2 minutes, and 72°C for 2 minutes. The reaction mixture is put through 22 cycles before the initial quantification check. The terminal extension cycle stays at 72°C for an additional 12 minutes. All DNA is placed at 4°C while PAGE quantification is carried out, using the molecular weight marker pBR322 Msp I Digest (New England Biolabs, Beverly, MA) as a guide. Five µl of each sample is mixed with loading buffer to a final composition of 10 mM EDTA, 10% glycerol, 0.1% SDS, 0.02% xylene cyanol, and 0.02% bromophenol blue. These prepared samples were run on a 0.8 mm polyacrylamide gel (37.5:1 acrylamide:bis ratio) in TBE buffer at 250 volts for 1 hour and 30 minutes. The quantity of PCR product present can be quantified by comparison to the marker ladder after EtBr staining and UV illumination. When adequate yield has been obtained, the PCR product is then ethanol precipitated by the method of Sambrook et al. [23] After the pellet is dried, the DNA is resuspended in 10 µl ddH₂O.

### 3.2.2 Additional Extension Procedure

The additional extension procedure occurs after PAGE purification of the PCR product to ensure complete removal of primers. Purified sample is precipitated, resuspended in 20 µl ddH₂O, and divided into two tubes- one for immediate electrophoretic analysis and one for additional extension. The terminal extension procedure mimics a PCR reaction, but no primers are placed in the reaction mixture. This mixture is then thermocycled as in PCR, but with no initiation or extension cycles. Samples were tested at 2, 4, and 6 additional extension thermocycles.
3.3 DNA Purification

3.3.1 Gel Electrophoresis

DNA purification was carried out using polyacrylamide gel electrophoresis. The samples were precipitated [23] and resuspended in 10 μl ddH₂O. The DNA was then mixed with a loading buffer solution to a final composition of 10 mM EDTA, 10% glycerol, 0.1% SDS, 0.02% xylene cyanol, and 0.02% bromophenol blue. These prepared samples were then run on a 0.8 mm polyacrylamide gel (37.5:1 acrylamide:bis ratio) in TBE buffer at 250 volts for 1 hour and 30 minutes.

For ³²P labeled DNA, the gel is wrapped in plastic film and exposed to an autoradiogram film which has been marked as per location adjacent to the gel. The film is developed and the DNA bands located by overlaying the autoradiogram. The bands are then excised with a razor blade.

For fluorescein labeled DNA, the molecular weight marker is run in outside lanes, and in the lane directly adjacent to the purified samples, a 2 μl aliquot of the DNA is run. By usage of the marker ladder and the known position of the DNA aliquot, an estimate is made of the position of the purified bands. The bands are excised and the remaining gel stained with EtdBr to ensure that the proper area of the gel was cut.

3.3.2 DNA Elution

The excised gel slice from either above procedure is placed on a microscope slide, and 1 μl of ddH₂O is added. A second microscope slide is placed over the top of the gel slice, and the gel is crushed between them. The gel is scraped into a 1.5 ml
Eppendorf tube with a razor blade, and 10 μl of 10X hybridization buffer (4.0 M NaCl, 100 mM Tris HCl pH 7.5, 20 mM EDTA) is added to induce osmotic shock. After allowing the gel to sit for one minute, 190 μl of ddH₂O is added. The sample is gently agitated three times over a 15 minute period, then centrifuged for 2 minutes. The supernatant is removed and 4 volumes of EtOH are added for precipitation of the DNA.

3.4 Denaturation and Rehybridization

Denaturation and rehybridization insure that mutants are in heteroduplex form with WT DNA, important for low frequency mutant detection by gel electrophoresis. The hybridization solution contains 400 mM NaCl, 10 mM Tris HCl pH 7.5, and 2 mM EDTA. The samples are placed at 94°C for 2 minutes, at 65°C for 30 minutes, and at room temperature for 2 hours. The DNA is then reprecipitated using the method of Sambrook et al. [23]

3.5 DNA Analysis

3.5.1 DGGE Technique

The rehybridized, radiolabeled DNA sample is ready to be analyzed on DGGE. First, however, a purified WT sample must be obtained in order to determine the background DNA nonspecific binding and handling-induced noise in the DGGE process. The denaturant gradient for purification and analysis conditions was 37.5% to 75% formamide/urea in an 8% polyacrylamide (37.5:1 acrylamide:bis) gel matrix.
A 100% denaturant mixture is defined as 6.7 M urea, 7.5 M formamide. The gel matrix contains Tris-acetate and EDTA in the same concentration as the running buffer- 40 mM Tris-acetate, 1 mM EDTA. The DNA is loaded into wells in a loading buffer concentration which has a final concentration of 10 mM EDTA, 10% glycerol, 0.1% SDS, 0.02% xylene cyanol, and 0.02% bromophenol blue; and is run at 150 volts for 16 hours at 40°C.

The finished gel is opened, soaked in ddH₂O to remove the denaturant, and dried under heat and vacuum onto a piece of filter paper. The dried gel is wrapped in plastic film and is exposed to an autoradiograph film to locate and visualize bands. To obtain purified WT DNA, the developed film is placed over the gel to locate the WT band, and that region of the gel is cut out with a razor blade. This gel slice is partially rehydrated with 20 μl ddH₂O, then 5 μl of 6X concentrated loading buffer is added and absorbed by the gel slice. During the next DGGE run, the gel slice is manipulated into a well and run as a control for background DNA nonspecific binding. In adjacent wells 5000 cpm (10¹⁰ copies of DNA) of the following samples were run: PFU amplified 20 doublings, PFU amplified 40 doublings, PFU amplified 60 doublings, Tli Vent™ amplified 20 doublings, Tli Vent™ amplified 40 doublings, and in addition in some experiments, terminal extension corrected PFU amplified 20, 40 and 60 doubling samples.

### 3.5.2 CDCE Technique

CDCE technique was as described by Khrapko et al [15]. The DNA sample was fluoroscein labeled and had been amplified through 60 duplications with PFU DNA polymerase. The sample was boiled for 1 minute to create all ss DNA, then 10^8
copies of this DNA was loaded (for 1 minute at 1 μA current) onto the capillary, which was run at 250 V/cm. Photomultiplier signal was amplified $10^8$ V/A by an Oriel current preamplifier and recorded by a MP100 16 bit data acquisition system (Biopac Systems, Goleta, CA).

3.5.3 Quantification of Noise

In order to quantify the heteroduplex fraction, the radioactive dried DGGE gel is exposed to a phosphor screen and scanned. The heteroduplex region is that which focuses above the WT band. Background nonspecific binding and handling-induced noise is determined as the percentage of the DGGE-purified WT sample that focuses above WT in the purified sample lane. This percentage is subtracted from the measured counts in the test sample lanes and the remaining counts are defined to be heteroduplex DNA.

Quantification of amplification-ascribed error is done using the following equation:

$$\frac{1}{2} HeF = b * f * d$$

(3.2)

where $HeF$ is the DGGE measured heteroduplex fraction, $b$ is the target lmd size, in bp, $f$ is enzyme fidelity expressed in errors per bp per duplication, and $d$ is number of DNA duplications.
Chapter 4

Results

4.1 Sequential DNA Amplification

During the first round of DNA duplication, $10^6$ copies of genomic DNA was used as template and amplified to approximately $10^{12}$ copies. Final product was from $9.7 \times 10^{11}$ to $2.0 \times 10^{12}$ copies. Enzyme efficiency was from 70% to 75%. Results follow in tabular form. See also figure 4.1a for visualization of final product.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>$N_o$</th>
<th>$N_f$</th>
<th>n</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kit #1</td>
<td>$10^6$</td>
<td>$1.5 \times 10^{12}$</td>
<td>26</td>
<td>73%</td>
</tr>
<tr>
<td>Kit #1</td>
<td>$10^6$</td>
<td>$1.5 \times 10^{12}$</td>
<td>26</td>
<td>73%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>$10^6$</td>
<td>$2.0 \times 10^{12}$</td>
<td>26</td>
<td>75%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>$10^6$</td>
<td>$9.7 \times 10^{11}$</td>
<td>26</td>
<td>70%</td>
</tr>
</tbody>
</table>

Table 4.1a. PCR results, first amplification

This DNA was diluted to $10^9$ copies per $\mu l$ in ddH$_2$O and used as a template for the next round of duplication. Product varied from $8.6 \times 10^{11}$ copies to $1.2 \times 10^{12}$ copies, and efficiency varied from 68% to 72%. Results follow in tabular form. See also Figure 4.1b.
Table 4.1b. PCR sequential amplification, round two results

<table>
<thead>
<tr>
<th>Buffer</th>
<th>N₀</th>
<th>Nᵣ</th>
<th>n</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kit #1</td>
<td>10⁹</td>
<td>8.6 x 10¹¹</td>
<td>13</td>
<td>68%</td>
</tr>
<tr>
<td>Kit #1</td>
<td>10⁹</td>
<td>1.0 x 10¹²</td>
<td>13</td>
<td>71%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>10⁹</td>
<td>1.2 x 10¹²</td>
<td>13</td>
<td>72%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>10⁹</td>
<td>1.2 x 10¹²</td>
<td>13</td>
<td>72%</td>
</tr>
</tbody>
</table>

This DNA was again diluted to 10⁹ copies per µl in ddH₂O and used as a template for round three of duplication. Final product amount varied from 9.2 x 10¹¹ copies to 2.0 x 10¹² copies per reaction, and efficiency varied from 69% to 79%. Results follow in tabular form. See also Figure 4.1c.

Table 4.1c. PCR sequential amplification, round three results

<table>
<thead>
<tr>
<th>Buffer</th>
<th>N₀</th>
<th>Nᵣ</th>
<th>n</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kit #1</td>
<td>10⁹</td>
<td>9.2 x 10¹¹</td>
<td>13</td>
<td>69%</td>
</tr>
<tr>
<td>Kit #1</td>
<td>10⁹</td>
<td>2.0 x 10¹²</td>
<td>13</td>
<td>79%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>10⁹</td>
<td>1.5 x 10¹²</td>
<td>13</td>
<td>76%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>10⁹</td>
<td>9.2 x 10¹¹</td>
<td>13</td>
<td>69%</td>
</tr>
</tbody>
</table>

This product was again diluted to 10⁹ copies per µl in ddH₂O and used as a template for round four of duplication. Final product amount from round four varied from 9.8 x 10¹⁰ copies to 2.0 x 10¹² copies per reaction, and efficiency varied from 56% to 79%. Results follow in tabular form. See also Figure 4.1d.
Figure 4a. PAGE of first PFU amplification

Figure 4b. PAGE of second round of amplification

Figure 4c. PAGE of third round of amplification

Figure 4d. PAGE of fourth round of amplification

Figure 4e. PAGE of fifth round of amplification
For the fifth and final round of duplication, only the two samples with higher efficiency were diluted and used as template DNA. In this round final product varied from $1.1 \times 10^{12}$ copies to $1.6 \times 10^{12}$ copies per reaction, and efficiency varied from 51% to 54%. Repeated attempts to obtain higher efficiency were unsuccessful, and sequential amplification was halted at 60 duplications. Results follow in tabular form. See also Figure 4.1e.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>$N_o$</th>
<th>$N_f$</th>
<th>n</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kit #1</td>
<td>$10^9$</td>
<td>$2.0 \times 10^{12}$</td>
<td>17</td>
<td>56%</td>
</tr>
<tr>
<td>Kit #1</td>
<td>$10^9$</td>
<td>$1.5 \times 10^{12}$</td>
<td>15</td>
<td>65%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>$10^9$</td>
<td>$5.9 \times 10^{11}$</td>
<td>17</td>
<td>46%</td>
</tr>
<tr>
<td>Kit #2</td>
<td>$10^9$</td>
<td>$2.0 \times 10^{12}$</td>
<td>13</td>
<td>79%</td>
</tr>
</tbody>
</table>

Table 4.1d. PCR sequential amplification, round four results

4.2 DGGE Sample Analysis

Autoradiogram visualization of dried DGGE gels clearly shows that the spectrum induced by PFU amplification is different from that induced by $Tli$ Vent\textsuperscript{TM} amplification. (Figure 4.2) The spectrum is also observed to have the same banding
Figure 4.2  DGGE of \(7h\) Vent<sub>1M</sub> versus PFU, 20 duplications

Figure 4.3  DGGE of PFU 20, 40, and 60 duplications
Figure 4.4. DGGE of terminal extension cycling and unaltered product. 20 duplications
pattern in independently prepared samples, and the entire spectrum appears to increase in intensity with number of duplications. (Figure 4.3) Autoradiographs also show that the terminal extension cycling produces a cleaner-looking spectrum than unaltered samples, as seen in Figure 4.4a and 4.4b.

Another surprising observation is the disappearance of PCR product following this extension process. After the procedure, a quantification PAGE was run to verify the existence of good quality product. Upon EtDBr staining and UV illumination, the marker ladder is clearly observed but no product bands appear at all. A geiger counter check of the area where the DNA normally focuses revealed counts at approximately 10% of expected levels, based on the known amount of DNA put into terminal extension cycling. These bands were excised and the DNA used for analysis, and the DNA obtained from them functioned normally during subsequent procedures.

4.3 CDCE Sample Analysis

A PFU amplified sample was boiled and not allowed to reanneal in order to determine ss velocity on the capillary. Results of a run conducted by Dr. John Hanekamp show multiple peaks, instead of the single sharp ss peak expected. (Figure 4.5 and 4.6)

4.4 Phosphorimager Quantification of Noise

Using a Molecular Dynamics phosphor imaging system and ImageQuant software, volume integration of the sample lanes was conducted. Background was subtracted and the percentage of total counts focusing above WT (heteroduplex fraction) determined. Correction is made for nonspecific binding and handling-induced noise
Figure 4.5. CDCE of PFU amplified DNA
Figure 4.6. CDCE of PFU amplified multiple ss peaks
by determining the percentage of counts which focus above WT in a lane of DGGE-purified DNA. This percentage is subtracted from the percentage of sample DNA focusing above WT, and from this corrected $HeF$, the fidelity is calculated from the following equation.

\[
\frac{1}{2} HeF = b * f * d
\]  \hspace{1cm} (3.2)

Results follow for a gel comparing separate preparations of sample of PFU and $Tli$ Vent amplification origin. WT nonspecific binding and handling-induced background signal for this gel was 1.4% See also figure 4.7.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Duplications</th>
<th>$HeF$</th>
<th>Fidelity</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFU 1.1</td>
<td>20</td>
<td>11%</td>
<td>2.7 x $10^{-5}$</td>
</tr>
<tr>
<td>PFU 1.2</td>
<td>40</td>
<td>12%</td>
<td>1.6 x $10^{-6}$</td>
</tr>
<tr>
<td>Vent 1.1</td>
<td>20</td>
<td>13%</td>
<td>3.2 x $10^{-5}$</td>
</tr>
<tr>
<td>PFU 2.1</td>
<td>20</td>
<td>6.1%</td>
<td>1.5 x $10^{-5}$</td>
</tr>
<tr>
<td>PFU 2.2</td>
<td>40</td>
<td>4.7%</td>
<td>5.9 x $10^{-6}$</td>
</tr>
<tr>
<td>PFU 2.3</td>
<td>60</td>
<td>8.4%</td>
<td>6.7 x $10^{-6}$</td>
</tr>
<tr>
<td>PFU 3.1</td>
<td>20</td>
<td>4.2%</td>
<td>1.0 x $10^{-5}$</td>
</tr>
<tr>
<td>Vent 2.1</td>
<td>20</td>
<td>19%</td>
<td>4.6 x $10^{-5}$</td>
</tr>
</tbody>
</table>

Table 4.2. PFU and $Tli$ Vent$_{TM}$ fidelity quantification.

Next, sample is compared before and after terminal extension cycling. The sample has been divided so that the only difference between the two lanes is the terminal extension cycling and one additional precipitation. Both samples have been through 20 duplications with PFU DNA polymerase, and the background purified WT which focuses above WT is 1.8%. For the untreated sample, the $HeF$ is 7.6%, corresponding to a fidelity of 1.9 x $10^{-5}$; for the treated sample the $HeF$ is 1.5%, corresponding to a fidelity of 3.8 x $10^{-6}$, a 20-fold improvement. See figure 4.8.
Figure 4.7. Phosphor scan of PFU versus *Tli* Vent<sup>TM</sup> amplified DNA
Figure 4.8. Phosphor scan of terminal extension and unaltered product, 20 duplications
Chapter 5

Discussion

5.1 Sequential Amplification

Sequential DNA amplification was completely successful in producing PCR product of good quality at 20, 30, 40, 50, and 60 duplications. The procedure was equally successful when amplification cycles consisted of 10 duplications, and when template DNA was more dilute and cycles consisted of 20 duplications each. Enzyme efficiency was very good and did not drop unacceptably until the fifth round of sequential 10-fold duplication. Product visualized by EtDBr staining and UV illumination was normal in appearance and did not show multiple bands or the 'smear' often associated with overamplification.

As no purification was necessary between rounds, the procedure progressed rapidly. One important aspect of simply diluting the DNA between rounds of replication is speed and ease of technique, another is that there are minimal opportunities to introduce contamination into the product.
5.2 PFU Noise Evaluation

The fidelity for PFU DNA polymerase was found to be as high as $1.6 \times 10^{-6}$ errors per base pair duplication, twenty times better than the highest measured Tli Vent\textsuperscript{TM} fidelity of $3.2 \times 10^{-5}$. Terminal extension cycling was found to improve the measured fidelity by a factor of twenty.

Lundberg et al [20] reported an average PFU fidelity of $1.6 \times 10^{-6}$ errors per base pair duplication, with the highest observed fidelity equal to $1.2 \times 10^{-6}$. The methods utilized in the study included resin-mediated polymerase removal and DNA elution. [20] DNA exposure to DEAE cellulose resin and DNA elution procedures are known to induce significant errors in DNA. [10] Due to this fact, it is probable that the methods employed by Lundberg et al induced some measurable DNA error and the observed enzyme fidelity is somewhat less than the actual inherent fidelity.

Similarly, the procedures utilized for DNA handling and analysis in this experiment are likely to have induced measurable error as well. The fluorescent lights in the laboratory emit 320 nm UV light, known to cause mutations. [10] Furthermore, the gel purification technique utilized, while it does not involve electrophoretic elution, has not been tested for noise induction. Exposure to even the lower heat conditions of 40\degree C DGGE may induce measurable noise, as may repeated DNA precipitation and exposure to heat during gel drying. For these reasons it is concluded that PFU fidelity is at least $1.6 \times 10^{-6}$, probably better.
Chapter 6

Conclusions

6.1 Direct Implications

The implications for future application of sequential PCR are very positive. The ease and rapidity of the procedure, as well as the high quality of amplified DNA allow for further study of enzyme fidelity and preferential error incorporation. In addition, this procedure may prove useful in enhancing the sensitivity of special application PCR, such as MAMA (3).

The implications for high fidelity PCR are already well known—higher fidelity takes research one step closer to the ultimate mutational spectrometry goal—observing unselected spectra in humans. The observed PFU $f$ in this study is $1.6 \times 10^{-6}$. In order to best grasp the implications of this, the search for a theoretical mutation will be considered. This mutation occurs at a frequency of $10^{-4}$; there are 100 copies present in a sample of $10^6$ copies of DNA. After a 20-fold duplication, $10^6$ copies of WT have become $10^{12}$; $10^2$ copies of mutant DNA have become $10^8$, sufficient for detection via DGGE analysis. For $Tli$ Vent$_{TM}$, with $f$ of $3 \times 10^{-5}$, the PCR-induced error will amount to ($10^{12}$ copies $\times 3 \times 10^{-5} \times 20$ duplications $\times 100$ bp target size), or $6 \times 10^{10}$ errors. A 10% hotspot of error induction would yield a signal of $6 \times 10^9$
copies, 1.8 log greater than the sought signal. Even perfect distribution of these errors (a statistically improbable event) would yield $3 \times 10^8$ errors at each detectable base site, masking the true mutant altogether. For PFU, however, with $f$ of $1.6 \times 10^{-6}$, the PCR-induced error will amount to $(10^{12} \text{ copies} \times 1.5 \times 10^{-6} \times 20 \text{ duplications} \times \text{100 bp target size})$, or $3 \times 10^9$ errors. A 10% hotspot of error induction would yield a signal of $3 \times 10^8$ copies, a magnitude on the order of the true signal. Once the PCR-induced spectrum for PFU has been determined, errors of this magnitude can be predicted and accounted for, making detection of the mutant signal possible. Even moderate fidelity improvements from this point would reduce PCR-induced error background to undetectable levels.

The terminal extension procedure, while clearly not yet optimized, has shown a significant improvement in the quality of enzyme-amplified DNA. Once optimized it may have the potential to permit visualization of previously masked mutant bands, and combined with sequential amplification under conditions of mutant enrichment, will make the pursuit and characterization of mutant DNA easier and more accurate.

6.2 Areas for Future Study

It will be very interesting to investigate the potential application of sequential DNA amplification to mutant enriching PCR such as MAMA. (3) A series of reconstruction experiments with spiked known numbers of mutants should elucidate quite clearly whether this method will increase sensitivity. In addition, as new high-fidelity amplification enzymes are discovered or engineered, sequential amplification may be the only method which is capable of duplicating DNA enough times to observe an amplification-induced mutant spectrum.
The observed PFU $f$ in this study is $1.6 \times 10^{-6}$. With a twenty-fold improvement from the terminal extension procedure, or from other fidelity optimization techniques such as those of Ling et al [19], this could potentially be improved to $3.2 \times 10^{-7}$, a goal which would permit observation of unselected spectra in multicopy genes with reasonably small tissue samples (a few grams).

The observed disappearance of PCR product after terminal extension cycling is troublesome and warrants further investigation. It appears that 90% of product was lost, and in some experiments the entire sample was lost and the procedure had to be repeated with fresh sample. The high losses and their seeming erratic nature are unacceptable. However, the success of the terminal extension procedure in improving observed enzyme fidelity is very significant. A carefully designed experiment may illuminate the cause of the DNA disappearance. A $10^{12}$ copy sample, placed in a reaction mixture, would be sufficient. An aliquot should be removed after each cycle, and the aliquot examined by PAGE for product visualization; the sample should then be precipitated, checked again for counts lost during precipitation, and analyzed via DGGE for lessening of PCR associated amplification error bands. Once the disappearance has been better characterized, the procedure may be optimized for routine laboratory use.
Bibliography


