MUTATIONAL SPECTRA OF MNNG AND ICR-191 IN HUMAN CELLS

by

Neal Foster Cariello

B.S., University of Massachusetts, Boston (1982)

SUBMITTED TO THE DEPARTMENT OF APPLIED BIOLOGICAL SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1988

Copyright (c) 1988 Massachusetts Institute of Technology

Signature of Author

Department of Applied Biological Sciences
September 9, 1988

Certified by

W.G. Thilly
Thesis Supervisor

Accepted by

Chairman, Department of Chemistry, on Graduate Students

SEP 08 1988
"This doctoral thesis has been examined by a Committee of the Department of Applied Biological Sciences as follows:"

Professor John Essigman

Chairman

Professor William Thilly

Thesis Supervisor

Professor Leonard Lerman

Dr. Samuel Latt

Professor Helmut Zarbl
Mutational spectra of MNNG and ICR-191 in human cells

by

Neal Foster Cariello

Submitted to the Department of Applied Biological Sciences on September 9, 1988 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

Denaturing gradient gel electrophoresis (DGGE) and polymerase chain reaction (PCR) have been combined to examine mutations in exon 3 of the human hypoxanthine guanine phosphoribosyltransferase (HPRT) gene. Mutations were induced in a human, male, lymphoblastoid, B cell line by N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) or ICR-191 and selected by resistance to 6-thioguanine.

PCR was used to (1) amplify exon 3 from genomic DNA and (2) to modify the melting characteristics of exon 3 such that mutations throughout the entire coding region could be detected by DGGE. The melting characteristics of exon 3 were altered by the addition of a GC-rich, high-temperature melting domain to the 5' and 3' ends of exon 3.

The pattern of mutations induced by MNNG and ICR-191 was examined in (1) a collection of independent, TG<sup>f</sup>, MNNG- and ICR-191-induced clones, and (2) a complex mutant population consisting of several thousand TG<sup>f</sup> MNNG- and ICR-191-induced mutants.

DGGE was used to identify exon 3 mutant sequences in the individual TG<sup>f</sup> clones. Individual clones showing a non-wild type pattern were sequenced.
A series of 29, ICR-191-treated, independent, mutant clones was examined from a population of $3.6 \times 10^4$ TG$^r$ mutants; 99% of the TG$^r$ mutants were induced by treatment. Thirteen of the ICR-191-induced mutants contained a point mutation detected in exon 3. Nine mutants contained a +1 frameshift in a GGGGGGG sequence (5p 292-297), 2 mutants contained a +1 frameshift in a CCC sequence, 1 mutant contained an AT → TA transversion, and one mutant contained a GC → TA transversion. One apparent deletion of the entire HPRT gene was confirmed by Southern blotting. Thirty one percent (9 of 29) of the clones showed an identical mutation, and thus the +1 frameshift in the GGGGGG sequence constituted a hotspot for ICR-191.

A series of 32, independent, MNNG-treated, mutant clones was examined from a population of 3080 TG$^r$ mutants; 90% of the TG$^r$ mutants were induced by treatment. Five of the MNNG-induced mutants had a point mutation detected in exon 3, and all mutations were base pair substitutions at GC base pairs. Four MNNG-induced mutations occurred in a GGGGGG (base pair location 292-297) sequence at base pair locations 293 and 294, and one MNNG-induced mutant occurred at base pair 328. MNNG treatment produced 3 GC → AT transitions, and 2 GC → TA transversions. Two apparent deletions of the entire HPRT gene were confirmed by Southern blotting. Thirteen percent (4 of 32) of the clones showed a base pair substitution at base pair positions 293 and 294 in the GGGGGG (base pair location 292-297) sequence. Furthermore, the MNNG-induced base pair substitutions in the GGGGGG sequence affected a single amino acid, Gly69, which is coded by a GGG sequence at base pair location 293-295. These mutations constitute a mutational hotspot for MNNG.

Untreated, spontaneous, independent, TG$^r$ clones were also examined for point mutations in exon 3. Gennett and Thilly (1988) showed that 39% (33 of 85) of the untreated TG$^r$ clones contained a large structural alteration, almost exclusively deletions, as determined by Southern blotting. Four point mutations in exon 3 were detected among 19 clones from the Gennett and Thilly (1988) study which possessed a wild type Southern blotting pattern for HPRT. One +1 frameshift in a TTTT sequence and one GC → TA transversion occurred. The sequence alteration of two spontaneous mutants was not determined.

DGGE was also used to identify mutational hotspots in several complex mutant populations. In this case, populations derived from about 2000-3000 MNNG-treated TG$^r$ cells or 10,000 ICR-191-treated TG$^r$ cells were examined for mutational hotspots in exon 3. A single novel band representing about 5%-15% of the TG$^r$ population was visible on a denaturing gradient gel. The novel band in the MNNG-treated population was at a different location than the novel band in the ICR-191-treated population. Importantly, the position of the novel band in both
the MNNG- and ICR-191-treated complex populations was identical to the position of most common mutation(s) found in the treated TG^R clones. No mutational hotspot was detected in the untreated complex population.

Thesis supervisor: W.G. Thilly
Title: Professor of Applied Biological Sciences
# Table of Contents

Abstract .................................................. 2
Table of Contents ....................................... 5
List of Figures ........................................... 8
List of Tables ............................................ 12
Acknowledgments ......................................... 13
List of abbreviations ..................................... 14

1. Introduction ........................................... 17

2. Literature Review ..................................... 22
   2.1 Mutational Specificity .......................... 22
   2.2 Spontaneous Mutation .......................... 40
   2.3 ICR-191 mutagenesis ......................... 45
   2.4 MNNG mutagenesis ............................. 47
   2.5 HPRT ............................................ 55
   2.6 Polymerase Chain Reaction .................. 65
   2.7 Denaturing gradient gel electrophoresis .... 69

3. Materials and Methods ............................... 78
   3.1 Human cell lines, bacterial strains, plasmids, vectors 78
   3.2 Isolation of independent, TG\(^r\), untreated, MNNG- and ICR-191-treated TK6 clones 80
   3.3 Generation of untreated, MNNG- and ICR-191-treated complex mutant populations 86
   3.4 Construction of vectors used to generate single-strand, radiolabeled DNA complementary to exon 3 91
      3.4.1 Vector used to produce 169 bp single-strand probe and conditions used to generate probe 95
      3.4.2 Vector used to make 224 bp single-strand probe and conditions for probe generation 100
   3.5 Digestion of human genomic DNA with TaqI and MboII, solution hybridization with single-strand, 169 bp exon 3 probe, and denaturing gradient gel electrophoresis 103
   3.6 PCR amplification of HPRT exon 3 from total genomic DNA 106
   3.7 Addition of a 54 bp GC-rich sequence to exon 3 using PCR 108
3.8 Hybridization of 332 bp, PCR-amplified DNA containing a GC-rich sequence to wild type, radiolabeled probe and analysis by DGGE

3.9 Direct dideoxy sequencing of TG\textsuperscript{f} MNNG- and ICR-191-induced mutants

4. Results

4.1 Calculated melting map and melting behavior of 169 base pair TaqI/MboII exon 3 fragment on perpendicular and parallel denaturing gradient gels

4.2 Resolution of methylated from unmethylated exon 3 sequence using denaturing gradient gel electrophoresis

4.3 Hybridization of 169 bp probe to TaqI/MboII-digested genomic DNA and analysis by conventional PAGE: the question of pseudogene interference

4.4 Resolution of HPRT\textsubscript{Munich} from the wild type sequence using denaturing gradient gel electrophoresis: a mutation in the low-temperature melting domain

4.5 Hybridization of 169 bp probe to TaqI/MboII-digested genomic DNA and examination by DGGE: three types of mutants

4.6 Hybridization of 169 bp probe to TaqI/MboII-digested genomic DNA and examination by DGGE: MNNG- and ICR-191-induced mutants

4.7 Modification of the calculated melting map of HPRT exon 3 by addition of a GC-rich sequence

4.8 Examination of MNNG- and ICR-191-induced TG\textsuperscript{f} clones by DGGE using the GC-clamp

4.9 Examination of untreated, independent, TG\textsuperscript{f} clones by DGGE using a GC-clamp

4.10 DNA sequencing of untreated, MNNG- and ICR-191-induced, TG\textsuperscript{f} clones

4.11 Analysis of complex mutant populations by DGGE

4.11.1 Sensitivity of DGGE

4.11.2 Reproducibility of PCR-amplification using wild type DNA containing 1% HPRT\textsubscript{Munich} DNA

4.11.3 Analysis of untreated, MNNG-, and ICR-191-treated complex mutant populations by DGGE

4.12 Mutational spectra using individual TG\textsuperscript{f} clones

4.13 Mutational spectrum using complex mutant populations

4.14 Denaturing gradient gel melting characteristics of several exon 3
mutants

5. Suggestions for future research

6. Appendix

6.1 Publications

6.1.1 Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT\textsubscript{Munich}

6.1.2 DNA damage produced by ethidium bromide staining and exposure to ultraviolet light

6.1.3 DNA amplification \textit{in vitro} using T4 DNA polymerase

6.1.4 Use of denaturing gradient gel electrophoresis to determine mutational spectrum in human cells

6.2 Melting programs

6.2.1 MELT program

6.2.2 SQHT program
List of Figures

Figure 2-1: Distribution of spontaneous and induced mutations in the lacI gene of E. coli 24
Figure 2-2: Normalized small marker mutational spectrum of human lymphoblasts. 37
Figure 2-3: Reactions catalyzed by HPRT 57
Figure 2-4: Sequence of the human HPRT cDNA 60
Figure 2-5: Genomic organization of human HPRT 63
Figure 4-1: Diagram showing XhoI/TaqI and MboII restriction sites in HPRT exon 3 120
Figure 4-2: Calculated melting map of the 169 bp TaqI/MboII fragment of HPRT exon 3 123
Figure 4-3: Autoradiogram of perpendicular denaturing gradient gel showing the double-strand and single-strand 169 base pair TaqI/MboII exon 3 fragment 127
Figure 4-4: Mobility of the double-strand and single-strand 169 base pair TaqI/MboII fragment on a parallel denaturing gradient gel 129
Figure 4-5: Resolution of unmethylated exon 3 DNA from exon 3 DNA methylated by bacterial dam methylase 133
Figure 4-6: Hybridization of TaqI/MboII-digested genomic DNA to 169 base pair probe and analysis by a conventional polyacrylamide gel: wild type, RJK853 (HPRT deletion), HPRT_{Munich} 138
Figure 4-7: Calculated destabilization produced by mismatches at each base pair location of the 169 base pair probe 142
Figure 4-8: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and resolution of HPRT_{Munich} from the wild type sequence by denaturing gradient gel electrophoresis 144
Figure 4-9: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by denaturing gradient gel electrophoresis: wild type, HPRT deletion line, line with 4 X chromosomes 146
Figure 4-10: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE (9 hour electrophoresis time): RM3, RM4, RM6, RI204, RI205, RI206, RI207, RI208 151
Figure 4-11: Hybridization of 169 base pair probe to 153 TaqI/MboII-digested genomic DNA and examination by PAGE: wild type, RM4, RI206, RI208

Figure 4-12: Hybridization of 169 base pair probe to 155 TaqI/MboII-digested genomic DNA and examination by DGGE (5 hour electrophoresis): wild type, RM4, RI206, RI208

Figure 4-13: Hybridization of 169 base pair probe to 160 TaqI/MboII-digested genomic DNA and examination by DGGE: RM1, RM2

Figure 4-14: Hybridization of 169 base pair probe to 162 TaqI/MboII-digested genomic DNA and examination by DGGE: RM6, RM8, RM9

Figure 4-15: Hybridization of 169 base pair probe to 164 TaqI/MboII-digested genomic DNA and examination by DGGE: RM11, RM12, RM13, RM14, RM15, RM16, RM18, RM19

Figure 4-16: Hybridization of 169 base pair probe to 166 TaqI/MboII-digested genomic DNA and examination by DGGE: RM20, RM21, RI200, RI202, RI203, RI209, RI210, RI213, RI214, RI215

Figure 4-17: Hybridization of 169 base pair probe to 168 TaqI/MboII-digested genomic DNA and examination by DGGE: RM10, RI216, RI217, RI219, RI220, RI221

Figure 4-18: Hybridization of 169 base pair probe to 170 TaqI/MboII-digested genomic DNA and examination by DGGE: RM23, RM24, RM25, RM26, RI228, RI229, RI230, RI231

Figure 4-19: Hybridization of 169 base pair probe to 172 TaqI/MboII-digested genomic DNA and examination by DGGE: RM28, RM30, RM31, RM32, RM33

Figure 4-20: Hybridization of 169 base pair probe to 174 TaqI/MboII-digested genomic DNA and examination by DGGE: RM34, RM35, RM37, RM38, RM39, RI242

Figure 4-21: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RI232, RI233, RI234, RI235, RI237, RI238, RI239, RI240, RI241

Figure 4-22: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM16, RM26, RI222, RI226, RI227. Re-examination of samples showing prior possible plasmid contamination RM12, RM30, RM31

Figure 4-23: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and re-examination of samples
showing prior possible plasmid contamination RI205, RI219.

**Figure 4-24:** Hybridization of 160 base pair probe to \( TagI/MboI\)-digested genomic DNA and examination by DGGE: RI224, RI225.

**Figure 4-25:** Southern blot of RM21, RM18, RI229, RI200

**Figure 4-26:** Modification of the calculated melting map of HPRT exon 3 by addition of a GC-rich sequence

**Figure 4-27:** Simulation of a mismatch at each base pair position of the \( HgiAI\) fragment containing exon 3 and a GC-clamp at the 5' end

**Figure 4-28:** Simulation of a mismatch at each base pair position of the \( HgiAI\) fragment containing exon 3 and a GC-clamp at the 3' end

**Figure 4-29:** Autoradiogram of electrophoretic time course showing the melting behavior of the two \( HgiAI\) fragments containing HPRT exon 3 with an added GC-clamp

**Figure 4-30:** Ethidium bromide-stained polyacrylamide gel showing the PCR-amplification of HPRT exon 3 from genomic DNA

**Figure 4-31:** Ethidium bromide-stained polyacrylamide gel showing PCR-amplification product with added GC-clamp

**Figure 4-32:** Autoradiogram showing DGGE of TG\(^f\) mutants with an added GC-clamp: HPRT\(\text{Munich}\), RM4, RM34, RI209

**Figure 4-33:** Autoradiogram showing DGGE of TG\(^f\) mutants with an added GC-clamp: RI202, RI204, RI206, RI207, RI208, RI210, RI213, RI214, RM2, RM3, RM6

**Figure 4-34:** Darker exposure of gel shown in Figure 4-33

**Figure 4-35:** Autoradiogram showing DGGE of TG\(^f\) mutants with an added GC-clamp: RM8, RM9, RM10, RM16, RM19, RM20

**Figure 4-36:** Autoradiogram showing DGGE of TG\(^f\) mutants with an added GC-clamp: RM1, RM11, RM12, RM13, RM14, RM15, RM23, RM24, RM25

**Figure 4-37:** Autoradiogram showing DGGE of TG\(^f\) mutants with an added GC-clamp: RI203, RI205, RI215, RI216, RI217, RI219, RI221, RI222, RI224

**Figure 4-38:** Autoradiogram showing DGGE of MNNG- and ICR-191-induced TG\(^f\) mutants with an added GC-clamp: RI225, RM26, RM28, RM30, RM31, RM32, RM33, RM35, RM37, RM38, RM39

**Figure 4-39:** Autoradiogram showing DGGE of MNNG- and ICR-191-induced TG\(^f\) mutants with an added GC-clamp: RI226, RI227, RI228, RI230, RI231, RI232, RI233, RI234, RI237
Figure 4-40: Autoradiogram showing DGGE of untreated TG\textsuperscript{r} mutants with an added GC-clamp: 11B, 29B, 37B, 30B, 19B, 8C, 46A, 5A-2, 28B-1, 16B, 20B

Figure 4-41: Autoradiogram showing DGGE of untreated TG\textsuperscript{r} mutants with an added GC-clamp: 26A, 57B, 38A, 18A, 25C, 33A, 52B-1, 53A

Figure 4-42: Sequencing gel showing RM4 and RM34

Figure 4-43: Sequencing gel showing RI209

Figure 4-44: Sequencing gel showing RI202 and RI206

Figure 4-45: Sequencing gel showing RI208 and RI213

Figure 4-46: Sequencing gel showing RM20

Figure 4-47: Sequencing gel showing RM16, RM19 and RI207

Figure 4-48: Sequencing gel showing RI207, RI219, RI222, RI205, RI224

Figure 4-49: Sequencing gel showing RI228, RI232, RI233

Figure 4-50: Sequencing gel showing spontaneous mutant 33A

Figure 4-51: Sequencing gel showing spontaneous mutant 30B

Figure 4-52: Base pair location of sequenced exon 3 mutants

Figure 4-53: Autoradiogram of denaturing gradient gel showing HPRT\textsubscript{Munich} DNA and wild type DNA in a 0.5, 10\textsuperscript{-1}, 10\textsuperscript{-2}, 10\textsuperscript{-3}, and 10\textsuperscript{-4} ratio

Figure 4-54: Autoradiogram showing three PCR-amplifications of DNA containing 1\% HPRT\textsubscript{Munich} and analysis by DGGE.

Figure 4-55: DGGE analysis of one untreated, one MNNG-treated, and one ICR-191-treated complex mutant population using a 224 bp probe

Figure 4-56: DGGE analysis of two MNNG-treated complex mutant populations using a 224 bp probe

Figure 4-57: DGGE using a GC-clamp to examine complex mutant populations: untreated, MNNG-treated, ICR-191-treated

Figure 4-58: Structure of ICR-191

Figure 4-59: Denaturing gradient gel showing a +1 frameshift and three different mutations in the GGGGGG sequence

Figure 4-60: Calculated destabilization produced by mismatches at each base pair location of the 169 base pair probe

Figure 4-61: Calculated destabilization produced by mismatches at each base pair location of the 224 base pair probe
List of Tables

Table 2-I: Specificity of induced mutations in prokaryotes. 28
Table 2-II: Distribution of MNNG-induced nonsense mutations in the lacI gene of E. coli. 50
Table 3-I: Summary of probes and their applications 93
Table 4-I: Flowchart showing the overall strategy 117
Table 4-II: Summary showing the results of examining MNNG- and ICR-191-induced TGf clones using a 169 bp exon 3 probe 189
Table 4-III: Summary showing the results of examining untreated, MNNG- and ICR-191-induced TGf clones using the GC-clamp 232
Table 4-IV: Summary showing the results of examining untreated, MNNG- and ICR-191-induced TGf clones using the GC-clamp 234
Table 4-V: Summary of sequence alterations found in exon 3 of untreated, MNNG- and ICR-191-induced TGf clones 258
Acknowledgments

I wish to thank everyone who helped.
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-AF</td>
<td>N-2-aminofluorene</td>
</tr>
<tr>
<td>AFT</td>
<td>aflatoxin B₁</td>
</tr>
<tr>
<td>AG</td>
<td>8-azaguanine</td>
</tr>
<tr>
<td>AP</td>
<td>2-aminopurine</td>
</tr>
<tr>
<td>APRT</td>
<td>adenosine phosphoribosyltransferase</td>
</tr>
<tr>
<td>B(a)P</td>
<td>benzo(a)pyrene</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BU</td>
<td>5-bromouracil</td>
</tr>
<tr>
<td>c/w</td>
<td>cells/well</td>
</tr>
<tr>
<td>cis-Pt</td>
<td>cis-diamminedichloroplatinum (II)</td>
</tr>
<tr>
<td>C(cd)P</td>
<td>cyclopenta(cd)pyrene</td>
</tr>
<tr>
<td>DGGE</td>
<td>denaturing gradient gel electrophoresis</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DRB</td>
<td>dichlororibofuranosylbenzimidazole</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>disodium ethylene diamine tetraacetate</td>
</tr>
<tr>
<td>EMS</td>
<td>ethyl methane sulfonate</td>
</tr>
<tr>
<td>gpt</td>
<td>xanthine-guanine phosphoribosyltransferase</td>
</tr>
<tr>
<td>HLA</td>
<td>human leukocyte antigen</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>HPRT</td>
<td>hypoxanthine guanine phosphoribosyltransferase</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base pair</td>
</tr>
<tr>
<td>MF</td>
<td>mutant fraction</td>
</tr>
<tr>
<td>MNNG</td>
<td>N-methyl-Ν'-nitro-N-nitrosoguanidine</td>
</tr>
<tr>
<td>N-AcO-AAF</td>
<td>N-acetoxy-N-2-acetylaminofluorene</td>
</tr>
<tr>
<td>NQO</td>
<td>4-nitro-quinoline-1-oxide</td>
</tr>
<tr>
<td>OUA</td>
<td>oubain</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PE</td>
<td>plating efficiency</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PPT</td>
<td>podophyllotxin</td>
</tr>
<tr>
<td>RF</td>
<td>replicative form</td>
</tr>
<tr>
<td>SA</td>
<td>specific activity</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SS</td>
<td>single-strand</td>
</tr>
<tr>
<td>SSC</td>
<td>150 mM NaCl, 15 mM sodium citrate</td>
</tr>
<tr>
<td>TAE</td>
<td>40 mM Tris (pH 8.0), 20 mM sodium acetate, 1 mM EDTA</td>
</tr>
<tr>
<td>TG</td>
<td>6-thioguanine</td>
</tr>
<tr>
<td>TG&lt;sup&gt;r&lt;/sup&gt;</td>
<td>6-thioguanine-resistant</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
</tbody>
</table>
V volts
XGAL 5-bromo-4-chloro-3-indolyl-β-D-galactoside
Chapter 1

Introduction

Mutagens do not act randomly (Benzer, 1961). They show remarkable specificity with regard to the nature and base pair position of the induced mutation (Coulondre and Miller, 1977b). The pattern of DNA alterations associated with a given mutagen is termed the mutational spectrum, and thus the mutational spectrum produced by a particular agent should be diagnostic for exposure to that agent. Therefore, it should be possible to link an environmental exposure to a given agent with a specific pattern of mutations. The work presented in this dissertation examines the feasibility of obtaining mutational spectra from human cells in vitro as a necessary step in the development of such a technology.

The patterns of mutations produced by two mutagens, MNNG and ICR-191, were analyzed in this dissertation. MNNG is a alkylating agent that overwhelmingly induces GC → AT base pair substitutions in prokaryotes (Coulondre and Miller, 1977b). ICR-191 is an intercalating agent that almost exclusively produces frameshift mutations in prokaryotes (Calos and Miller, 1981). By choosing these two mutagens, I was able to compare the data for human cells with data from many prior studies in both prokaryotic and eukaryotic systems.

The technical objective of this work was to devise a means to discriminate between two mutagen-treated populations based on the specific nature of the induced mutations. Simply stated, if one were to treat human cells with MNNG or
ICR-191 and then present these cells for analysis, would it be possible to determine which cells had been treated with MNNG and which cells had been treated with ICR-191? Is there a pattern of mutation that is diagnostic for exposure to each mutagen?

The pattern of mutation in an exon of an endogenous human gene is examined, and the techniques developed in this work can be directly applied to the examination of mutations which occur in people. Thus a cohort of individuals heavily mutated by the same agent may show a similar pattern of mutation, and the link between environmental exposure and mutation may be established.

The region of the human genome that my work has focused on is the third exon of the HPRT gene. The HPRT locus spans about 44 kilobases on the X chromosome and the 654 base pair coding region is contained in 9 exons. Exon 3 was chosen for study because, (1) exon 3 contains the largest contiguous coding region and offers a sizeable mutational target (28% of the coding region is present in the 184 base pair exon 3 sequence), and (2) exon 3 has been reported to code for the catalytic sites of the enzyme and mutations in the catalytic site may severely impair the function of the enzyme.

A subset of all possible mutations at the HPRT locus confers resistance to the purine analog 6-thioguanine (TG), the selective agent used in this study. TG is a stringent selective agent, in that only cells essentially devoid of HPRT enzyme activity are resistant to TG. Most frameshift mutations and those base pair substitutions that produce a stop codon are expected to abolish the activity of the
enzyme and thus produce a TG\textsuperscript{r} cell. However, cells that possess a variety of HPRT missense mutations are expected to have a wide range of enzyme activities; only those cells with missense mutations that result in essentially no HPRT enzyme activity are expected to be TG\textsuperscript{r}. Therefore, only a subset of all induced mutations can be examined, namely, those mutations in HPRT exon 3 that confer resistance to TG.

The separation technique used in this work, denaturing gradient gel electrophoresis, is fundamental to my analytical strategy. Denaturing gradient gel electrophoresis separates DNA duplexes on the basis of sequence and physically separates mutant from wild type DNA (Fischer and Lerman, 1983). Furthermore, the sensitivity of denaturing gradient gel electrophoresis to DNA sequence is so great that different mutations can often be resolved from one another based on their migration in a denaturing gradient gel. Denaturing gradient gel electrophoresis is well-suited to the study of mutational spectra because it is sensitive to both the specific nature of a mutation as well as to the base pair location at which this mutation occurs.

DNA is organized into cooperative, isomelting-temperature domains. An original limitation of denaturing gradient gel electrophoresis was that mutations could not be detected in the highest-temperature melting domain of a DNA molecule. Furthermore, denaturing gradient gel electrophoresis could reveal that a mutation had occurred, but it provided little information about the specific nature of the mutation. Another technique, called polymerase chain reaction (Saiki et al.,
1985), was used to overcome these limitations. Polymerase chain reaction was used to (1) modify the DNA that contained exon 3 such that mutations at all base pair locations could be detected by denaturing gradient gel electrophoresis, and (2) rapidly sequence any mutants that were detected by denaturing gradient gel electrophoresis.

Human B cell were treated in vitro with either MNNG or ICR-191 and HPRT mutants were selected with 6-thioguanine. The mutational spectra of MNNG- and ICR-191 was derived in two ways, (1) by a clone-by-clone examination of independent, TG\(^f\) mutants, and (2) by an examination of complex mutant populations derived from several thousand TG\(^f\) mutants.

The analysis of a complex mutant population permits an experimenter to rapidly obtain mutational spectra. To generate these complex mutant populations, (1) greater than $10^8$ cells were treated with either MNNG or ICR-191, (2) 6-thioguanine was added to the cultures en masse, and (3) the TG\(^f\) mutants continued to grow and eventually overgrew the culture. 2000-10,000 TG\(^f\) cells were present immediately after treatment, thus the complex mutant populations were derived from these original TG\(^f\) mutants.

Approximately 80 independent, TG\(^f\) clones were examined to obtain the mutational spectra of MNNG and ICR-191. Much of the same information can be obtained by examining only three complex mutant populations: untreated, MNNG-treated, ICR-191-treated. The labor and time required to obtain mutational spectra from a complex mutant population is roughly 1 to 2 orders of
magnitude less than that required to derive the same information from an
examination of independent clones.

The goal of this work was to develop a means to identify mutational spectra
in human cells. The mutational spectra of both MNNG and ICR-191 have been
established with a sufficient level of precision such that an experimenter can
identify a population of cells which have been exposed to either MNNG or
ICR-191.
Chapter 2

Literature Review

2.1 Mutational Specificity

The phenomenon of mutational specificity was first observed in 1958 by Benzer and Freese (1958). The mutations induced by 5-bromouracil were shown to be different than the spontaneously occurring mutations. They mapped mutations to a given region of the rII locus of the bacteriophage T₄ by a series of genetic crosses with deletion mutants in the rII locus. The rII locus controls the ability of T₄ phage to grow on E. coli strain K, and rII mutants are unable to grow on this E. coli strain. Mutants were selected by the inability to grow on E. coli strain K, and any mutant that did not score positively was missed. Benzer and Freese examined a subset of the mutations produced by 5-bromouracil, and this subset possessed sufficient information to discriminate between the spontaneous and chemically-induced mutations. This system was then used to investigate the genetic effects of 2-aminopurine, 2,6-diaminopurine, 5-bromodeoxyctydine, proflavin, ethyl methane sulfonate, and nitrous acid (Benzer, 1961). Two key observations were made:

- The distribution of mutations as a function of base pair position was different for spontaneous and chemically-induced mutations, and the distributions were different among the chemically-induced mutations.

- For a given agent, several sites in the gene displayed much higher mutation frequencies than other sites. These sites were termed "hotspots" and furthermore, different mutagens produced different sites
of mutational hotspots.

The site-specificity of mutagenesis has also been extensively studied in the lacI gene of *E. coli* by Miller and coworkers (Coulondre and Miller 1977a; Coulondre and Miller, 1977b). All detectable lacI mutants were initially selected, then only those mutants which possessed a suppressible amber or ochre mutation were further characterized. A total of 72 amber and ochre sites could be produced from a single base pair change in the lacI gene. The system therefore monitored the pattern of mutation at 72 noncontiguous nucleotides in the 1080 base pair lacI gene. The system mapped amber and ochre mutations in the lacI gene to within about 5 base pairs by genetic crosses with a series of deletion mutants (Miller *et al.*, 1977; Schmeissner *et al.*, 1977). Since the protein sequence was known, the exact nature and position of the majority of the mutants could be determined. Only AT-to-GC transitions were not detectable using this system, since this single base pair substitution does not result in a nonsense codon; however, some AT-to-GC transitions were monitored in a separate reversion assay.

These studies again demonstrated the specificity of the induced mutations. Figure 2-1 summarizes the GC-to-AT transitions in untreated *E. coli* and with populations treated with the mutagens NQO, 2-AP, UV, EMS, and MNNG. This figure shows that the distribution of mutations was not random with regard to the position of the mutation and that the distribution of mutations was particular for each mutagen studied.

It was found that (1) the mutagens MNNG and EMS induced GC-to-AT
Figure 2-1: Distribution of spontaneous and induced mutations in the lacI gene of E. coli

This table summarizes the results of 3738 mutations arising from the GC-to-AT transition. Suppressor amber and ochre mutants were selected and the location in the gene was determined by crosses with deletion mutants. The number of independent occurrences is indicated by the bar height. Data is taken directly from Coulondre and Miller (1977b).
almost exclusively (502 out of 518 occurrences for MNNG and 680/691 for EMS), (2) NQO induced predominately GC-to-AT (900/1013) with some GC-to-TA (92/1013), (3) 2-AP produced mainly AT-to-GC (about 90%) with some GC-to-AT (about 10%) (4) the pattern of mutation produced spontaneously and with UV light was more complex. The mutational spectra of treated and untreated populations were easily differentiated. Importantly, the pattern of mutation produced by EMS and MNNG was different, even though both mutagens are methylating agents that produced over 97% GC-to-AT transitions. Certain sites were several-fold more mutable by either methylating agent, and it should be possible to discriminate between the two mutagens based on the pattern and frequency of induced mutations.

In addition to the work of Benzer and Miller, many systems have been developed to examine the specificity of various mutagens. Bruce Ames has developed a bacterial reversion test system that has been widely employed (Ames et al., 1975). Ames selected a set of three mutants which contained a known point mutation in the his operon of S. typhimurium. Mutagens are evaluated by the ability to revert the mutant strain to the wild type phenotype. These assays permit analysis at a single base pair or at a restricted number of nucleotides. Reversion tests are useful in the determination of mutagenicity of an uncharacterized compound, however, these test systems generally do not monitor a variety of transitions and transversions over a sizeable region of the genome.

Table 2-I reviews some of the major studies that have examined mutational spectra using systems which can detect a wide range of base pair substitutions in a
target gene. These studies have all been done in prokaryotes and have used the following DNA as a mutational target: lacI gene of *E. coli*, tetracycline-resistance gene coded for by the plasmid pBR322, cl gene of bacteriophage lambda, M13-lacZ construction in *E. coli*.

Mutational spectra is the pattern of mutations found after treatment with a mutagen, and this pattern is the product of many components. Initially, a mutagen must damage the DNA, and in the case of many mutagens, this involves binding to the DNA. Fuchs has examined the binding spectrum and mutational spectrum of AAF and has found that AAF binding was not sufficient to explain the pattern of mutation produced by AAF (Fuchs, 1983). The 3'-to-5' exonuclease activity of T₄ DNA polymerase was used to examine the binding of AAF, since an AAF lesion constituted a absolute block for the enzyme; in this way, the location and frequency of AAF lesions was determined. It was found that all guanine residues reacted with the mutagen and that the chemical reactivity varied about 40-fold. When the pattern of mutations was examined, it was found that 49/57 mutations clustered at eight guanine hotspots, however, no preferential AAF binding occurred at the mutational hotspots. The authors concluded that mutational specificity was not dictated by binding alone and that subsequent processing of the DNA reaction product played a central role in the final mutational spectra.

The work of Fuchs and co-workers showed that mutational spectra are influenced by the ability of the organism to repair DNA damage. This phenomenon has also been observed by comparing the *lacI* UV mutational spectra
Table 2-I: Specificity of induced mutations in prokaryotes.

This table summarizes studies which have examined the phenomenon of mutational specificity in prokaryotes. Tet means that tetracycline-sensitivity conferred by the plasmid pBR322 in E. coli was examined, lacI indicates that mutagenesis in the lacI gene of E. coli was studied, cl means that mutations in the cl gene of bacteriophage lambda were examined, and M13lacZ indicates that the mutational target was an M13lacZ hybrid phage in which 789 bp of the lacZ gene from E. coli was inserted into bacteriophage M13mp2. The nature of the principal induced mutation is also given; the occurrences of the principal mutation is shown over the total number of mutants examined.
<table>
<thead>
<tr>
<th>Agent</th>
<th>System</th>
<th>Principal mutation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-AcO-AAF</td>
<td>Tet</td>
<td>Deletion of GC (9/9)</td>
<td>Fuchs et al., 1981</td>
</tr>
<tr>
<td>N-AcO-AAF</td>
<td>Tet</td>
<td>Deletion of GC (53/57)</td>
<td>Koffel-Schwartz et al., 1984</td>
</tr>
<tr>
<td>2-AF</td>
<td>Tet</td>
<td>GC → TA (14/17)</td>
<td>Bichara and Fuchs 1985</td>
</tr>
<tr>
<td>AFT</td>
<td>lacI</td>
<td>GC → TA (169/189)</td>
<td>Foster et al., 1983</td>
</tr>
<tr>
<td>B(a)P</td>
<td>lacI</td>
<td>GC → TA (93/155)</td>
<td>Eisenstadt et al., 1982</td>
</tr>
<tr>
<td>BU</td>
<td>cl</td>
<td>AT → GC (80/83)</td>
<td>Skopek and Hutchinson, 1982</td>
</tr>
<tr>
<td>cis-Pt</td>
<td>lacI</td>
<td>GC → TA and GC → AT (455/650)</td>
<td>Brouwer et al., 1981</td>
</tr>
<tr>
<td>cis-Pt</td>
<td>Tet</td>
<td>AT → TA (12/24)</td>
<td>Burnouf et al., 1987</td>
</tr>
<tr>
<td>C(cd)P</td>
<td>lacI</td>
<td>GC → TA (150/186)</td>
<td>Eisenstadt et al., 1982</td>
</tr>
<tr>
<td>Agent</td>
<td>System</td>
<td>Principal mutation</td>
<td>Reference</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>EMS</td>
<td>lacI</td>
<td>GC → AT (102/105)</td>
<td>Burns et al., 1986</td>
</tr>
<tr>
<td>UV</td>
<td>clI</td>
<td>CG → TA (23/62)</td>
<td>Wood et al., 1984</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TA → CG (18/62)</td>
<td></td>
</tr>
<tr>
<td>UV</td>
<td>lacI</td>
<td>GC → AT</td>
<td>Miller, 1985</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(662/1046)</td>
<td></td>
</tr>
<tr>
<td>UV</td>
<td>M13lacZ</td>
<td>TA → CG (29/84)</td>
<td>LeClerc et al., 1984</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG → TA (22/84)</td>
<td></td>
</tr>
<tr>
<td>UV</td>
<td>clI</td>
<td>CG → TA (16/56)</td>
<td>Wood and Hutchinson, 1987</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TA → CG (8/55)</td>
<td></td>
</tr>
</tbody>
</table>
in \textit{uwr}^+ \text{ and } \textit{uwr}B \text{ strains of } \textit{E. coli} \text{ (Todd and Glickman, 1982). The spectra were generally similar with the exception of a new hotspot in the } \textit{uwr}B \text{ strain; this new hotspot accounted for almost 20\% of the amber mutations in the } \textit{uwr}B \text{ cells while comprising only 0.4\% in the wild-type strain. Likewise, the mutational spectra of 8-methoxypsoralen in the \textit{lacI} gene of } \textit{uwr}^+ \text{ and } \textit{uwr}B \text{ strains of } \textit{E. coli} \text{ (Yatagai \textit{et al.}, 1987) were examined. The proportion of base pair substitutions, frameshifts and deletions was altered between the repair-proficient and repair-deficient strains, and importantly, several new hotspots for base pair substitution occurred in the } \textit{uwr}B \text{ strain. In the } \textit{uwr}^+ \text{ strain, 25/34 base pair substitutions occurred at one site, while the frequency of mutation at this site dropped to 22/89 in the } \textit{uwr}B \text{ strain, with a rise in the number of mutations at another site (53/89).}

Fuchs and coworkers (Koeffel-Schwartz \textit{et al.}, 1984) examined the mutational spectra of N-AcO-AAF in wild-type, \textit{uwr}A and \textit{umu}C \textit{E. coli}. The spectrum in the wild-type was similar to the \textit{uwr}A mutants, but the \textit{umu}C spectrum revealed that mutagenesis was occurring at a subset of the wild-type mutations, in particular at several GGC GCC sequences. The authors concluded that \textit{umu}C-dependent and \textit{umu}C-independent pathways existed for the processing of DNA damage.

In a similar fashion, Sedwick \textit{et al.} (1986) looked at the mutational spectrum in wild-type and \textit{E. coli} mutant at the deoxyuridinetriphosphatase (\textit{dut}) locus. The \textit{dut} mutants showed an overall increase in base pair substitution, duplication and deletions; a new base pair substitution hotspot occurred in the \textit{dut} mutants, such that a single GC-to-AT transition site now accounted for 11/23 of the
mutants. In another study, EMS mutagenesis was examined in urv+ and urv- strains of E. coli (Burns et al., 1986); certain hotspots were found in the mutant strain that did not correspond to hotspots in the wild-type strain.

The phenomenon of mutational spectra has been more difficult to study in mammalian cells. Recent studies suggest that many of the same parameters that affect mutational spectra in prokaryotes will also affect mutational spectra in mammalian cells. Mutagens often interact directly with the DNA, which is organized into chromatin in eukaryotic cells. This additional level of organization may also affect the formation of premutagenic lesions. Gale et al. (1987) has demonstrated that UV-induced pyrimidine dimers in the nucleosome core showed a distinct periodicity, and that this periodicity was obtained when the nucleosome core DNA was isolated from UV-treated chromatin fibers or from UV-treated cells. Furthermore, the periodicity was found to be 10.3 bases, which is the average repeat for the DNA helix in core particles. The positions of maximum dimer formation occurred where the DNA backbone was predicted to be farthest from the nucleosome surface.

Consistent with the observations of Fuchs and co-workers, the knowledge that a UV-induced lesion existed was not sufficient to predict whether a particular base pair location would be a hotspot for mutation in human cells (Brash et al., 1987). A shuttle vector system was employed to measure the frequencies of cyclobutane dimers and pyrimidine-pyrimidone (6,4) photoproducts before transfection into human cells. Thus the pattern of DNA modification could be
compared with the pattern of mutation found after the shuttle vector was recovered from human cells. Mutational hotspots occurred at sites with low and high photoproduct formation and coldspots occurred at sites that suffered considerable damage. A similar phenomenon has been observed in prokaryotic cells, in that the frequency of modification at a particular base pair was not sufficient to predict the level of mutagenicity at that base pair.

Shuttle vectors, which can replicate in both bacteria and mammalian cells, have been used in an attempt to provide information about mutational spectra in mammalian cells. In some systems, the lacI gene was the mutational target so a direct comparison between bacterial lacI mutagenesis and human lacI mutagenesis was possible. The shuttle vector systems are all artificial to some degree, in that they do not study mutagenesis of an endogenous gene in its native chromosomal location. In SV40 and Epstein-Barr virus shuttle systems, the DNA replicates extrachromosomally, and there is no assurance that mutational spectra in the chromosome will be identical to spectra obtained in these shuttle systems.

M. Calos has examined the specificity of UV (Lebkowski et al., 1985), EMS (Lebkowski et al., 1986), and N-nitroso-N-methylurea (NMU) (DuBridge et al., 1987) mutagenesis in human cells using shuttle vector systems which employed the lacI gene as the target. Thus a direct comparison between mutation observed in E. coli and mutation observed in human cells was possible. The nature of base pair substitutions for human UV mutagenesis was not significantly different than that of UV mutagenesis in E. coli. Eight one percent (43 of 53) of human mutations in the lacI gene were found to be GC-to-AT transitions, while 63% (662 of 1046) of
the UV-induced mutations in *E. coli* were found to be GC-to-AT; the 95% confidence limits for the proportions 43/53 and 662/1046 show that the ratios are not significantly different. Moreover, two of the three sites best induced by UV light in *lacI* in human cells were two of the sites best induced in *lacI* in *E. coli*.

The shuttle vector system showed that 98% (53/54) of the EMS-induced mutations in human cells were to be GC-to-AT transitions, a figure which agrees very well with the *lacI* *E. coli* data in which 98% (680/691) of the EMS-induced nonsense mutations were GC-to-AT substitutions; however, the hotspots for human mutation were not the same hotspots for bacterial mutation. NMU-induced mutagenesis in human cells showed that 97% (33/34) of the mutations were GC-to-AT transitions, and a hotspot existed at a run of three consecutive guanines; no data is available for NMU mutagenesis using the *lacI* system in *E. coli*.

Another shuttle vector system used a bacterial suppressor tRNA, *supF*, as the mutational target. The mutational spectrum of B(a)P (Yang et al., 1987) and UV (Bredberg et al., 1986) in human cells has been examined. B(a)P produced 61/71 base pair substitutions, and of these base pair substitutions 45/61 were GC-to-AT, with two hotspots in two different GGG sequences accounting for 11/61 and 10/61 of the base pair substitutions. This agrees with the prokaryotic data in which 93/155 B(a)P-induced mutations were GC-to-AT (see above).

The repair capacity of the human cells may alter the mutational specificity; this phenomenon has been observed in prokaryotic systems. Bredberg et al. (1986) has used the *supF* shuttle vector system to study the UV mutational spectrum in
repair-proficient human cells and in repair-deficient xeroderma pigmentosum (XP) human cells. 93% (66/71) of the base pair substitutions in the XP line were found to be GC-to-AT as compared to 73% (59/81) in the normal cell line; the 95% confidence limits for the proportions 66/71 and 59/81 show that these figures are significantly different. The pattern of mutations were similar with the exception of a new GC-to-AT hotspot in the XP line. This is similar to the results in *E. coli* found by Todd and Glickman (1982) in which a single new UV hotspot was found in the repair-deficient strain. Interestingly, the new hotspot in both *E. coli uvrB* and human XP cells was in the sequence TCC.

Certain forward selection systems exist which monitor mutations at relatively few nucleotide pairs. These systems are all similar in that the selective agent seems to be toxic to wild-type cells because it binds to, and inactivates, essential proteins. Resistance presumably occurs from a specific modification of the target protein which renders it functional but refractory to selective agent binding and/or inactivation. Such modifications in protein structure would be expected to arise from the set of base substitution mutations which would change the amino acids principally responsible for the binding of and/or protein inactivation site of the toxic agent. In these types of assays, the genetic target is putatively limited to a small number of codons which can mutate to prevent toxin binding but which have little effect on the enzyme's catalytic activity. Mutations affecting toxin binding but also resulting in loss of enzyme activity would not be detected because the enzyme's activity is essential for cell survival. OUA, PPT and DRB appear to act via this type of mechanism. TG', however, is not mediated by the same
mechanism. Loss of functional hypoxanthine guanine phosphoribosyl transferase (HPRT) will result in the TG\textsuperscript{r} phenotype. Since the HPRT gene product is nonessential for cell survival, it was expected that this locus will be sensitive to all forms of gene mutation, including deletions and frameshifts.

Phaiik-Mooi Leong, a former student in Professor William Thilly's laboratory, examined mutational spectrum in a human lymphoblast line, TK6 (Leong, 1984). She examined the ability of different mutagens to induce resistance to the drugs 6-thioguanine (TG), ouabain (OUA), dichlororibofuranosylbenzimidazole (DRB) and podophyllotoxin (PPT).

In Leong's assay, the responses were normalized by expressing the ratio of DRUG\textsuperscript{r} to TG\textsuperscript{r}. The normalization was necessary since 1) mutant frequencies varied at the specific markers 2) spontaneous mutation frequency was much lower than chemically-induced mutant frequency. The TG\textsuperscript{r} fraction was a reflection of the frequencies of all types of mutation, while the DRUG\textsuperscript{r} fraction only revealed that portion of cells that have suffered a missense mutation at the drug binding site. The TG\textsuperscript{r} fraction was taken as a general indication of the frequency of mutation, and if a mutational hotspot existed for a particular drug-binding site, then an elevated DRUG\textsuperscript{r}/TG\textsuperscript{r} ratio was expected. In this system, the "window" for mutational spectrum was the different drug-binding sites; the results were normalized with respect to TG\textsuperscript{r}. She has calculated the OUA\textsuperscript{r}/TG\textsuperscript{r}, DRB\textsuperscript{r}/TG\textsuperscript{r} and PPT\textsuperscript{r}/TG\textsuperscript{r} ratios induced by the mutagens NQO, EMS, MNNG and ICR-191. The results are summarized in Figure 2-2.
**Figure 2-2:** Normalized small marker mutational spectrum of human lymphoblasts.

Spectra generated from an untreated, EMS-, MNNG-, NQO-, and ICR-191-treated population are shown. The mutant fraction at thioguanine-resistance (TG), ouabain-resistance (OUA), dichlororibofuranosylbenzimidazole-resistance (DRB), and podophyllotoxin-resistance (PPT) was determined. The ratios of mutant fraction at each small marker to mutant fraction at TG is shown. The hatched area represents one standard deviation and the error bars represent the 99% confidence intervals. Data is from Leong (1984).
The results showed that the spontaneous mutant fraction was different than the EMS-, MNNG- or NQO-induced mutant fraction. Furthermore, the system discriminated between EMS- and MNNG-induced mutants and between EMS- and NQO-induced mutants. The system did not discriminate between spontaneous and ICR-191-induced mutants. This latter finding is particularly noteworthy, since it indicated that the system was behaving as expected. ICR-191 generated essentially only frameshift mutations in bacteria (Calos and Miller, 1982), but the "small marker" assay used here was not expected to be sensitive to frameshift mutations since these mutations would destroy the essential gene product. The results suggested that ICR-191 was not causing base pair substitution mutations in the human cell line TK6. The results suggest that EMS, MNNG and NQO were causing base pair substitution mutations and furthermore, it appeared that the specificity of mutation induced by EMS and MNNG were different.

In summary, the phenomenon of mutational spectrum has been demonstrated in a variety of systems, from bacteriophage to human cells. The pattern of base pair mutagenesis which emerges after treatment with a given agent represents a complex set of events including DNA-sequence-specific modification by a mutagen, repair capability at a given DNA base pair, and very importantly, the effect of the selection system. The nature of mutations produced by the same agent often is similar, even with different organisms, different genes, and different selection systems.
2.2 Spontaneous Mutation

The spectrum of spontaneous mutations found in prokaryotes is complex, including deletions, insertions, base pair substitutions, and frameshifts. I shall review some of the salient studies.

The lacI suppressible nonsense system was described in the previous section. In this test system, deletions and duplications at repeated sequences were found to be the predominant mode of spontaneous mutation (Farabaugh et al., 1978); in fact, one particular genetic event dominated the entire spectrum. Deletion or addition of a CTGG sequence in the triple tandem repeat CTGGCTGGCTGG sequence accounted for 94/140 of the spontaneous mutations. In 7/12 of the other sequenced deletions, repeats of five to eight base pairs were found at the end in the wild type sequence; the deletions removed the intervening DNA and one of the repeated sequences. The importance of sequence homology to large deletion mutagenesis was further examined by constructing a lacI-Z construction in which deletions from 700-1000 bp could be monitored (Albertini et al., 1982). The major hotspot (68/113) consisted of two partially homologous sequences (14/17 base pairs) separated by 759 base pairs; other frequent deletions included repeated 6-8 base pair sequences separated by 700-1000 base pairs. Interestingly, when the homology between the two repeated sequences at the major hotspot was reduced to 13/17 base pairs, the incidence of deletion was reduced by an order of magnitude.

In an effort to compare the contribution of deletions to spontaneous and induced mutagenesis in the nonsense lacI system, over 800 chemically-induced
mutants were compared to the spontaneous collection (Farabaugh et al., 1978). When the contribution of the CTGG hotspot was subtracted from the mutational spectrum, it was found that about 14% (19/139) of the spontaneous mutants carried a deletion; importantly, the mutagens NQO and 2-AP did not stimulate deletions, yielding only 0.8% (1/87) and 2% (9/334) deletions, respectively. The mutagens NQO and 2-AP appeared to induce predominately point mutations, while deletions between repeated DNA sequences was a significant mode of spontaneous mutagenesis.

Spontaneous base pair substitution has been examined in the lacI gene using the supressible nonsense system (Coulondre et al., 1978). 5-methyl-cytosine can spontaneously deaminate to thymine, forming a mismatch. E. coli K12 will methylate the second cytosine in the sequence CCAGG, and it was a mutation in precisely this sequence that comprised the two major spontaneous base pair substitution hotspots in the lacI system. Eighty eight of 222 GC-to-AT amber mutations occurred at the methylated base. These hotspots disappeared in an E. coli strain deficient in the methylase activity, suggesting a role for cytosine deamination in spontaneous mutation.

Schaaper et al. (1986) has sequenced 174 spontaneous lacI mutants. All missense mutations (estimated to be 200) that could produce a lacI− phenotype were examined in this study. The system used in Schaaper et al. differed from the system used in the Miller studies (Coulondre and Miller, 1977b; Farabaugh et al., 1978). Only those lacI− mutants which yielded a supressible nonsense mutation were examined in Miller's studies, 72 single base pair substitutions could produce a
suppressible nonsense mutation in Miller's system. All studies of mutational spectrum use a phenotypic screening process, and differences in the screening process can yield differences in mutational spectra.

The nature and distribution of spontaneous mutations in the Schaaper et al. (1986) study were similar to the results of the Miller studies discussed above. About two-thirds of the mutations were found at the triple tandem repeat hotspot noted above and the 57 remaining mutants consisted of 22 deletions, 20 base pair substitutions, 7 insertions, 5 single base pair frameshifts, and 3 duplications. The deletions were found to be bounded by short repeats, while the spectrum of base pair substitutions was found to be different than in the nonsense spectrum of Miller. While the sample size in this study was small, transversions predominated; in Miller's nonsense spectrum, transitions outnumbered transversions after correcting for the spontaneous deamination hotspot.

Mammalian spontaneous mutagenesis at the endogenous Chinese hamster ovary adenosine phosphoribosyltransferase (APRT) locus has been studied by M. Meuth for several years. The alterations of Southern blotting patterns of spontaneously occurring APRT− mutants was examined (Nalbantoglu et al., 1983). This study could only reveal large alterations, and it was observed that 16/119 mutants had altered patterns, which implied that large structural alterations were not the predominate mode of spontaneous mutagenesis at this locus. Later studies have examined the molecular nature of the deletions which produced an altered Southern blotting pattern (Nalbantoglu et al., 1986). The distribution of deletion endpoints was found to be non-random within the 4 kb gene, clustering within a 40
bp region which was rich in dyad symmetry and short direct repeats. Short repeats (less than 7 bp) were found at the deletion endpoints, many of the deletions appeared to be simple cross-over events at the homologous bases which deleted the intervening sequence. The sizes of the deletions varied considerably, ranging from 37-1942 bp. Twelve mutations could be traced by loss or gain of a restriction site (Nalbantoglu et al., 1987), thus base pair substitution mutagenesis could be studied in this collection. All base pair mutations occurred in coding regions and with the exception of one nonsense mutant, all base pair substitution mutants produced a missense mutation. Transitions and transversion were found with no obvious hotspot.

Spontaneous mutagenesis of the E. coli xanthine-guanine phosphoribosyltransferase (gpt) gene integrated into mouse DNA has been studied (Ashman and Davidson, 1987). TG\textsuperscript{r} clones were isolated and the gpt gene was recovered; the vector containing the gpt gene was part of an SV40 retroviral shuttle vector so that the proviral form could be recovered by fusion to COS-1 cells. Twenty nine of 43 TG\textsuperscript{r} mutants contained small deletions, large deletions could not be recovered in this system since a large deletion would eliminate the SV40 sequences necessary to rescue the proviral form. Sixteen of 29 deletions were identical, eliminating one TGA sequence in the tandem repeat TGATGA. 11 mutations were base pair substitutions which were about equally divided between transversions and transitions, with no apparent hotspot.

In summary, deletion events appear to contribute significantly to spontaneous mutagenesis. In particular, the presence of short DNA repeat elements often
-44-

stimulates spontaneous mutagenesis in which the DNA between the repeated
elements, and one repeat element itself, is deleted. Chemical mutagens such as
ICR-191, MNNG, NQO and AP do not appear to induce deletion mutagenesis with
a significant frequency.
2.3 ICR-191 mutagenesis

Acridine derivatives are generally thought to produce frameshift mutations (Nasim and Brychcy, 1979, for review). Mutagenesis by ICR-191, an acridine with a half-mustard side chain, has been investigated using the nonsense system in *E. coli* (Calos and Miller, 1981). It was found that ICR-191 overwhelmingly produced +1 frameshifts in runs of three or more consecutive guanines; 365/378 ICR-191-induced mutations were of this nature. Twenty three different GGG runs existed in the *lacI* gene and mutations occurred at all GGG sites, however the frequency of induction at a particular site varied more than 30-fold. To a first approximation, when the GGG run was flanked by AT base pairs mutations were poorly induced. When both nearest neighbors were CG base pairs, mutations were induced at a higher frequency, although clear exceptions to this generalization existed. At some sites +1 frameshifts were favored by a 10:1 ratio, while at other sites -1 frameshifts predominated by a 7:1 ratio. The most mutable site (64/378) was a GGGG run, leading the authors to suggest that even higher runs of guanines may be more mutable.

Skopek and Hutchinson (1984) examined ICR-191 mutagenesis in the *cl* gene of lambda prophage and the results were very similar to those of Calos and Miller (1981). Seventeen mutations were sequenced and 16/17 involved a ±1 frameshift in a run of four or more guanines. Again, certain sites were found to be particularly mutable, with a GATGGGGGCAG site incurring 10/17 frameshifts, while the sequence TGTGGGGAAAA only produced a single frameshift.
Similar results were found using another acridine half-mustard, ICR-170, in the cyc I gene in the yeast Saccharomyces cerevisiae (Ernst et al., 1985). Twelve of 33 mutations occurred at the same site, and consisted of a +1 frameshift in a GGG run. Not all GGG runs were equally mutable and the predominate mode of mutagenesis was +1 frameshifts.

A model mechanism for spontaneous and induced frameshift mutagenesis has been proposed by Streisinger and co-workers (Streisinger et al., 1966; Streisinger and Owen, 1985) which suggests that frameshift mutations are generated at DNA gaps produced during replication, DNA repair, or recombination. Streisinger proposed that DNA structures form which have the capability of opening base pairs, slipping, and re-forming improper base pairs. If the DNA is misaligned at the time of replication, addition or deletion of bases in one strand is predicted to occur. The model predicts frameshift mutations should occur frequently by addition or removal of single bases from monotonous runs. Streisinger suggested that frameshift mutagens intercalate into the DNA helix and promote the slippage of one strand.

In summary, ICR-191 appears to induce predominately frameshift mutations. The frameshifts are ±G in runs of three or more consecutive guanines, with longer runs of guanines showing very high mutability.
2.4 MNNG mutagenesis

MNNG is an N-nitroso alkylating agent which is a potent mutagen and carcinogen. MNNG is capable of reacting directly with DNA in neutral aqueous solution, and all nitrogens and oxygens of DNA can be modified by N-nitroso alkylating agents (Singer, 1979). It is believed that the diazonium ion $\text{CH}_3\text{N}_2^+$ represents the alkylating intermediate which rapidly yields the corresponding carbonium cation (Gichner and Veleminsky, 1982).

In the late 1960's it was found that MNNG could methylate nucleic acids (Singer et al., 1968). Loveless (1969) showed that O$^6$-MeG could be produced by treating dG with methyl-nitrosourea; he also made the observation that ethyl methane sulfonate, a mutagen, produced O$^6$-MeG while methyl methane sulfonate, not a mutagen in his experiments, did not produce O$^6$-MeG. This lead him to suggest that O$^6$-MeG was disrupting the base pairing of guanine and was responsible for the mutagenic and carcinogenic effects of alkylating agents. O$^6$-MeG was soon identified in MNNG-treated E. coli (Lawley and Orr, 1970) and in MNNG-treated mouse cells (Lawley and Thatcher, 1970). In MNNG-treated mouse cells, O$^6$-MeG comprised about 7% of the total methylated DNA, but in mouse cells treated with dimethyl sulfate, no O$^6$-MeG was found. The amount of O$^6$-MeG produced correlated well with the biological activity of the two methylating agents; MNNG was a potent carcinogen, while dimethyl sulfate was a weak carcinogen, again suggesting the involvement of O$^6$-MeG in mutation and carcinogenesis.
Goth and Rajewsky (1974) radiolabeled ethyl-nitrosourea and administered a dose to rats that led to a high incidence of brain tumors in these animals. They isolated DNA from the brain and other tissues and analyzed the amounts of alkylated DNA present. It was found that O\textsuperscript{6}-EtG was not rapidly removed from the brain, the target organ, while O\textsuperscript{6}-EtG was greatly reduced in other tissues. This led the authors to suggest that it was not the extent of O\textsuperscript{6}-alkylation that was specifically important in carcinogenesis, but rather that O\textsuperscript{6}-alkyl groups were not removed enzymatically from a given cell population and would remain in the DNA and mispair during replication. O\textsuperscript{6}-MeG was shown to cause miscoding when included in templates for RNA polymerase (Gerchman and Ludlum, 1973) and also to cause misincorporation when included in a DNA template (Abbott and Saffhill, 1979). In the latter study, a poly(dC-dG) template containing O\textsuperscript{6}-MeG was used as a substrate for *E. coli* polymerase I and it was found that increased levels of O\textsuperscript{6}-MeG caused increasing misincorporation of dTMP but not dAMP, suggesting that O\textsuperscript{6}-MeG could base pair with thymine. The miscoding potential of O\textsuperscript{6}-MeG was demonstrated directly by chemically synthesizing a tetranucleotide with a single O\textsuperscript{6}-MeG. This tetranucleotide was then ligated into the *PstI* site of the M13mp8 viral genome and *PstI*-insensitive mutants were analyzed at the sequence level (Loechler *et al.*, 1984). Sixty mutants were sequenced and it was found that O\textsuperscript{6}-MeG produced exclusively GC-to-AT transitions. Thus chemical analysis of alkylated DNA has led to the conclusion that O\textsuperscript{6}-MeG, while not the major DNA lesion produced by simple alkylating agents, was the lesion predominately responsible for the mutagenesis and carcinogenesis observed with these compounds;
furthermore, O\(^6\)-MeG appeared to mispair with thymine, causing GC-to-AT transitions.

Examination of the specificity of MNNG mutagenesis has revealed that MNNG produces almost exclusively GC-to-AT transitions. MNNG mutagenesis was examined by Prakash and Sherman (1973) in the iso-1-cytochrome c gene of yeast by using a series of 11 cyl mutants that could be reverted to the wild type by a known base pair substitution. Seventeen mutagens were tested with this system, and it was found that MNNG induced 1755/1769 GC-to-AT transitions.

Nonsense mutations in the lacI gene of E. coli induced by MNNG were found to be predominately GC-to-AT. Fourteen sites in the gene could produce an amber mutation by this transition and all sites were mutable; likewise, all of the 12 ochre sites were mutable. Two hundred and ninety two of 303 amber mutations were GC-to-AT, and 210/215 ochre mutations were also GC-to-AT. The distribution of mutations was found to be site-dependent, with certain sites accumulating more mutations than others. The results are summarized in Table 2-II. The mutability of a particular site ranged from 1-13% of the total scored mutations. If GC-to-AT transitions were distributed randomly, each amber site would be expected to accumulate 20.9 ± 9.1 mutations, while each ochre site would receive 17.5 ± 8.4 mutations (± represents two standard deviations). Three amber sites and two ochre sites incurred mutations at a rate greater than chance, and these can be simplistically considered as hotspots for MNNG mutagenesis. Therefore 3/14 (21%) of the amber sites and 2/12 (17%) of the ochre were hotspots. MNNG
Table 2-II: Distribution of MNNG-induced nonsense mutations in the lacI gene of E. coli.

Fourteen amber sites and 12 ochre sites in the lacI gene were capable of detecting GC-to-AT transitions. The absolute number of mutations at each site is given, this number is also given as a percentage of the total number of amber or ochre mutations. 292/303 amber mutations were GC-to-AT, 210/215 ochre mutations were GC-to-AT. Data is from Coulondre and Miller (1977b).
<table>
<thead>
<tr>
<th>Amber site</th>
<th>Number of mutants</th>
<th>Percent of total</th>
<th>Ochre site</th>
<th>Number of mutants</th>
<th>Percent of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A5</td>
<td>13</td>
<td>4</td>
<td>O10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>A6</td>
<td>20</td>
<td>7</td>
<td>O9</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>A9</td>
<td>34</td>
<td>11</td>
<td>O11</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>A15</td>
<td>4</td>
<td>1</td>
<td>O13</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>A16</td>
<td>21</td>
<td>7</td>
<td>O17</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>A19</td>
<td>37</td>
<td>12</td>
<td>O21</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>A21</td>
<td>22</td>
<td>7</td>
<td>O24</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>A23</td>
<td>32</td>
<td>11</td>
<td>O27</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>A24</td>
<td>34</td>
<td>11</td>
<td>O28</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>A26</td>
<td>12</td>
<td>4</td>
<td>O29</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>A31</td>
<td>18</td>
<td>6</td>
<td>O34</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>A33</td>
<td>24</td>
<td>8</td>
<td>O35</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>A34</td>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A35</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
mutagenesis essentially was monitored at 26 sites throughout the lacI gene which is roughly 1000 bp in length. If we consider these sites uniformly distributed, on average a site to monitor MNNG mutagenesis existed every 39 bp and that approximately 1 out of 5 sites was a hotspot. Thus a MNNG hotspot can be said to occur about every 200 bp.

The specificity of MNNG mutagenesis was also examined for a different locus, the bacteriophage P22 mnt repressor gene (Lucchesi et al., 1986). The system was a forward assay which examined mutations which inactivated the 254 bp repressor. Twenty nine of 30 MNNG-induced mutations were GC-to-AT substitutions, with two sites accumulating 9/30 and 7/30 mutations.

The effect of neighboring base pairs on MNNG mutagenesis has been examined by Burns et al. (1987) using a forward assay in the E. coli lacI gene. Their system did not involve the nonsense suppressible system of Miller and co-workers, but rather involved selecting all detectable lacI− clones. Sequencing the mutants permitted analysis of all types of mutations which produced a lacI− phenotype, including base pair substitutions, frameshifts, additions, and deletions. MNNG-induced mutants were sequenced and 164/167 were GC-to-AT transitions at 24 different sites. Considerable site variation was found in the mutation frequency, and a high rate of mutation at a particular guanine was associated with a 5' purine. A guanine one base pair upstream from the mutagenic target appeared to make the site particularly mutable. A guanine residue was eight-fold more mutable when preceded 5' by a purine than a pyrimidine. All of the mutational hotspots involved 5′ Pu-G, while all of the cold sites consisted of 5′ Py-G.
Interestingly, the alkylating agent EMS, while showing the same strong bias towards GC-to-AT transitions, did not reveal a bias towards 5' Pu-G mutagenesis (Burns et al., 1986).

The mutagenic specificity of \textit{N}-methyl-\textit{N}-nitrosourea (MNU) and \textit{N}-ethyl-\textit{N}-nitrosourea (ENU) was examined by Richardson et al. (1987). Mutations in \textit{gpt} gene of \textit{E. coli} were selected by resistance to 6-thioguanine. Sequencing revealed that all (39/39) of the MNU-induced mutations were GC → AT transitions, while 24/33 point mutations induced by ENU treatment were GC → AT substitutions. Mutational hotspots were associated with the middle guanine in the sequence 5' GG(A or T), 32/39 MNU and 17/24 ENU GC → AT transitions occurred at this position. Both Richardson et al. (1987) and Burns et al. (1987) found that the second guanine in a 5' GG sequence was particularly mutable by the \textit{N}-nitroso alkylating agents MNNG, MNU and ENU.

Mutagenesis of the Ha-ras-1 oncogene induced by \textit{N}-methyl-\textit{N}-nitrosourea (NMU) has been reported (Zarbl et al., 1985). Mammary carcinomas were induced in rats by a single dose of NMU. DNA from 36 of the tumors produced malignant transformation of NIH 3T3 cells in gene transfer experiments. DNA from the NIH 3T3 transformants was examined for mutations in the 12th codon of the Ha-ras-1 oncogene by hybridization to oligonucleotide probes. Each of the 36 transformants carried an identical GC → AT transition at the second position of the 12 codon, coded by a GGA sequence. As discussed previously, the second guanine in a 5' GG sequence was often highly mutable by \textit{N}-nitroso alkylating agents.
In summary, MNNG induces mainly GC → AT transitions, and the most likely mutagenic lesion produced by this agent is O\textsuperscript{6}-MeG. A 5'-Pur-G sequence appeared to be more mutable than a 5'-Pyr-G sequence, and a G-G sequence appeared to be especially mutable.
2.5 HPRT

Hypoxanthine guanine phosphoribosyl transferase (HPRT) is a purine salvage enzyme that converts the purine bases, hypoxanthine and guanine, to their respective nucleotides. HPRT provides an alternative pathway to de novo synthesis of IMP and GMP. The gene is X-linked (Shapiro et al., 1966; Nyhan et al., 1967; Ricciuti and Ruddle, 1973), therefore only one functional copy exists per male cell. Forward and reverse selection systems exist, that is, both HPRT\(^{-}\) and HPRT\(^{+}\) cells can be selected in culture (Szybalski and Szybalski, 1962; Szybalski, 1958). HPRT is a nonessential enzyme for cells in culture, allowing the accumulation of mutations without killing the cells. In humans, severe HPRT deficiency can produce the Lesch-Nyhan syndrome (Seegmiller et al., 1967) and a partial HPRT deficiency may lead to severe gout and overproduction of uric acid (Kelley et al., 1967).

What can be said about the nature of mutations expected at the HPRT locus in cell culture? Selection for HPRT\(^{-}\) cells will yield those cells with reduced or no HPRT activity. Null HPRT activity can be produced by insertion/deletion, nonsense and frameshift mutants; missense mutations can also produce no HPRT activity if the mutation is in a critical portion of the enzyme. Cells with reduced HPRT activity may be produced by promotor mutants or by a set of missense mutations that impair, but do not destroy, the enzyme’s function.

The selective agent can also affect the nature of mutations recovered. Typically, 6-thioguanine (TG) or 8-azaguanine (AG) is used to select for HPRT\(^{-}\)
cells. As stated above, HPRT is part of the purine salvage pathway and the enzyme scavenges guanine and hypoxanthine and transfers the ribose phosphate moiety to the purine to form the corresponding nucleotide as is shown in Figure 2-3. Both 8-azaguanine and 6-thioguanine are recognized by the enzyme and are metabolized to the triphosphates, these toxic metabolites are eventually incorporated into DNA and/or RNA. It has been suggested that 8-azaguanine is mainly incorporated into RNA (Nelson et al., 1975). Incorporation into nucleic acids results in cell death by an as yet undetermined mechanism, however, HPRT- cells cannot process the analogs and therefore escape the cytotoxic action of the purine analogs. The counterselection system (HAT) uses aminopterin to block the de novo synthesis pathway and only HPRT+ cells, those with a functional purine salvage pathway, will survive.

6-thioguanine is a much better substrate for the HPRT enzyme than 8-azaguanine, the apparent $K_m$ in vitro for 8-azaguanine (300 x $10^{-6}$ M) was over 300-fold greater than that of 6-thioguanine (0.9 x $10^{-6}$ M); 6-thioguanine showed about the same $K_m$ as the natural substrate, hypoxanthine (0.7 x $10^{-6}$ M) (vanDiggelen et al., 1979). Azaguanine-resistant clones have been isolated that contain significant levels of HPRT activity (Littlefield, 1963; Gillen et al., 1972; Sharp et al., 1973), while TG-resistant clones generally show drastically reduced levels of HPRT activity (Sharp et al., 1973; Hsie et al., 1975; Cox and Masson, 1976). Furthermore, AG-resistant mutants able to grow in the counterselective HAT medium were found (vanDiggelen et al., 1979; Simili et al., 1983), suggesting that these mutants still possessed significant HPRT activity.
**Figure 2-3:** Reactions catalyzed by HPRT
Hypoxanthine + PRPP \rightarrow \text{inosinate} + \text{PP}_1 \\
\text{Guanine} + \text{PRPP} \rightarrow \text{guanylate} + \text{PP}_1
It appears that 6-thioguanine is a more stringent selective agent than 8-azaguanine, that is, cells with reduced HPRT activity will be TG-sensitive but may be AG-resistant. Thilly et al. (1978) has suggested that 6-thioguanine is extremely stringent, that is, essentially all HPRT activity must be lost before the cell becomes TG-resistant. TG may select for total loss of functional enzyme, as would be produced by additions/deletions, frameshift and nonsense mutations. Base pair substitution mutants which possess partial HPRT activity are not expected to be TG-resistant. Only the subset of base pair substitution mutants which are essentially devoid of HPRT activity are expected to be TG-resistant.

The cDNA coding for human HPRT has been cloned and sequenced (Jolly et al., 1983). Full-length mRNA is about 1600 bp (Jolly et al., 1982), and there is a single open reading frame of 654 bp, a 5' untranslated sequence of about 90 bp and a single polyadenylation signal. The base pair composition of the cDNA is given in Figure 2-4. It should be noted that all base pair positions in this dissertation refer to the numbering in this figure.

The HPRT locus consists of nine exons spanning about 44 kb on the X chromosome. The organization of the gene is given in Figure 2-5. Four distinct HPRT cross-hybridizing sequences are dispersed on human chromosomes 3, 5 and 11 (Patel et al., 1984). These autosomal sequences could represent pseudogenes or genes with sequence homology to HPRT. Hybridization of specific cDNA fragments to the autosomal sequences, as well as partial sequence analysis, has led to the belief that the HPRT cross-hybridizing sequences represent intronless pseudogenes (Patel et al., 1984) although the cross-hybridizing species have not
Figure 2-4: Sequence of the human HPRT cDNA

The first base pair of the start codon ATG is at position 86, the first base pair of the stop codon TAA is at position 740. The base pair positions of the nine exons are indicated below the figure.
Exon 1 1-112; exon 2 113-219; exon 3 220-403; exon 4 404-469; exon 5 470-487; exon 6 488-570; exon 7 571-617; exon 8 618-694; exon 9 695-1331.
been sequenced.
Figure 2-5: Genomic organization of human HPRT

The coding sequences are shown as blocks, the length of each exon is given in base pairs. The intervening sequences are shown as lines, the length of each intron is in kilobase pairs.
2.6 Polymerase Chain Reaction

Polymerase chain reaction (PCR) permits specific *in vitro* synthesis of a desired DNA sequence. The starting material may be a crude cell lysate, genomic DNA, double-strand or single-strand plasmid DNA. Two oligonucleotide primers, one complementary to the (+) strand and the other complementary to the (-) strand of the DNA to be amplified, are hybridized in solution to heat-denatured target DNA. This forms a template suitable for DNA chain elongation which occurs upon addition of a DNA polymerase. The cycle of denaturation, hybridization, and DNA polymerization is repeated, which results in an exponential amplification of the DNA which is spanned by the primers.

PCR was first used to achieve a 220,000-fold amplification of a 110 bp region of the human β-globin locus using 1 μg of total genomic DNA as the starting material (Saiki *et al*., 1985; Mullis and Faloona, 1987). Two 20 bp oligonucleotides were used as primers, the DNA polymerase used was the Klenow fragment of *E. coli* DNA polymerase I. The amplified DNA was analyzed for the presence of the β^a^ allele which is diagnostic for sickle cell anemia. Hybridization of the amplified, wild type, β^a^ locus to a radiolabeled, single-strand probe formed a unique *DdeI* restriction site. The amplified, sickle cell β^a^ allele did not form a *DdeI* site, and the presence of mutant DNA was revealed by the inability to restrict the DNA with *DdeI*.

The nucleotide misincorporation rate of the Klenow fragment during the PCR process was examined by cloning and sequencing the amplified products, this
was done for the 110 bp fragment amplified from the \( \beta \)-globin locus and for a 242 bp fragment of the second exon of the human leukocyte antigen (HLA) DQ\( \alpha \) locus (Scharf et al., 1986). The \( \beta \)-globin locus was amplified for 20 cycles, 10 clones were sequenced and no base pair substitutions were found; the frequency of nucleotide misincorporation was therefore less than 1 in 700. Likewise, the HLA DQ\( \alpha \) locus was amplified for 27 cycles and three clones were sequenced. One base pair substitution was found, indicating an error rate of about 1/600 after 27 cycles of amplification. The error rate can also be estimated by

\[
\text{(mutant fraction)} = \frac{(cycles) \times (bp \ length \ of \ amplified \ DNA)}{(error/bp/cycle)}.
\]

Using a mutant fraction of .33, 27 cycles, and 202 as the length of the amplified HLA DQ\( \alpha \) DNA (not including the primers) the error rate for the Klenow fragment under PCR conditions was \( 6 \times 10^{-5} \) error/bp/cycle. This figure is in agreement with the literature values for the fidelity of the Klenow fragment (Loeb and Kunkel, 1982).

A thermostable DNA polymerase from *Thermus Aquaticus* (*Taq*) was next introduced for use in PCR (Saiki et al., 1988). This enzyme is stable at the temperatures needed for PCR, thus eliminating the need to add fresh enzyme after each cycle. Use of Klenow fragment in PCR often resulted in a broad molecular size distribution of amplification products, presumably the result of nonspecific annealing and extension of primers to unrelated genomic sequences under nonstringent \( (37^\circ \ C) \) hybridization conditions. The *Taq* DNA polymerase
hybridization temperature was increased to 55º C which resulted in amplification of only the desired β-globin locus (Saiki et al., 1988).

* Taq* DNA polymerase may be able to amplify a single target molecule in 10^5-10^6 cells (Saiki et al., 1988). This was demonstrated by adding varying amounts of wild type cells to cells deleted for the β-globin locus, *Taq* DNA polymerase was able to successfully amplify a single β-globin target in the presence of the DNA equivalent of 5 x 10^5 cells deleted for this locus.

The authors estimated the fidelity of *Taq* DNA polymerase to be 2 x 10^-4 error/bp/cycle (Saiki et al., 1988). 239 bp of DNA from the HLA DPβ locus was amplified 30 cycles and 28 clones were sequenced. Seventeen base pair substitutions were found, and the authors calculated the error frequency by the formula:

\[ m = 2(f/d) \]

where \( m \) is the error frequency, \( f \) is the error frequency in the PCR product, and \( d \) is the number of doublings. Assuming that the 17 base pair substitutions occurred in different clones, the percentage of molecules which contained a PCR-induced base pair substitution was 17/28, or about 60%. Ten of 28 base pair substitutions were AT → GC, 3/28 were GC → AT, 2/28 were AT → TA, and 2/28 were GC → TA.

PCR has been used to amplify and directly sequence a wide variety of human genetic loci including mitochondrial DNA (Wrischrik et al., 1987), activated
c-Ki-ras oncogenes (McMahon et al., 1987), hemophilia factor VIII (Levinson et al., 1987), n-ras, k-ras and h-ras oncogenes (Farr et al., 1988), β-thalassemia (Wong et al., 1987), β- and γ-globin (Engelke et al., 1988), and HPRT (Keohavong et al., 1988; Cariello et al., 1988). In addition, the extraordinary efficiency of PCR has been employed to detect mitochondrial and HLA DQα sequences from a single human hair follicle (Higuchi et al., 1988), as well as to identify human immunodeficiency virus sequences from infected individuals (Kwok et al., 1987; Ou et al., 1988) where the frequency of infection may be $10^{-4}$ per cell.

PCR is capable of DNA sequence modification by selection of appropriate primers. The oligonucleotide primers become incorporated into the amplified DNA at the 5' and 3' ends, and a novel restriction enzyme site has been added by using a primer which contains the novel restriction site sequence at the 5' end (Scharf et al., 1986). During the first PCR cycle, the 5' end of the primer was not base paired with the target DNA, however, this did not interfere with the amplification process. During subsequent cycles, a blunt-ended DNA fragment was produced which contained the new restriction site.

In summary, PCR is capable of providing microgram quantities of a desired DNA sequence directly from total human genomic DNA, this greatly facilitates DNA sequencing. The sensitivity of PCR is such that a few DNA molecules may be sufficient to permit amplification. PCR also permits the addition of a DNA sequence at the 5' and/or 3' ends of the DNA molecule by amplification with oligonucleotides which contain the desired sequence alterations.
2.7 Denaturing gradient gel electrophoresis

Denaturing gradient gel electrophoresis (DGGE) can separate DNA molecules on the basis of sequence. The central principle is that the electrophoretic mobility in a polyacrylamide gel of a partially melted DNA molecule containing both helical and random coil segments is considerably less than that of a fully helical molecule; the mobility declines rapidly as the length of the melted portion increases. As a helical DNA molecule migrates into an increasing concentration of denaturant, portions of the molecule begin to melt and the mobility of the molecule is sharply reduced. A sequence alteration can change the concentration of denaturant at which the molecule melts, permitting the molecule to travel for either a longer or shorter time at high mobility before the melting occurs and its mobility becomes reduced. Furthermore, the melting behavior of DNA can be predicted solely from the primary sequence.

The statistical-mechanical analysis of the equilibrium representing the change in DNA structure as a function of temperature from the orderly helix to a disordered, unstacked structure has been presented by Wartell and Montroll (1972) and Poland (1974), and recently reviewed by Gotoh (1983) and Wartell and Benight (1985). A more complete discussion may be found in Poland (1978) or Poland and Scheraga (1970).

Poland (1974) developed an algorithm which calculates, from the nucleotide sequence, the temperature at which each base pair is in a 50:50 equilibrium between the helical and the melted state. However, the computer time for the
Poland algorithm becomes prohibitive for long DNA sequences, since the computational time is proportional to $N^2$, where $N$ is the number of base pairs. Fixman and Freire (1977) have developed an approximation for the loop entropy term of the Poland algorithm which reduced the calculation time to the order of $N$.

Gotoh and Tagashira (1981) have examined 43 cooperatively melting domains in several DNAs to evaluate the role of nearest neighbor interaction in the melting of a given base pair. Unlike the original Poland algorithm, which assigned the same stability value to a given AT or GC base pair, Gotoh and Tagashira assigned a stability parameter to a base pair doublet. There are 10 doublets, and Gotoh and Tagashira assigned a melting temperature in degrees Centigrade to each doublet.

The melting algorithm used in this thesis was obtained from L. Lerman and is given in the Appendix. It uses the Fixman-Freire modification of Poland’s algorithm with the Gotoh-Tagashira nearest neighbor stability values. The algorithm was used with a cooperativity parameter, $\sigma$, of $2.0 \times 10^{-5}$, $\Delta S / R = 12.25$, a starting temperature of $60^\circ$ C, and a temperature increment of $0.2^\circ$ C. Use of this algorithm generates a melting map showing the temperature at which each base pair is in 50:50 equilibrium between the helical and melted states.

The electrophoretic mobility of DNA in polyacrylamide gels is sensitive to the secondary structure of the DNA molecule. A DNA molecule that is partially melted, containing both helical and random coil segments, shows a drastically reduced mobility compared to the totally helical form. As the temperature
increases, DNA begins to melt in a stepwise fashion with a contiguous region melting as a single unit, this contiguous region is termed a domain. Domains generally are from 50-300 bp in length (Myers et al., 1987). DGGE employs a linear gradient of urea/formamide in a polyacrylamide gel, this gradient is the equivalent of a shallow temperature ramp.

The estimated pore size in a typical polyacrylamide gel is thought to be only several times the mean radius of a double-stranded DNA molecule (Lerman et al., 1984). Lerman and Frisch (1982) proposed that partial DNA melting effects the electrophoretic mobility by changing a helical molecule, which has two ends, into a molecule with more than two ends. If melting occurs from one end only, the configuration can be regarded as a "star polymer" with three ends, if both ends are melted, the configuration can be regarded as a star polymer with four ends. Each arm of a partially melted molecule can enter several pores, but it is unlikely that two arms will follow the same pore path. The authors suggest that the kinetics of branched polymer diffusion can explain the reduced mobility which accompanies partial melting. The relative mobility can be expressed by

$$\mu = \mu_0 e^{-(melted/L_r)}$$

where $\mu$ is the melted mobility, $\mu_0$ is the mobility of the helical form, melted represents the total number of melted base pairs, and $L_r$ is the length of a flexible DNA unit. The flexible DNA unit length is 70-90 base pairs, and is regarded as the length of a DNA statistical segment. The behavior of DNA in solution can be predicted from polymer chemistry, in which a molecule is regarded as a linked chain of rigid statistical segments.
Fischer and Lerman (1983) examined the behavior of homoduplex, perfectly base paired DNA using DGGE. Sixteen known mutants in a 536 base pair fragment of bacteriophage lambda were examined. Mutants in the low-temperature melting domain were displaced from the wild type position, while mutations in the higher-temperature melting domains of the molecule were not displaced from the wild type position. Importantly, in two mutant DNA's, GC → AT transitions at different base pair locations were resolved from one another, as well as resolved from the wild type. Also, mutant DNA's with a TA → GC and a TA → CG substitution at the same base pair location were separated by DGGE. A double mutant containing an AT → GC and a GC → TA, which was identical to the wild type DNA in gross base composition was resolved from the wild type. DGGE was sensitive only to mutations which occurred in the lower-temperature melting domains. A mutation in a higher-temperature melting domain did not affect the melting of a lower-temperature domain. If the denaturant composition of the gel and the electrophoresis time are such that only the lowest-temperature domain melts, then only homoduplex mutations in the lowest-temperature domain will be detected. This work demonstrates the remarkable sensitivity of the gel system.

The possibility of using DGGE to directly obtain a mutational spectrum was first discussed in Thilly (1985). The melting characteristics of HPRT exon 3 were shown to be compatible for use with DGGE and the feasibility of obtaining mutational spectra from a set of individual TG\textsuperscript{r} mutants, as well as from a complex population of TG\textsuperscript{r} mutants was discussed. These ideas are also put
forward in Cariello and Thilly (1986).

DGGE was used to detect single base pair substitutions in total human genomic DNA (Myers et al., 1985a). A single-strand, radiolabeled, 272 bp, wild type probe complementary to a region of the \( \beta \)-globin gene was hybridized in solution to genomic DNA from patients with \( \beta \)-thalassemia. The DNA of these patients contained known base pair substitutions. Use of a wild type probe ensured that a mismatched heteroduplex formed between the wild type probe and the target DNA, and the mismatched heteroduplex was resolved from the wild type homoduplex by DGGE. Four \( \beta \)-thalassemia alleles were examined and all four alleles were resolved from one another as well as resolved from the wild type. Patients heterozygous for the various alleles could also be distinguished. The sensitivity of DGGE was illustrated by the fact that all alleles were distinguishable, and by the fact that two GC \( \rightarrow \) AT substitutions at different base pair location were resolved from one another.

A mismatched heteroduplex always destabilizes a domain, so all mutants melt at a lower concentration of denaturant than does the wild type. This is not the case when homoduplexes are examined, since some substitutions stabilize a domain and hence the DNA will melt at a higher concentration of denaturant. In the case of homoduplexes, some base pair substitutions are predicted to not alter the melting characteristics of a domain, and the mutations are predicted to not be detected. Gel positions for all possible homoduplex base pair substitutions were calculated for the mouse \( \beta \) globin promoter (Myers et al., 1985c), and about 6\% (23/399) of the base pair substitutions were predicted to be undetectable.
However, if a mismatch can be introduced by hybridization to a wild type probe, all base pair substitutions in the low-temperature melting domain are predicted to be detected (Lerman and Silverstein, 1987); in addition, the destabilization may be so great that it alters domain boundaries such that the region of the molecule sensitive to base pair substitutions may be increased (Myers et al., 1987).

In an effort to determine what fraction of base pair substitutions may be detected by DGGE, Lerman et al. (1986) analyzed the entire continuous 35,400 bp region of the human β-globin cluster. The simulation entailed (1) fragmenting the gene with a restriction endonuclease which recognized a four-base pair sequence, (2) introduction of a mismatch every 10 bp which represented hybridization to a wild type probe, and (3) calculation of the destabilization produced by the mismatch. The mismatch was simulated by a decrement of 50°C from the Gotoh-Tagashira stability values of the nearest neighbors that included the mismatch. If a mismatch produced a 0.1°C alteration, the mismatch was scored as positive; the 0.1°C may be conservative, in that this represents about a 3 mm displacement with an appropriate gradient. Fifty percent of all possible base changes for AluI fragments were predicted to be detected, likewise, 53% of all changes for HaeIII were predicted to be detected. Some sites that could not be seen in the HaeIII fragments could be detected in the AluI fragments and vice versa. Adding those additional base pair substitutions from the second restriction enzyme increased the overall fraction of detectable site to about 66.5%. These calculations required over four continuous days of computer time.

The melting characteristics of a DNA fragment may be altered by the
addition of a GC-rich sequence, termed a GC-clamp (Myers et al., 1985b; (Myers et al., 1985c), at one end of the molecule. As mentioned above, DGGE is not sensitive to mutations which occur in the highest-temperature melting domain, but when a GC-rich sequence is added to a DNA molecule the GC-clamp becomes the highest-temperature melting domain, and mutations in the previously inaccessible region of the DNA molecule become detectable. Myers et al. (1985b, 1985c) has used a 300 bp GC-rich region (80% GC) from the human α1-globin gene as a GC-clamp. A 135 bp fragment of the mouse β-globin gene was used as a test molecule and the 300 bp GC-clamp was essentially ligated to the target molecule. When the hybrid molecule was examined by DGGE, mutations in all domains of the original molecule were detected. If the 135 bp fragment was examined as a homoduplex without a GC-clamp, about 40% of all possible base pair substitutions were predicted to be detectable, but when a GC-clamp was added, 95% of all possible base pair substitutions were predicted to be resolved from the wild type. These studies examined homoduplex melting behavior, it seems reasonable to assume that very close to 100% of the mutations in the 135 bp target molecule could be detected by heteroduplex formation using a wild type probe with a GC-clamp.

DNA methylation states can alter the mobility of a DNA fragment on a denaturing gradient gel (Collins and Myers, 1987). In this study, the authors examined the effects of 5-Me-C and 6-Me-A and found that 6-Me-A destabilized melting domains, while 5-Me-C stabilized melting domains. Furthermore, the DGGE effects of methylation appeared additive, in that the effects of full methylation were always greater, roughly two-fold, than the DGGE shift produced
by a hemimethylated molecule. The destabilization produced by 6-Me-A was calculated to be approximately 0.2° C per methylated base (using 0.313° C = 1% denaturant).

DGGE has been used to examine human DNA polymorphisms within a 1.2 kilobase region of chromosome 20 (Noll and Collins, 1987). The sequence of the probe, D20S4, was not determined; the probe was chosen on the basis of its behavior on a denaturing gradient gel. The DNA from 41 individuals from two family studies was examined and it was found that the DGGE polymorphisms were inherited in a Mendelian fashion. The authors applied DGGE to the screening of human genomic DNA for sequence polymorphisms in regions where the base sequence had not been determined.

One instance of using DNA:RNA and RNA:RNA duplexes for DGGE has been reported (Smith et al., 1986). Six point mutations in the influenza virus were distinguishable by analysis of RNA:RNA hybrids using DGGE. The DNA:RNA hybrids were also found to yield satisfactory results.

DGGE has been used to generate and isolate random single base substitutions in cloned DNA (Myers et al., 1985d). Mutations were generated by treatment of single-stranded DNA with chemical mutagens, followed by enzymatic synthesis of the complementary strand. Misincorporation often occurred when the polymerase encountered a damaged base in the template. The duplex DNA fragments, containing a variety of single base substitutions was cloned en masse and the cloned DNA population was run on a denaturing gradient gel and the DNA in the
area of the gel above and below the wild type position was recovered. In this way DGGE was used to physically separate wild type from mutant DNA species, eliminating a time-consuming screening procedure to locate the mutants.

DGGE separates DNA molecules on the basis of sequence. This technique is sensitive to both the nature and location of mutation. Mutational spectra has two components, namely, the nature and location of mutations, and DGGE seems well-suited to the study of the phenomenon of mutational specificity.
Chapter 3

Materials and Methods

References to my laboratory notebooks are found in this dissertation. For example, NFC9-101 means laboratory notebook number 9, page 101.

3.1 Human cell lines, bacterial strains, plasmids, vectors

TK6, a human diploid male lymphoblast line (Skopek et al., 1978), was the cell line predominately used in this study; a clonal isolate of TK6, THP1, was also employed. TK6 was used to generate untreated, MNNG- and ICR-191-induced independent TG$^+$ clones as a part of Monica Cahilly Bretzin's masters thesis (Bretzin, 1985). THP1, the clonal isolate of TK6, was used as a source of wild type DNA, as well as for the generation of untreated, MNNG- and ICR-191-treated complex mixed mutant populations.

Other human cell lines used include, (1) RJK853, a lymphoblast male cell line totally deleted for HPRT which was obtained from C.T. Caskey (Baylor College of Medicine, Houston, TX), (2) HPRT$_{Munich}$, a male lymphoblast cell line with a known amino acid substitution in HPRT exon 3, which was obtained from J.M. Wilson (Massachusetts General Hospital, Boston, MA), (3) GM1416B, a human lymphoblast line containing four X chromosomes which was obtained from the Human Genetic Mutant Cell Repository in Camden, NJ. All cell lines were grown in RPMI 1640 medium (GIBCO, Grand Island, NY) with 10% horse serum
(GIBCO, Grand Island, NY) in a 5% CO$_2$ humidified atmosphere.

Both (1) plasmid pHPT30, a plasmid bearing the entire human HPRT cDNA on a 1 kb PstI fragment, and (2) plasmid pλ3001, which contains human HPRT exon 3 and 1.2 kb of flanking sequence on a PstI fragment, were obtained from C.T. Caskey (Baylor College of Medicine, Houston, TX).

*E. coli* bacterial strain HB101 (a dam$^+$ strain) and *E. coli* strain GW5100 (a dam$^+$ strain) was obtained from J. Essigmann (Massachusetts Institute of Technology, Cambridge, MA), GW3810 (a dam$^-$ strain) was obtained from G. Walker (Massachusetts Institute of Technology, Cambridge, MA), JM101 (a dam$^+$ strain) was obtained from Bethesda Research Laboratories (Gaithersburg, MD), JM110 (a dam$^-$ strain) was a gift from J. Messing (Waksman Institute, Rutgers University, Piscataway, NJ). All RF M13 vectors were obtained from New England Biolabs (Beverly, MA) and SP6 vectors were from Promega Biotec (Madison, WI).
3.2 Isolation of independent, TG¹, untreated, MNNG- and ICR-191-treated TK6 clones

Mutagenesis was performed by Monica Cahilly Bretzin (Bretzin, 1985). Briefly, 200 TK6 cells were seeded in 200 ml of medium and allowed to grow for 11 days. Cells were dosed and immediately after treatment 2 – 5 x 10⁵ surviving cells were seeded into 105 separate small cultures. After 10 days of growth for phenotypic expression, cells were plated in the presence of 1 μg/ml TG, one clone from each small culture was recloned and characterized. All mutants are therefore independent.

At the time of treatment, cells were plated to determine the background mutant fraction; cells were also grown for 10 days after treatment and again plated to determine the background MF. Thus the background MF was determined at two points in time. However, the MF at the time of treatment was determined with greater precision because 20 microtiter dishes were used to determine the background MF this time, while 2 microtiter dishes were used to determine the background MF 10 days after treatment.

The background MF was thus determined at two points in time, however, the two values were not in agreement for the MNNG-treated culture. I will assume that the background MF determined with 20 microtiter dishes at the time of treatment was the best value. MNNG-treated cells were grown for 10 days after treatment and the expected increase in the background MF as a function of time is given below. Thus the increase in background MF after 10 days of growth was
added to the background MF determined at the time of treatment. This is explained more fully below.

The background MF for the MNNG-treated cells was determined at the time of treatment. 31,140 cells/well (c/w) were plated in 20 microtiter dishes in the presence of 1 µg/ml TG; plating efficiency (PE) was determined by plating 2 c/w in two microtiter dishes. The PE was 0.53 and the background MF was $4.1 \times 10^{-6}$ (see Furth et al. [1981] for discussion of MF calculation).

After MNNG treatment, $5 \times 10^7$ untreated cells were grown for 10 days and plated to determine the background MF, however, only two microtiter dishes were used to determine the background MF at this point. The background MF for the MNNG-treated culture was $1.8 \times 10^{-5}$, and the PE was 0.25; the PE was determined by seeding 2 microtiter dishes at a density of 2 c/w. Fourteen days after treatment, the background MF was again determined as described immediately above, the background MF was $1.6 \times 10^{-4}$ based on a PE of 0.098; this PE is extremely low and the background MF of $1.6 \times 10^{-4}$ will be discarded.

As discussed above, the background MF of $4.1 \times 10^{-6}$ was determined with a high level of accuracy at the time of MNNG treatment, as 20 microtiter dishes were used to determine the MF at this point. If the spontaneous increase in background MF is assumed to be $1.5 \times 10^{-7}$/day (Penman et al., 1983), then 10 days after treatment the MF should be

$$(1.5 \times 10^{-7}/\text{day}) \times (10 \text{ days}) + 4.1 \times 10^{-6} = 5.6 \times 10^{-6}.$$

I shall assume that the background MF calculated with 20 microtiter dishes
was more accurate than that calculated with 2 microtiter dishes, therefore the background MF in the MNNG experiment was $5.6 \times 10^{-6}$.

About $5 \times 10^8$ TK6 cells were treated for 60 min with 15 ng/ml MNNG (Sigma, St. Louis). Survival was determined by growback experiments to be 0.11. Immediately after treatment, $4.5 \times 10^6$ cells were seeded into 105 individual flasks; this inoculation consisted of $5 \times 10^5$ surviving cells (33 TG$^r$ mutants). After 10 days of growth for phenotypic expression, about 20,000 c/w were plated in microtiter dishes in the presence of 1 μg/ml TG. Clones were visible after 10 to 11 days and one clone per plate was selected for recloning at 0.1 c/w. Only one clone per original small individual culture was selected for further characterization.

To determine the mutant fraction induced by treatment, about $5 \times 10^7$ MNNG-treated cells were grown to allow phenotypic expression of HPRT mutants then 20,000 c/w were plated in 1 μg/ml TG. Ten days after treatment, the MF was $6.5 \times 10^{-5}$, this value was based on plating 2 microtiter dishes to determine the MF and plating two microtiter dishes to determine the PE (PE = 0.24). Fourteen days after treatment, the treated cells were again plated as described immediately above and the MF was $4.7 \times 10^{-5}$ (PE = 0.37). Assuming that the average of both MF's is the best value, the total MF for the MNNG-treated cells was $5.6 \times 10^{-5}$. Thus 3080 TG$^r$ cells were present after treatment { 3080 = ($5 \times 10^8$ cells) (0.11 survival) ($5.6 \times 10^{-5}$ MF) }.

About $10^8$ TK6 cells were treated for 24 hr in the dark with 500 ng/ml ICR-191 (Polysciences Inc., Warrington, PA). Survival was determined by
growback experiments to be 0.7. Immediately after treatment, $2.8 \times 10^5$ cells, which consisted of $2 \times 10^5$ surviving cells ($60 \text{ TG}^r$ mutants), were seeded into 105 individual flasks. After 10 days of growth for phenotypic expression, about 20,000 c/w were plated in microtiter dishes with 1 $\mu$g/ml TG. Clones were visible after 10 to 11 days and one clone per plate was selected for cloning at 0.1 c/w. Only one clone per original small individual culture was selected for further characterization.

At the time of treatment, the background ICR-191 MF was determined to be $2.0 \times 10^{-6}$; this value was based on plating 20 microtiter dishes to determine the MF and plating 2 microtiter dishes to determine the PE ($\text{PE} = 0.67$). As discussed above, assuming an increase in background MF of $1.5 \times 10^{-7}$/day, the expected background MF 10 days later was $3.5 \times 10^{-6}$. In the case of ICR-191, the expected value agreed with the observed value. About $3 \times 10^7$ untreated cells were grown for 10 days after ICR-191 treatment and the MF was determined to be $4.1 \times 10^{-6}$; this calculation was based on plating 2 microtiter dishes for the MF and plating 2 microtiter dishes to determine the PE ($\text{PE} = 0.51$). 14 days after treatment the untreated MF was again determined as described immediately above to be $4.3 \times 10^{-6}$. The average of these two values will be used as the best estimate of the background MF, thus the background MF for the ICR-191-treated cells was $4.2 \times 10^{-6}$.

About $3 \times 10^7$ ICR-191-treated cells were grown in two separate, non-independent cultures (cultures A and B) for 10 days to determine the induced mutant fraction. The mutant fraction was $4.5 \times 10^{-4}$ in culture A (2 microtiter
dishes for MF, 2 microtiter dishes for PE, PE = 0.52), and 8.6 x 10^{-5} in culture B
(2 microtiter dishes for MF, 2 microtiter dishes for PE, PE = 0.29). 14 days after
treatment the two cultures were again plated to determine the MF. The MF 14
days after treatment was 2.4 x 10^{-4} for culture A (2 microtiter dishes used for the
MF, two microtiter dishes used for PE, PE = 0.31). The PE determined for
culture B was extremely low, 0.04, and the value for the MF of culture B 14 days
after treatment will be discarded. The average of the three platings will be taken
as the best value, the total MF in the ICR-191-treated population was 5.2 x 10^{-4}.
Thus 36,400 TG^r cells were present after treatment \{ 36,400 = (10^8 cells) (0.7
survival) (5.2 x 10^{-4} MF) \}.

The fraction of TG^r cells induced by treatment is given by

\[(fraction\ induced) = \frac{[ (total\ MF) - (background\ MF) ]}{(total\ MF)}\].

Thus 90\% of the TG^r clones were induced by treatment with MNNG, and 99\% of
the TG^r clones were induced by treatment with ICR-191. The probability that any
single mutant was of spontaneous origin was 10\% for the MNNG-treated cells and
1\% for the ICR-191-treated cells.

As discussed above, one TG^r clone per individual small culture was selected
and recloned at a cell density of 0.1 c/w. The recloned line was expanded, live
cells were frozen in 10\% DMSO, and genomic DNA was isolated by the method of
Blin and Stafford (1976). All MNNG- and ICR-191-treated clones are of
independent origin.
Untreated, independent, TG<sup>r</sup> mutants were generated by inoculating individual cultures with about 100 TK6 cells. The cultures were grown for 48-58 generations to allow spontaneous mutagenesis to occur. 20,000 c/w were plated in microtiter dishes in the presence of 1 μg/ml TG, and the average MF was 1 x 10<sup>-5</sup>. Several clones per individual culture were selected for recloning and further characterization. Clones from the same individual culture were designated 52A and 52B, for example. While TG<sup>r</sup> clones from separate individual cultures are independent, clones taken from the same individual culture may not be independent.
3.3 Generation of untreated, MNNG- and ICR-191-treated complex mutant populations

Two methods of obtaining mutational spectrum are present in this thesis, (1) a clone-by-clone analysis of individual TG\(^r\) colonies, and (2) an analysis of a complex population consisting of several thousand different mutants. This section will describe the generation of the complex mutant populations.

Briefly, about 5 x 10\(^8\) cells were untreated or treated with MNNG or ICR-191. TG was added directly to the population en masse; the TG\(^r\) mutants continued growing and eventually overgrew the culture. The culture then consisted of a complex population of TG\(^r\) mutants which were derived from the original TG\(^r\) mutants.

A clonal isolate of TK6, THP1, was used for the generation of the complex mutant populations. 1.7 x 10\(^7\) THP1 cells were placed into 200 ml of RPMI1640 (GIBCO, Grand Island, NY), 10\% horse serum (GIBCO, Grand Island, NY), and CHAT (1X CHAT = 10\(^{-5}\) M cytidine, 2 x 10\(^{-4}\) M hypoxanthine, 2 x 10\(^{-7}\) M aminopterin, 1.75 x 10\(^{-5}\) M thymidine). CHAT treatment reduced the background mutant fraction by killing TG\(^r\) cells. After 48 hours, the cells were centrifuged and resuspended in 500 ml of fresh medium with THC (THC is CHAT without aminopterin). THC enabled the cells to recover more rapidly from the CHAT treatment. This culture was expanded for three days to provide sufficient cells for the experiment.

MNNG (Sigma Chemical Co., St. Louis, MO) was diluted to 100 \(\mu\)g/ml in
DMSO. 50 µl of 100 µg/ml MNNG was added to two separate 500 ml cultures, called A and B; the final MNNG concentration was 10 ng/ml. The cell density at the time of treatment was about 10^6/ml. After 1 hour of exposure to MNNG, 250 ml of culture A and B was centrifuged and resuspended in 500 ml of fresh medium. Thus about 2.5 x 10^8 cells were exposed to MNNG.

ICR-191 (Polyscience Inc., Warrington, PA) was diluted to 2.5 mg/ml in DMSO. 100 µl of 2.5 mg/ml ICR-191 was added to two 500 ml cultures, called C and D, at a cell concentration of about 5 x 10^5 cells/ml. Cells were exposed to 500 ng/ml ICR-191 for 24 hours in the dark. After exposure, 300 ml of culture C and D was centrifuged and resuspended in 500 ml of fresh medium, after resuspension the cells were at a concentration of about 5 x 10^5/ml; thus about 2.5 x 10^8 cells were treated with ICR-191.

MNNG treatment produced very little toxicity, the doubling times for the MNNG-treated cultures were only reduced slightly after treatment. One day after treatment the doubling times were 18-19 hours, two days after treatment the doubling times were 19-22 hours, and three days after treatment the doubling times returned to the normal value of 15-16 hours. ICR-191 treatment produced some toxicity, one day after treatment the doubling times of both ICR-191-treated cultures were about 30 hours, on the second day after treatment the doubling times were about 19 hours, and three days after treatment the doubling times returned to the normal value of 15-16 hours.

Toxicity was not monitored by plating the cultures, so the toxicity must be
estimated from the doubling times. ICR-191 treatment produced a 30 hour doubling time after treatment which is consistent with 50% survival, and MNNG treatment produced an 18 hour doubling time after treatment which is consistent with 85% survival. This is the minimum survival value since without plating there was no way to distinguish cell death from a reduction in doubling times. Therefore, a minimum of \(2.1 \times 10^8\) surviving cells were treated with MNNG, and \(1.3 \times 10^8\) surviving cells were treated with ICR-191.

Four treated 500 ml cultures and two untreated 500 ml cultures, called \(E\) and \(F\), were diluted daily for 6 days to allow phenotypic expression of HPRT mutations. Six days after treatment, a portion of the cultures was plated in four microtiter dishes in the presence of 1 \(\mu g/ml\) TG to determine the mutant fraction as described in Section 3.2. After plating to determine the mutant fraction, 1 \(\mu g/ml\) TG was added to the 500 ml cultures. On the second, fourth and sixth day after adding TG to the 500 ml cultures, one-half the culture volume was discarded and fresh medium containing 1 \(\mu g/ml\) TG was added. After about a week, the \(TG^r\) cells had overgrown the culture and the cultures exhibited normal doubling times. Cell pellets were frozen at this time for DNA isolation.

The mutant fraction in untreated culture \(E\) was \(3.21 \times 10^6\) with a PE of 79%, the MF in MNNG-treated culture \(A\) was \(9.27 \times 10^6\) with a PE of 77%, and the MF in ICR-191-treated culture \(C\) was \(7.8 \times 10^5\) with a PE of 108%. The absolute number of spontaneous mutants in culture \(E\) was

\[
(3.21 \times 10^6 \text{ mutants/cell}) \times (2.5 \times 10^8 \text{ cells}) = 803 \text{ mutants.}
\]
Therefore, the population that eventually overgrew the untreated culture after TG was added was derived from about 800 original TG\textsuperscript{r} mutants. The spontaneous mutants are a subset of the mutants in the MNNG- and ICR-191-treated populations, the proportion of spontaneous mutants drops as the level of mutagenesis increases.

MNNG-treated population \( A \) consisted of \( 2.1 \times 10^8 \) surviving cells. Therefore the absolute number of spontaneous mutants in this population was about 675 \( \{ (2.1 \times 10^8) (3.21 \times 10^{-6}) \} \), while the absolute number of MNNG induced mutants was about 1275 \( \{ (2.1 \times 10^8) (6.06 \times 10^{-6}) \} \). \( 1.3 \times 10^8 \) cells survived ICR-191 treatment, so the absolute number of spontaneous mutants in culture \( C \) was about 420 \( \{ (1.3 \times 10^8) (3.21 \times 10^{-6}) \} \) and the absolute number of ICR-191-induced mutants was about 9725 \( \{ (1.3 \times 10^8) (7.50 \times 10^{-5}) \} \). The complex mutant populations and the number of cells from which they were derived were:

- Spontaneous population \( E \): 800 TG\textsuperscript{r} cells
- MNNG-treated population \( A \): 675 spontaneous and 1275 induced TG\textsuperscript{r} cells
- ICR-191-treated population \( C \): 420 spontaneous and 9725 induced TG\textsuperscript{r} cells

This experiment is in NFC5-105.

It should be noted that the mutant fraction in cultures \( B, D, \) and \( F \) was not determined. The portion of the cultures which was to be used to determine the mutant fraction in these three cultures was contaminated and therefore unusable. However, after treatment with MNNG, the doubling times of cultures \( A \) and \( B \)
were the same, as were the doubling times for ICR-191-treated cultures C and D. The mutant fraction of MNNG-treated culture B is probably close to that of culture A. Likewise, the mutant fraction of ICR-191-treated culture D is probably close to that of culture C, and the mutant fraction of untreated culture F is probably close to that of culture E.

Monica Cahilly treated TK6 cells with MNNG and individual TG\textsuperscript{r} clones were eventually isolated from her population. However, cells from her MNNG-treated population were frozen, and I used these cells to produce another MNNG-treated population. As described previously, Monica Cahilly's MNNG-treated culture contained 3080 TG\textsuperscript{r} cells, 2750 of which were induced by treatment. Monica Cahilly's treated culture was retrieved from the freezer, and TG was added as described above. After about a week, the TG\textsuperscript{r} cells overgrew the culture and the cells were frozen for later DNA isolation. The induced MF in her culture was 5.04 \times 10^{-5} and the untreated MF was 5.6 \times 10^{-6}. Thus the MNNG-treated, complex mutant population was derived from 2750 induced and 330 untreated TG\textsuperscript{r} cells.
3.4 Construction of vectors used to generate single-strand, radiolabeled DNA complementary to exon 3

Several different probes were used to examine HPRT exon 3 mutations. The probes and their applications are summarized in Table 3-I. Two different single-strand probes were used, both probes were capable of detecting mutations only in the low-temperature melting domain of exon 3; these probes are called the 169 bp and the 224 bp probe.

Exon 3 is 184 bp in length and comprises base pairs 220-403 in the HPRT cDNA (base pair numbering is as in Nucleic Acid Sequence Database, Release 32.0, National Biochemical Research Foundation, Washington, DC; the first nucleotide of the start ATG codon is at position 86). The 169 bp probe was generated from a XhoI/MboII exon 3 fragment, the XhoI site is at bp 234 and the MboII site is at bp 403. Thus the 169 bp probe does not contain the 5' 14 bp of exon 3. The 224 bp single-strand probe contained the entire 184 bp coding region of exon 3 and 20 bp of both 5' and 3' intron sequence. The construction of the probes as well as generation of single-strand probe is given in Sections 3.4.1 and 3.4.2.

Polymerase chain reaction was used to (1) amplify exon 3 from genomic DNA, which produced a 224 bp fragment, and (2) add a GC-rich sequence to exon 3, producing a 332 bp fragment. The GC-rich sequence, termed a GC-clamp, modified the melting characteristics such that mutations at all exon 3 base pair locations could be detected by denaturing gradient gel electrophoresis.

Amplified DNA with and without a GC-clamp was examined for mutations.
A 224 bp, single-strand probe was used to analyze amplified DNA which did not contain a GC-clamp. If a GC-clamp was added to exon 3, then a double-strand, radiolabeled, 332 bp sequence which contained a GC-clamp was used as a probe. This is summarized in Table 3-I.
**Table 3-I:** Summary of probes and their applications

DNA was hybridized in solution to a radiolabeled exon 3 probe and mutations were revealed by DGGE. The 169 base pair probe and the 224 base pair probe could only detect mutations which occurred in the low-temperature melting domain of exon 3, while use of the GC-clamp permitted the resolution of mutations at all exon 3 base pair locations.
**Independent, TG^r clones**

1) Genomic DNA was restricted with TaqI and MboII which produced a 169 bp exon 3 fragment. Restricted genomic DNA was hybridized in solution to a 169 bp single-strand, radiolabeled probe. Mutations in only the low-temperature melting domain of the 169 bp, exon 3 fragment can be detected by DGGE.

2) PCR was used to add a GC-clamp to exon 3. A double-strand, radiolabeled, exon 3 sequence was hybridized in solution to DNA amplified from the mutants. Both the wild type probe and the mutant sequences contained a GC clamp. Mutations at all exon 3 base pair locations are expected to be resolved by DGGE.

**Complex, TG^r populations**

1) PCR was used to amplify exon 3 without a GC-clamp, which produced a 224 bp fragment. Amplified DNA was hybridized in solution to a single-strand, radiolabeled, 224 bp exon probe. The 224 bp fragment is sensitive to mutations which occur in low-temperature melting domain.

2) PCR was used to add a GC-clamp to exon 3. A double-strand, radiolabeled, exon 3 sequence was hybridized in solution to DNA amplified from the complex, TG^r populations. Both the wild type DNA and the mutant sequences contained a GC clamp. Mutations at all exon 3 base pair locations are expected to be resolved by DGGE.
The independent, TG\textsuperscript{r} clones were analyzed by (1) restriction of genomic DNA and solution hybridization to the 169 bp probe, and (2) PCR amplification of exon 3 to add a GC-clamp. The complex mutant populations were examined by (1) PCR amplification and hybridization to a 224 bp single-strand probe, and (2) PCR amplification to add a GC-clamp.

3.4.1 Vector used to produce 169 bp single-strand probe and conditions used to generate probe

This work was done by Jamie K. Scott and can be found in her notebook on pages 11-38. The protocols for growth and isolation of single-strand, recombinant M13 vectors were exactly as described in Messing (1983).

The 1 kb \textit{PstI} fragment of vector pHPT30 containing the entire HPRT cDNA was ligated into the \textit{PstI} site of RF M13mp18 using T4 DNA ligase following the manufacturer’s protocol (Bethesda Research Laboratories, Gaithersburg, MD). The recombinant DNA was then transfected into competent \textit{E. coli} strain GW5100 and plated. Five individual plaques were selected and a 1.5 ml virion culture was made from each plaque. Five \(\mu\)l of each virion preparation was dotted onto two sheets of GeneScreen Plus (New England Nuclear, Boston, MA) and baked in a vacuum for 2 hr at 80\(^\circ\) C.

Both (+) and (-) strand, full-length, RNA probes complementary to the HPRT cDNA were made using the SP64 system (Promega Biotec, Madison, WI) following the manufacturer’s protocols. The filters were prehybridized for 4 hr at 60\(^\circ\) C in 10 ml of 50\% formamide, 50 \(\mu\)M NaPO\(_4\) (pH 6.5), 800 \(\mu\)M NaCl, 1 \(\mu\)M
EDTA, 0.5% SDS, 2.5X Denhart's solution (50X = 1% Ficoll, 1% polyvinylpyrrolidione, 1% bovine serum albumin), 250 μg/ml salmon sperm DNA, 500 μg/ml yeast RNA. 6 x 10^6 cpm of probe was added and hybridized for 12 hr at 58° C. After hybridization, the filters were washed twice for 15 min at 58° C in 50 mM NaCl, 20 mM NaPO_4 (pH 6.5), 1 mM EDTA, 0.1% SDS. The filters were exposed to Kodak XAR-5 film, and recombinant M13 clones complementary to either the (+) or (-) RNA probe were detected. One M13 clone containing the (+) strand and one clone containing the (-) strand of HPRT was selected and a 1 l virion culture was prepared as described in Messing (1983). dam^- E. coli strain GW3810 or JM110 was used to obtain unmethylated, single-strand DNA containing the HPRT (-) strand. dam^+ E. coli strain JM101 was used to obtain methylated, single-strand DNA containing the HPRT (+) strand. The virions containing single-strand DNA were recovered by PEG/NaCl precipitation (PEG was molecular weight 8,000; Sigma, St. Louis, MO) and the virion protein coat was removed by phenol/chloroform extraction as described in Messing (1983).

Single-strand recombinant M13mp18 DNA propagated in a dam^- E. coli line bearing the (-) strand of HPRT cDNA was used as a template to prepare radiolabeled, 169 bp, single-strand, (+) strand, exon 3 probe. The recombinant M13mp18 vector with the HPRT (-) strand was hybridized to an oligonucleotide to form a template suitable for extension with a DNA polymerase. Two hundred μl of a mixture containing 200 nM recombinant ss M13mp18 with the HPRT (-) strand, 1 μM 15-mer primer (Sequencing primer; New England Biolabs, Beverly, MA), 50 mM NaCl, 3 mM Tris (pH 7.8), and 3 mM EDTA was boiled for 5 min and placed
at 60° for 1 hr.

The radiolabeled (+) strand was generated by adding 6 μl Klenow fragment (New England Biolabs, Beverly, MA; 5 U/μl) to a 54 μl reaction volume containing 53 nM primed template, 50 mM NaCl, 10 mM MgCl₂, 10 mM DTT, 30 mM Tris (pH 7.5), 50 μM each dCTP, dGTP, dTTP and a total concentration of dATP of 4 μM. After incubation for 12 hr at 30° C, the polymerase was inactivated by heating for 20 min at 70° C. Ten μl of 10X medium salt restriction buffer (1X = 50 mM NaCl, 10 mM Tris [pH 7.8], 10 mM MgCl₂, 1 mM DTT), 5 μl MboII (New England Biolabs, Beverly, MA; 5 U/μl), 5 μl XhoI (New England Biolabs, Beverly, MA; 15 U/μl), and 80 μl of water was added and the mixture was incubated for 3 hr at 37° C.

The average length of DNA synthesis was controlled by the total concentration of dATP, DNA synthesis stopped when the dATP was depleted. A dATP concentration of 4 μM was empirically determined to provide sufficient DNA synthesis through HPRT exon 3, a distance of about 400 bp from the priming site (NFC 4-195). The specific activity of the synthesized (+) DNA strand was controlled by varying the ratio of radiolabeled with non-radiolabeled dATP. For example, to synthesize DNA with a specific activity of 1000 Ci/mmol/adenosine, α-³²P-dATP of specific activity 3000 Ci/mmol (3.25 μM from supplier) would be diluted with non-radioactive dATP. The final concentration of radiolabeled dATP in the reaction mixture was calculated by
\[
\frac{\text{(desired SA) (total dATP concentration in reaction)}}{\text{(SA from supplier)}} = \frac{\text{(final concentration of } ^{32}\text{P—dATP)}}
\]

and using the above example,

\[
\frac{(1000 \text{ Ci/mmol}) (4 \mu M \text{ dATP})}{(3000 \text{ Ci/mmol})} = 1.33 \mu M \ ^{32}\text{P—dATP}
\]

To obtain a 1.33 \mu M \ ^{32}\text{P-dATP concentration in a 60 } \mu l \text{ reaction volume,}

\[
\left[ \frac{(1.33 \mu M)/(3.25 \mu M)}{(60 \mu l)} \right] = 24.6 \mu l
\]

Therefore 24.6 \mu l \text{ of radiolabeled dATP (New England Nuclear, Boston, MA) was}

dried under vacuum. The radionucleotide was suspended in the reagents described

above with sufficient nonradiolabeled dATP to adjust the total concentration of

dATP to 4 \mu M. DNA polymerase was added last.

After DNA synthesis and restriction enzyme digestion, the DNA was

precipitated in ethanol and resuspended in 21 \mu l \text{ of water and } 4 \mu l \text{ of loading}

buffer. The DNA was loaded onto a 10\%-35\% \text{ denaturing gradient gel as described}

in Section 3.5. Electrophoresis was for 5 hr at 150 V at 60\degree \text{ C in TAE buffer. The}

169 bp, double-stranded, \text{Xhol/MboII fragment melted at about 23\% denaturant,}

the band was located by autoradiography and cut from the wet gel. DNA was

electroeluted onto dialysis membranes using DNA sample concentration cups from

Isco, Inc (Lincoln, NE). DNA was recovered by precipitation with ethanol.

The radiolabeled DNA recovered at this point was an unmethylated, double-
stranded, 169 bp, *XhoI/MboI* fragment of exon 3. A small aliquot was saved for later use and the remainder was hybridized to a molar excess of recombinant M13mp18 bearing HPRT sequences complementary to the non-radiolabeled strand of the 169 bp, double-strand fragment. This procedure was a simple competition reaction designed to produce the radiolabeled (+) strand free in solution.

Typically, 5–10 x 10^6 cpm of ds probe was recovered. The radiolabeled, 169 bp, (+) strand contained 48 adenines. For example, the total yield was calculated by:

\[
\frac{7.5 \times 10^6 \text{ cpm}}{(2.22 \times 10^{12} \text{ cpm Ci}^{-1}) (1000 \times 10^3 \text{ Ci mol}^{-1} \text{ A}^{-1}) (48 \text{A})} = 7.0 \times 10^{-14} \text{ mol}.
\]

A 50 to 100-fold molar excess of recombinant, single-strand, M13mp18 with the (+) HPRT strand was added to the radiolabeled double-strand fragment, and the mixture was brought to 300 mM NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA. The mixture was boiled for 4 min, then placed at 65° C for 2 hr. After hybridization was complete, the mixture was ethanol precipitated, dried, and resuspended in water. The DNA was loaded onto an 8% polyacrylamide gel and the single-strand, radiolabeled (+) strand was located by autoradiography, excised from the gel, and electroeluted onto dialysis membrane. The DNA was precipitated in ethanol and resuspended in water. An example of 169 bp, single-strand probe production can be found on NFC9–219.
3.4.2 Vector used to make 224 bp single-strand probe and conditions for probe generation

This probe contains the entire 184 bp HPRT exon 3 coding sequence and 20 bp of intron sequence 5' and 3'. PCR was used to add unique restriction sites at the 5' and 3' ends of the 224 bp molecule and these restriction sites were used to clone the fragment into an M13 vector. In this way, only exon 3 and 20 bp of both flanking sequence was cloned.

The oligonucleotide primers (Synthetic Genetics, San Diego, CA) containing the BamHI and HindIII sites were (NFC 9-415):

\[ 5' \text{ATATGGATCC ATATATTTAAA TATACTCAC 3'} \text{ (BamHI primer)} \]
\[ 5' \text{ATATAAGCTT CCTGATTTTA TTTCTGTAG 3'} \text{ (HindIII primer)} \]
The restriction sites are underlined.

The oligonucleotides were purified on a 15% polyacrylamide/7 M urea gel. The gel was placed on a thin layer chromatography plate and only one lane was not covered with aluminum foil, the foil blocked the UV light from reaching the DNA. A hand-held, 254 nm UV light source was used to visualize the DNA in the single lane that was not covered with aluminum foil, the location of the DNA was marked. The foil was removed and the DNA in the remaining lanes was excised. The DNA in the single lane exposed to UV was not excised, in this way, oligonucleotides not exposed to UV light were recovered. DNA was recovered by electroelution onto DEAE cellulose using a D-GEL apparatus (Epigene Corporation, Baltimore, MD), recovered with a 1.0 M NaCl wash, and precipitated with ethanol.
Plasmid pλ3001 was used as a template for PCR amplification with the
BamHI and HindIII primers described in the preceding paragraph. A 100 μl
reaction volume contained 1 nM pλ3001, 33 μM each of dATP, dCTP, dGTP,
dTTP, 1 μM primers, 16.6 mM (NH₄)₂SO₄, 67 mM Tris-HCl (pH 8.6 at 22°C), 6.7
mM MgCl₂, and 10 mM βME. 1 μl of *Thermus aquaticus* DNA polymerase (New
England Biolabs, Beverly, MA) was added and 10 cycles of amplification was
performed. One cycle consisted of a 1 min denaturation at 93°C, 2 min
hybridization at 53°C, and a 2 min polymerization at 70°C. The DNA was
precipitated with ethanol and resuspended in 85 μl of medium salt restriction
buffer (50 mM NaCl, 10 mM Tris [pH 7.8], 10 mM MgCl₂, 1 mM DTT). The DNA
was digested with HindIII and BamHI, ethanol precipitated, resuspended in water,
and loaded onto a polyacrylamide gel. The DNA was located by reference to
molecular weight markers and excised with no exposure to UV light or exposure to
ethidium bromide and recovered by electroelution. This procedure is in
NFC10–407.

RF M13mp19 was digested with BamHI and HindIII (NFC10–423) and
loaded onto a 1% agarose gel. DNA was recovered by electroelution. The volume
for ligation was 20 μl which contained 2 x 10⁻¹⁰ M BamHI/HindIII-digested RF
M13mp19, 10⁻⁹ M BamHI/HindIII-digested insert, ligation buffer (Bethesda
Research Laboratories, Gaithersburg, MD), and 1 μl T4 DNA ligase (Bethesda
Research Laboratories, Gaithersburg, MD). The ligation mixture was incubated at
16°C for 2.5 hr. The DNA was transfected into competent GW5100 and plated on
IPTG and XGAL. One clear plaque was selected and single-strand DNA was
isolated from a 1 l culture as described in Messing (1983). The ligation is described in NFC10–391.

The (+) strand of exon 3 and 20 bp of 5′ and 3′ intron sequence was cloned into M13mp19 by the procedure described immediately above. Single-strand M13mp19 DNA with the (+) strand of exon 3 was recovered and was used as a primed template to prepare radiolabeled, single-strand, (-) strand exon 3 probe. A 20 bp oligonucleotide of sequence CATATATTAAATATACCTAC was used as a primer, this primer is complementary to the 3′ 20 bp of the (+) strand exon 3 intron sequence. The conditions to make the primed template and conditions for primer extension with radiolabeled dATP were as described in 3.4.1. After DNA synthesis, the Klenow fragment was heat-inactivated as described in 3.4.1 and the DNA was digested with HindIII. The DNA was brought to a concentration of 5 x 10⁻⁹ M with water and loading buffer, boiled for 5 min, and loaded onto a 6% polyacrylamide gel. The single-strand exon 3 DNA was located by autoradiography and the DNA was recovered as described in 3.4.1. The (-) strand, single-strand, 224 bp probe contained 20 bp of 5′ synthetic DNA and 65 radiolabeled adenines. An example of probe preparation is given in NFC10–257.
3.5 Digestion of human genomic DNA with *Taq*I and *Mbo*II, solution hybridization with single-strand, 169 bp exon 3 probe, and denaturing gradient gel electrophoresis.

Human genomic DNA was isolated by the method of Blin and Stafford (1976). Sixty μg of genomic DNA was brought to a concentration of about 200 ng/μl in low salt restriction enzyme buffer (10mM Tris [pH 7.5], 10 mM MgCl₂, 1 mM DTT) and 150 U of *Mbo*II (New England Biolabs, Beverly, MA) was added. The DNA was incubated at 37°C for 12 - 20 hr. 1/10 volume of 10X high salt restriction enzyme buffer (10X = 0.5 M Tris [pH 7.5], 1 M NaCl, 100 mM MgCl₂, 10 mM DTT) and 150 U of *Taq*I (New England Biolabs, Beverly, MA) was added and the mixture was incubated at 65°C for 12 - 20 hr. The concentration of the DNA was determined by spectrophotometry (1 OD = 50 ng DNA/μl), an aliquot containing 20 μg was removed and precipitated in ethanol. This procedure can be found in NFC8–203.

Twenty μg of *Taq*I/*Mbo*II-digested genomic DNA were resuspended in a 30 μl volume containing 300 mM NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA and radiolabeled, single-strand exon 3 probe. Twenty μg of human genomic DNA was assumed to have 3 x 10⁶ copies of HPRT based on 3 x 10⁹ bp per haploid genome and an average molecular weight of 660 for all base pairs (6.58 pg DNA per diploid genome). Four HPRT cross-hybridizing species have been reported (Patel et al., 1984) so it was assumed that 1.5 x 10⁷ copies per 20 μg of genomic DNA would hybridize to an exon 3 probe. 5–10 x 10⁷ copies of single strand probe was added
per 20 μg genomic DNA, the solution was boiled for 5 min, and placed at 65°C for 20–60 hr. The number of copies of probe added and the hybridization times for each experiment are given in the Results section.

After the genomic hybridization was complete, 100 ng of ss M13 DNA bearing the (−) strand of HPRT was added and hybridization was continued for another hour. The ss M13 DNA was complementary to the unbound single-strand probe and retained most of the unbound probe at the origin of the gel.

The DNA was precipitated in ethanol and resuspended in 10 μl of water and 2 μl of 6X loading buffer (1X loading buffer = 5% glycerol, 0.1% SDS, 10 mM EDTA, 0.1% bromophenol blue). Generally, a 1 mm, 12.5% polyacrylamide (37.5% polyacrylamide/1% bisacrylamide), 10%–35% denaturing gradient gel was used (100% denaturant = 7 M urea/40% formamide [vol/vol]). The gradient was formed using a 50 ml Hoefer gravity flow gradient maker (Hoefer Scientific, San Francisco, CA). Gradient solutions were prepared immediately before use, deionized for 30 min using Bio-Rad mixed bed resin AG 501-X8(D) (Bio-Rad Laboratories, Richmond, CA) and degassed. 15.5 ml of denaturant solution was added per chamber with 100 μl 10% AP and 3.5 μl TEMED. A small stir bar was placed in the chamber nearest the exit port and the solution was stirred as the gradient was formed. 10 μl of AP and 1.5 μl of TEMED was added when the level of the gradient was about 4 cm from the top of the glass plate. The gradient was poured between 1/4 inch notched glass plates. Equipment for the denaturing gradient gels was purchased from Green Mountain Lab Supply in Waltham, MA.
The gradient was allowed to polymerize for at least two hours before further handling. The comb was removed, the wells were flushed immediately with water, and the gel was immersed in a tank containing 12.5 l of TAE at 60° C. A carbon electrode was placed in the upper gel chamber, and the buffer was recirculated using a XX80 Millipore pump (Millipore Corp., Bedford, MA). The gel was allowed to equilibrate for 30 min at 150 V before flushing the wells and loading the samples. Electrophoresis was at 150 V for 4–10 hr, the electrophoresis time and gradient conditions for each experiment are given in the Results section.
3.6 PCR amplification of HPRT exon 3 from total genomic DNA

Polymerase chain reaction (PCR) permits in vitro synthesis of a desired DNA sequence by repeated hybridization of oligonucleotide primers to target DNA sequences and DNA synthesis from the primed DNA template. The quality and quantity of the DNA produced by PCR allows direct sequencing of the amplified DNA.

PCR was used to amplify HPRT exon 3 from total human genomic DNA. A 100 µl reaction volume contained 1 µg of genomic DNA, 33 µM each of dATP, dCTP, dGTP, dTTP, 1 µM oligonucleotide primers (Synthetic Genetics, San Diego, CA), 16.6 mM (NH₄)₂SO₄, 67 mM Tris-HCl (pH 8.6 at 22° C), 6.7 mM MgCl₂, and 10 mM βME. The first denaturation consisted of a 5 min incubation at 93° C, followed by a 2 min incubation at 53° C. The tube was cooled to room temperature and the polymerase was added. Subsequent cycles consisted of a 1 min denaturation at 93° C, 2 min hybridization at 53° C, and a 2 min polymerization at 70° C. The primers used to amplify exon 3 from genomic DNA were:

5' CATATATTAA ATATACTCAC 3' (downstream primer)
5' TCCTGATTTT ATTTCTG1AG 3' (upstream primer)

These primers are complementary to the intron sequence immediately flanking human HPRT exon 3. Thirty two cycles of amplification from 1 µg of genomic DNA using 1 µl of Thermus aquaticus DNA polymerase (New England Biolabs, Beverly, MA) added at cycles 0 and 15 yielded 250-500 ng of the 224 bp exon 3 fragment that contained the entire 184 bp coding region and 20 bp of both
5' and 3' intron sequence. A single band of size 224 bp was observed on an ethidium bromide-stained gel as shown in Figure 4-30.

The dNTP's were obtained in powder form (Sigma, St. Louis, MO) and neutralized with NaOH to pH 7.5. After neutralization, the concentration of sodium ion in the dNTP's was measured by DC plasma emission spectrometry, a solution of 1500 mM dNTP's was found to contain 316 mM sodium. Thus about 7 mM sodium ion was carried with the dNTP's into the final PCR reaction. Primers were HPLC-purified by Synthetic Genetics (San Diego, CA) without any exposure to UV light. A brief exposure of DNA to UV light can induce photoproducts that can be detected by DGGE; this is shown in the Appendix.
3.7 Addition of a 54 bp GC-rich sequence to exon 3 using PCR

PCR was also used to add a 54-bp GC-rich sequence, termed a GC-clamp, to the 5' and 3' ends of exon 3. The GC-clamp modified the melting characteristics of the molecule such that mutations throughout the coding region of exon 3 could be detected.

The 224 bp product containing exon 3 and 20 bp of 5' and 3' flanking sequence was used as a template for a second PCR amplification with 74 bp GC-rich primers. After amplification from genomic DNA as described in Section 3.6, 2 µl of the total reaction mix was used for the second amplification; the concentration of the 224 bp product after 32 cycles of amplification from genomic DNA was generally 1-5 × 10¹⁰ copies/µl. The conditions for the second amplification to add the GC-clamp were as described in Section 3.6 using the following primers:

5' GCCGCCTGCA GCCCGCGCCC CCCGTGCCCC
   CGCCCCGCCG CCGC CGGCCGCCGG CGCCCATATA
   TTTAAATATAC TCAC 3' (downstream primer)

5' GCCGCCTGCA GCCCGCGCCC CCCGTGCCCC
   CGCCCCGCCG CCGG CGGCCGCCGG CGCCTCTCTGA
   TTTTTTTTCT GTAG 3' (upstream primer)

The 5' 54 bp sequence is the GC-rich clamp, the 3' 20 bp are identical to the primer sequence used in the first amplification from genomic DNA. The second amplification added a 54 bp GC-clamp to the 5' and 3' ends of the 224 bp DNA, producing a 332 bp fragment. An ethidium bromide-stained gel showing the
addition of the GC-clamp is shown in Figure 4-31. After amplification, the DNA was precipitated in ethanol, resuspended in water, and loaded onto an 8% polyacrylamide gel. The 332 bp DNA product from the TGF clones was located in the gel by ethidium bromide-staining and a brief exposure to UV light. The DNA from the complex mixed mutant populations was given no exposure to UV light; the lane containing the molecular weight markers was cut from the gel and stained with ethidium bromide, the molecular weight markers were used to approximate the position of the 332 bp product in the gel. The DNA was recovered by electroelution and used for DGGE analysis.
3.8 Hybridization of 332 bp, PCR-amplified DNA containing a GC-rich sequence to wild type, radiolabeled probe and analysis by DGGE

After addition of the GC-clamp to exon 3, the total length of the molecule was 332 bp which consisted of the 184 bp coding region of HPRT exon 3, 20 bp of both 5' and 3' intron flanking sequence, and a 54 bp GC-clamp at both the 5' and 3' ends of the molecule. This molecule contained a unique HgiA1 restriction site at base pair position 288 (position 288 refers to the cDNA numbering sequence). The 332 bp product was restricted at the unique HgiA1 site and hybridized in solution to HgiA1-restricted, wild type, double stranded, radiolabeled probe.

A wild-type, double-stranded, radiolabeled 332 bp exon 3 fragment containing GC-clamps at the 5' and 3' ends was made by PCR amplification of a 224 bp exon 3 fragment using Taq DNA polymerase. The 224 bp exon 3 fragment was produced by 10 cycles of amplification of plasmid pλ3001 using Sequenase™. The conditions for amplification using Sequenase™ were: 2.25 mM dNTP's, 10 mM Tris (pH 8.0), 5 mM MgCl₂, 1 μM primers. One cycle consisted of a 45 s incubation at 100° C, 15 s at room temperature, addition of 0.5 μl enzyme, and a 4 min incubation at 37° C. The DNA was loaded onto a polyacrylamide gel, the 224 bp fragment was located without exposure to UV light and electroeluted from the gel. The same stock of 224 bp fragment was used to generate radiolabeled probe with the GC-clamps.

Double strand, radiolabeled probe with GC-clamps was made by amplification of the 224 bp fragment as described in Section 3.6 using both 74 bp
GC-rich primers. The concentration of $\alpha^{32}$P dATP (3000 Ci/mmol; New England Nuclear, Boston, MA) in the amplification reaction was such that the specific activity of the DNA produced by amplification was 20-75 Ci/mmol/adenine. Amplification was performed for 15 cycles as described in Section 3.6. The DNA was precipitated with ethanol, loaded onto an 8% polyacrylamide gel, located by autoradiography, and recovered by electroelution.

One thousand cpm (about $10^8$ copies) of radiolabeled probe was added to about a 5-fold molar excess of target DNA and the volume was brought to 30 $\mu$l with restriction enzyme buffer (buffer = 100 mM NaCl, 50 mM Tris [pH 7.7], 10 mM MgCl$_2$, 1 mM DTT). One $\mu$l of $H_{gi}AI$ (5U/$\mu$l; New England Biolabs, Beverly MA) was added and the mixture was incubated for 2 hr at 37° C. After restriction was complete, the DNA was boiled for 5 min and hybridized at 65° C for 4 hr in restriction buffer. The DNA was precipitated with ethanol and resuspended in water and loading buffer. A 1 mm thick, 12.5% polyacrylamide (37.5:1 polyacrylamide:bisacrylamide), 10%-40% linear denaturing gradient gel was used. The gel was run submerged at 60° C for 6 hr at 150 V, fixed in 40% MeOH/5% glycerol, dried, and exposed to Kodak XAR-5 film.
3.9 Direct dideoxy sequencing of TG<sup>F</sup> MNNG- and ICR-191-induced mutants

All mutations in the coding region of HPRT exon 3 should be detected by DGGE using the GC-clamp. The identity of DGGE mutants was determined by direct dideoxy sequencing of the PCR-amplified 224 bp product.

Sequencing was performed using the reagents and the protocols supplied in the Sequenase<sup>TM</sup> DNA Sequencing Kit (United States Biochemical Corp., Cleveland OH). 66 µl of sequencing reaction buffer was made by adding the following:

- 19.2 µl Sequenase<sup>TM</sup> 5X buffer (5X concentrate = 200 mM Tris [pH 7.5], 100 mM MgCl<sub>2</sub>, 250 mM NaCl)
- 6.5 µl 100 mM DTT
- 1.0 µl of a 1:20 dilution of labeling mix (original concentration of labeling mix was 7.5 µM each of dCTP, dGTP and dTTP)
- 1.0 µl of 2.0 µM 20 bp downstream primer (see Section 3.6 for sequence)
- 3.0 µl<sup>32</sup>P α-dATP (3000 Ci/mmol; New England Nuclear, Boston, MA)
- 35 µl water

Eleven µl of sequencing reaction buffer was added to 0.17 pmol (about 25 ng) of the gel-purified, PCR-amplified, 224 bp product and the reaction volume was brought to 16 µl. The mixture was boiled for 3 min and cooled at room temperature for 3 minutes. Two µl of Sequenase<sup>TM</sup> diluted 1:8 in Sequenase<sup>TM</sup> buffer was added and the reaction was allowed to continue for 3 min at room
temperature. Three and one half μl of the reaction mixture was added to each of four separate tubes containing the four dideoxy termination mixtures. The composition of the dideoxy termination mixture was 50 mM NaCl, 80 μM each dATP, dCTP, dGTP, dTTP and 8 μM of the appropriate dideoxy substrate. The reaction mixtures were incubated at 37° C for 3 min and terminated by adding 4 μl of stop solution (stop solution = 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol).

Two μl of the reaction mixture was boiled, chilled on ice and loaded onto a prewarmed 8% polyacrylamide/8.3 M urea sequencing gel (polyacrylamide:bisacrylamide = 19:1). The sequencing gel was formed using 0.4-0.8 mm wedge spacers. The gel was run for 3-6 hours at 80 watts using the constant power setting. After electrophoresis, the gel was fixed in 10% methanol/10% glacial acetic acid, dried, and exposed to Kodak XAR 5 film at -70° C with an intensifying screen.

The DNA from one mutant, RI207, could not be sequenced directly. Direct sequencing consistently yielded a wild type pattern, even though DGGE electrophoresis showed that this mutant contained an alteration in the 5' portion of exon 3. The sequencing gels were readable with the exception of about the last 5 nucleotides at the 5' end of exon 3.

To sequence mutant RI207, the 224 bp amplification product was reamplified with the BamHI and HindIII primers as described in Section 3.4.2. The reamplified DNA was restricted with BamHI and HindIII, cloned into RF
M13mp19, and single-strand DNA was isolated as described in Messing (1983). DNA was prepared from three individual plaques and was sequenced as described above. All three DNA's showed an identical base pair substitution at the last 5' base pair in the coding sequence of exon 3.
Chapter 4

Results

References to my laboratory notebooks appear in this dissertation. For example, NFC10-312 indicates notebook 10, page 312.

Several strategies were employed to obtain the mutational spectra of MNNG and ICR-191. Table 4-I presents a flowchart summarizing the overall approach.

A restriction fragment of exon 3 was found to be suitable for use with DGGE. Mutations in the low-temperature melting domain of the restriction fragment could be resolved by DGGE. Unfortunately, very few mutations occurred in the low-temperature melting domain of the restriction fragment. Therefore, it was necessary to modify the melting characteristics of exon 3 by adding a GC-rich sequence. The GC-rich sequence, termed a GC-clamp, permitted the resolution of mutations in the original high-temperature melting domain of exon 3.

Nearly all of the MNNG- and ICR-191-induced mutations occurred in the original high-temperature melting domain and could only be resolved by use of the GC-clamp. The pattern of mutation produced by ICR-191 and MNNG were found to be different, such that an MNNG-treated and an ICR-191-treated population could be recognized.

The mutational spectra of MNNG and ICR-191 was investigated in two ways, by (1) an examination of independent, TG$^+$ clones, and (2) an examination of
complex, TG<sup>r</sup> populations. The complex, TG<sup>r</sup> populations were derived by (1) treating 2-5 x 10<sup>8</sup> cells with either MNNG or ICR-191, (2) adding TG to the treated culture en masse, and (3) allowing the TG<sup>r</sup> cells to continue to grow, eventually overgrowing the culture. The complex, TG<sup>r</sup> populations were derived from 2,000-10,000 original TG<sup>r</sup> cells.
Table 4-I: Flowchart showing the overall strategy
Can the 169 bp TaqI/MboII restriction fragment be used with DGGE? YES Will human genomic methylation patterns confound the analysis of mutations in exon 3? NO

Will human pseudogenes confound the analysis of mutations in exon 3? NO Does the low-temperature melting domain of exon 3 contain sufficient information to discriminate between MNNG- and ICR-191-treated cells? NO

Modify the melting properties of exon 3 by using PCR to add a GC-clamp Does the high-temperature melting domain of exon 3 contain sufficient information to discriminate between MNNG- and ICR-191-treated cells? YES

Examine independent, TG⁻ clones and complex, TG⁻ populations using the GC-clamp
4.1 Calculated melting map and melting behavior of 169 base pair

*TaqI/MboII* exon 3 fragment on perpendicular and parallel denaturing gradient gels

Mutations in HPRT exon were examined by denaturing gradient gel electrophoresis in this dissertation. Exon 3 was selected for study because (1) it contains the largest portion of the coding region, 184 base pairs, and offers a sizeable mutational target, and (2) exon 3 is reported to code for both catalytic sites of the enzyme and mutations in the catalytic sites may produce a TG-resistant cell.

HPRT exon 3 is 184 base pairs in length, and consists of base pairs 220 – 403 in the HPRT cDNA (numbering as in Nucleic Acid Sequence Database, Relcasc 32.0, National Biochemical Research Foundation, Washington, DC; the first base pair of the start codon ATG is at position 86). This base pair numbering scheme will be used throughout the thesis.

A *XhoI/TaqI* site near the 5' end of exon 3 and a *MboII* site at the extreme 3' end, permitted the generation of a 169 bp exon 3 fragment from both genomic DNA and from plasmid DNA containing the HPRT cDNA. These restriction sites are shown in Figure 4-1.

The 5' site can be cut with either *TaqI* or *XhoI*. *TaqI* was used to restrict human genomic DNA because *TaqI* is not inhibited by 5-methyl cytosine, while *XhoI* is inhibited by 5-methyl cytosine (Kessler *et al.*, 1985). 5-methyl cytosine is found almost exclusively at CpG sequences in mammalian DNA (Wyatt, 1951;
**Figure 4-1:** Diagram showing *XhoI/TaqI* and *MboII* restriction sites in HPRT exon 3

Human HPRT exon 3 is 184 bp in length. Restriction enzyme digestion with *XhoI* or *TaqI* and *MboII* produced a 169 bp fragment. Also shown is the location of the DNA base pair substitution in HPRT<sub>Munich</sub>, an HPRT variant isolated from a patient with gout.
Vanyushin *et al.*, 1970), and one CpG sequence is present in the *XhoI/TaqI* recognition site. *XhoI* was used to restrict plasmid DNA for the generation of single-stranded exon 3 probe because the presence of multiple *TaqI* sites in the plasmid DNA precluded use of *TaqI* to generate probe.

The restriction sites are as follows:

(+) 5' 234 C↑TCGA G......TG T↓ 403 3'
(-) G AGCT↑C......AC↓C

*XhoI*  *MboII*

The *TaqI* recognition site is contained within the *XhoI* recognition site, the *TaqI* recognition sequence is TCGA and *TaqI* cleaves between the T and C.

The calculated melting map of the 169 bp *TaqI/MboII* human HPRT exon 3 fragment is given in Figure 4-2. The melting map was generated using the algorithm developed by L. Lerman, and the fortran program for the melting algorithm is given in the Appendix. The melting map showed (1) a high-temperature melting domain of about 65 base pairs, from base pair position 234 to about base pair position 300, and (2) a low-temperature melting domain of about 100 base pairs, from about base pair position 301 to 403. The low-temperature melting domain is expected to melt between 66° C and 70° C.

Denaturing gradient gel electrophoresis is sensitive to base pair substitutions in the low-temperature melting domain, while mutations in the high-temperature melting domain are not expected to be resolved. Therefore, using the 169 base pair probe, mutations which occur from base pair 300 to 400 are expected to be
Figure 4-2: Calculated melting map of the 169 bp TaqI/MboII fragment of HPRT exon 3

The melting map represents the temperature at which each base pair is in a 50:50 equilibrium between the melted and helical state. The high-temperature melting domain is from the 5' end of the molecule to about base pair 300, the low-temperature melting domain extends from about base pair 301 to the 3' end of the molecule. DGGE is sensitive to mutations which occur in the low-temperature melting domain, therefore it is expected that mutations in about 100 base pairs of exon 3 will be detected (base pair position 300-403).
CALCULATED MELTING PROFILE
Human cDNA HPRT 234 - 403 bp
detected, these 100 base pairs represent about 15% of the coding frame of HPRT.

A perpendicular denaturing gradient gel can be used to examine the physical melting characteristics of a DNA molecule. A perpendicular gel showing the mobility of the single-strand form and double-strand form of the 169 bp exon 3 fragment is shown in Figure 4-3. As expected, the mobility of the single-strand form was unaffected by the denaturant, while the double-strand form showed a reduction in mobility at concentrations of denaturant greater than 20%. The midpoint of the transition from the helical to the partially melted state occurred at about 25% denaturant.

The relationship between temperature and concentration of denaturant was determined by Lerman et al. (1984) to be 1% denaturant = 0.313°C. The perpendicular denaturing gradient gel showed that the molecule melted at about 25% denaturant, which corresponds to 67.8°C (the ambient gel temperature of 60°C was added to the temperature value calculated for 25% denaturant). The calculated melting map showed that the low-temperature melting domain was expected to melt between 66°C and 70°C, thus the calculated melting map and the empirical data from the perpendicular gradient denaturing gel were in agreement.

The number of base pairs which are still helical when the molecule has assumed the partially melted state can be determined by comparing the mobility of the helical state to the mobility of the partially melted state (Lerman et al., 1984). From Lerman et al. (1984), we have

$$\mu = \mu_0 e^{-(\text{melted} / L_r)}$$

where $$\mu_0$$ is the mobility of the helical form, $$\mu$$ is the mobility of the partially
melted form, melted is the number of base pairs which have melted, and \( L_r \) is the length of a statistical DNA segment which is taken as 75. From the perpendicular gel, using values of \( \mu_0 = 57 \) mm and \( \mu = 10 \) mm, we find that 130 base pairs are melted when the molecule has assumed the partially melted configuration, and 40 base pairs remain helical. The empirical data from the perpendicular denaturing gradient gel was in rough agreement with the calculated melting map which showed a high-temperature melting domain of about 65 base pairs.

Analytical work is generally performed with a parallel denaturing gradient gel, in which the DNA migrates into an increasing concentration of denaturant. An experiment showing the mobility of the exon 3 fragment in a parallel gel as a function of electrophoretic time is given in Figure 4-4. The mobility of the single-strand form was unaffected by the increasing concentration of denaturant, while the double-strand form showed a sharp reduction in mobility at about 25% denaturant. The double-strand molecule was relatively immobile after it assumed the partially melted state. Electrophoresis was continued for 2, 4, and 6 hours after the molecule had become partially melted, and the intensity of the signal of the partially melted molecule decreased in relation to the electrophoretic time. This suggested that the partially melted form was not infinitely stable and that increasing the electrophoretic time past the point where the molecule became partially melted resulted in the transition from the partially melted form to complete strand dissociation. To achieve maximum sensitivity using the 169 base pair probe, the minimum electrophoretic time which allows the molecule to reach about 25% denaturant should be used, increasing the electrophoresis time past this
**Figure 4-3:** Autoradiogram of perpendicular denaturing gradient gel showing the double-strand and single-strand 169 base pair *Taql/MboII* exon 3 fragment

SS is the single-strand exon 3 fragment, DS is the double-strand exon 3 fragment. $\mu_0$ is the distance traveled by the helical form, 57 mm, and $\mu$ is the distance traveled by the partially melted form, 10 mm. About $3 \times 10^4$ cpm of both SS and DS were loaded in a single well and electrophoretic travel was perpendicular to the increasing concentration of denaturant. The concentration of denaturant increased from 0% at the left side of the gel to 40% at the right side of the gel. A 14% polyacrylamide gel was used and electrophoresis time was 3 hours.

The midpoint of the transition from the helical state to the partially melted state occurred at about 25% denaturant. The faint multiple bands beneath the SS form are DNA molecules which are less than full-length; these are SS breakdown products. A schematic representation of the fully helical state and partially melted state is illustrated in the figure. This experiment is in NFC5-192.
**Figure 4-4:** Mobility of the double-strand and single-strand 169 base pair TaqI/MboI fragment on a parallel denaturing gradient gel

DS is double-stranded DNA and SS is single-stranded DNA. Electrophoresis was for 2, 4, 6, 8, 10 and 12 hours at 150 volts. 1000 cpm of both DS and SS were loaded for each of six time points, the electrophoresis time is given at the bottom of the figure. A 12.5% polyacrylamide, 10%-35% parallel denaturing gradient gel was used. The experiment is in NFC8-53.
TIME COURSE DOUBLE-STRAND AND SINGLE-STRAND

DENATURANT

- 15%
- 20%
- 25%
- 30%
- 35%

SS DS SS DS SS DS SS DS SS DS
12 10 8 6 6 4 2
point will result in a diminution of radiolabeled signal.
4.2 Resolution of methylated from unmethylated exon 3 sequence using denaturing gradient gel electrophoresis

Denaturing gradient gel electrophoresis is sensitive to the methylation state of the DNA (Collins and Myers, 1987; Cariello et al., 1988). *dam* methylase will produce an N\(^6\)-methyl adenine residue in the sequence GATC (Hattman et al., 1978), and a single bacterial *dam* methylation site exists at base pair 357 in exon 3. HPRT cDNA propagated in a *dam*\(^-\) *E. coli* strain can be resolved from HPRT cDNA propagated in a *dam*\(^+\) strain as shown in lanes 1 and 2 of Figure 4-5. The presence of a single methyl group destabilized the molecule by about 0.25\(^\circ\) C, using a value of 1\% denaturant = 0.313\(^\circ\) C (Lerman et al., 1984). The molecule was actually hemimethylated, since unmethylated single-strand probe was hybridized to fully methylated target DNA. To confirm that the destabilization was due to bacterial *dam* methylation, the cDNA was propagated in a *dam*\(^-\) strain and then methylated *in vitro* using *dam* methylase. The DGGE displacement of the cDNA methylated *in vitro* was identical to that of cDNA propagated in a *dam*\(^+\) strain, as is shown in lane 3 of Figure 4-5.

Since denaturing gradient gels are sensitive to the methylation state of the DNA, a methylated molecule could be mistaken for a mutation. In mammalian DNA, the only methylated base is 5-methyl-cytosine, which is found almost exclusively at the sequence CpG (Wyatt, 1951; Vanyushin et al., 1970). However, the 169 base pair *TaqI/MboII* fragment of exon 3 has no CpG sequences, and thus no 5-methyl cytosine residues are expected to confound the results.
Figure 4-5: Resolution of unmethylated exon 3 DNA from exon 3 DNA methylated by bacterial dam methylase

HPRT cDNA was propagated in both dam\(^+\) and dam\(^-\) E. coli strains and digested with TaqI and MboII. One \(\mu\)g of unmethylated cDNA was methylated \textit{in vitro} using dam methylase following the manufacturer's instructions (New England Biolabs, Beverly, MA). 1000 cpm of single-strand, 169 bp probe was hybridized in solution to a molar excess of methylated or unmethylated cDNA and the species were resolved using a 10\%–35\% denaturing gradient gel. The methylated species was displaced 5 mm from the unmethylated DNA. This experiment is in NFC7-184.

Lane 1: cDNA propagated in dam\(^+\) E. coli strain HB101 (NFC6-275); lane 2: cDNA propagated in dam\(^-\) E. coli strain JM110 (NFC6-228); lane 3: cDNA propagated in dam\(^-\) E. coli strain JM110 and methylated \textit{in vitro} using dam methylase (NFC6-229).
In summary, the 169 base pair fragment of exon 3 appeared to be suitable for use with denaturing gradient gel electrophoresis. A single transition from the helical to the partially melted state occurred at about 25% denaturant. The low-temperature melting domain was about 100 base pairs, and it was expected that mutations in this domain would be detected by DGGE; these 100 base pairs represent about 15% of the coding region of HPRT. No sites for the formation of 5-methyl cytosine were present in the 169 base pair fragment of exon 3, and variability in genomic methylation patterns should not be a confounding factor in the analysis of mutations.
4.3 Hybridization of 169 bp probe to *TaqI/MboII*-digested genomic DNA and analysis by conventional PAGE: the question of pseudogene interference

The previous sections showed that the 169 base pair *XhoI/MboII* fragment of HPRT exon 3 appeared suitable for use with denaturing gradient gel electrophoresis. In the next two sections, the feasibility of using this probe to detect base pair mutations in human genomic DNA will be examined. The approach was to (1) restrict human genomic DNA with *TaqI* and *MboII* to produce a 169 base pair fragment of exon 3, (2) hybridize a radiolabeled, single-strand, 169 base pair, exon 3 probe to restricted genomic DNA in solution, (3) analyze the radiolabeled duplex by denaturing gradient gel electrophoresis.

Four autosomal HPRT cross-hybridizing species have been reported (Patel et al., 1984) and it was possible that these pseudogenes could have created undesirable background problems. To investigate possible interference from exon 3 pseudogenes, two cell lines were examined, (1) a wild type cell line, and (2) a cell line totally deleted for the HPRT gene. The cell line deleted for X-linked HPRT possessed only the pseudogenes.

Genomic DNA from a wild type and HPRT deletion cell line was (1) restricted with both *TaqI* and *MboII*, (2) hybridized in solution to a single-strand, radiolabeled, exon 3 probe, and (3) loaded onto a conventional polyacrylamide gel. This experiment is shown in Figure 4-6. The polyacrylamide gel showed that the cell line with HPRT deleted had no visible hybridizing species; any cross-
hybridizing exon 3 species between 100–1000 base pairs would have been visible on this gel. A single very faint band was present at the position of the single-strand probe, and it was assumed that this band was unbound single strand probe. The wild type cell line showed the expected single species of size 169 base pairs.

This experiment showed that the HPRT pseudogenes, when digested with TaqI/MboII, had no exon 3-hybridizing species in the size range of 100-1000 base pairs. Thus, there was no interference from the HPRT pseudogenes and the X-linked HPRT sequence appeared to be the sole hybridizing species.
Figure 4-8: Hybridization of TaqI/MboII-digested genomic DNA to 169 base pair probe and analysis by a conventional polyacrylamide gel: wild type, RJK853 (HPRT deletion), HPRT\textsubscript{Munich}

Twenty $\mu$g of TaqI/MboII-digested genomic DNA was hybridized for 62 hours to $4.5 \times 10^7$ copies of single-strand, 169 base pair probe of specific activity 1234 Ci/mmol. Ten $\mu$g of genomic DNA was loaded per lane. Lane 1: 1000 cpm ss probe; lane 2: 1000 cpm ds probe; lane 3: 1000 cpm ss probe hybridized to a molar excess of XhoI/MboII-digested HPRT cDNA; lane 4: wild type (THP1, NFC6-169); lane 5: HPRT\textsubscript{Munich}' (NFC6-169); HPRT deletion (RJK853, NFC4-101).

This experiment is in NFC8-215.
1 SS
2 DS
3 DNA
4 WT
5 MUNCH
6 DELETION
4.4 Resolution of \( \text{HPRT}_{\text{Munich}} \) from the wild type sequence using denaturing gradient gel electrophoresis: a mutation in the low-temperature melting domain

The previous section showed that there was no interference from the HPRT pseudogenes. In this section, the resolution of a known HPRT exon 3 mutant from the wild type sequence is demonstrated.

\( \text{HPRT}_{\text{Munich}} \), a cell line isolated from a male patient with gout, has been examined at the amino acid level and a single amino acid substitution was found (Wilson and Kelly, 1984). A serine → arginine transition at amino acid 103 placed the corresponding DNA mutation at the extreme 3' end of exon 3 (bp 395-397) as shown in Figure 4-1.

Detection of mutations by denaturing gradient gel electrophoresis will involve the hybridization of a wild type probe to mutant DNA, forming a mismatched heteroduplex. The formation of a mismatch will reduce the stability of a mismatched heteroduplex compared to the wild type homoduplex, thus all mutations are expected to melt at a concentration of denaturant less than that of the wild type.

The effect of a mismatch can be modeled using a program called SQHT (Myers et al., 1987). The fortran code for SQHT is given in the Appendix. This program simulates the effect of a mismatch by reducing two Gotoh-Tagashira values at the base pair location of the mismatch. The program gives the expected difference in gradient level for a single base mismatch at every position along a
fragment as a function of electrophoresis time. In this way, the portion of the molecule expected to be sensitive to base pair substitutions can be discerned.

The effect of a single mismatch at each base pair location of the 169 bp Taq1/MboII exon 3 fragment is given in Figure 4-7. A mismatch at the expected location of HPRT$_{Munich}$, base pair 395-397, should be detected as this mismatch is predicted to alter the melting temperature about 0.75° C compared to the wild type homoduplex.

The algorithm predicts that only mutations from about base pair 255-300 will not be detected. Mutations at the 5' 20 base pairs, which appear to be part of the high-temperature melting domain, should be detected. A mismatch may cause a perturbation of the helix which is sufficient to alter the melting domains, thus allowing detection of mutations in a portion of the high-temperature domain (Myers et al., 1987). Therefore, the 120 base pair section of the molecule sensitive to base pair substitutions will be considered the low-temperature melting domain.

The autoradiogram of the denaturing gradient gel showing HPRT$_{Munich}$ and the wild type is given in Figure 4-8. HPRT$_{Munich}$ was displaced 14 mm from the wild type unmethylated sequence. This corresponded to a reduction in the melting temperature of about 0.65° C using a value of 1% denaturant = 0.313° C (Lerman et al., 1984). The observed value of the alteration in heteroduplex melting temperature was in agreement with the predicted calculated value of 0.75° C.

HPRT$_{Munich}$ was sequenced several years later and was found to have a CG → AT transversion at base pair location 397 (Cariello et al., 1988; this
**Figure 4-7:** Calculated destabilization produced by mismatches at each base pair location of the 169 base pair probe

The wild type homoduplex melting map is shown as a solid line (Y axis), and the difference in gradient level between the wild type homoduplex and a mismatched heteroduplex (in degrees Centigrade) is shown as a dotted line (Y2 axis). The coding portion of exon 3 in the TaqI/MboII fragment is from base pair position 234-403. A difference of 0.1° C should produce a 1-2 mm shift using typical denaturing gradient gel conditions of a 30% change in gradient over a 17 cm gel, and 1% denaturant = 0.313° C. The single amino acid substitution detected in HPRT\textsubscript{Munich} places the corresponding DNA mutation at base pair location 395-397. A mismatched wild type:HPRT\textsubscript{Munich} heteroduplex is predicted to be detected.

The program SQHT was run with the following values: destabilization = 50° C, retardation length = 75, velocity zero = 1.8. The value for the heteroduplex is for a 5 hour electrophoresis time.
Calculated melting profile of TaqI/HpaII exon 3 fragment and calculated heteroduplex destabilization at each base pair location.

Retardlength=75, Velzero=1.8, Destab=50, t=5 hr.
Figure 4-8: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and resolution of HPRT\textsubscript{Munich} from the wild type sequence by denaturing gradient gel electrophoresis.

Thirty \( \mu g \) of TaqI/MboII-digested genomic DNA was hybridized for 47 hours to 6.8 x \( 10^7 \) copies of single-strand, 169 base pair probe of specific activity 1178 Ci/mmol. Fifteen \( \mu g \) of genomic DNA was loaded on a 10\% - 35\% denaturing gradient gel and the electrophoresis time was 5 hours. Two hundred cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA which was propagated in both a \( dam^+ \) and \( dam^- \) \textit{E. coli} strain.

Lane 1: SS, \( 10^3 \) cpm single-strand (ss) probe; lane 2: \( DAM^+ \), \( 10^3 \) cpm ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a \( dam^+ \) \textit{E. coli} strain; lane 3: MUNICH, HPRT\textsubscript{Munich} genomic DNA (lymphoblast line, NFC6-169); lane 4: WT, THP1 genomic DNA (wild type lymphoblast line, NFC6-169); lane 5: \( DAM^- \), \( 10^3 \) cpm ss probe hybridized to HPRT cDNA propagated in a \( dam^- \) strain.

This experiment is in NFC8-101.
Figure 4-9: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by denaturing gradient gel electrophoresis: wild type, HPRT deletion line, line with 4 X chromosomes

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 38 hours to 6.0 x 10^7 copies of single-strand, 169 base pair probe of specific activity 907 Ci/mmol. Twenty μg of genomic DNA was loaded on a 5% - 30% denaturing gradient gel and the electrophoresis time was 5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA which was propagated in both a dam^+ and dam^- E. coli strain.

Lane 1: SS, 10^3 cpm ss probe; lane 2: DAM^-, 10^3 cpm ss probe hybridized to HPRT cDNA propagated in a dam^- strain; lane 3: DAM^+, 10^3 cpm ss probe hybridized to HPRT cDNA propagated in a dam^+ strain; lane 4: Δ, RJK853 genomic DNA (HPRT deletion lymphoblast line, NFC 9-234); lane 5: WT, THP1 genomic DNA (wild type lymphoblast cell line, NFC 8-169); lane 6: 4X, GM1416B genomic DNA (lymphoblast line with four X chromosomes, NFC 9-63).

This experiment is in NFC9-238.
publication is given in the Appendix). A (+) strand probe was used in Figure 4-8, so the displacement seen in the denaturing gradient gel was caused by a single C:T mismatch at base pair location 397.

The cell line deleted for HPRT, RJK853, showed no signal on a denaturing gradient gel, as is shown in Figure 4-9. This was expected since analysis of this cell line using a conventional polyacrylamide gel (see Figure 4-6) revealed that no hybridizing exon 3 pseudogene species were present. Also shown in Figure 4-9 is a cell line with 4 X chromosomes, GM1416B. As expected, the cell line with 4 X chromosomes showed a more intense signal compared to the wild type.

In summary, the 169 base pair fragment of exon 3 appeared well-suited for the detection of base pair mutations by DGGE. An HPRT mutant, HPRT_{Munich}, was resolved from the wild type sequence.
4.5 Hybridization of 169 bp probe to *TaqI/MboII*-digested genomic DNA and examination by DGGE: three types of mutants

The previous section showed the resolution of an HPRT mutant, HPRT<sub>Munich</sub>, from the wild type sequence by DGGE; no interference from the four HPRT cross-hybridizing species was observed. In this section and in the following section, MNNG- and ICR-191-treated TG<sup>r</sup> clones will be examined for mutations in HPRT exon 3 by (1) digesting the genomic DNA with *TaqI* and *MboII*, (2) hybridizing the restricted genomic DNA to 169 bp, single-strand exon 3 probe, and (3) resolving the mutant sequence from the wild type sequence by DGGE. It is expected that mutations will be detected in about 120 bp of exon 3, from base pair 300-400 and from base pair 235-255 (see Figure 4-7).

MNNG- and ICR191-induced, independent, TG<sup>r</sup> clones were generated by Monica Cahilly Bretzin as described in Section 3.2. In the case of MNNG, the induced MF was 5.0 x 10<sup>-5</sup> and the untreated MF was 5.6 x 10<sup>-6</sup>; for ICR-191 the induced MF was 5.2 x 10<sup>-4</sup> and the untreated MF was 4.2 x 10<sup>-6</sup>. Thus about 90% of the MNNG-treated TG<sup>r</sup> clones were induced by treatment, while about 99% of the ICR-191-treated clones were induced by treatment. In the following figures, the prefix RM indicates an MNNG-induced clone, and RI indicates an ICR-191-induced clone. The clones were isolated at random, so there was no significance to the numbering scheme.

The autoradiogram in Figure 4-10 shows the results of a denaturing gradient gel examining several MNNG- and ICR-191-induced mutants. Several wild type
bands, several apparent deletions, and one apparent mutant were visible. The apparent mutant, RI205, was later determined to be an artifact caused by HPRT cDNA plasmid contamination of genomic DNA; note that the position of RI205 (lane 9) was identical to the position of the cDNA propagated in a dam$^+$ E. coli strain (lane 7). Nearly all of the double-stranded plasmid bearing the cDNA for HPRT was propagated in a dam$^+$ E. coli strain, and the single dam site in exon 3 destabilized the 169 bp exon 3 fragment as described in Section 4.2. The genomic DNA from RI205 was re-isolated and the newly isolated DNA produced a wild type band as is shown in Figure 4-23. Therefore all mutants that showed a non-wild type band at the gel position of the cDNA propagated in a dam$^+$ E. coli strain were considered to be contamination artifacts; the DNA was re-isolated and re-examined by DGGE.

Figure 4-10 also shows three apparent deletions for exon 3; RM4, RI206 and RI208 showed no signal on the denaturing gradient gel. However, when analyzed by conventional PAGE, it was apparent that these mutants were not deleted for HPRT exon 3. RM4, RI206 and RI208 appeared wild type by PAGE as is shown in Figure 4-11.

The mutants that appeared to be deleted in the denaturing gradient gel shown in Figure 4-10 were examined on another denaturing gradient gel, as is shown in Figure 4-12. Mutant RM4 showed a distinct unfocused smear, while RI206 and RI208 appeared wild type. The denaturing gradient gel that showed these clones to be deleted was run for 9 hours, while the gel that showed RM4 as a smear and RI206 and RI208 as wild type was run for 5 hours.
Figure 4-10: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE (9 hour electrophoresis time): RM3, RM4, RM6, RI204, RI205, RI206, RI207, RI208

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 23 hours to 7.2 x 10^7 copies of single-strand, 169 base pair probe of specific activity 800 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10%–35% denaturing gradient gel and the electrophoresis time was 9 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA which was propagated in both a dam^+ and dam^- E. coli strain.

Lane 1: SS, 1200 cpm ss probe; lane 2: DS, 1200 cpm ds probe; lane 3: RM3; lane 4: RM4; lane 5: RM6; lane 6: RI204; lane 7: DAM^+, 10^3 cpm ss probe hybridized to HPRT cDNA propagated in a dam^+ strain; lane 8: DAM^-, 10^3 cpm ss probe hybridized to HPRT cDNA propagated in a dam^- strain; lane 9: RI205; lane 10: RI206; lane 11: RI207; lane 12: RI208; lane 13: HPRT_Munich (NFC6-169). All RM and RI clones were prepared on NFC4-101.

RM4, RI206 and RI208 appear deleted for exon 3. RI205 appears contaminated with plasmid DNA. RM3, RM6, RI204 and RI207 appear wild type.

This experiment is in NFC7-184, gel 1.
DENATURED

20%  25%

13 MUNICH
12 RI208
11 RI207
10 RI206
9  RI205
8  DAM^-
7  DAM^+
6  RI204
5  RM6
4  RM4
3  RM3
2  DS
1  SS
Figure 4-11: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by PAGE: wild type, RM4, RI206, RI208

Thirty μg of TaqI/MboII-digested genomic DNA was hybridized for 47 hours to 6.8 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 1178 Ci/mmol. Fifteen μg of genomic DNA was loaded on a conventional 8% polyacrylamide, 1.5 mm gel and electrophoresis time was 3 hours at 150 V.

Lane 1: WT (THP1, NFC8-169); Lane 2: RM4; Lane 3: RI206; Lane 4: RI208. The DNA for the RM and RI mutants was isolated in NFC8-69. This experiment is in NFC8-92.
Figure 4-12: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE (5 hour electrophoresis): wild type, RM4, RI206, RI208

Thirty μg of TaqI/MboII-digested genomic DNA was hybridized for 47 hours to 6.7 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 1178 Ci/mmol. Fifteen μg of genomic DNA was loaded on a 10%–35% denaturing gradient gel and the electrophoresis time was 5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ E. coli strain.

Lane 1: SS, 1000 cpm ss probe; Lane 2: DAM⁺, 1000 cpm of hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; Lane 3: MUNICH, HPRTₘᵤₜₜₛ (NFC6-169); Lane 4: WT, wild type (THP1, NFC6-169); Lane 5: RM4; Lane 6: RI206; Lane 7: RI208. The DNA for the mutants was prepared in NFC8-69.

RM4 showed a unfocused, non-wild type band. RI206 and RI208 appeared wild type.

This experiment is in NFC8-101.
What could cause this behavior? The stability of the partially melted form of mutant RM4 has been decreased, the DNA apparently made the transition from the partially melted form to complete strand dissociation at an increased rate compared to the wild type. As was shown in Figure 4-4, the melted form of the 169 bp exon 3 fragment was not indefinitely stable; increasing the electrophoretic time after the molecule melted resulted in decreased signal intensity, presumably because some fraction of the melted form was undergoing complete strand dissociation. The apparent deletions seen using DGGE with an extended electrophoretic time are consistent with a mutation in the high-temperature melting domain, a mutation in the high-temperature melting domain could reduce the stability of the partially melted form thus increasing the rate of strand dissociation. Later sequence analysis confirmed that these mutants did in fact have a mutation in the high-temperature melting domain.

Thus four classifications of the TG" mutants can be made: (1) wild type for exon 3 (2) deletion for exon 3 (for example, RJK853) (3) focused non-wild type band (for example, HPRT_{Munich}) (4) unfocused non-wild type band (for example, RM4). Deletions for exon 3 were confirmed by Southern blotting. However, the classification of unfocused non-wild type mutants at this point in the thesis was quite subjective; it was often difficult to distinguish between the wild type and a mutant with an unfocused, smeared band. Subsequent sections will clarify this issue.

Focused DGGE mutants were assumed to have a base pair substitution in the low-temperature melting domain, unfocused DGGE mutants were assumed to have
a mutation in the high-temperature melting domain.
4.6 Hybridization of 169 bp probe to TaqI/MboII-digested genomic DNA and examination by DGGE: MNNG- and ICR-191-induced mutants

Section 4.5 showed that three types of mutants can be seen using DGGE: focused DGGE mutant, non-focused DGGE mutant, deletion mutant. In this section, 32 MNNG- and 36 ICR-191-induced TG^f clones will be examined for mutations in exon 3. All apparent exon 3 deletions were confirmed by Southern blotting. All mutants showing a non-wild type band at the position of the dam^+ species were considered artifacts of plasmid contamination; the DNA was re-isolated and re-examined by DGGE.

Figures 4-13, 4-14, 4-15, 4-16, 4-17, 4-18, 4-20, 4-21, 4-22, 4-23 and 4-24 show the analysis of MNNG- and ICR-191-induced TG^f clones by (1) restriction enzyme digestion of genomic DNA with TaqI and MboII, (2) hybridization of a 169 bp single-strand exon 3 probe to genomic DNA, and (3) examination by DGGE. The text continues on page 185.
Figure 4-13: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM1, RM2

Twenty µg of TaqI/MboII-digested genomic DNA was hybridized for 23 hours to 7.2 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 800 Ci/mm mol. Twenty µg of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 9 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ E. coli strain.

Lane 1: SS, 1200 cpm ss probe; lane 3: DS, 1200 cpm ds probe; lane 5: MUNICH, HPRTMunich (NFC6-169); lane 7: WT, wild type (THP1, NFC6-169); lane 9: DAM⁺, 10³ cpm ss probe hybridized to HPRT cDNA propagated in a dam⁺ strain; lane 14: RM1; lane 15: RM2. The remainder of the lanes contain other samples. The mutant DNA was isolated in NFC4-101. This experiment is in NFC7-184, gel 2.

RM1 and RM2 appear wild type.
DENATURANT

25%  20%

1  SS
2  DS
3  MUNCH
4  WLT
5  7
6  8
7  DAM
8  9
9  10
10 11
11 12
12 13
13 14 RM1
14 15 RM2

7-184, gel 2
Figure 4-14: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM6, RM8, RM9

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 35 hours to 4.5 x 10^7 copies of single-strand, 169 base pair probe of specific activity 1500 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 5.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam^+ E. coli strain.

Lane 1: WT, wild type (THP1, NFC6-169); lane 2: DAM^+, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam^+ E. coli strain; lane 3: MUNICH, HPRT_Munich (NFC6-169); lanes 4-6: other samples; lane 7: 0, ss probe hybridization with no target DNA; lane 8: RM6, lane 9: RM8; lane 10: RM9; lane 11: DS, double-strand probe. The DNA for the mutants was prepared in NFC4-101.

RM6, RM8 and RM9 appear wild type.

This experiment is in NFC8-92, gel 1.
Figure 4-15: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM11, RM12, RM13, RM14, RM15, RM16, RM18, RM19

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 35 hours to 4.5 x 10^7 copies of single-strand, 169 base pair probe of specific activity 1500 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10%—35% denaturing gradient gel and the electrophoresis time was 5.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam^+ E. coli strain.

Lane 1: DS, double-strand probe; lane 2: DAM^+, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam^+ E. coli strain; lane 3: MUNICH, HPRT_Munich (NFC6-169); lane 4: RM11; lane 5: RM12; lane 6: RM13; lane 7: RM14; lane 8: RM15; lane 9: RM16; lane 10: RM18; lane 11: RM19; lane 12: SS, single-strand probe. The DNA for mutants RM11 and RM12 was prepared in NFC4-101, DNA for the remainder of the mutants was prepared in NFC8-69.

RM16 and RM19 appear to yield an unfocused, non-wild type band. RM12 appears contaminated with plasmid DNA. RM18 appears deleted for exon 3. RM11, RM13, RM14, and RM15 appear wild type.

This experiment is in NFC8-92, gel 2.
Figure 4-16: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM20, RM21, RI200, RI202, RI203, RI209, RI210, RI213, RI214, RI215

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 35 hours to 4.5 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 1500 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 5.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ E. coli strain.

Lane 1: DS, double-strand probe; lane 2: DAM⁺, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 3: MUNICH, HPRT_Munich (NFC6-169); lane 4: RM20; lane 5: RM21; lane 6: RI200; lane 7: RI202; lane 8: RI203; lane 9: RI209; lane 10: RI210; lane 11: RI213; lane 12: RI214; lane 13: RI215; lane 14: SS, single-strand probe. The DNA for the mutants was prepared in NFC8-69.

RI209 shows a focused, non-wild type band. RM20, RI202 and RI213 appear to yield an unfocused, non-wild type band. RM21 and RI200 appear to be deleted for exon 3. RI203, RI210, RI214 and RI215 appear wild type.

This experiment is in NFC8-92, gel 3.
Figure 4-17: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM10, RI216, RI217, RI219, RI220, RI221

Twenty $\mu$g of TaqI/MboII-digested genomic DNA was hybridized for 35 hours to 4.5 x $10^7$ copies of single-strand, 169 base pair probe of specific activity 1500 Ci/mmol. Twenty $\mu$g of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 5.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a $dam^+$ E. coli strain.

Lane 1: DS, double-strand probe; lane 2: DAM$, ss$ probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a $dam^+$ E. coli strain; lane 3: MUNICH, HPRT$_{Munich}$ (NFC6-169); lane 4: RI216; lane 5: RI217; lane 6: RI219; lane 7: RI220; lane 8: RI221; lane 9: SS, single-strand probe, lane 10: RM10. The DNA for mutant RM10 was prepared in NFC4-101, DNA for the remainder of the mutants was prepared in NFC8-69.

RI220 appears to yield an unfocused, non-wild type band. RI219 appears contaminated with plasmid DNA. RI216, RI217, RI221 and RM20 appear wild type.

This experiment is in NFC8-92, gel 4.
Figure 4-18: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM23, RM24, RM25, RM26, RI228, RI229, RI230, RI231

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 62 hours to 4.5 \times 10^7 copies of single-strand, 169 base pair probe of specific activity 1234 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 4.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ E. coli strain.

Lane 1: SS, single-strand probe; lane 2: DS, double-strand probe; lane 3: DAM⁺, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 4: WT, wild type, (THP1, NFC6-169); lane 5: MUNICH, HPRT_{Munich} (NFC6-169); lane 6: RI228; lane 7: RI229; lane 8: RI230; lane 9: RI231; lane 10: RM23; lane 11: RM24; lane 12: RM25; lane 13: RM26.

RI229 appears deleted for exon 3. RI228, RI230, RI231, RM23, RM24, RM25 and RM26 appear wild type.

The DNA for the mutants was prepared in NFC8-206. This experiment is in NFC8-214, gel 2.
Figure 4-19: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM28, RM30, RM31, RM32, RM33

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 62 hours to 4.5 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 1234 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10% – 35% denaturing gradient gel and the electrophoresis time was 4.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ E. coli strain.

Lane 1: SS, single-strand probe; lane 2: DS, double-strand probe; lane 3: DAM⁺, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 4: WT, wild type, (THF1, NFC6-169); lane 5: MUNCH, HPRTMunich (NFC6-169); lane 6: RM28; lane 7: RM30; lane 8: RM31; lane 9: RM32; lane 10: RM33.

RM30 and RM31 appear contaminated with plasmid DNA. RM28, RM32 and RM33 appear wild type.

The DNA for the mutants was prepared in NFC8-206. This experiment is in NFC8-214, gel 3.
**Figure 4-20:** Hybridization of 169 base pair probe to *TaqI/MboII*-digested genomic DNA and examination by DGGE: RM34, RM35, RM37, RM38, RM39, RI242

Twenty µg of *TaqI/MboII*-digested genomic DNA was hybridized for 32 hours to 6.0 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 822 Ci/mmol. Twenty µg of genomic DNA was loaded on a 10%–35% denaturing gradient gel and the electrophoresis time was 4.0 hours. About 1000 cpm of ss probe was hybridized to a molar excess of *TaqI/MboII*-digested HPRT cDNA that was propagated in a dam⁺ *E. coli* strain.

The probe preparation was contaminated with some complementary exon 3 DNA. HPRT<sub>Munich</sub> in lane 4 shows two bands, one at the wild type position which is contamination, and another band which is the HPRT<sub>Munich</sub> wild type mismatch. RM 34 in lane 6 also shows two bands and was considered a focused DGGE mutant. Contamination of the probe preparation made it difficult to determine if a deletion had occurred since there was always a band at the wild type position.

Lane 1: SS, single-strand probe; lane 2: DS, double-strand probe; lane 3: DAM⁺, ss probe hybridized to a molar excess of *TaqI/MboII*-digested HPRT cDNA propagated in a dam⁺ *E. coli* strain; lane 4: MUNICH, HPRT<sub>Munich</sub> (NFC9-92); lane 5: RI242; lane 6: RM34; lane 7: RM35; lane 8: RM37; lane 9: RM38, lane 10: RM39.

RM34 yielded a focused, non-wild type band. RI242, RM35, RM37, RM38 and RM39 appear wild type.

The DNA for the mutants was prepared in NFC9-245. This experiment is in NFC9-252, gel 2.
NFC9-252, gel 2

10 RM39
9 RM38
8 RM37
7 RM35
6 RM34
5 RI242
4 MUNICH
3 DAM+
2 DS
1 SS

DENATURANT

20%
25%
**Figure 4-21:** Hybridization of 169 base pair probe to *Taql/MboII*-digested genomic DNA and examination by DGGE: RI232, RI233, RI234, RI235, RI237, RI238, RI239, RI240, RI241

Twenty µg of *Taql/MboII*-digested genomic DNA was hybridized for 32 hours to 6.0 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 822 Ci/mmol. Twenty µg of genomic DNA was loaded on a 10% - 35% denaturing gradient gel and the electrophoresis time was 4.0 hours. About 1000 cpm of ss probe was hybridized to a molar excess of *Taql/MboII*-digested HPRT cDNA that was propagated in a *dam*⁺ *E. coli* strain.

The probe preparation was contaminated with some complementary exon 3 DNA as explained in the legend to Figure 4-20.


RI232, RI233, RI234, RI235, RI237, RI238, RI239, RI240 and RI241 appear wild type.

The DNA for the mutants was prepared in NFC9-245. This experiment is in NFC9-252, gel 1.
NFC9-252, gel 1

<table>
<thead>
<tr>
<th>20%</th>
<th>25%</th>
</tr>
</thead>
</table>

1 SS
2 DS
3 DAM+
4 MUNICH
5 RI232
6 RI233
7 RI234
8 RI235
9 RI237
10 RI238
11 RI239
12 RI240
13 RI241

DE Nugaturant
Figure 4-22: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and examination by DGGE: RM16, RM26, RI222, RI226, RI227. Re-examination of samples showing prior possible plasmid contamination RM12, RM30, RM31

Twenty µg of TaqI/MboII-digested genomic DNA was hybridized for 38 hours to 6.0 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 907 Ci/mmol. Twenty µg of genomic DNA was loaded on a 10% - 35% denaturing gradient gel and the electrophoresis time was 4.0 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ and dam⁻ E. coli strain. The probe preparation was contaminated with some complementary exon 3 DNA as explained in the legend to Figure 4-20.

Mutants RM12, RM30 and RM31 showed a band at the position of the dam⁺ methylated species in Figures 4-15 and 4-19 that could have been caused by HPRT cDNA plasmid contamination. DNA was re-isolated from these mutants and no band at the dam⁺ position was visible, all mutants appeared wild type for exon 3 in this gel.

Lane 1: SS, single-strand probe; lane 2: DS, double-strand probe; lane 3: DAM⁻, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 4: DAM⁺, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 5: MUNICH, HPRT_Munich (NFC9-92); lane 6: RI226; lane 7: RI227; lane 8: RM12; lane 9: RM16, lane 10: RM26; lane 11: RM30; lane 12: RM31; lane 13: RI222.

Legend continued on the next page.
RM12, RM16, RM26, RM30, RM31, RI222, RI226 and RI227 appear wild type. RM16 appeared to give an unfocused, non-wild type band in the gel shown in Figure 4-15, however, RM16 appears wild type in this gel. RM16 showed an unfocused, non-wild type band with an electrophoresis time of 5.5 hours (Figure 4-15), while it showed a wild type band with an electrophoresis time of 4 hours (this gel).

The DNA for mutants RM12, RM30, RM31, RM26 prepared in NFC9-63. The DNA for mutants RI222, RI226 and RI227 was prepared in NFC8-206. This experiment is in NFC9-238, gel 2.
Figure 4-23: Hybridization of 169 base pair probe to TaqI/MboII-digested genomic DNA and re-examination of samples showing prior possible plasmid contamination RI205, RI219.

Twenty μg of TaqI/MboII-digested genomic DNA was hybridized for 38 hours to 6.0 x 10⁷ copies of single-strand, 169 base pair probe of specific activity 907 Ci/mmol. Twenty μg of genomic DNA was loaded on a 10% - 35% denaturing gradient gel and the electrophoresis time was 4.0 hours. About 1000 cpm of ss probe was hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA that was propagated in a dam⁺ and dam⁻ E. coli strain.

The probe preparation was contaminated with some complementary exon 3 DNA as explained in the legend to Figure 4-20.

Mutants RI205 and RI219 showed a band at the position of the dam⁺ methylated species in Figures 4-10 and 4-17 that could have been caused by HPRT cDNA plasmid contamination. DNA was re-isolated from these mutants and no band at the dam⁺ position was visible, all mutants appeared wild type for exon 3 on this gel.

Lane 1: SS, single-strand probe; lane 2: DAM⁻, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 3: DAM⁺, ss probe hybridized to a molar excess of TaqI/MboII-digested HPRT cDNA propagated in an dam⁺ E. coli strain; lane 6: MUNICH, HPRTMunich (NFC9-92); lane 9: WT, wild type (THP1, NFC8-169); lane 10: RI205; lane 12: RI219. Lanes 4, 5, 7, 8, 11: other samples.

RI205 and RI219 appear wild type.

The DNA for mutants RI205 and RI219 was prepared in NFC9-63. This experiment is in NFC9-238, gel 1.
Figure 4-24: Hybridization of 169 base pair probe to Taql/MboII-digested genomic DNA and examination by DGGE: RI224, RI225.

Twenty μg of Taql/MboII-digested genomic DNA was hybridized for 38 hours to 6.0 x 10^7 copies of single-strand, 169 base pair probe of specific activity 907 Ci/mmol. Twenty μg of genomic DNA was loaded on a 5% – 30% denaturing gradient gel and the electrophoresis time was 5.5 hours. About 1000 cpm of ss probe was hybridized to a molar excess of Taql/MboII-digested that was propagated in a dam⁺ and dam⁻ E. coli strain.

The probe preparation was contaminated with some complementary exon 3 DNA as explained in the legend to Figure 4-20.

Lane 1: SS, single-strand probe; lane 2: DAM⁻, ss probe hybridized to a molar excess of Taql/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 3: DAM⁺, ss probe hybridized to a molar excess of Taql/MboII-digested HPRT cDNA propagated in a dam⁺ E. coli strain; lane 6: MUNICH, HPRT_Munich (NFC9-92); lane 9: WT, wild type (THP1, NFC8-169); lane 10: RI224; lane 12: RI225. Lanes 4, 5, 7, 8: other samples.

RI224 and RI225 appear wild type.

The DNA for mutants RI205 and RI219 was prepared in NFC9-63. This experiment is in NFC9-238, gel 3.
NFC9-238, gel 3

11 RI225
10 RI224
9 WT
8
7
6 MUNICH
5
4
3 DAM+
2 DAM-
1 SS

DENATURANT
The data in figures 4-15, 4-16, and 4-18 suggested that RM18, RM21, RI200 and RI229 were deleted for exon 3. A Southern blot was performed to confirm the deletions, this is shown in Figure 4-25. RM18, RM21 and RI200 were deleted for the entire HPRT coding sequence. RI229 appeared wild type on the Southern blot but appeared deleted in a denaturing gradient gel. The genomic DNA concentrations of the mutants were checked before the Southern blot was performed and it was found that the DNA concentration of RI229 was about 25% of the recorded concentration (NFC9-55). Thus only about 5 µg of RI229 genomic DNA was used for the denaturing gradient gel shown in Figure 4-18 which made this mutant appear deleted; 20 µg of genomic DNA were typically used for DGGE analysis. No information was obtained about this mutant.

The DNA was re-isolated from samples which showed a band at the position of the dam+ exon 3 species and these mutants were re-examined by DGGE. It was assumed the band at the dam+ position was HPRT cDNA plasmid contamination. Re-isolated DNA from samples RM12, RM30, RM31, RI205 and RI219 appeared wild type as shown in Figures 4-23 and 4-22, suggesting that the bands seen in previous gels were plasmid artifacts.

Summarizing this data, we have:

**MNNG mutagenesis**


- **FOCUSED DGGE MUTANT**: RM34
• **UNFOCUSED DGGE MUTANT**: RM4, RM16, RM19, RM20,

• **DELETION**: RM18, RM21

**ICR-191 mutagenesis**

• **WILD TYPE FOR EXON 3**: RI203, RI204, RI205, RI206, RI207, RI208, RI210, RI214, RI215, RI216, RI217, RI219, RI221, RI222, RI224, RI225, RI226, RI227, RI228, RI230, RI231, RI232, RI233, RI234, RI235, RI237, RI238, RI239, RI240, RI241, RI242

• **FOCUSED DGGE MUTANT**: RI209,

• **UNFOCUSED DGGE MUTANT**: RI202, RI213, RI220

• **DELETION**: RI200

A summary of the data is given in Table 4-II. One focused DGGE mutant was observed for both the MNNG- and ICR-191-treated mutants, thus mutations in the low-temperature melting domain occurred at a frequency of about 3% in the collection of MNNG- and ICR-191-treated TG^f^ clones. 4 MNNG- and 3 ICR-191-induced clones produced non-focused DGGE bands, suggesting that mutations in the high-temperature melting domain were occurring more frequently than mutations in the low-temperature melting domain.
Figure 4-25: Southern blot of RM21, RM18, RI229, RI200

Ten µg of genomic DNA was digested with 60 U of PstI (New England Biolabs; Beverly, MA) for 5 hours at a concentration of 100 ng/µl in 100 mM NaCl, 50 mM Tris (pH 7.5), 10 mM DTT. The DNA was precipitated with ethanol, resuspended in water and loaded onto a 0.8% agarose gel. The gel was run for 15 hours at 30 V and the DNA was transferred to GeneScreen Plus (New England Nuclear; Boston, MA) according to manufacturer’s instructions. The 1 kb PstI fragment of pHPT30 which contained the entire HPRT cDNA was labeled to a specific activity of about 10^8 cpm/µg using the reagents and following the instructions provided in the IBI random priming kit (International Biotechnologies, Inc.; New Haven, CT). About 2 x 10^6 cpm of radiolabeled probe were added and the hybridization time was 19 hours at 65°C. Hybridization conditions were as described in Maniatis et al. (1982). The membrane was washed (1) twice for 5 minutes at room with 2X SSC, (2) twice for 30 minutes at 65°C with 2X SSC, 1% SDS, and (3) twice for 30 minutes at room temperature with 0.2X SSC. The gel was exposed to Kodak XAR 5 film for 2 days at -70°C with an intensifying screen.

Lane 1; WT, wild type, TK6; lane 2; DELETION, RJK853, total deletion for HPRT; lane 3; RM21; lane 4; RM18; lane 5; RI229; lane 6; RI200.
Table 4-II: Summary showing the results of examining MNNG- and ICR-191-induced TG\textsuperscript{r} clones using a 169 bp exon 3 probe

A focused mutant produced a sharp non-wild type band on a denaturing gradient gel. An unfocused mutant produced a diffuse smear below the wild type position. The 169 bp probe was expected to resolve mutations occurring in the low-temperature melting domain of exon 3.
<table>
<thead>
<tr>
<th>Classification</th>
<th>MNNG</th>
<th>ICR-191</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Unfocused</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Deletion</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Wild type</td>
<td>25</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>36</td>
</tr>
</tbody>
</table>
4.7 Modification of the calculated melting map of HPRT exon 3 by
addition of a GC-rich sequence

The previous section showed that the 169 bp *TaqI/MboII* fragment of HPRT
exon 3 can be used to examine MNNG- and ICR-191-induced TG\(^r\) clones for
mutations using DGGE. About 3% of the MNNG- and ICR-191-induced clones
had a mutation in the low-temperature melting domain of exon 3. However, the
data suggested that mutations may be occurring in the high-temperature melting
domain of the molecule. 10%-15% of the clones showed a non-wild type, smeared
DGGE pattern which could be caused by a mutation in the high-temperature
melting domain which destabilized the melted form of the molecule. In an effort to
detect mutations throughout the entire coding region of exon 3, a GC-rich sequence
was added to exon 3 and the alteration in melting profile was modeled.

A GC-rich sequence was added to exon 3 by using 74 base pair, GC-rich,
oligonucleotide primers in the polymerase chain reaction. The GC-rich
oligonucleotide primers contained 54 bp of a GC-rich sequence at the 5' and 20 bp
at the 3' which were complementary to the intron sequence immediately flanking
exon 3. During the first amplification cycle, the GC-rich portion of the
oligonucleotide primer was not base paired with the template DNA of exon 3; this
did not interfere with the amplification process. During subsequent cycles, the GC-
rich portion of the oligonucleotide was copied by the DNA polymerase, resulting in
the addition of a 54 bp, double-stranded, GC-rich sequence to the 5' and 3' ends of
exon 3.
PCR was used for two separate amplifications, (1) amplification of exon 3 from genomic DNA which produced a 224 bp product, and (2) re-amplification of the 224 bp product with GC-rich, oligonucleotide primers which produced a 332 bp product with 54 bp GC-rich sequences at both the 5' and 3' ends.

The calculated melting map of the 224 bp exon 3 sequence produced by PCR-amplification from genomic DNA using 20 bp intron primers is given in Figure 4-26. This molecule contains the entire 184 bp coding region of exon 3 and 20 bp of both 5' and 3' flanking sequence. The calculated melting map of the 224 bp fragment was essentially the same as the calculated melting map of the 169 bp TaqI/MboII exon 3 fragment (See Figure 4-2) in that the low-temperature melting domain of the molecule was from about base pair position 300 to the 3' end of the molecule. Therefore, use of the 224 bp exon 3 PCR product for DGGE analysis was not expected to reveal new mutations which were not detected by the 169 bp exon 3 probe. Any mutation occurring in the 20 bp intron sequence immediately flanking exon 3 will not be detected because the mutation will be converted to the wild type sequence by the PCR process. The DNA sequence of the 224 bp molecule produced by PCR using the 20 bp intron primers described on page 106 is:

TCCTGATTTT ATTTCTGTAG GACTGAACGT CTTGCTCGAG ATGTGATGAA
GGAGATGGGA GGCCATCACA TTGTAGCCCT CTGTGTGCTC AAGGGGGGCT
ATAAATTCTT TGCTGACCTG CTGGATTACA TCAAGCACT GAATAGAAAT
AGTGATAGAT CCATTCCTAT GACTGATAG TTTATCAGAC TGAAGAGCTA
TTGTGTGAGT ATATTTAATA TATG

A second PCR-amplification using the 224 bp sequence as the template and
using the 74 bp GC-rich primers described on page 108 produced a 332 bp fragment. The 332 bp product contained the entire 184 bp exon 3 coding region, 20 bp of both 5' and 3' intron sequence, and a 54 bp GC-clamp at both the 5' and 3' ends. The 332 bp molecule contained a unique HgiAI site at base pair position 288 (position 288 refers to the cDNA numbering sequence) as shown in Figure 4-26. The sequence of the 332 bp fragment is:

```
GCCGCCTGCA GCCCGCCGCC CCCGTGCCCC CGCCCCCCCGG CCGGCCCGGG
CGCCTCTGTA TTTATTTTCT GTAGGACTGA ACGTCTTGCT CGAGATGTGA
TGAAGGAGAT GGGAGGGCAT CACATTTGAT CCCTCTGTGT GCTCAAGGGG
GGCTATAAAAT TCTTTGCTGA CCTGCTGGAT TACATCAAAG CACTGAATAG
AAATAGTGAT AGATCCATTC CTATGACTGT AGATTTTATC AGACTGAAGA
GCTATTGTGT GAGTATATTT AATATATGGA CGCAGGGGCC GCCGCGCGGG
CGGCCGCAACG GGCGGGCGCG GCTGCAGGCG GC
```

The calculated melting maps of the two fragments produced by restricting the 332 bp product with HgiAI are given in Figure 4-26. The 54 bp GC-clamp was now the high-temperature melting domain of both molecules, and the coding portion of exon 3 was present in the low-temperature melting domains of the hybrid molecules. The $T_m$ of the low-temperature melting domain of the hybrid molecule with a GC-clamp at the 5' end was 72\(^\circ\) C, while the molecule with a GC-clamp at the 3' end was predicted to melt at about 66\(^\circ\) C. Using a value of 1% denaturant = 0.313\(^\circ\) C (Lerman et al., 1984), the molecule with the GC-clamp at the 5' end was expected to melt at about 38% denaturant, and the molecule with the GC-clamp at the 3' end should melt at about 19% denaturant.

The effect of a base pair mismatch can be modeled using a program called SQHT (Myers et al., 1987), the fortran code for SQHT is given in the Appendix.
THIS COPY MAY NOT BE FURTHER REPRODUCED OR DISTRIBUTED IN ANY WAY WITHOUT SPECIFIC AUTHORIZATION IN EACH INSTANCE, PROCURED THROUGH THE DIRECTOR OF LIBRARIES, MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

24 : 1
Figure 4-26: Modification of the calculated melting map of HPRT exon 3 by addition of a GC-rich sequence

The unmodified melting map of the 224 bp fragment containing the 184 bp coding region of exon 3 and 20 bp of both 5' and 3' sequence is shown with a solid line. The high-temperature melting domain of the 224 bp molecule is from the 5' end to about base pair position 300, mutations in the high-temperature melting domain are not expected to be resolved by DGGE. The dotted lines represent the calculated melting maps produced by (1) addition of a 54 bp GC-rich sequence to the 5' and 3' ends of the 224 bp fragment, and (2) cutting the 332 bp molecule at the unique HgiA1 site. The entire coding region of exon 3 is now in the low-temperature melting domains of the two hybrid molecules.
Modification of melting map of HPRT exon 3
by addition of GC-rich sequences

[HglAI site]

bp position

- 224 bp pcr product
- GC-clamp added
This program simulates a mismatch at each base pair location of a given sequence, and gives the expected difference in gradient level of the mismatched heteroduplex and the wild type homoduplex. In this way, the portion of the molecule expected to be sensitive to base pair substitutions can be discerned.

Computer simulation of a mismatch at each every base pair position of the fragments with a GC-clamp is given in Figure 4-27 and 4-28; hybridization of a radiolabeled wild type probe to mutant DNA would ensure the formation of a mismatch. The calculated difference in gradient level caused by a mismatch at any exon 3 base pair location was 0.4-0.8°C. This corresponds to a minimum gel shift of 7-8 mm using standard gradient conditions (gradient difference of 30% in a 17 cm gel, 1% denaturant = 0.313°C). The coding region of exon 3 was present in the low-temperature melting domain of the hybrid molecules, and thus all base pair substitutions in the coding region of exon 3 are predicted to be detecting using the GC-clamp.

The behavior of the HgiAl-cut, 332 bp fragment with the GC-clamps on a parallel denaturing gradient gel is shown in Figure 4-29. The experiment showed that the two HgiA1 fragments melted at about 23% and 33% denaturant. This result is in basic agreement with the calculated melting maps of these fragments, which predicted that the molecules would melt at about 19% and 38% denaturant.

The molecule which melted at 23% denaturant contained the exon 3 coding region from base pair location 288-403, this molecule contained about 12 base pairs of the original high-temperature melting domain. The molecule which melted at
Figure 4-27: Simulation of a mismatch at each base pair position of the HgiA1 fragment containing exon 3 and a GC-clamp at the 5' end

The wild type homoduplex melting map is shown as a solid line (Y axis), and the difference in gradient position in degrees Centigrade between the wild type homoduplex and a mismatched heteroduplex is shown as a dotted line (Y2 axis). The coding region of HPRT exon 3 is from base pair position 220-403. A difference of 0.1° C should produce a 1-2 mm shift using typical denaturing gradient gel conditions of a 30% change in gradient over a 17 cm gel, and 1% denaturant = 0.313° C.

The program SQHT was run with the following values: destabilization = 50° C, retardation length = 75, velocity zero = 2. The value for the heteroduplex is for a 6 hour electrophoresis time.
Calculated melting map of wild type homoduplex and destabilization produced by a mismatch in the HgiAI fragment with a GC-clamp at the 5' end.
**Figure 4-28:** Simulation of a mismatch at each base pair position of the HgiAI fragment containing exon 3 and a GC-clamp at the 3' end

The wild type homoduplex melting map is shown as a solid line (Y axis), and the difference in gradient position in degrees Centigrade between the wild type homoduplex and a mismatched heteroduplex is shown as a dotted line (Y2 axis). The coding region of HPRT exon 3 is from base pair position 220-403. A difference of $0.1^\circ C$ should produce a 1-2 mm shift using typical denaturing gradient gel conditions of a 30% change in gradient over a 17 cm gel, and 1% denaturant $= 0.313^\circ C$.

The program SQHT was run with the following values: destabilization $= 50^\circ C$, retardation length $= 75$, velocity zero $= 2$. The value for the heteroduplex is for a 6 hour electrophoresis time.
Calculated melting map of wild type homoduplex and destabilization produced by a mismatch in the HgiAI fragment with a GC-clamp at the 3' end.
Figure 4-29: Autoradiogram of electrophoretic time course showing the melting behavior of the two HgiAI fragments containing HPRT exon 3 with an added GC-clamp.

The unmethylated DNA is labeled $U$ and the methylated DNA is labeled $M$. The electrophoretic time is shown in hours. The unique $dam$ site at base pair location 357 was methylated in vitro with $dam$ methylase and the unique $TaqI$ methylation site at base pair 234 was methylated in vitro with $TaqI$ methylase. Both $HgiAI$ fragments have a GC-clamp.

The solid boxes show the fragment which melts at about 23% denaturant. This molecule has a GC-clamp at the 3' end and contains the exon 3 sequence from the $HgiAI$ site at base pair location 288 to the 3' end of the coding region. The dashed boxes show the fragment which melts at about 33% denaturant. This molecule has a GC-clamp at the 5' end and contains the exon 3 sequence from the $HgiAI$ site at base pair location 288 to the 5' end of the coding region.

The 332 bp fragment was produced by PCR-amplification of plasmid pλ3001 with both GC-rich primers as described in Section 3.7. The 332 bp fragment was not gel-purified, resulting in other amplification products which were visible on this gel. The DNA from the PCR-amplification was restricted with $HgiAI$, methylated with both $dam$ and $TaqI$ methylase following the manufacturer's protocols (New England Biolabs; Beverly, MA). About 1000 cpm per lane was loaded onto a 12.5% polyacrylamide, 0-50% denaturing gradient gel.

This experiment is in NFC11-21.
about 33% denaturant contained the exon 3 coding region from base pair 220-287, this molecule contained the remainder of the original high-temperature melting domain. Together, the two molecules contained the entire coding region of exon 3, and since the molecules melted at different concentrations of denaturant, it was possible to examine the entire coding region using a single denaturing gradient gel.

Also shown in Figure 4-29 are methylated and unmethylated *HgiAI* fragments. As shown in Section 4.2, methylation at the N\(^6\) position of adenine will decrease the stability of the homoduplex, and methylated DNA served as a positive control in this experiment. Figure 4-29 shows that methylated DNA was separated from unmethylated DNA for both of the *HgiAI* fragments.

In summary, use of the GC-clamps altered the melting characteristics of exon 3 such that all base pair substitutions in the coding region were predicted to be detected. The entire coding region was examined on a single denaturing gradient gel, and DNA methylated *in vitro* was separated from unmethylated DNA.
4.8 Examination of MNNG- and ICR-191-induced TG<sup>f</sup> clones by DGGE using the GC-clamp

32 MNNG- and 36 ICR-191-induced clones were examined with a 169 bp, single strand, exon 3 probe as discussed in Section 4.6. The 169 bp probe clearly resolved two, well-focused mutations that putatively occurred in the low-temperature melting domain of exon 3. Six unfocused, non-wild type bands were also observed, which suggested that mutations may be occurring in the high-temperature melting domain of exon 3.

Section 4.7 showed that all base pair substitutions in exon 3 should be detected using the GC-clamp. In this section, most of the MNNG- and ICR-191-treated TG<sup>f</sup> clones previously examined in Section 4.6 with the 169 bp probe will be re-examined using the GC-clamp.

Exon 3 was amplified from genomic DNA by PCR using two 20 bp primers complementary to the intron sequence immediately flanking HPRT exon 3. The first PCR-amplification produced a 224 bp DNA fragment which contained the 184 bp exon 3 coding region and 20 bp of both 5′ and 3′ intron sequence. 32 cycles of amplification using the thermostable Taq DNA polymerase produced a major band of size 224 bp. Other higher molecular weight species were often also observed. The results of a typical amplification of exon 3 from genomic DNA is shown in Figure 4-30.

The concentration of the 224 bp exon 3 fragment in the reaction mixture after 32 cycles of PCR-amplification using genomic DNA was typically 10<sup>10</sup>
Figure 4-30: Ethidium bromide-stained polyacrylamide gel showing the PCR-amplification of HPRT exon 3 from genomic DNA.

PCR-amplification was performed using 20 bp intron primers as described in Section 3.6. The arrow shows the 224 bp DNA fragment. 10% of the reaction mixture was loaded onto an 8% polyacrylamide gel. A typical yield of the 224 bp exon 3 fragment was 100-500 ng.

Leftmost lane: 200 ng molecular weight marker pBR322 digested with MspI; others lanes: one-tenth of the PCR amplification produced after 32 cycles of amplification using 1 μg of human genomic DNA.

This experiment is in NFC10-161.
copies/μl. 2 μl of this PCR-amplification mixture was used without further purification as a template for another PCR-amplification which added the 54 bp GC-clamps to both ends of the 224 bp molecule, producing a 332 bp molecule. An ethidium bromide-stained gel showing the 332 bp product is given in Figure 4-31. The 332 bp product was loaded onto a polyacrylamide gel, electroeluted from the gel, restricted with HgiAI, hybridized to wild type radiolabeled probe, and loaded onto a denaturing gradient gel.

Thirty two MNNG-induced mutants were originally examined with the 169 bp probe. One focused putative low-temperature melting domain mutant, 4 unfocused putative high-temperature melting domain mutants, 2 total gene deletions, and 25 wild type clones were found. Thirty of the MNNG-induced mutants were re-examined by DGGE with the GC-clamp; the two deletion mutants were not re-examined.

The single MNNG-induced mutant which showed a focused band and the 4 mutants which showed an unfocused band when using the 169 bp probe all showed a mutation when using the GC-clamp. All of the 24 MNNG-induced which showed a wild type pattern using the 169 bp probe showed a wild type pattern when examined with the GC-clamp.

The 36 ICR-191 induced mutants that were examined using the 169 bp probe showed 1 focused mutant, 3 unfocused mutants, 1 total gene deletion, and 31 mutants that were wild type for exon 3. Twenty eight ICR-191-induced mutants were re-examined using the GC-clamp, the deletion mutant and mutants RI229,
Figure 4-31: Ethidium bromide-stained polyacrylamide gel showing PCR-amplification product with added GC-clamp

Two μl of the unpurified amplification mixture which produced the 224 bp PCR product (see Section 3.6) was used as a template for a second PCR-amplification with GC-rich primers. After the second PCR-amplification, 10% of the reaction volume was loaded onto an 8% polyacrylamide gel to monitor the efficiency of the second PCR-amplification. The arrow shows the 332 bp product containing the 184 bp coding region of HPRT exon 3, 20 bp of both 5' and 3' intron sequence, and 54 bp GC-clamps at both the 5' and 3' ends.

Rightmost lane: 200 ng molecular weight marker pPB322/MspI; other lanes: one-tenth of the PCR re-amplification (10 cycles) that added the GC-clamps to the 224 bp amplification product.

This experiment is in NFC12-307.
RI235, RI238, RI239, RI240, RI241, and RI242 were not re-examined. Mutant RI220 showed an unfocused band when analyzed with the 169 base pair probe and this mutant was an obvious candidate for re-examination using the GC-clamp; failure to do so was an error. The remaining 6 mutants appeared wild type using the 169 bp probe and were not re-examined with the GC-clamp.

The single ICR-191-induced mutant which showed a focused band and 2 mutants which showed an unfocused band when using the 169 bp probe all showed a mutation when analyzed with the GC-clamp. In addition, 10 of 25 mutants which appeared wild type when using the 169 bp probe were found to have a mutation when examined with the GC-clamp.

Figures 4-32, 4-33, 4-35, 4-36, 4-38, 4-39 and 4-37 show the examination of the TGF MNNG- and ICR-191-treated mutants. The results are summarized in Table 4-III on page 232 and Table 4-IV on page 234. All mutants which showed a non-wild type pattern were sequenced, the sequencing gels will be shown in subsequent sections.
4.9 Examination of untreated, independent, TG\textsuperscript{f} clones by DGGE using a GC-clamp

Untreated, independent, TG\textsuperscript{f} mutants were also examined for point mutations in exon 3. Our laboratory has previously shown that 39\% (33 of 85) of the untreated TG\textsuperscript{f} mutants contained large structural gene alterations, almost exclusively deletions, as determined by Southern blotting (Gennett and Thilly, 1988). Nineteen of the untreated mutants which showed a wild type Southern blotting pattern were examined for point mutations in exon 3 using the GC-clamp. Each untreated mutant was from a different individual starting culture and was therefore independent (see Section 3.2 for generation of untreated clones). Exon 3 was PCR-amplified from genomic DNA and a GC-clamp was added as discussed previously. Figures 4-40 and 4-41 show the denaturing gradient gels of the untreated mutants. The results of the screening are summarized in Table 4-III on page 232.
Figure 4-32: Autoradiogram showing DGGE of TG\textsuperscript{r} mutants with an added GC-clamp: HPRT\textsubscript{Munich}, RM4, RM34, RI209

About 600 cpm of ds, 8.2 Ci/mmol, 332 bp probe with GC-clamps was hybridized to the 332 bp PCR-amplification product which was generated from the TG\textsuperscript{r} mutant clones. After hybridization, the DNA was restricted with HgiAI and loaded onto a 5-35% denaturing gradient gel, electrophoresis time was for 5.5 hr.

The 54 bp GC-clamps at both ends of the 332 bp molecule are completely complementary; thus a single-stranded, intramolecularly-base paired, hairpin structure may form. The GC-clamps may also form intermolecularly-base paired structures. It is thought that the DNA at the gel origin is the intermolecularly-base paired form, and the DNA found at about 30% denaturant represents the hairpin form. The formation of these species can be nearly totally eliminated by restricting the molecule with HgiAI prior to hybridization as is shown in Figure 4-36.

Lane 1: WT, wild type (NFC11-23); Lane 2: MUNICH, HPRT\textsubscript{Munich} (NFC11-23); Lane 3: RM4 (NFC11-23); Lane 4: RM34 (NFC11-23); Lane 5: RI209 (NFC-23).

HPRT\textsubscript{Munich}, RM4, and RM34 show a mutation in the HgiAI fragment which melts at about 23% denaturant, while RI209 shows a mutation in the HgiAI fragment which melts at about 33% denaturant.

This experiment is in NFC11-78.
Figure 4-33: Autoradiogram showing DGGE of TG\(^r\) mutants with an added GC-clamp: RI202, RI204, RI206, RI207, RI208, RI210, RI213, RI214, RM2, RM3, RM6

About 4000 cpm of ds, 13.6 Ci/mmol, 332 bp probe with GC-clamps was hybridized to the 332 bp PCR-amplification product which was generated from the TG\(^r\) mutant clones. After hybridization, the DNA was restricted with HgiAI and loaded onto a 10-40% denaturing gradient gel, electrophoresis time was for 6 hr.

The 54 bp GC-clamps at both ends of the 332 bp molecule are completely complementary; thus a single-stranded, intramolecularly-base paired, hairpin structure may form. The GC-clamps may also form intermolecularly-base paired structures. It is thought that the DNA at the gel origin is the intermolecularly-base paired form, and the DNA found at about 30% denaturant represents the hairpin form. The formation of these species can be nearly totally eliminated by restricting the molecule with HgiAI prior to hybridization as is shown in Figure 4-36.

Lane 1: RI202; Lane 2: RI204; Lane 3: RI206; Lane 4: RI207; Lane 5: RI208; Lane 6: RI210; Lane 7: RI213; Lane 8: RI214; Lane 9: RM2; Lane 10: RM3; Lane 11: RM6.

RI202, RI206, RI208, and RI213 show a mutation in the HgiAI fragment which melts at about 23% denaturant. The HgiAI fragment which melts at about 33% denaturant was not visible in this exposure. A longer exposure is shown immediately following this figure. The DNA was prepared in NFC12-255. RI204, RI210, RI214, RM2, RM3 and RM6 appear wild type.

This experiment is in NFC11-101, gel 1.
Figure 4-34: Darker exposure of gel shown in Figure 4-33

This is a darker exposure of the gel shown in Figure 4-33. This exposure shows that mutant RI207 has a mutation in the DNA fragment which melts at about 33% denaturant.

Lane 1: RI202; Lane 2: RI204; Lane 3: RI206; Lane 4: RI207; Lane 5: RI208; Lane 6: RI210; Lane 7: RI213; Lane 8: RI214; Lane 9: RM2; Lane 10: RM3; Lane 11: RM6.

RI207 shows a mutation in the HgiAI fragment which melts at about 33% denaturant. A long film exposure was necessary to see the mutation in RI207. The DNA was prepared in NFC12-255.

This experiment is in NFC11-101, gel 1.
Figure 4-35: Autoradiogram showing DGGE of TG^f mutants with an added GC-clamp: RM8, RM9, RM10, RM16, RM19, RM20

4000 cpm of ds, 13.6 Ci/mmol, 332 bp probe with GC-clamps was hybridized to the 332 bp PCR-amplification product which was generated from the TG^f mutant clones. After hybridization, the DNA was restricted with HgiAI and loaded onto a 10-40% denaturing gradient gel, electrophoresis time was for 6 hr.

The 54 bp GC-clamps at both ends of the 332 bp molecule are completely complementary; thus a single-stranded, intramolecularly-base paired, hairpin structure may form. The GC-clamps may also form intermolecularly-base paired structures. It is thought that the DNA at the gel origin is the intermolecularly-base paired form, and the DNA found at about 30% denaturant represents the hairpin form. The formation of these species can be nearly totally eliminated by restricting the molecule with HgiAI prior to hybridization as is shown in Figure 4-36.

Lane 1: RM8; Lane 2: RM9; Lane 3: RM10; Lane 4: RM16; Lane 5: RM19; Lane 6: RM20;

RM16, RM19, and RM20, show a mutation in the HgiAI fragment which melts at about 23% denaturant. RM8, RM9 and RM10 appear wild type. A long film exposure was necessary to see the HgiAI fragment which melted at about 33% denaturant, which resulted in an overexposure of the other mutants. The DNA was prepared in NFC12-255.

This experiment is in NFC11-101, gel 2.
Figure 4-36: Autoradiogram showing DGGE of TG\textsuperscript{r} mutants with an added GC-clamp: RM1, RM11, RM12, RM13, RM14, RM15, RM23, RM24, RM25

About 1000 cpm of ds, 28.6 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with HgiAI. After restriction, the DNA was boiled and hybridized at 65\degree C. The DNA was loaded onto a 10%-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI206; Lane 2: RI209; Lane 3: RM11; Lane 4: RM12; Lane 5: RM13; Lane 6: RM14; Lane 7: RM15; Lane 8: RM23; Lane 9: RM24; Lane 10: RM25; Lane 11: RM1.

RI206 (lane 1) and RI209 (lane 2) were mutants previously found to have an alteration in the HgiAI fragments which melt at 23\% and 33\% denaturant, respectively; these mutants were used as a positive control.

RM1, RM11, RM12, RM13, RM14, RM15, RM23, RM24 and RM25 appear wild type. The DNA was prepared in NFC12-282.

This experiment is in NFC11-157, gel 1.
Figure 4-37: Autoradiogram showing DGGE of TG\textsuperscript{r} mutants with an added GC-clamp: RI203, RI205, RI215, RI216, RI217, RI219, RI221, RI222, RI224

About 1000 cpm of ds, 28.6 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with *HgiAI*. After restriction, the DNA was boiled and hybridized at 65°C. The DNA was loaded onto a 10-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI206; Lane 2: RI209; Lane 3: RI203; Lane 4: RI205; Lane 5: RI215; Lane 6: RI216; Lane 7: RI217; Lane 8: RI219; Lane 9: RI221; Lane 10: RI222; Lane 11: RI224.

RI206 (lane 1) and RI209 (lane 2) were mutants previously found to have an alteration in the *HgiAI* fragments which melt at 23% and 33% denaturant, respectively; these mutants were used as a positive control.

RI205, RI219, and RI222 have a mutation in the *HgiAI* fragment which melts at about 23% denaturant, while RI224 has a mutation in the fragment which melts at about 33% denaturant. RI203, RI215, RI216, RI217 and RI221 appear wild type. The DNA was prepared in NFC12-282.

This experiment is in NFC11-157, gel 2.
NFC11-157, gel 2
Figure 4-38: Autoradiogram showing DGGE of MNNG- and ICR-191-induced TG\textsuperscript{r} mutants with an added GC-clamp: RI225, RM26, RM28, RM30, RM31, RM32, RM33, RM35, RM37, RM38, RM39

About 1000 cpm of ds, 7.7 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with \textit{HgiAI}. After restriction, the DNA was boiled and hybridized at 65°C. The DNA was loaded onto a 10%-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI206; Lane 2: RI224; Lane 3: RM26; Lane 4: RM28; Lane 5: RM30; Lane 6: RM31; Lane 7: RM32; Lane 8: RM33; Lane 9: RM35; Lane 10: RM37; Lane 11: RM38; Lane 12: RM39; Lane 13: BLANK; Lane 14: RI225.

RI206 (lane 1) and RI224 (lane 2) were mutants previously found to have an alteration in the \textit{HgiAI} fragments which melt at 23% and 33% denaturant, respectively; these mutants were used as a positive control.

RM26, RM28, RM30, RM31, RM32, RM33, RM35, RM37, RM39 and RI225 appear wild type. The DNA was prepared in NFC12-318.

This experiment is in NFC11-194, gel 1.
Figure 4-39: Autoradiogram showing DGGE of MNNG- and ICR-191-induced TG<sup>r</sup> mutants with an added GC-clamp: RI226, RI227, RI228, RI230, RI231, RI232, RI233, RI234, RI237

About 1000 cpm of ds, 7.7 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with HgiAI. After restriction, the DNA was boiled and hybridized at 65<sup>°</sup> C. The DNA was loaded onto a 10-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI206; Lane 2: RI224; Lane 3: RI226; Lane 4: RI227; Lane 5: RI228; Lane 6: RI230; Lane 7: RI231; Lane 8: RI232; Lane 9: RI233; Lane 10: BLANK; Lane 11: RI234; Lane 12: RI237.

RI206 (lane 1) and RI224 (lane 2) were mutants previously found to have an alteration in the HgiAI fragments which melt at 23% and 33% denaturant, respectively; these mutants were used as a positive control.

RI228 contains a mutation in the fragment which melts at 33% denaturant, RI232 and RI233 contain a mutation in the fragment which melts at 23% denaturant. RI226, RI227, RI230, RI231, RI234 and RI237 appear wild type. The DNA was prepared in NFC12-318.

This experiment is in NFC11-194, gel 2.
Figure 4-40: Autoradiogram showing DGGE of untreated $\text{TG}^r$ mutants with an added GC-clamp: 11B, 29B, 37B, 30B, 19B, 8C, 46A, 5A-2, 28B-1, 16B, 20B

About 1000 cpm of ds, 21.4 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with $Hgl\text{AI}$. After restriction, the DNA was boiled and hybridized at 65° C. The DNA was loaded onto a 10-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI209; Lane 2: RI206; Lane 3: 11B; Lane 4: 29B; Lane 5: 37B; Lane 6: 30B; Lane 7: 19B; Lane 8: 8C; Lane 9: 46A; Lane 10: 5A-2; Lane 11: 28B-1; Lane 12: 16B; Lane 13: 20B;

RI209 (lane 1) and RI206 (lane 2) were mutants previously found to have an alteration in the $Hgl\text{AI}$ fragments which melt at 33% and 23% denaturant, respectively; these mutants were used as a positive control.

30B, 5A-2, and 16B have a mutation in the $Hgl\text{AI}$ fragment which melts at about 23% denaturant. 11B, 29B, 37B, 19B, 8C, 46A, 28B-1 and 20B appear wild type. The DNA was prepared in NFC12-309.

This experiment is in NFC11-188, gel 1.
Figure 4-41: Autoradiogram showing DGGE of untreated TG<sup>+</sup> mutants with an added GC-clamp: 26A, 57B, 38A, 18A, 25C, 33A, 52B-1, 53A

About 1000 cpm of ds, 21.4 Ci/mmol, 332 bp probe with GC-clamps was added to the 332 bp PCR-amplification product and restricted with HgiAI. After restriction, the DNA was boiled and hybridized at 85<sup>0</sup> C. The DNA was loaded onto a 10-40% denaturing gradient gel, and electrophoresis time was for 6 hr.

Lane 1: RI209; Lane 2: RI206; Lane 3: 26A; Lane 4: 57B; Lane 5: 38B; Lane 6: 18A; Lane 7: 25C; Lane 8: 33A; Lane 9: 52B-1; Lane 10: 53A;

RI209 (lane 1) and RI206 (lane 2) were mutants previously found to have an alteration in the HgiAI fragments which melt at 33% and 23% denaturant, respectively; these mutants were used as a positive control.

33A has a mutation in the HgiAI fragment which melts at about 23% denaturant. 26A, 57B, 38B, 18A, 25C, 52B-1 and 53A appear wild type. The DNA was prepared in NFC12-309.

This experiment is in NFC11-188, gel 2.
Table 4-III: Summary showing the results of examining untreated, MNNG- and ICR-191-induced TG\textsuperscript{r} clones using the GC-clamp

Mutations were detected in both \textit{HgiAI} fragments containing exon 3. The fragment containing a GC-clamp at the 5' end melted at about 33\% denaturant and the fragment with the GC-clamp at the 3' end melted at about 23\% denaturant. The exon 3 sequence from base pair location 220-288 is present in the fragment that melts at about 33\% denaturant, and the exon 3 sequence from base pair location 289-403 is present in the fragment which melts at about 23\% denaturant.
<table>
<thead>
<tr>
<th>Mutagen</th>
<th>Mutation in fragment that melted at 23% denaturant</th>
<th>Mutation in fragment which melted at 33% denaturant</th>
<th>wild type</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICR-191</td>
<td>RI202, RI205 RI206, RI208 RI213, RI219 RI222, RI232 RI233</td>
<td>RI207, RI209 RI224, RI228</td>
<td>RI203, RI204 RI210, RI214 RI215, RI216 RI217, RI221 RI225, RI226, RI227 RI230, RI231 RI234, RI237</td>
</tr>
<tr>
<td>(Spontaneous)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4-IV: Summary showing the results of examining untreated, MNNG- and ICR-191-induced TG<sup>f</sup> clones using the GC-clamp

Mutations were detected in both HgiAI fragments containing exon 3. The fragment containing a GC-clamp at the 5' end melted at about 33% denaturant and the fragment with the GC-clamp at the 3' end melted at about 23% denaturant. Deletions that were confirmed by Southern blotting in Figure 4-25 were added to the total.
<table>
<thead>
<tr>
<th>Mutagen</th>
<th>Mutation in fragment that melts at 23% denat.</th>
<th>Mutation in fragment that melts at 33% denat.</th>
<th>deletion</th>
<th>wild type</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNNG</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>25</td>
<td>32</td>
</tr>
<tr>
<td>ICR-191</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td>NONE</td>
<td>4</td>
<td>0</td>
<td>-</td>
<td>15</td>
<td>19</td>
</tr>
</tbody>
</table>

* Nineteen untreated mutants that showed a wild type Southern blotting pattern were examined. Previous work by Gennett and Thilly (1988) showed that about 40% of the untreated, TG<sup>f</sup> clones contained a large structural gene alteration, almost exclusively deletions, as determined by Southern blotting.

The H<sub>gi</sub>AI fragment that melted at 23% denaturant contained the exon 3 coding region from base pair location 288-403 and has a GC-clamp at the 3' end. The H<sub>gi</sub>AI fragment that melted at 33% denaturant contained the exon 3 coding region from base pair location 220-287 and has a GC-clamp at the 5' end.
4.10 DNA sequencing of untreated, MNNG- and ICR-191-induced, TG\textsuperscript{5} clones

Most of the mutants in Sections 4.8 and 4.9 that showed a non-wild type pattern on a denaturing gradient gel were sequenced. All of the MNNG-induced (5) and ICR-191-induced (13) exon 3 mutants were sequenced. Two of four spontaneous mutants that showed a ncn-wild type band could not be sequenced by the September dissertation deadline.

The double-stranded, 224 bp PCR-amplification product was sequenced using the 20 bp downstream primer (the primer sequence is given in Section 3.6); the (-) strand was sequenced. Each mutant was loaded twice on a single sequencing gel and the electrophoretic migration time was generally 2 and 4 hours; in this way, the entire coding region of exon 3 was visible on a single gel. It was generally possible to read the sequence from the downstream primer to within about 5 bp of the upstream end of the coding region. The section of the gel that most clearly showed the mutation was selected for the figures, in some cases different exposures of the same gel were necessary.

The DNA from one mutant, RI207, could not be sequenced directly. Direct sequencing consistently yielded a wild type pattern, even though DGGE electrophoresis showed that this mutant contained an alteration in the 5' portion of exon 3. The sequencing gels were readable with the exception of about the last 5 nucleotides at the 5' end of exon 3.

To sequence mutant RI207, the 224 bp amplification product was reamplified
with the \textit{BamHI} and \textit{HindIII} primers as described in Section 3.4.2. The reamplified DNA was restricted with \textit{BamHI} and \textit{HindIII}, cloned into RF M13mp19, and single-strand DNA was isolated as described in Messing (1983). DNA was prepared from three individual plaques and was sequenced as described above. All three DNA’s showed an identical base pair substitution at the extreme 5' base pair in the coding sequence of exon 3. A photograph showing the mutation in RI207 could not be prepared in time for the September thesis deadline, however, the sequence data for RI207 in included in this dissertation.

A summary of the sequencing results is given in Table 4-V on page 258 and in Figure 4-52 on page 262.
Figure 4-42: Sequencing gel showing RM4 and RM34

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RM4, from left to right: A, C, G, T; Lanes 9-12 are RM34, from left to right: A, C, G, T.

RM4 contains a C → T base pair substitution as shown which corresponds to a G → A at base pair position 293 of the (+) strand. RM34 contains a G → T base pair substitution as shown which corresponds to a C → A substitution at base pair position 328 of the (+) strand.

This experiment is given in NFC Seq-1.
Figure 4-43: Sequencing gel showing RI209

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RI209, from left to right: A, C, G, T.

RI209 contains a T → A base pair substitution as shown which corresponds to an A → T at base pair position 248 of the (+) strand.

This experiment is in NFC Seq-2.
Figure 4-44: Sequencing gel showing RI202 and RI206

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RI202, from left to right: A, C, G, T; Lanes 9-12 are RI206, from left to right: A, C, G, T.

RI202 and RI206 both contain a +C addition as shown, which increased the wild type run of 6 C's to 7 C's in the mutants. This corresponds to a +G addition in the GGGGGGG sequence from base pair position 292-297 in the (+) strand.

This experiment is given in NFC Seq-3-1.
Figure 4-45: Sequencing gel showing RI208 and RI213

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RI202, from left to right: A, C, G, T; Lanes 9-12 are RI206, from left to right: A, C, G, T; Lanes 13-16 are RI208, from left to right: A, C, G, T; Lanes 17-20 are RI213, from left to right: A, C, G, T.

This is a longer exposure of Figure 4-44, the wild type sequence is overexposed.

RI213 is mislabeled as RI212 in the photograph.

RI208 and RI213 both contain a +C addition as shown, which increased the wild type run of 6 C's to 7 C's in the mutants. This corresponds to a +G addition in the GGGGGGGG sequence from base pair position 292-297 in the (+) strand.

This experiment is given in NFC Seq-3-2.
Figure 4-46: Sequencing gel showing RM20

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 17-20 are RM20, from left to right: A, C, G, T.

RM20 contains a C → T base pair substitution as shown, which corresponds to a G → A substitution at base pair position 294 of the (+) strand.

This experiment is given in NFC Seq-4-1.
**Figure 4-47:** Sequencing gel showing RM16, RM19 and RI207

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RM16, from left to right: A, C, G, T; Lanes 9-12 are RM19, from left to right: A, C, G, T; Lanes 13-16 are RI207, from left to right: A, C, G, T.

This is a longer exposure of the gel shown in Figure 4-46, the wild type sequence is overexposed.

RM16 contains a C → A base pair substitution as shown, which corresponds to a G → T substitution at base pair position 294 of the (+) strand. RM19 contains a C → T base pair substitution as shown, which corresponds to a G → A substitution at base pair position 293 of the (+) strand.

RI207 shows no visible sequence alteration. RI207 is mislabeled as R0207 in the photograph.

This experiment is given in NFC Seq-4-2.
Figure 4-48: Sequencing gel showing RI207, RI219, RI222, RI205, RI224

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are RI207, from left to right: A, C, G, T; Lanes 9-12 are RI219, from left to right: A, C, G, T; Lanes 13-16 are RI222, from left to right: A, C, G, T; Lanes 17-20 are RI205, from left to right: A, C, G, T; Lanes 17-20 are RI224, from left to right: A, C, G, T.

RI219, RI222 and RI205 all contain a +C addition as shown, which increased the wild type run of 6 C's to 7 C's in the mutants; this corresponds to a +G addition in the GGGGGGG sequence from base pair position 292-297 in the (+) strand. RI224 contains a +G addition as shown, which increased a wild type run of 3 G's to 4 G's in the mutant; this corresponds to a +C addition in a CCC run at base pair position 276-278 in the (+) strand.

No sequence alteration in RI207 is visible.

This experiment is given in NFC Seq-5.
**Figure 4-49:** Sequencing gel showing RI228, RI232, RI233

The wild type sequence of the (−) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are RI228, from left to right: A, C, G, T; Lanes 5-8 are RI232, from left to right: A, C, G, T; Lanes 9-12 are RI233, from left to right: A, C, G, T.

RI232 is mislabeled as RI223 in the left side of the figure. The wild type sequence was not run on this gel.

RI232 and RI233 contain a +C addition as shown, which increased the wild type run of 6 C's to 7 C's in the mutants; this corresponds to a +G addition in the GGGGGGGG sequence from base pair position 292-297 in the (+) strand. RI228 contains a +G addition as shown, which increased a wild type run of 3 G's to 4 G's in the mutant; this corresponds to a +C addition in a CCC run at base pair position 276-278 in the (+) strand.

This experiment is given in NFC11-208.
Figure 4-50: Sequencing gel showing spontaneous mutant 33A

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are 33A, from left to right: A, C, G, T.

33A contains a +A addition as shown, which increased the wild type run of 3 A's to 4 A's in mutant 33A; this corresponds to a +T addition in the TTT sequence from base pair position 308-310 in the (+) strand.
**Figure 4-51:** Sequencing gel showing spontaneous mutant 30B

The wild type sequence of the (-) strand is shown; several runs of consecutive nucleotides are also shown which facilitate the reading of the sequencing gel.

Lanes 1-4 are wild type, from left to right: A, C, G, T; Lanes 5-8 are 30B, from left to right: A, C, G, T.

30B contains a C→A transversion as shown; this corresponds to a G→T substitution at base pair position 377 in the (+) strand.
Table 4-V: Summary of sequence alterations found in exon 3 of untreated, MNNG- and ICR-191-induced TG^f clones

The nature of the mutation, amino acid substitution, and base pair location is given (base pair numbering as in Nucleic Acid Sequence Database, Version 32.0, National Biomedical Research Foundation, Washington, DC; the first nucleotide of the ATG start codon is at position 86). The nature of the mutation is given in reference to the (+) strand.

Most of the ICR-191-induced mutations were +1 frameshift events in a GGGGGG sequence (292-297). Most of the MNNG-induced mutations were base pair substitutions at base pair locations 293 and 294 in the same GGGGGG sequence (292-297). MNNG-induced mutations primarily effected the following underlined base pairs in the run of six guanines—GGGGGG.
<table>
<thead>
<tr>
<th>Mutagen</th>
<th>Mutant</th>
<th>Base pair substitution and location</th>
<th>Amino acid substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNNG</td>
<td>RM4, RM19</td>
<td>GC → AT bp 293</td>
<td>Gly69 → Arg GGG → AGG</td>
</tr>
<tr>
<td></td>
<td>RM16</td>
<td>GC → TA bp 294</td>
<td>Gly69 → Val GGG → GTG</td>
</tr>
<tr>
<td></td>
<td>RM20</td>
<td>GC → AT bp 294</td>
<td>Gly69 → Glu GGG → GAG</td>
</tr>
<tr>
<td></td>
<td>RM34</td>
<td>CG → AT bp 328</td>
<td>Tyr80 → Stop TAC → TAA</td>
</tr>
<tr>
<td>ICR-191</td>
<td>RI202, RI205</td>
<td>+G bp 292-297</td>
<td>frameshift in GGGGGGGG sequence</td>
</tr>
<tr>
<td></td>
<td>RI206, RI208</td>
<td>+G bp 292-297</td>
<td>frameshift in CCC sequence</td>
</tr>
<tr>
<td></td>
<td>RI213, RI219</td>
<td>+C bp 276-278</td>
<td>frameshift in CCC sequence</td>
</tr>
<tr>
<td></td>
<td>RI222, RI232</td>
<td>AT → TA bp 248</td>
<td>Lys54 → Stop AAG → TAG</td>
</tr>
<tr>
<td></td>
<td>RI232</td>
<td>GC → TA bp 220</td>
<td>Probable mRNA splicing mutant*</td>
</tr>
<tr>
<td>SPONTANEOUS</td>
<td>30B</td>
<td>GC → TA bp 377</td>
<td>Asp97 → Tyr GAT → TAT</td>
</tr>
<tr>
<td></td>
<td>33A</td>
<td>+T bp 308-310</td>
<td>frameshift in TTT sequence</td>
</tr>
</tbody>
</table>

* The base pair substitution in RI207 occurred at the extreme 5' base pair in exon
3. This position has been shown to be part of the mRNA splicing consensus sequence (Breathnach and Chambon, 1981).
**Figure 4-52:** Base pair location of sequenced exon 3 mutants

The coding region of HPRT exon 3 is from base pair 220-403. The location of 2 untreated, 13 ICR-191-treated, and 5 MNNG-treated mutants are given.
LOCATION OF MUTATIONS IN HUMAN HPRT EXON 3

- ICR-191
  13 MUTANTS IN 29 TG' CLONES

- MNNG
  5 MUTANTS IN 32 TG' CLONES

- SPONTANEOUS
  2 MUTANTS
4.11 Analysis of complex mutant populations by DGGE

The clone-by-clone analysis presented in section 4.8 suggested that mutational hotspots for both MNNG and ICR-191 existed in exon 3 at a GGGGGGG sequence (base pair 292-297). The results presented in this section will demonstrate that a population derived from several thousand TG\textsuperscript{F} mutants can be analyzed simultaneously by DGGE. First the ability of DGGE to detect a minor mutant species will be examined, then the accuracy of PCR in amplifying a mutant present at 1% of the total population will be presented, and finally the MNNG- and ICR-191-treated populations will be examined.

4.11.1 Sensitivity of DGGE

This section shows that mutant DNA present at a frequency of 1% in a wild type background can be detected by DGGE. Figure 4-8 showed that HPRT\textsubscript{Munich} contained a mutation detectable by DGGE. HPRT\textsubscript{Munich} exon 3 was amplified from genomic DNA and was added to wild type, amplified exon 3 DNA in the following proportions: 0.5, 10\textsuperscript{-1}, 10\textsuperscript{-2}, 10\textsuperscript{-3}, 10\textsuperscript{-4}. The DNA was hybridized in solution to a high specific activity, single strand probe and the results are shown in Figure 4-53.

The limit of detection of HPRT\textsubscript{Munich} in a wild type background was 1%. Therefore, a mutant species present after PCR-amplification at a frequency of 1% should be detected.
Figure 4-53: Autoradiogram of denaturing gradient gel showing $\text{HPRT}_{\text{Munich}}$ DNA and wild type DNA in a $10^{-1}$, $10^{-2}$, $10^{-3}$, and $10^{-4}$ ratio.

About $10^{10}$ total copies of DNA containing different proportions of $\text{HPRT}_{\text{Munich}}$ in a wild type background was hybridized to $1.3 \times 10^5$ cpm ($5.5 \times 10^8$ copies) of single strand, 224 bp exon 3 probe of specific activity 960 Ci/mmol. Single strand probe was made as described in Section 3.4.2.

The 224 bp fragment of $\text{HPRT}_{\text{Munich}}$ was amplified from genomic DNA using Taq DNA polymerase as described in Section 3.6. The concentration of amplified DNA was estimated by comparison of ethidium bromide fluorescence with that of known quantities of DNA. Wild type 224 bp exon 3 DNA was produced by 10 cycles of amplification with the DNA polymerase Sequenase™ (United States Biochemical Corp., Cleveland, OH), using plasmid pλ3001 as a template. The conditions for amplification with Sequenase are given on page 110. Wild type DNA was purified on a polyacrylamide gel and quantitated as described immediately above.

After hybridization, 50% of the sample was loaded onto a 10-30% denaturing gradient gel and electrophoresis time was for 16 hours.

Lane 1: 100 cpm ss probe; Lane 2: 1000 cpm ss probe; Lane 3 and 4: other samples; Lane 5: WT, wild type; Lane 6: MUNICH, $\text{HPRT}_{\text{Munich}}$; Lane 7: 0.5, 50% $\text{HPRT}_{\text{Munich}}$ and 50% wild type; Lane 8: $10^{-1}$, 10% $\text{HPRT}_{\text{Munich}}$ and 90% wild type; Lane 9: $10^{-2}$, 1% $\text{HPRT}_{\text{Munich}}$ and 99% wild type; Lane 10: $10^{-3}$, 0.1% $\text{HPRT}_{\text{Munich}}$ and 100% wild type; Lane 11: $10^{-4}$, 0.01% $\text{HPRT}_{\text{Munich}}$ and 100% wild type.

This experiment is in NFC10-229.
4.11.2 Reproducibility of PCR-amplification using wild type DNA containing 1% HPRT<sub>Munich</sub> DNA

Figure 4-53 showed that a mutant species present at a fraction of 1% after PCR-amplification can be detected. 1 μg of genomic DNA was typically used for a PCR amplification, which represented about 1.5 x 10<sup>5</sup> copies of HPRT exon 3. A mutant present at 1% will represent about 1500 HPRT exon 3 copies. In this section, HPRT<sub>Munich</sub> cells at a fraction of 1% in a wild type cell background were used as a template for PCR amplification.

One million HPRT<sub>Munich</sub> cells were added to 10<sup>8</sup> wild type cells, the DNA was isolated, and 1 μg of DNA was used in three separate PCR amplifications with Taq DNA polymerase. The PCR-amplified DNA was hybridized to a high specific activity, single strand probe and the results are given in Figure 4-54. The 1% HPRT<sub>Munich</sub> signal was visible as a faint band which was not present in the wild type. Importantly, the relative intensity of the 1% HPRT<sub>Munich</sub> signal was similar in the three independent PCR-amplifications, indicating that PCR was capable of accurately amplifying 1500 copies of HPRT<sub>Munich</sub> in a background of 1.5 x 10<sup>5</sup> wild type copies. The limit of resolution of a mutant species in a wild type background is therefore about 1%.

Figure 4-54 also shows genomic DNA amplified 32 cycles with Taq DNA polymerase (lane 3), and plasmid DNA amplified 10, 20, and 30 cycles with Sequenase<sup>TM</sup> (lanes 8-10). Increasing the number of amplification cycles from 10 to 20 increased the background noise, however, increasing the number of cycles
from 20 to 30 did not significantly increase the background noise. The background noise produced by amplification was expected to be directly proportional to the number of amplification cycles, so the ratio of background noise in DNA amplified 10, 20 and 30 cycles was expected to be 1:2:3; the reason for the observed discrepancy is unknown. A direct comparison between Taq polymerase (lane 3) and Sequenase™ (lane 10) indicated that the background noise produced by Taq was slightly higher than the background noise produced by Sequenase™.

4.11.3 Analysis of untreated, MNNG-, and ICR-191-treated complex mutant populations by DGGE

Figure 4-54 showed that a mutant present at a frequency of 1% in a wild type cell background will be detected by PCR/DGGE. In this section, the untreated, MNNG- and ICR-191-treated populations will be examined. The complex mutant populations will be analyzed in two ways, (1) by DGGE using the 224 bp exon 3 fragment, and (2) by DGGE using GC-clamps. DGGE will detect mutations only in the low-temperature melting domain of the 224 bp fragment, while DGGE with the GC-clamps should be sensitive to any mutation in exon 3.

Several complex mutant populations will be analyzed. The complex populations were generated by simply adding TG to 1-2 x 10^6 treated or untreated cells. Since the mutant fraction was determined, the original number of TG^r in the culture could be determined. These original TG^r cells continued to grow and in about a week overgrew the culture. The culture then consisted of a complex population of TG^r cells which were derived from the original TG^r population.
**Figure 4-54:** Autoradiogram showing three PCR-amplifications of DNA containing $1\%$ HPRT$_{Munich}$ and analysis by DGGE.

One million HPRT$_{Munich}$ cells and $10^8$ wild type cells were mixed, DNA was extracted, and 1 µg of DNA (about $1.5 \times 10^5$ HPRT copies) was PCR-amplified in three separate experiments. Amplification was performed using 32 cycles of *Taq* DNA polymerase as described in Section 3.6. The arrow indicates the position of HPRT$_{Munich}$.

Plasmid pλ3001 was also amplified 10, 20 and 30 cycles using Sequenase$^\text{TM}$ as described on page 110.

About $10^{10}$ copies of DNA was hybridized to $1 \times 10^5$ cpm (4.2 $\times 10^8$ copies) of single strand 224 bp exon 3 probe of specific activity 1000 Ci/mmol. Single strand probe was made as described in Section 3.4.2.

DNA was loaded onto a 15-30% denaturing gradient gel and electrophoresis time was 15.5 hours.

Lane 1: 100 cpm ss probe; Lane 2: 1000 cpm ss probe; Lane 3: WT, wild type (THP1 B, #7, NFC10-248); Lane 4: MUNICH, HPRT$_{Munich}$ (NFC10-144); Lane 5: 1% MUNICH, 1% HPRT$_{Munich}$ (NFC10-323, #7); Lane 6: 1% MUNICH, 1% HPRT$_{Munich}$ (NFC10-323, #11); Lane 7: 1% MUNICH, 1% HPRT$_{Munich}$ (NFC10-323, #9); Lane 8: 10 CYCLES, 10 PCR cycles with Sequenase$^\text{TM}$; Lane 9: 20 CYCLES, 20 PCR cycles with Sequenase$^\text{TM}$; Lane 10: 30 CYCLES, 30 PCR cycles with Sequenase$^\text{TM}$;

This experiment is in NFC11-57.
Most of the complex mutant populations were generated by myself, however, one complex population was generated by Monica J. Cahilly. It should be recalled that the individual, TG$^r$ clones analyzed in previous sections were produced by Monica Cahilly. Monica Cahilly had frozen down about $10^8$ cells from the MNNG-treated culture which was used to produce the individual clones. I retrieved her MNNG-treated culture from the freezer and added TG, thus one MNNG-induced complex population was produced by Monica Cahilly.

I will summarize below the conditions used to generate the complex mutant populations and the number of original TG$^r$ mutants present in each culture. Section 3.3 in Material and Methods also describes the generation of the complex mutant populations. The reader is advised to remove the pages with the summary so he/she can refer to the summary as necessary.

The complex mutant populations are characterized as follows:

**Untreated populations**

- Untreated cultures labeled $E$ and $F$ each containing $2.5 \times 10^8$ cells were generated by myself.

- These cultures are the control populations for my MNNG-treated complex mutant population (labeled $A$ and $B$) and my ICR-191-treated complex mutant population (labeled $C$).

- The mutant fraction in culture $E$ was $3.2 \times 10^{-6}$. The mutant in culture $F$ could not be determined.

- The complex mutant population in culture $E$ was derived from 800 TG$^r$ cells.
MNNG-treated populations

- Cultures labeled A and B were generated by myself.
  - 2.5 x 10^8 cells were treated for 1 hour with 10 ng/ml MNNG. The survival was 0.85, thus 2.1 x 10^8 cells survived treatment.
  - The induced mutant fraction in culture A was 6.1 x 10^{-6}. The untreated, background mutant fraction was 3.2 x 10^{-6}. The mutant fraction in culture B could not be determined, but the survival after treatment was the same in both cultures.
  - The complex mutant population in culture A was derived from 1275 induced and 675 untreated TG^r cells. 65% of the TG^r cells in the culture were induced by treatment, 35% of the TG^r cells were not induced by treatment with MNNG.

- Culture labeled MJC was generated by Monica J. Cahilly.
  - 5 x 10^8 cells were treated for 1 hour with 15 ng/ml MNNG. Survival was 0.11, thus 5.5 x 10^7 cells survived treatment.
  - The induced mutant fraction in culture MJC was 5.0 x 10^{-5} and the untreated mutant fraction was 5.6 x 10^{-6}.
  - The complex mutant population in culture MJC was derived from 2750 induced and 330 untreated TG^r cells. 90% of the TG^r clones in the culture were induced by MNNG treatment.

ICR-191-treated population

- ICR-191-treated population labeled C was generated by myself.
  - 2.5 x 10^8 cells were treated for 24 hours with 500 ng/ml ICR-191. Survival was 0.5, thus 1.3 x 10^8 cells survived treatment.
  - The induced mutant fraction in culture C was 7.5 x 10^{-5}. The untreated, background mutant fraction was 3.2 x 10^{-6}.
  - ICR-191-treated population C was derived from 9725 induced and
420 untreated TG\textsuperscript{r} cells. 96\% of the TG\textsuperscript{r} cells in the culture were induced by treatment.

The first experiment is given in Figure 4-55. DNA from MNNG-treated culture A, ICR-191-treated culture C, and untreated culture E was PCR-amplified and hybridized to high specific activity, single strand, 224 bp exon 3 probe. Mutations in the low-temperature melting domain can be detected using the 224 bp probe.

The DNA from one of the MNNG-treated cultures, culture A, was amplified in three separate experiments, the DNA from ICR-191-treated culture C was amplified in two separate experiments, and the DNA from untreated culture E was amplified in two separate experiments. Multiple amplifications from the same DNA was performed to verify the reproducibility of the system, if any mutational hotspots were observed.

Also shown in Figure 4-55 is a untreated, unselected, wild type population. This serves as a control for the amplification process, any hotspot for polymerase-induced errors should be visible in all amplified populations.

MNNG-treated culture A showed a novel band not present in the untreated TG\textsuperscript{r} population (culture E) or in the wild type, unselected control. This novel band was present in three independent amplifications, indicating that the novel band was not an artifact of the amplification process. The ICR-191-treated TG\textsuperscript{r} population (culture C) and the untreated TG\textsuperscript{r} population (culture E) showed no distinct novel bands compared to the wild type control.
The novel band in MNNG-treated population A comprised about 1% of the population. This was estimated by comparing the intensity of the novel band to the intensity of radiolabeled standards (lanes 1 and 2). The novel band was about as intense as the signal which contained $10^3$ cpm. $1.25 \times 10^5$ cpm of single-strand probe was hybridized to the complex mutant population, so the novel band represented about 1% of the complex mutant population.

Also shown in Figure 4-55 is plasmid pλ3001 DNA amplified 30 cycles with Taq DNA polymerase. The pattern of background bands seen when Taq DNA polymerase was used to amplify plasmid DNA (lane 9) was identical to the pattern of background bands produced when Taq DNA polymerase was used to amplify genomic DNA (lane 8). This indicated that the background noise produced by Taq polymerase was due to the amplification process itself, and that the background noise seen using genomic DNA was not due to amplification of other genomic species.

Also shown in Figure 4-55 is human genomic DNA amplified with T4 DNA polymerase (lane 13). The absence of background noise was striking, unfortunately, variation in the manufacturer's lots of T4 polymerase precluded general use of this enzyme.

Figure 4-55 showed that a novel band was present in MNNG-treated, complex mutant population A. The experiment shown in Figure 4-56 shows the results of analyzing the other independent, MNNG-treated, complex mutant population (culture B). The DNA from the MNNG-treated populations (A and B)
Figure 4-55: DGGE analysis of one untreated, one MNNG-treated, and one ICR-191-treated complex mutant population using a 224 bp probe

Triplicate amplifications of MNNG-treated culture A (lanes 3-5), and duplicate amplifications of ICR-191-treated culture C (lanes 11 and 12) and untreated culture E (lanes 6 and 7) were performed. The arrow indicates the location of the novel band in the MNNG-treated complex mutant population A.

DNA in lanes 3-12 was produced by 32 cycles of amplification using Taq DNA polymerase as described in Section 3.6. The DNA shown in lane 13 was produced by 30 amplification cycles using T4 DNA polymerase as described in Keohavong et al. (1988).

About $10^{10}$ copies of DNA was hybridized to $1.25 \times 10^5$ cpm ($5.3 \times 10^8$ copies) of single strand, 224 bp probe of specific activity 1000 Ci/mmol; about $10^6$ copies of HPRT-Munich DNA was hybridized to 1000 cpm of probe. See Section 3.4.2 for probe generation.

DNA was loaded onto a 10%-30% denaturing gradient gel. The electrophoresis time was 16 hours.

Lane 1: 100 cpm ss probe; Lane 2: 1000 cpm ss probe; Lane 3: MNNG-treated complex population A (NFC10-160); Lane 4: MNNG-treated complex population A (NFC10-248, #1); Lane 5: MNNG-treated complex population A (NFC10-248, #2); Lane 6: untreated complex population E (NFC10-160); Lane 7: untreated complex population E (NFC10-248); Lane 8: TK6, wild type (THP1 B, #7, NFC10-248); Lane 9: P, plasmid p3001 amplified 30 cycles with Taq polymerase; Lane 10: M, HPRT-Munich (the heteroduplex contains a G:A mismatch); Lane 11: ICR-191-treated complex population C (NFC10-160); Lane 12: ICR-191-treated complex population C (NFC10-248); Lane 13: T4, TK6 amplified 30 cycles with T4 DNA polymerase (NFC10-360); Lane 14: 0, hybridization control, no DNA.

This experiment is in NFC10-285.
was PCR-amplified and hybridized to a high specific activity, 224 bp, single strand exon 3 probe.

Figure 4-56 shows that the novel band present in MNNG-treated culture A was not reproducible, the novel band was not present in independent MNNG-treated culture B. Therefore, no reproducible mutational hotspots were detected in untreated culture E, MNNG-treated culture A, or ICR-191-treated culture C using the 224 base pair exon 3 fragment. A mutant in the low-temperature melting domain of exon 3 present at a frequency of about 1% should have been detected by this technique.

The novel band present in MNNG-treated culture A comprised about 1% of the population. The novel band was present after three separate amplifications of MNNG-treated culture A but not in the corresponding controls, indicating that the novel band was not an artifact of the amplification process.

Why was the corresponding novel band not seen in MNNG-treated culture B? About 2000 original TG^R mutants were present after MNNG treatment, so the novel 1% band in culture A represented about 20 original TG^R mutants. It is possible that MNNG-treated culture B contained less than 20 mutants due to the normal statistical variation associated with a small number of events. If culture B contained 10 mutants (0.5% of the population), it is likely that this signal would not be visible on the gel, since a single mutant must comprise about 1% of the population to be detected. Thus the normal variation associated with 20 events may explain the failure to reproduce the novel band in MNNG-treated culture B.
Figure 4-56: DGGE analysis of two MNNG-treated complex mutant populations using a 224 bp probe

Triplicate amplifications of MNNG-treated culture B (lanes 5-7), and untreated culture F (lanes 8-10) were performed. MNNG-treated complex population A is shown in lane 4, the novel band seen in culture A in Figure 4-55 is present, however no novel band is present in independent MNNG-treated culture B.

DNA shown in lanes 3-12 was produced by 32 cycles of amplification using Taq DNA polymerase as described in Section 3.6. \(10^{10}\) copies of DNA was hybridized to \(1 \times 10^5\) cpm (5.3 \(\times 10^8\) copies) of single strand, 224 bp probe of specific activity 1000 Ci/mmol; about \(10^9\) copies of HPRT\(_{\text{Munich}}\) DNA was hybridized to 1000 copies of probe. See Section 3.4.2 for probe generation.

DNA was loaded onto a 13%-28% denaturing gradient gel, electrophoresis time was for 16 hours.

Lane 1: 100 cpm ss probe; Lane 2: 1000 cpm ss probe; Lane 3: WT, wild type (THP\(_1\) B, #7, NFC10248); Lane 4: MNNG A, complex MNNG-treated population A (NFC10-248, #1); Lane 5: MNNG B, complex MNNG-treated population B (NFC11-64, #3); Lane 6: MNNG B, complex MNNG-treated population B (NFC11-64, #4); Lane 7: MNNG B, complex MNNG-treated population B (NFC11-64, #5); Lane 8: SPONT F, complex untreated population F (NFC10-322, #4); Lane 9: SPONT F, complex untreated population F (NFC10-322, #5); Lane 10: SPONT F, complex untreated population F (NFC10-322, #6); Lane 11: SPONT E, complex untreated population E (NFC10-248); Lane 12: MUNICH, HPRT\(_{\text{Munich}}\); Lane 13: 0, hybridization control, no DNA.

This experiment is in NFC11-90.
The complex mutant populations did not possess a sufficient number of mutants in the low-temperature melting domain to allow discrimination by DGGE. This was not totally unexpected, since the clone-by-clone analysis of both MNNG- and ICR-191-induced TG^r mutants showed that the GGGGGG sequence in the high-temperature melting domain was the predominate site of mutations. The complex mutant populations were then analyzed using the GC-clamp which allowed resolution of mutations throughout exon 3.

Figure 4-57 shows the examination of the complex mutant populations with the GC-clamp. Genomic DNA was amplified with 20 bp intron primers to produce a 224 bp exon 3 fragment. The 224 bp fragment was re-amplified with GC-rich primers to yield a 332 bp fragment with a GC-clamp at both ends, the 332 bp fragment was restricted at the unique HgiAI site. Restriction with HgiAI yielded two fragments which melted at about 23% and 33% denaturant. The fragment which melted at 23% denaturant contained the GGGGGG sequence and denaturing gradient conditions were such that the fragment which melted at about 33% denaturant was run off the gel.

An unselected, wild type sample is shown in lane 13, this is the amplification control. The complex mutant populations are shown in lanes 1, 5, 6, 7, 9 and 10; note that there is a band 14 mm above the wild type position in all samples. This band seems to be an artifact of the amplification process and may represent a hotspot for polymerase-induced errors.

The wild type, unselected control (lane 13) appeared identical to the
untreated, complex TG⁷ populations E and F. Thus the untreated, complex TG⁷ populations did not exhibit a mutational hotspot.

Three different MNNG-treated cultures were examined in Figure 4-57; my A and B cultures, as well as Monica Cahilly’s MJC culture, are shown. All three MNNG-treated cultures showed a novel band not present in the untreated, complex, TG⁷ populations E and F; likewise, the novel bands present in the MNNG-treated populations were not present in the wild type unselected control. Furthermore, the three MNNG-treated populations (lanes 1, 5 and 6) all showed a novel band at the same gel position. The novel band in the MNNG-treated populations was estimated to represent 5%-15% of the total population. The fraction in the novel band can be estimated by comparing the intensity of the novel band with the intensity of the MNNG-induced clones in lanes 2-4. Ten thousand cpm of double strand radiolabeled probe was hybridized to the complex mutant populations, while 1000 cpm was hybridized to the individual clones. The intensity of the novel band in culture A (lane 8) was about half the intensity of the individual mutant shown in lane 2, thus about 500 cpm was present in the novel band in culture A; the novel band in culture A represented about 5% of the total population. Likewise, the novel bands in culture B (lane 5) and culture MJC (lane 6) were estimated to contain about 1500 cpm each; the novel band in both of these cultures represented about 15% of the population. Clearly, the estimations of the fraction contained in the novel band are rather crude.

The results of the clone-by-clone analysis (Section 4.8) suggested that MNNG-
induced mutations were primarily occurring at base pair locations 293 and 294. Three individual MNNG-induced mutants containing different base pair substitutions at base pair locations 293 and 294 are shown in lanes 2-4 of Figure 4-57. A double strand, radiolabeled, wild type probe was used in this experiment, and both heteroduplexes were visible for mutant RM16 (lane 3). A single band was visible for mutant RM4 (lane 2) and mutant RM20 (lane 4), indicating that both heteroduplexes were at the same location in the denaturing gradient gel.

Importantly, the novel band in the complex mutant populations was aligned with several of the individual clones, which suggested that the novel band in the complex mutant populations may contain the same mutation(s) as have been discovered in the individual clones. The novel band in the complex mutant populations was at the same gel location as the GC → AT substitution at base pair location 293 (mutant RM4, lane 2) and the lower heteroduplex of the GC → TA substitution at base pair location 294 (mutant RM16, lane 3).

The putative PCR noise band common to all PCR amplifications appeared at the same location as the GC → AT substitution at base pair location 294 (mutant RM20, lane 4) and the upper heteroduplex of the GC → TA substitution at base pair location 294 (mutant RM16, lane 3). Therefore, an additional novel band in the MNNG-treated complex mutant populations may be obscured by the PCR noise.

The complex, ICR-191-treated population C is shown in lane 7 of Figure 4-57 and it also shows a novel band not present in the untreated TG^r populations E and
F (lanes 9 and 10); the novel band present in ICR-191-treated culture C is not present in the wild type, unselected amplification control (lane 13). As described in the preceding paragraphs, the fraction of DNA in the novel band may be estimated by comparison to the intensity of the signal in the radiolabeled clones. The novel band seen in the ICR-191-treated population represented about 15% of the population.

The ICR-191 clone-by-clone data suggested that a +G frameshift in the GGGGGGG sequence (base pair location 292-297) was the mutational hotspot for ICR-191. An ICR-191-induced clone, RI202, with a +G frameshift in the GGGGGGG sequence is shown in lane 8. Importantly, the novel band in the ICR-191-treated complex mutant population was at the same location as the +G frameshift in mutant RI202, suggesting that the novel band in the complex mutant population may contain the same +G frameshift.

Also shown in Figure 4-57 is DNA with a 1% HPRT\textsubscript{Munich} fraction (lane 12). One million HPRT\textsubscript{Munich} cells were mixed with \(10^8\) wild type cells, the DNA was extracted and used as a template for amplification. The 224 bp fragment containing exon 3 was subjected to a second round of amplification to add the GC-clamp. 32 cycles of amplification were used to amplify exon 3 from genomic DNA, and another 12 cycles were used to add the GC-clamp; thus the DNA containing a 1% fraction of HPRT\textsubscript{Munich} was amplified 44 cycles. HPRT\textsubscript{Munich} present at a fraction of 1% was not detected with the GC-clamp (lane 12). The double strand, radiolabeled probe with a GC-clamp was made by 13 cycles of amplification of a 224 bp exon 3 sequence; the 224 bp sequence was itself generated by 10 cycles of
amplification. Thus both the DNA containing HPRT\textsubscript{Munich} and the probe itself were subjected to extended cycles of amplification. It is expected that the PCR-generated noise will be directly proportional to the number of cycles, and that the failure to detect a mutant fraction of 1% in this gel was due to the extended number of amplification cycles used to prepare the target DNA and the probe itself.

This thesis presents the first data concerning mutational spectra in an endogenous human gene at the nucleotide sequence level. Mutations in exon 3 of the HPRT gene were examined and the mutations that occurred in this exon possessed sufficient specificity to allow discrimination between MNNG- and ICR-191-treated human cells. The mutational spectra in exon 3 of both MNNG and ICR-191 were dominated by one mutational hotspot. The mutational spectra of these compounds were discerned (1) by an examination of individual, independent, TG\textsuperscript{r} clones, and (2) by an examination of a complex TG\textsuperscript{r} population that consisted of several thousand mutants.
Figure 4-57: DGGE using a GC-clamp to examine complex mutant populations: untreated, MNNG-treated, ICR-191-treated

The DNA shown in this gel has a GC-clamp at the 3' end. $5 \times 10^9$ copies of the DNA shown in lanes 1, 5, 6, 7, 9, 10, 12, and 13 was hybridized to $10^4$ cpm ($3 \times 10^8$ copies) of double strand, radiolabeled probe of specific activity 68 Ci/mmol. About $10^9$ copies of DNA from the individual MNNG- and ICR-191-induced mutants (lanes 2, 3, 4, and 8), and HPRT$_{Munich}$ (lane 11) were hybridized to $10^3$ cpm ($3 \times 10^7$ copies) of double strand probe.

The nature and base pair location of the mutations of the individual MNNG- and ICR-191-induced clones is given. The composition of the complex mutant populations is described in the text.

DNA was loaded onto a 15%-30% denaturing gradient gel and electrophoresis was for 12 hours. Under these conditions, the HgiAl fragment which melts at 33% denaturant was run off the gel. The fragment seen on this gel melts at 23% denaturant and contains the DNA from the HgiAl site at base pair location 288 to the 3' end of exon 3.

Lane 1: MNNG A, MNNG-treated complex mutant population A; Lane 2: RM4, MNNG-induced clone RM4; Lane 3: RM16, MNNG-induced clone RM16; Lane 4: RM20, MNNG-induced clone RM20; Lane 5: MNNG B, MNNG-treated complex mutant population B; Lane 6: MNNG MJC, MNNG-treated complex mutant population generated by Monica J. Cahilly; Lane 7: ICR-191 C, ICR-191-treated complex mutant population C; Lane 8: RI202, ICR-191-induced clone RI202; Lane 9: SPONT E, untreated complex mutant population E; Lane 10: SPONT F, untreated complex mutant population F; Lane 11: MUNICH, HPRT$_{Munich}$; Lane 12: 1% MUNICH, HPRT$_{Munich}$ present at a fraction of 1% in a wild type population; Lane 13: WT, wild type (THP1).

This experiment is in NFC11-150.
4.12 Mutational spectra using individual TG\textsuperscript{r} clones

The set of mutations observed in a treated population consists of two components, the mutants induced by the treatment and the spontaneous mutants that were not induced by the treatment. The final mutational spectrum is the combination of the induced population and the spontaneous population. Therefore the contribution of the spontaneous mutants to the mutational spectrum must be evaluated.

Our laboratory has previously shown that 39\% (33 of 85) of the spontaneous TG\textsuperscript{r} mutants occurring \textit{in vitro} in the TK6 cell line contained large structural gene alterations, almost exclusively deletions, as determined by Southern blotting (Gennett and Thilly, 1988). Sixty one percent (52 of 85) of the spontaneous mutants had a wild type Southern blotting pattern and presumably contained a point mutation or a deletion of less than approximately 50 base pairs; the 95\% confidence limits for the proportion 52/85 are 0.50-0.72.

Nineteen of the spontaneous mutants that showed a wild type Southern blotting pattern were examined by PCR/DGGE, and four point mutants were detected in exon 3. Therefore the overall frequency of spontaneous point mutations occurring in exon 3 was 13\% \[13\% = (4/19) (61\%)\].

The MNNG treatment used for the isolation of the individual TG\textsuperscript{r} clones induced a mutant fraction \(5.0 \times 10^{-5}\) 9-fold greater than the spontaneous background mutant fraction \(5.6 \times 10^{-6}\), while ICR-191 induced a mutant fraction \(5.2 \times 10^{-4}\) about 100 times greater than the background mutant fraction \(4.2 \times \)}
Therefore the probability that an exon 3 mutant in the treated culture was of spontaneous origin was $1.3\% \times 13\% = 13\% \times 0.1$ and $0.13\% \times (13\% \times 0.01)$ for MNNG and ICR-191, respectively, and the contribution of spontaneous mutations to the pattern of point mutations observed after treatment with MNNG and ICR-191 was relatively small. Deletions of spontaneous origin are expected to comprise about 4% and 0.4% of the clones in the MNNG- and ICR-191-induced collection, respectively.

**ICR-191**

Thirteen of 29 ICR-191-treated mutants contained a point mutation detected in HPRT exon 3. Therefore 45% of the ICR-191-induced, TG$^r$, point mutations occurred in exon 3, and the 95% confidence limits for the proportion 13/29 are 0.27-0.64.

Nine of the ICR-191-treated, TG$^r$ clones contained the same mutation, a +G frameshift in a GGGGGG sequence (base pair position 292-297). Also observed were 2 +C frameshifts in a CCC sequence, 1 AT → TA transversion, 1 GC → TA transversion, and one complete gene deletion. Fifteen TG$^r$ mutants were wild type for exon 3. The sequence data is summarized in Table 4-V on page 258.

Calos and Miller (1981) investigated the mutational spectrum of ICR-191 in the lacI gene of *E. coli* using the suppressible nonsense system and discovered that ICR-191 overwhelmingly produced ±1 frameshifts in runs of three or more consecutive guanines, 365 of 378 ICR-191-induced mutations were of this nature. Twenty-three different GGG runs existed in the lacI gene and mutations occurred at all GGG sites, however the frequency of induction at a particular site varied
more than 30-fold. Generally, when a GGG run was flanked by AT base pairs, mutations were poorly induced; when both nearest neighbors were CG base pairs, mutations were well induced. However, clear exceptions existed to this generalization. Some sites showed about the same frequency of +1 and -1 events, while other sites favored +1 frameshifts by a 10:1 ratio, and other sites favored -1 frameshifts by a 7:1 ratio. The most mutable site incurred 17% of the total mutations (64 of 378), and the most mutable site was a GGGG run, leading the authors to suggest that even higher runs of guanines may be more mutable.

Skopek and Hutchinson (1984) examined ICR-191 mutagenesis in the cl gene of lambda prophage and the results were very similar to those of Calos and Miller (1981). Seventeen mutations were sequenced and 16 of 17 involved a +1 frameshift in a run of four or more guanines. Again, certain sites were found to be particularly mutable, with a GATGGGGGCAG site incurring 10 of 17 frameshifts, and a CTTGGGGGGTGA sequence yielding 5 of 17 frameshift mutations; both +1 and -1 frameshift events occurred at these sequences.

The nature of the ICR-191-induced mutations in human HPRT exon 3 is consistent with the mutagenic specificity of ICR-191 observed in prokaryotes. Eleven of 12 point mutations detected in HPRT exon 3 were frameshifts at runs of 3 or more guanines. In particular, a GGGGGG sequence appeared to be highly mutable by ICR-191, this site incurred 31% (9 of 29) of the total mutations; the 95% confidence limits for the proportion 9/29 are 0.16-0.50. The run of six consecutive guanines in exon 3 is the longest monotonous run of GC base pairs in the entire coding frame.
One complete gene deletion, one \( AT \to TA \) and one \( GC \to TA \) base pair substitution were found in the collection of ICR-191-induced mutants. These types of mutations are not associated with ICR-191 mutagenesis in studies to date. As discussed above, the probability that the deletion is of spontaneous origin is about 0.4%, while the probability that the base pair substitutions are of spontaneous origin is about 0.1%.

The ICR-191-induced frameshift events in human HPRT exon 3 showed a strong preference for +1 additions, 11 of 11 frameshifts were of this nature. The reason for this +1 frameshift preference is unknown. The ICR-191 data from prokaryotic studies showed that certain sites demonstrated a strong preference for either +1 or -1 frameshifts, but the reason for the bias at a particular site was unclear; no sequence specificity could be correlated with +1 or -1 events in the Calos and Miller (1981) study. It is possible that the frameshift hotspots in human exon 3 are in a DNA sequence context that favors the +1 event, or it is possible that frameshifts in human cells all have a bias towards the +1 insertion.

ICR-191 contains a substituted acridine ring system, it a planar quinacrine ring with a half-mustard side chain, the structure of ICR-191 is shown in Figure 4-58. Acridine derivatives such as proflavin and 9-aminoacridine bind double-stranded DNA by two different modes (Porumb, 1978 for review). One mode is a strong binding process, that becomes saturated when one acridine molecule is bound per 4-5 base pairs. The other type of interaction is a weak binding mechanism that occurs at high ratios of acridine to DNA, and saturates when one acridine molecule is bound per base pair.
THIS COPY MAY NOT BE FURTHER REPRODUCED OR DISTRIBUTED IN ANY WAY WITHOUT SPECIFIC AUTHORIZATION IN EACH INSTANCE, PROCURED THROUGH THE DIRECTOR OF LIBRARIES, MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

24 : 1
Studies of acridine-DNA complexes have revealed that as the amount of acridine bound to DNA increased, (1) the mass per unit length of DNA decreased (Lerman, 1961; Luzzati et al., 1961), (2) the radius of gyration around the DNA axis increased (Luzzati et al., 1961), that corresponded to increased DNA rigidity, and (3) the viscosity and linear length of the DNA increased (Cohen and Eisenberg, 1969; Drummond et al., 1966). These observations are consistent with the strong binding model in which the acridine molecules intercalate between two adjacent DNA base pairs (Lerman, 1961). The weak binding interaction is thought to involve electrostatic attraction between the positively charged acridines and the negatively charged phosphates of the DNA backbone.

A model for frameshift mutation has been suggested by Streisinger (1966) that (1) strand breakage, (2) localized melting, and (3) slipped mispairing of the DNA. The model proposes that frameshift mutations are generated in gaps in the DNA, and that the gaps might be produced during DNA repair, recombination, or replication. Such structures have the capability of opening base pairs, slipping, and re-forming base pairs improperly. Deletions would be caused by slippage of the template strand, while additions would be caused by slippage of the growing daughter strand. This model was formulated by Streisinger based on the fact that an addition or a deletion of a tandemly repeated sequence often occurred in bacteriophage T₄. The model predicts that frameshift mutations should occur by addition or deletion of single bases from monotonous runs.

Brenner et al. (1961) first demonstrated that acridine compounds caused frameshift mutations by analyzing reversions at the rII locus of bacteriophage T₄.
Streisinger suggested that intercalation of acridines into mispaired regions of the DNA might increase the half-life of the structures arising by slippage of one strand. Lerman (1961, 1963) has shown that intercalation stabilized the double helix with regard to thermal denaturation. Thus intercalation could increase the probability that the mispaired configuration would exist at the moment of DNA synthesis.

However, there exists a poor correlation between intercalation ability and mutagenicity. The original Streisinger model predicted that any intercalating agent that stabilized the helix should promote mutagenesis, but many compounds that strongly intercalate into DNA have not been found to be mutagenic (Drake, 1970). This suggests that intercalation per se may not be sufficient for mutagenesis.

Streisinger and Owen (1985) have suggested a modified version of the frameshift model in which the mutagen promotes stabilization of an extrahelical base by stacking on one or both sides of the extrahelical base, in effect forming a sandwich around the looped out extrahelical base. This model predicts that agents that cannot intercalate into DNA, but that can stack, should still exhibit frameshift mutagenesis. The authors suggest that the greater than first order dependence of mutagen concentration may be due to the participation of several mutagen molecules in the extrahelical stabilization.

Newton et al. (1972) showed that ICR-191 mutagenesis was associated with DNA replication. *E. coli* cultures were synchronized and treated at various times after the start of a new round of DNA synthesis. Extensive mutagenesis was
observed in a particular gene when ICR-191 was added at about the time a replication fork was passing through the gene. *Uvr*+ *E. coli* showed a sharp peak in mutagenesis when ICR-191 was present as the replication fork was passing through the gene scored for mutation, however, ICR-191 was equally mutagenic to *uvr*− cells at all times during the cell cycle. *Uvr*− cells are defective in excision repair and cannot remove bulky DNA adducts. This result is consistent with a model that assumes that ICR-191 binds covalently to the DNA. The *uvr*− cells may be unable to remove the bound ICR-191 and the adduct would be present when the next replication fork passes through the modified DNA; the *uvr*+ cells may remove the adduct by excision repair such that the mutagenicity is related to the DNA repair time available before the next replication fork passes the adduct. Covalent binding of ICR-191 may also explain the decrease in mutagenicity observed when the reactive -Cl group in the half-mustard side chain was replaced by a nonreactive -OH group (Creech et al., 1972).

Mattes et al. (1986) recently investigated the alkylation patterns of several nitrogen mustards, including quinacrine mustard. The structure of quinacrine mustard is closely related to that of ICR-191, ICR-191 is essentially the half-mustard derivative of the full-mustard quinacrine mustard. The side chain of quinacrine mustard is -NHCH(CH$_3$)$_2$(CH$_2$)$_2$N(CH$_2$CH$_2$Cl)$_2$, while the side chain of ICR-191 is -NH(CH$_2$)$_3$NHCH$_2$CH$_2$Cl. Mattes et al. (1986) found that all eight nitrogen mustards examined primarily alkylated the N$^7$ position of guanine and furthermore, quinacrine mustard showed the strongest alkylation patterns at runs of consecutive guanines. Interestingly, uracil mustard did not show the same
Figure 4-58: Structure of ICR-191
preference for alkylation of consecutive guanines which suggested that non-covalent interactions were occurring between the non-alkylating moieties of the drugs and the DNA prior to covalent reaction. The authors found that the non-alkylating moiety to which the mustard group is attached strongly affected the sequence selectivity of covalent binding to the \( N^7 \) position of guanine.

It is probable that both the alkylating half-mustard side chain and the quinacrine ring of ICR-191 are responsible for the mutagenic specificity and mutagenic potency of ICR-191. The quinacrine ring may preferentially intercalate into guanine runs and the reactive side chain may then alkylate the DNA, or the side chain may preferentially alkylate the \( N^7 \) position of guanine and the quinacrine ring may then intercalate. Either reaction mechanism would present the same adduct to the DNA polymerase or DNA repair enzymes.

All frameshift mutations in exon 3 are expected to confer \( TG^f \) since these mutations are expected to completely abolish the activity of the HPRT enzyme. If ICR-191 is overwhelming producing frameshift mutations, then nearly every ICR-191-induced mutation should produce a \( TG^f \) phenotype. Thus the relationship between ICR-191-induced mutations and observed \( TG^f \) mutants may be nearly one-to-one.

**MNNG**

The examination of the independent, \( TG^f \), MNNG-induced clones indicated that 16% (5 of 32) of the mutants had a point mutation in exon 3; the 95% confidence limits for the proportion 5/32 is 0.05-0.33. All mutations were base pair substitutions at GC base pairs (see Table 4-V on page 258). Four MNNG-induced
mutations occurred in the GGGGGG (bp 292-297) sequence at base pair locations
293 and 294, and one MNNG-induced mutant was observed at bp 328. MNNG
treatment produced 3 GC → AT transitions and 2 GC → TA transversions.

Studies that have examined the specificity of MNNG mutagenesis in
prokaryotes have found that MNNG produced almost exclusively GC-to-AT
transitions. MNNG mutagenesis is believed to involve the formation of O\(^6\)-methyl
guanine, which mispairs with thymine during replication, leading to GC → AT
substitutions. Nonsense mutations in the lacI gene of E. coli induced by MNNG
were found to be predominately GC-to-AT (Coulondre and Miller, 1977b). 292/303
of the amber mutations were GC-to-AT, and 210/215 ochre mutations were also
GC-to-AT. The distribution of mutations was found to be site-dependent.

The specificity of MNNG mutagenesis was examined in the bacteriophage
P22 mnt repressor gene (Lucchesi et al., 1986). Twenty nine of 30 MNNG-induced
mutations were GC-to-AT substitutions, with two sites accumulating 9 of 30 and 7
of 30 mutations.

The effect of neighboring base pairs on MNNG mutagenesis has been
examined by Burns et al. (1987). One hundred and sixty seven MNNG-induced
mutants were sequenced and 164/167 were found to be GC-to-AT transitions.
Considerable site variation was found in the mutation frequency, and a high rate of
mutation at a particular guanine was associated with a 5' purine. A guanine
residue was eight-fold more mutable when preceded 5' by a purine rather than a
pyrimidine. All of the mutational hotspots involved 5' Pu-G, while all of the cold
sites consisted of 5' Py-G.

Sixty percent (3 of 5) of the MNNG-induced mutations detected in exon 3 were GC → AT base pair substitutions. However, the number of exon 3 mutants examined was small and the 95% confidence limits for the proportion 3/5 range from about 0.20-0.95. Clearly, more mutants would have to be examined to estimate the fraction of GC → AT transitions with reasonable degree of certainty. The nature of mutations induced by MNNG in human HPRT exon 3 is not inconsistent with the prokaryotic data for MNNG.

Forty percent (2 of 5) of the MNNG-induced mutations detected in exon 3 were GC → TA transversions. This transversion has not been associated with MNNG mutagenesis in previous studies. It is possible that the transversions are of spontaneous origin. As discussed above, the probability that an exon 3 point mutation is of spontaneous origin is about 1%.

Two complete gene deletions were found in the collection of 31 MNNG-induced mutants, however, MNNG has not been reported to cause deletions. As discussed previously, the probability that each MNNG-induced deletion mutant is of spontaneous origin is about 4%.

The MNNG-induced mutations reported in this thesis were found to exhibit a mutational hotspot, about 13% (4 of 32) of the TG' mutations occurred in the GGGGGG sequence at base pair 293 and 294 (base pair position 292-297) and the mutations affected a single amino acid, Gly60. The 95% confidence limits for the fraction 4/32 are 0.03-0.31.
The amino acids coded by the GGGGGG sequence are

68 69 70  \text{ amino acid }
\text{ LysGlyGly }
\text{ caaGGGGGcta (codon for Gly69 is underlined) }
\text{ 292 297  DNA sequence }

Thus 4 of the 5 MNNG-induced mutants that were detected in exon 3 occurred affected Gly69. Substitutions of Gly69 by arginine, valine or glutamic acid produced a TG\textsuperscript{r} phenotype, suggesting that Gly69 may be critical for HPRT activity. This amino acid residue is strongly conserved among phosphoribosyltransferases. Davidson \textit{et al.} (1988) aligned (1) human, hamster and mouse HPRT, (2) human, mouse and \textit{E. coli} APRT, and (3) \textit{E. coli} XGPRT to give the maximal overlap of predicted secondary structure by the method of Chou and Fasman (Chou and Fasman, 1978), and a glycine was present in all seven enzymes at the equivalent position of human HPRT Gly69. Glycine is often present in \( \beta \)-turns, and when amino acid residues 59-79 were analyzed for \( \beta \)-turns by the method of Chou and Fasman (Chou and Fasman, 1978), the tetrapeptide Leu67 Lys68 Gly69 Gly70 showed the highest \( \beta \)-turn probability. It is possible that Gly69 helps form a \( \beta \)-turn, and that disruption of this \( \beta \)-turn destroys the catalytic ability of the HPRT enzyme.

As we have seen, MNNG appears to be inducing mainly base pair substitutions in HPRT. However, only a subset of the MNNG mutations are expected to confer the TG\textsuperscript{r} phenotype. TG appears to be an extremely stringent selective agent, in that only cells essentially devoid of HPRT activity will survive
Thus all base pair substitution mutations that produce a stop codon in exon 3 should produce a $T^r$ cell, while only a restricted set of missense mutations, those mutations that severely impair the catalytic ability of the enzyme, are expected to confer $T^r$.

Mutations were found only at two guanines in the GGGGGGG sequence, however it is probable that MNNG is producing other mutations in this sequence that do not confer $T^r$. For example, any mutation occurring at the third codon position of Gly69 will not be detected since the codon is completely degenerate at this position (GGX = Gly). The relationship between MNNG-induced mutations and observed $T^r$ mutants is not expected to be one-to-one.

**Spontaneous Mutations**

Spontaneous mutations were examined so that their contribution to the mutational spectra of MNNG and ICR-191 could be evaluated. About 39% (33 of 85) of the spontaneous $T^r$ mutants occurring in vitro in the TK6 cell line contained large deletions (Gennett and Thilly, 1988). Thus 61% of the spontaneous mutants presumably contained a point mutation.

Nineteen of the spontaneous mutants that presumably contained a point mutation were examined by PCR/DGGE, and four point mutants were detected in exon 3. Therefore the overall frequency of spontaneous point mutations occurring in exon 3 was 13% [$13\% = (4/19) (61\%)$]. The spontaneous $T^r$ population comprised about 10% and 1% of the MNNG- and ICR-191-treated populations, respectively. Therefore, the probability that a point mutation in exon 3 is of
spontaneous origin is about 1% and 0.1% for the MNNG- and ICR-191-treated populations, respectively.

Two of the four spontaneous exon 3 mutants could not be sequenced by the September dissertation deadline. Two mutants were sequenced, and one GC → TA transversion at base pair position 378 which produced an Asp → Tyr amino acid substitution, and one +1 frameshift in a TTT at base pair location 308-310 was found. No mutations in the MNNG- and ICR-191-treated populations were found at the base pair location of the two spontaneous mutations, again suggesting that the contribution of the spontaneous TG\textsuperscript{r} mutants to the pattern of chemically-induced mutations is small.

Spontaneous point mutagenesis in \textit{E. coli} and mammalian cells has shown a complex pattern of base pair substitutions and frameshifts (Ashman and Davidson, 1987; Schaaper \textit{et al.}, 1986; Coulondre and Miller, 977b; Coulondre \textit{et al.}, 1978). The original Streisinger model (Streisinger \textit{et al.}, 1966) predicted that frameshift mutations would occur at monotonous runs of nucleotides, and it is interesting to note that the +1 spontaneous frameshift mutation occurred in a TTT sequence. The sample size is too small to draw any firm conclusions concerning the nature of spontaneous mutations in exon 3.

\textbf{Point mutations that produce the TG\textsuperscript{r} phenotype}

Exon 3 of HPRT is reported to code for both the catalytic sites of the enzyme (Wilson \textit{et al.}, 1983). It is very probable that any frameshift mutation in exon 3 will abolish the activity of the enzyme and thus produce a TG\textsuperscript{r} phenotype. Twelve
frameshifts at three different locations were observed in the examination of the TG$^r$ clones. All frameshift were one base pair additions in runs of three or more monotonous nucleotides.

Eight base pair substitutions were found in the collection of TG$^r$ clones. Two base pair substitutions produced a stop codon, and one base pair substitution probably affected an RNA splice site. The probable splicing mutant contained a base pair substitution at the extreme 5' nucleotide of exon 3; the 5' nucleotide of the exon often forms part of the consensus sequence for RNA splicing (Breathnach and Chambon, 1981).

The five remaining base pair substitutions affected only two amino acids, Gly69 and Asp97. Substitution of Gly69 by arginine, valine or glutamic acid produced a TG$^r$ phenotype, as did substitution of Asp97 by tyrosine. As discussed above, Gly69 is highly conserved in phosphoribosyltransferases and may form part of a $\beta$-turn. Substitution of Asp97 by tyrosine also produced the TG$^r$ phenotype. Aspartic acid is a negatively charged amino acid with a small side chain while tyrosine is an uncharged amino acid with a large side chain, and this substitution apparently greatly reduced the catalytic ability of the enzyme.

Frameshift mutations, base pair substitutions which produced a stop codon, and a mutation that probably affected an RNA splicing site produced a TG$^r$ cell. These mutations are predicted to completely abolish the activity of the HPRT enzyme. Base pair substitutions which affected Gly69 and Asp97 also produced a TG$^r$ phenotype. Missense mutations at a restricted set of nucleotides are predicted
to confer TG\textsuperscript{r}. Only those missense mutations that greatly reduce the catalytic
ability of the enzyme are predicted to confer TG\textsuperscript{r}. 
4.13 Mutational spectrum using complex mutant populations

The results from the clone-by-clone analysis indicated that a mutational hotspot for both ICR-191 and MNNG existed in the GGGCGGG sequence (base pair location 292-297). Thirteen percent (4 of 32) of the MNNG-induced mutations occurred at base pair locations 293 and 294, the 95% confidence limits for the proportion 4/32 are 0.03-0.31. Thirty one percent (9 of 28) of the ICR-191 induced mutations were +1 frameshifts in the GGGGG sequence, the 95% confidence limits for the proportion 9/28 are 0.17-0.52.

Reconstruction experiments shown in Figures 4-53 and 4-54 demonstrated that HPRT<sub>Munich</sub> could be detected by PCR/DGGE if it comprised 1% of the population; a 1% HPRT<sub>Munich</sub> fraction was detected using 224 bp probe with no GC-clamp. The sensitivity was reduced when the GC-clamp was added (see Figure 4-57), HPRT<sub>Munich</sub> could not be detected when present at a 1% fraction. As discussed previously, the addition of the GC-clamp required addition amplification cycles, which presumably increased the background noise. Although the absolute sensitivity was not determined for the GC-clamp construction, it seemed possible that a mutant present at several percent could be detected. Therefore, the sensitivity and expected mutant fraction at the GGGGG sequence indicated that the analysis of a complex mutant population was feasible.

The denaturing gradient gel with the complex mutant populations is shown in Figure 4-57 on page 285. I will summarize the conditions used to generate the complex mutant populations and the number of original TG<sup>r</sup> mutants present in
each culture. The complex mutant populations are characterized as follows:

**Untreated populations**

- Untreated cultures labeled $E$ and $F$ each containing $2.5 \times 10^8$ cells were generated by myself.

- These cultures are the control populations for my MNNG-treated complex mutant population (labeled $A$ and $B$) and my ICR-191-treated complex mutant population (labeled $C$).

- The mutant fraction in culture $E$ was $3.2 \times 10^{-6}$. The mutant in culture $F$ could not be determined.

- The complex mutant population in culture $E$ was derived from 800 TG^R cells.

**MNNG-treated populations**

- Cultures labeled $A$ and $B$ were generated by myself.

  - $2.5 \times 10^8$ cells were treated for 1 hour with 10 ng/ml MNNG. The survival was 0.85, thus $2.1 \times 10^8$ cells survived treatment.

  - The induced mutant fraction in culture $A$ was $6.1 \times 10^{-6}$. The untreated, background mutant fraction was $3.2 \times 10^{-6}$. The mutant fraction in culture $B$ could not be determined, but the survival after treatment was the same in both cultures.

  - The complex mutant population in culture $A$ was derived from 1275 induced and 675 untreated TG^R cells. 65% of the TG^R cells in the culture were induced by treatment, 35% of the TG^R cells were not induced by treatment with MNNG.

- Culture labeled $MJC$ was generated by Monica J. Cahilly.

  - $5 \times 10^8$ cells were treated for 1 hour with 15 ng/ml MNNG. Survival was 0.11, thus $5.5 \times 10^7$ cells survived treatment.
The induced mutant fraction in culture MJC was $5.0 \times 10^{-5}$ and the untreated mutant fraction was $5.6 \times 10^{-6}$.

The complex mutant population in culture MJC was derived from 2750 induced and 330 untreated TG$^r$ cells. 90% of the TG$^r$ clones in the culture were induced by MNNG treatment.

**ICR-191-treated population**

- ICR-191-treated population labeled C was generated by myself.

- $2.5 \times 10^8$ cells were treated for 24 hours with 500 ng/ml ICR-191. Survival was 0.5, thus $1.3 \times 10^8$ cells survived treatment.

- The induced mutant fraction in culture $C$ was $7.5 \times 10^{-5}$. The untreated, background mutant fraction was $3.2 \times 10^{-6}$.

- ICR-191-treated population C was derived from 9725 induced and 420 untreated TG$^r$ cells. 96% of the TG$^r$ cells in the culture were induced by treatment.

Culture MJC was used to generate the independent, TG$^r$ clones, and culture MJC was also examined as a complex, TG$^r$ population. Thus a direct comparison between the independent, MNNG-induced, TG$^r$ clones and MNNG-induced, TG$^r$, complex mutant population MJC is possible.

The intensity of the novel band in MNNG-treated culture A was greater than the intensity of the novel band in duplicate culture B. The reason for the discrepancy is not known. Both cultures were treated at the same time with the same stock solution of MNNG. Unfortunately, as discussed previously, the portion of culture B to be used to determine the MF was contaminated, and the MF could
not be determined.

The novel bands were estimated to represent 5%-15% of the MNNG-treated, complex TG\(^{r}\) population. MNNG-treated complex populations A and B had about 1275 induced and 675 spontaneous TG\(^{r}\) mutants (1950 total). Therefore, 100-300 mutations occurred at the MNNG hotspot in MNNG-treated culture A. Likewise, 150-450 mutations occurred at the MNNG hotspot in Monica Cahilly's population; her culture had 2750 induced and 330 spontaneous mutants (3080 total).

The same MNNG hotspot was visible in two independent experiments performed by two investigators; this is always reassuring. Two different levels of MNNG mutagenesis also produced the same hotspot. The induced mutant fraction in my culture A was 6.2 \(\times\) 10\(^{-6}\), while the induced mutant fraction in Monica Cahilly's culture MJC was 5.0 \(\times\) 10\(^{-5}\). Thus nearly a 10-fold difference in mutation induction produced the same mutational hotspot.

The induced mutant fraction in culture A was only two-fold above the background mutant fraction of 3.2 \(\times\) 10\(^{-6}\). The total mutant fraction observed in culture A was about 9 \(\times\) 10\(^{-6}\), which is typical of the mutant fractions found in peripheral T-lymphocytes isolated from individuals. It is possible that the analysis of complex mutant populations can be performed at levels of mutagenesis that are relevant to human exposures, unlike many experiments in which the cells are given massive mutagenic doses.

The results from the MNNG-induced complex mutant populations are consistent with the data from the clone-by-clone analysis. The clone-by-clone
analysis indicated that 13% of the MNNG-induced mutants occurred in the GGGGGG sequence, and the 95% confidence limits were 0.03-0.31. The complex mutant analysis indicated that the novel band represented 5%-15% of the TG\(^r\) mutants.

It is possible that another hotspot is present in the MNNG-treated, complex mutant populations. As discussed previously, the putative Taq DNA polymerase PCR noise is at the same gel location as the G \(\rightarrow\) A transition at base pair 294; the upper heteroduplex of the G \(\rightarrow\) T transversion at base pair 294 is also at the same gel location as the PCR noise. Thus it is possible the PCR noise is masking some of the MNNG-induced mutations.

The clone-by-clone analysis indicated a strong hotspot for ICR-191-induced mutations existed in the GGGGGG sequence, almost one-third (9 of 29) of all the clones had an identical +G frameshift. The TG\(^r\) clones were generated by Monica Cahilly, while the complex, ICR-191-treated population \(C\) shown in Figure 4-57 was generated by myself. The level of mutation in Monica Cahilly's ICR-191-treated culture that was used to generate the TG\(^r\) clones was different than the level of mutation in my ICR-191-treated, complex population \(C\).

Monica Cahilly's ICR-191-treated culture that produced the TG\(^r\) clones had an induced MF of 5.2 x 10\(^{-4}\) with a background MF of 4.2 x 10\(^{-6}\). ICR-191-treated, complex population \(C\) had an induced MF of 7.5 x 10\(^{-5}\) with a spontaneous MF of 3.2 x 10\(^{-6}\). The induced mutant fraction ICR-191-treated, complex population \(C\) was about 7-fold less than the induced mutant fraction in Monica Cahilly's culture
that was used to generate the TG\textsuperscript{r} clones.

The results from the analysis of the ICR-191-treated complex mutant culture is in basic agreement with the data from the clone-by-clone analysis. The novel band in the ICR-191-treated culture in Figure 4-57 was estimated to represent roughly 15\% of the population. Data from the clone-by-clone analysis indicated that 31\% of the population contained a +G frameshift, the 95\% confidence limits are 0.16-0.50. The novel band in the complex mutant population was not measured with a sufficient degree of precision to conclude that the mutational spectrum seen in clone-by-clone analysis was in fact different from the spectrum seen in the complex mutant population. Further experiments are necessary to determine if the fraction of mutations at the ICR-191 hotspot in the clone-by-clone collection is different from the fraction of mutations at the hotspot in the complex mutant population.

It should be recalled that the level of mutation in Monica Cahilly's ICR-191-treated culture that produced the TG\textsuperscript{r} clones was 7-fold higher than the level of mutation in my complex, ICR-191-treated population C. It is possible that the level of induced mutation may affect the mutational spectrum. The phenomenon of nonlinear dose-mutation curves have been widely observed. Induction of repair enzymes and saturation of repair capacity have been suggested as possible causes of the nonlinear dose-mutation curves. Different mutagenic doses could affect the proportion of mutations at a given base pair and could explain the fact that the observed fraction in the novel band in complex mutant population C was about
15%, while the observed fraction of +G frameshifts in the clone-by-clone analysis was 31%. Further experiments are necessary to clarify this point.

The complex mutant populations were previously examined with a probe that could detect mutations in the low-temperature melting domain, and no reproducible novel bands were seen. Only when the mutant populations were examined with the GC-clamp were the hotspots visible. Therefore, the mutations that gave rise to the novel bands in both the MNNG- and ICR-191-treated complex mutant populations must lie in the original high-temperature melting domain of exon 3 from base pair 288 to about base pair 300. The HgiAI fragment that revealed the novel bands contained the entire original low-temperature melting domain (base pair 301-403) and about 13 base pairs of the original high-temperature melting domain (base pair 288-300). The novel band was visible using a probe that could detect mutations in the original low-temperature melting domain and 12 base pairs of the original high-temperature melting domain. It is probable that the novel bands in the complex mutant populations contain the same mutations detected in the clone-by-clone approach.

The ability to examine an entire complex TG\textsuperscript{T} population derived from several thousand mutants greatly facilitates the speed with which a mutational spectrum can be obtained. The clone-by-clone mutational spectra for MNNG and ICR-191 was obtained by screening about 80 individual samples- 20 spontaneous, 30 MNNG-treated, and 30 ICR-191-treated clones. Much of the same information can be obtained by screening three complex mutant populations: untreated, MNNG-treated, ICR-191-treated. Dr. Phouthone Keohavong, a postdoctoral
associate in our laboratory, has been able to sequence DNA from minor bands in a denaturing gradient gel. Thus with the ability to sequence the novel bands and with the ability to detect minor hotspots by using a high fidelity DNA polymerase, it should be possible to rapidly obtain a mutational spectrum of a variety of agents.

The level of precision increases dramatically when analyzing several thousand mutants. 4 of 32 MNNG-induced mutations discovered in the clone-by-clone analysis affected the same amino acid and this was considered to be a mutational hotspot. The statistical error associated with 4 events in a sample size of 32 is large, the 95% confidence limits are 0.3-0.31. However, if the novel band in a complex mutant population can be measured with a reasonable degree of accuracy, the novel band should represent 100-1000 mutants and the statistical error associated with sampling a small number of mutants would be greatly reduced.

This dissertation has demonstrated that a mutant present at a frequency of about 1% in a complex population can be detected. This level of sensitivity should be sufficient to discover the TG7 mutational spectrum of a variety of mutagenic agents. Different mutational hotspots would be revealed by their location on a denaturing gradient gel, and the frequency of a mutant species would be proportional to the intensity of its autoradiographic band. The DNA could be recovered from the denaturing gradient gel and the sequence alteration at the mutational hotspots could be determined.

The sensitivity could be further enhanced by use of a high fidelity DNA
polymerase in the PCR process. The DNA polymerase used in this work, Taq DNA polymerase, has a rather high error rate (Saiki et al., 1988), about $2 \times 10^{-4}$/error/bp/cycle. Use of this enzyme yielded a rather high background, that was visible as a series of faint, non-wild type bands on a denaturing gradient gel. Use of a DNA polymerase with a lower error rate should reduce the noise associated with the PCR process and thus increase the sensitivity of the system.

**4.14 Denaturing gradient gel melting characteristics of several exon 3 mutants**

Denaturing gradient gel electrophoresis showed remarkable sensitivity in discriminating between different mutations that occurred in the GGGGGG sequence in exon 3 (base pair 292-297). 3 different base pair substitutions and a +1 frameshift were resolved from one another as shown in Figure 4-59. A double-strand, radiolabeled, wild type probe was hybridized to mutant DNA, forming two mismatched heteroduplexes; in one case, both heteroduplexes were visible.

The base pair substitutions that occurred in the GGGGGG sequence were a GC → AT transition at base pair 293, a GC → AT transition at base pair 294, and a GC → TA transversion at base pair 294. The GC → AT transitions at adjacent base pair locations were separated by about a millimeter. Both heteroduplexes of the GC → TA transversion were visible, while a single band was seen for both of the GC → AT transitions.

The destabilization of the heteroduplex produced by a +1 frameshift in the GGGGGG sequence was less than the destabilization produced by any other
Figure 4-59: Denaturing gradient gel showing a +1 frameshift and three different mutations in the GGGGGG sequence

The base pair location of the GGGGGG sequence is 292-297. The mutation and base pair location of several MNNG- and ICR-191-induced mutants are given. The mutants in lanes 2, 3, 4 and 6 contain a +1 frameshift. The mutants in lanes 7-10 contain a single base pair substitution at position 293 or 294.
mutation detected in this study. The $6G \rightarrow 7G$ frameshift lowered the melting temperature of the mismatched heteroduplex about $0.10^\circ C$ (1.8 mm shift, 30% change in denaturant in a 17 cm gel, 1% denaturant = $0.313^\circ C$), while the base pair substitutions often produced a gel shift 3 to 5 times greater than the $6G \rightarrow 7G$ frameshift.

Two other frameshift mutations were also detected, (1) a $3C \rightarrow 4C$ frameshift at base pair location 276-278 that lowered the melting temperature of the heteroduplex by $0.32^\circ C$ (5.5 mm gel shift), and (2) a $3T \rightarrow 4T$ mutation at base pair location 308-310 that lowered the melting temperature of the heteroduplex by a $0.38^\circ C$ (6.5 mm gel shift).

The GGGGGGGG sequence at base pair location 292-297 is in a portion of the molecule that is predicted to be relatively sensitive to base pair substitutions. The calculated alteration in the melting temperature for a mismatch at base pair 292-297 is about $0.50^\circ C$ (see Figure 4-27). The calculated alteration in the melting temperature for a mismatch at the other locations at which frameshifts were detected, base pair 276-278 and 308-310, is about $0.65^\circ C$ and $0.53^\circ C$, respectively. Therefore, the small perturbation in heteroduplex melting behavior seen for the $6G \rightarrow 7G$ frameshift is not due to a localized region of greater heteroduplex stability. It is possible that a frameshift mutation will destabilize a shorter run of monotonous nucleotides to a greater extent than a longer run of monotonous nucleotides.

It is interesting to note that the 169 base pair probe resolved a base pair
substitution at the 5' end of the 169 bp sequence, in what appears to be the high-temperature melting domain. As described in Myers et al. (1987), the formation of a mismatch can alter the melting domains such that mutations may be detected in a portion of the high-temperature melting domain.

The melting algorithm predicted that mismatches at the 5' end of the 169 base pair probe would be detected, this is shown in Figure 4-60. Base pair substitutions from the 5' end to about position 255 should be resolved.

The algorithm predicted that a base pair substitution at base pair 248 would be detected with the 169 base pair probe. Mutant RI209 contains an AT → TA transversion at base pair 248 and this mutant produced a well-focused band on a denaturing gradient gel when the 169 base pair probe was used. The base pair substitution in mutant RI209 occurred 9 base pairs from the last base paired nucleotides of the mismatched heteroduplex.

The algorithm also predicted (see Figure 4-61) that mutant RI209 which contains a base pair substitution at location 248 will not be detected if the 224 bp exon 3 fragment was used as a probe. Mutant RI209 was PCR-amplified and examined as a 224 base pair fragment, and RI209 appeared wild type, that is, there was no shift seen in the denaturing gradient gel (data not shown). This occurred because the mismatch was now 49 base pairs from the 5' end of the 224 bp molecule and did not destabilize the high-temperature melting domain.
**Figure 4-80:** Calculated destabilization produced by mismatches at each base pair location of the 169 base pair probe

The wild type homoduplex melting map is shown as a solid line (Y axis), and the difference in gradient level in degrees Centigrade between the wild type homoduplex and a mismatched heteroduplex is shown as a dotted line (Y2 axis). The coding region of HPRT exon 3 is from base pair position 220-403. A difference of 0.1°C should produce a 2 mm shift using typical denaturing gradient gel conditions of a 30% change in gradient over a 17 cm gel, and 1% denaturant = 0.313°C.

The program SQHT was run with the following values: destabilization = 50°C, retardation length = 75, velocity zero = 1.8. The value for the heteroduplex is for a 5 hour electrophoresis time.
Calculated melting profile of TaqI/HkiI exon 3 fragment and calculated heteroduplex destabilization at each base pair location.

Retardlength=75, Velzero=1.8, Destab=50, t=5 hr
Figure 4-61: Calculated destabilization produced by mismatches at each base pair location of the 224 base pair probe

The wild type homoduplex melting map is shown as a solid line (Y axis), and the difference in gradient level in degrees Centigrade between the wild type homoduplex and a mismatched heteroduplex is shown as a dotted line (Y2 axis). The coding region of HPRT exon 3 is from base pair position 220-403. A difference of 0.1°C should produce a 2 mm shift using typical denaturing gradient gel conditions of a 30% change in gradient over a 17 cm gel, and 1% denaturant = 0.313°C.

The program SQHT was run with the following values: destabilization = 50°C, retardation length = 75, velocity zero = 2.0. The value for the heteroduplex is for a 5 hour electrophoresis time.
Calculated melting map of wild type homoduplex and difference in melting temperature between the wild type homoduplex and a mismatched heteroduplex.
Chapter 5

Suggestions for future research

The ability to obtain mutational spectra from a TG^† complex mutant population is extremely powerful, it decreases the experimental time and increases the precision of the system by several orders of magnitude. It is now feasible to quickly examine the pattern of in vitro mutations produced by a variety of mutagens. Furthermore, Dr. Phouthone Keohavong, a postdoctoral associate in our laboratory, has shown that it is possible to sequence directly the minor bands seen in a denaturing gradient gel. Thus it is possible to obtain sequence information concerning the novel DGGE bands which may appear in a treated population.

A variety of mutagens could be examined for the nature of mutations produced in HPRT exon 3, and the techniques developed in this thesis could also be applied directly to the other HPRT exons. Nine amplifications would be required to examine the entire coding region, however, it may be possible to carry out several amplifications simultaneously by adding two or three sets of primers. It would be desirable to use primers which are complementary to the intron sequence 10-30 base pairs from the end of each exon, in this way, mutations in the intron which affect the splicing of the mRNA could be detected. The primers used in this thesis were complementary to the intron sequence immediately flanking exon 3 and any intron mutation would not have been detected.

Another alternative would be to use HPRT mRNA as a substrate for PCR,
the first DNA strand is synthesized by reverse transcriptase and then PCR proceeds normally (Simpson et al., 1988). Mutants which produce little or no mRNA, mutants that have a greatly altered mRNA size, and mutants that have an unstable mRNA would probably not be detected. The entire coding region could be obtained as a single PCR-product, and the coding region could be re-amplified with three sets of different GC-rich primers, thus generating three 200 base pair molecules which contain the entire HPRT coding region. In this way, the entire coding region could be quickly examined for mutations.

The use of reverse transcriptase would be expected to introduce additional noise into the system, since the error rate for this enzyme is quite high, about $10^{-3}$ error/bp (Loeb and Kunkel, 1982). Reverse transcriptase has about a 10-fold higher error rate than the DNA polymerase used in this study, Taq polymerase (Saiki et al., 1988). Thus one PCR cycle using reverse transcriptase is expected to produce about as much polymerase error noise as 10 cycles using Taq DNA polymerase. The high error rate of reverse transcriptase may be partially overcome by the fact that there are 2-5 copies of HPRT mRNA per cell (Jolly et al., 1983), and therefore fewer PCR cycles would be needed to achieve $10^{10}$-$10^{11}$ copies of PCR product. Thus the noise produced by 30 cycles of amplification with Taq polymerase may about equal the noise produced by using reverse transcriptase for one cycle followed by Taq polymerase for 20 cycles.

TG$^T$ T-lymphocytes can be cloned from human peripheral blood (Albertini et al., 1985; Turner et al., 1985), and it would be possible to examine these cells for patterns of HPRT mutations. Individuals heavily mutated by the same agent may
show a similar pattern of mutation. However, due to cloning efficiencies of 10%-50%, a blood sample of 50 ml generally yields only 20-30 TG$^r$ clones (R. Albertini, personal communication), and discerning a mutational hotspot from such a small number of mutants would be difficult. Therefore, it may be necessary to devise different methods to obtain mutational spectra from people.

Very recent work by Dr. Phouthone Keohavong has demonstrated that PCR/DGGE can detect 100 HPRT$_{Munich}$ mutant cells in a background of $10^8$ wild type cells. The ability to detect a mutant fraction of $10^{-6}$ may permit the study of unselected mutagenesis. For example, the MNNG hotspot presented in this thesis is expected to be present at a frequency of $10^{-5}$-$10^{-6}$ if the MNNG-induced mutant fraction is $10^{-4}$-$10^{-5}$; this is within the limit of detection and the MNNG hotspot should be detected in an unselected population. The TG$^r$ mutant fraction in humans is often greater than $10^{-5}$ and it may be possible to determine unselected mutational spectra in people.

HPRT is a single copy gene. It may be possible to examine the unselected mutational spectra of a gene which is repeated several hundred times in the human genome. Use of a repeated sequence would reduce the sample size necessary to obtain a given number of copies of the target DNA. It is also possible that a cell may be able to tolerate more mutations in the repeated sequences than in single copy sequences since a single mutation in one of the repeated genes is not expected to be a lethal event. A sequence which is repeated 100-1000 times per genome and which is highly conserved may serve as mutational targets; mitochondrial genes, ribosomal RNA genes, and histone genes could be investigated.
The word count of this thesis is 50,384 (not including the Appendix), which translates to 22.39 words per day of graduate school. A frightening statistic.
Chapter 6

Appendix

6.1 Publications

6.1.1 Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT<sub>Munich</sub>


**Resolution of a Missense Mutant in Human Genomic DNA by Denaturing Gradient Gel Electrophoresis and Direct Sequencing Using In Vitro DNA Amplification: HPRT<sub>Munich</sub>**

Neal F. Cariello, Jamie K. Scott, T Alexandra G. Kat, William G. Thilly, and Phouonthone Keohavong

Department of Applied Biological Sciences, Massachusetts Institute of Technology, Cambridge

**Summary**

The combination of denaturing gradient gel electrophoresis (DGGE) and in vitro DNA amplification has allowed us to (1) localize a DNA mutation to a given 100-bp region of the human genome and (2) rapidly sequence the DNA without cloning. DGGE showed that a mutation had occurred, but the technique revealed little about the nature or position of that mutation. The region of the genome containing the mutation was amplified by the polymerase chain-reaction technique, providing DNA of sufficient quality and quantity for direct sequencing. Amplification was performed with a 32P end-labeled primer that allowed direct Maxam-Gilbert sequencing of the amplified product without cloning. HPRT<sub>Munich</sub> was found to contain a single base-pair substitution, a C to A transition at base-pair position 397. We report the generation of a 169-bp, wild-type DNA probe that encompasses most of exon 3 of the human hypoxanthine guanine phosphoribosyltransferase (HPRT) gene and contains a low-temperature melting domain of ~100 bp. HPRT<sub>Munich</sub>, an HPRT mutant isolated from a patient with gout, has a single amino acid substitution; the corresponding DNA sequence alteration must lie within the low-temperature melting domain of exon 3.

We report the separation of HPRT<sub>Munich</sub> from the wild-type sequence using DGGE. In addition to base-pair substitutions, DGGE is also sensitive to the methylation state of the molecule. The cDNA for HPRT was cloned into a vector and propagated in Escherichia coli dam" and dam" strains; thus, methylated and unmethylated HPRT cDNA was obtained. A single dam methylation site exists within exon 3, and we have separated the methylated and unmethylated sequences by DGGE. To rapidly sequence HPRT<sub>Munich</sub>, we used the polymerase chain reaction. Two 20-bp oligonucleotide primers complementary to the intron sequence immediately 5' and 3' to exon 3 were hybridized in solution to wild-type or HPRT<sub>Munich</sub> genomic DNA, and the primed template was extended with Taq DNA polymerase. The process was repeated 30 times, yielding ~150 ng of amplified product starting from 5 μg of genomic DNA. A combination of DGGE and polymerase chain reaction should permit rapid identification and sequencing of base-pair-substitution mutations in human genomic DNA. Both techniques are expected to play a significant role in the analysis of base-pair mutations and in the study's relationship to human disease.

**Introduction**

Denaturing gradient gel electrophoresis (DGGE) can demonstrate that a mutation has occurred in genomic DNA, but the technique does not provide precise information about the nature or location of the mutation. This information can be revealed by sequencing, which, until recently, involved construction of genomic libraries and screening for the desired clone. The polymerase chain reaction (PCR) permits rapid in vitro amplification of a given region of the genome. The DNA produced by PCR is often of sufficient quality and quantity for direct DNA sequencing. We have identified a human hypoxanthine guanine phosphoribosyltransferase (HPRT) mutant by using DGGE, amplified the region of the genome contain-
ing the mutation, and directly sequenced the amplified region to determine the precise nature of the mutation.

DGGE can separate DNA sequences differing by a single nucleotide substitution (Fischer and Lerman 1983; Lerman et al. 1984, 1986; Myers and Maniatis 1986). The technique utilizes an increasing concentration of denaturant in a polyacrylamide gel. When the DNA has migrated through the gel to a certain concentration of denaturant, the molecule undergoes an abrupt transition from the totally helical state to a partially melted state. Importantly, the point in the denaturing gel at which this transition occurs is extremely sequence dependent. Although DGGE can identify a mutant, it does not reveal the nature or exact location of the mutation.

DGGE has been used to resolve single-base-pair substitutions in the β-globin gene in total human genomic DNA (Myers et al. 1985c), as well as to identify polymorphic sites within a 1.2-kb region of chromosome 20 (Noll and Collins 1987). In both cases, a radiolabeled wild-type probe was hybridized in solution to a mutant locus, thus forming a heteroduplex containing a mismatched base. This heteroduplex was then resolved on a denaturing gradient gel. We have employed this method to resolve a base-pair mutant at the human HPRT locus.

The HPRT gene spans ~44 kb in genomic DNA and is split into nine exons (see Stout and Caskey 1985 for review). We have examined exon 3 of HPRT because this exon (1) contains the largest continuous portion of the coding frame (28%) and thus offers a sizable region for detection of mutations and (2) is reported to code for both catalytic sites of the enzyme (Wilson et al. 1983) (and mutations within this region may impair the function of the enzyme such that a patient suffers Lesch-Nyhan syndrome or gouty arthritis). To discover whether any mutant sequences can be separated from the wild-type sequence by using DGGE, we examined the melting behavior of HPRT<sub>Munch</sub>, an HPRT mutant isolated from a patient with gout. HPRT<sub>Munch</sub> has a serine-to-arginine transition at amino acid 103 (Wilson and Kelley 1984); this mutant has not been characterized at the DNA-sequence level. From the position of the amino acid substitution, we know that the DNA mutation must be located near the 3' end of exon 3 as shown in figure 1. We report the resolution of this mutant from the wild-type sequence and discuss the advantages and limitations of this technique.

To quickly determine the nature of the mutation in HPRT<sub>Munch</sub> we used the PCR (Saiki et al. 1985; Mullis and Falcooa 1987), which permits a rapid synthesis of a particular sequence directly from genomic DNA. Two oligonucleotide primers flanking the desired region of DNA are hybridized in solution to genomic DNA, a DNA polymerase is added, and synthesis occurs. The cycle of denaturation, hybridization, and polymerization is repeated 20–30 times, resulting in an exponential increase of blunt-ended DNA fragments delimited by the primers; a 220,000-fold increase in a 110-bp fragment of the human β-globin locus has been reported (Saiki et al. 1985). The quantity and quality of amplified DNA is such that direct sequencing of the amplified product is possible (McMahon et al. 1987; Wrischnik et al. 1987).

We have used two 20-bp oligonucleotide primers complementary to the intron sequence immediately 5' and 3' to HPRT exon 3, such that the entire coding region of exon 3 was amplified, producing a 224-bp fragment. Amplification was performed using Taq DNA polymerase (Koehavong et al. 1988) with a single radiolabeled primer so that the amplified material could be directly sequenced by the Maxam-Gilbert method. A single-base-pair substitution, a C-to-A transversion at base-pair position 397, was found.

Material and Methods

DGGE

TK6 is a diploid human male lymphoblast line that is HPRT<sup>-</sup> (Skopek et al. 1978). HPRT<sub>Munch</sub>, a human male lymphoblast line isolated from a patient with gout, was a gift from James M. Wilson (Massachusetts General Hospital, Boston). GM1416B, a hu-
man lymphoblast cell line containing four X chromosomes, was obtained from the Human Genetic Mutant Cell Repository in Camden, NJ. RJK833, a human male lymphoblast Lesch-Nyhan cell line that is totally deleted for HPRT, was a gift from C. T. Caskey (Baylor College of Medicine, Houston). Cell lines were grown in RPMI 1640 with 10% horse serum (GIBCO, Grand Island, NY) in a 5% CO₂ humidified atmosphere.

High-molecular-weight genomic DNA was isolated essentially as described by Blin and Stafford (1976). Plasmid pHPT30, bearing the cDNA for human HPRT, was a gift from C. T. Caskey. The vector was propagated in Escherichia coli strain HB101, a dam­+ strain. Plasmid DNA was isolated using an alkaline lysis method (Maniatis et al. 1982) and purified by CsCl centrifugation. This plasmid cDNA was used as a source of wild-type dam­+ HPRT sequence.

Human genomic DNA was digested with a 2.5-fold excess of MboII and TaqI (see § 1) according to manufacturer’s specifications (New England Biolabs, Beverly, MA). The 5’ site may be cut by either XhoI or TaqI. Genomic DNA was digested with TaqI, because XhoI may be inhibited by the 5-methyl cytosine present in the genomic DNA; TaqI is not inhibited by 5-methyl cytosine (Kesseler et al. 1985). Plasmid DNA containing the cDNA for HPRT was also digested with MboII and TaqI.

The PstI fragment of vector pHPT30, which contains the entire coding region of HPRT, was (1) subcloned in both orientations in M13mp18, (2) propagated in JM110, a dam­- dam­+ E. coli strain, and (3) used to isolate both double-strand (ds) and single-strand (ss) DNA (Messing 1983). ss M13 DNA with the (-) strand of HPRT was used as a template to synthesize the radiolabeled (+) strand, which was used as a probe. ds DNA with HPRT was used as a source of unmethylated cDNA. Growth of M13 template was done in JM110 to eliminate methylation-induced mobility differences on the denaturing gradient gels; a single dam methylation site exists in exon 3. We found that using a dam­- bacterial strain to prepare the phage template simplified later analysis of denaturing gradient gels.

A 2.5-fold molar excess of 15-mer primer (New England Biolabs) was hybridized to the ss M13 template. The radiolabeled (+) strand was generating by adding 30 U Klenow fragment (5 U/µl; New England Biolabs), 53 nM ss primed template, 50 mM NaCl, 10 mM MgCl₂, 10 mM dithiothreitol, 30 mM Tris (pH 7.5), 4 µM 32P dATP (800 Ci/mmol; Amersham, Arlington Heights, IL), and 50 µM each dCTP, dGTP, and dTTP in a reaction volume of 60 µl. After incubation for 12 h at 30 °C, the polymerase was inactivated by phenol extraction and the DNA was ethanol precipitated. The DNA was then digested with MboII and XhoI, ethanol precipitated, resuspended in water, and electrophoresed in a parallel denaturing gradient gel. The double-strand exon 3 fragment was located by autoradiography and cut from the wet denaturing gradient gel; the DNA was recovered by electroelution, ethanol precipitated, and resuspended in water.

Purification of the ds probe on a parallel denaturing gradient gel was necessary to obtain a homogeneous probe preparation. If a normal polyacrylamide gel was used to purify the ds exon 3 fragment, this DNA gave several bands when analyzed on a denaturing gradient gel.

The gel-purified, ds 169 fragment was hybridized to an excess of ss M13 bearing the (+) strand of HPRT; this ss M13 DNA is complementary to the nonradiolabeled strand of exon 3. After a 10-fold molar excess of the ss M13 was added and the NaCl concentration was adjusted to 300 mM, the solution was boiled for 4 min and incubated at 65 °C for 1 h. The DNA was ethanol precipitated, resuspended in water, and run on a conventional polyacrylamide gel. The ss exon 3 fragment was visualized in the wet gel by autoradiography and was isolated as described above.

Twenty micrograms of MboII/TaqI-digested genomic DNA was resuspended in 30 µl of 300 mM NaCl, 30 mM Tris (pH 8), 1 mM EDTA with 1.7 × 10⁶ cpm of radiolabeled ss probe (1.2 × 10⁸ copies). The mixture was boiled for 4 min and placed at 65 °C for 18 h. One hundred nanograms of ss M13 DNA containing the (-) strand of HPRT was added, and hybridization was continued another hour; this DNA is complementary to the ss probe and retains most of the unbound probe at the origin of the gel. The DNA was ethanol precipitated, washed twice with 20 C 75% ethanol, and resuspended in 8 µl of water and 1.6 µl of loading buffer. One thousand counts per minute of ss probe was also hybridized to a molar excess of TaqI/MboII-digested plasmid containing the cDNA for HPRT.

A denaturing gradient gel was used to purify the ds probe and to resolve mutants from the wild-type DNA sequence. A 1-mm thick, 12.5% polyacrylamide gel (37.5% polyacrylamide/1% bis) with
either a 5%–30% or a 10%–35% denaturing gradient was used (100% denaturant = 7 M urea/40% formamide). Use of a 5%–30% gel gave better separation of the unbound ss probe from the wild-type sequence; otherwise the behavior of the DNA in both gels is identical (data not shown). Denaturant was prepared fresh, and the gradient was formed with a Hoeffer gravity-flow gradient maker. The gels were run submerged at 60 C at 150 V (7.5 V/cm) in TAE buffer (40 mM Tris, 20 mM NaOAc, 1 mM EDTA [pH 8.2]); the 5%–30% gel was run for 5.5 h, and the 10%–35% gel was run for 4 h. If DNA was to be recovered from the gel, it was not fixed; otherwise the gel was fixed for 2 h in 40% methanol/5% glycerol, dried, and then exposed to Kodak XAR-5 film for 2 days at -70 C with DuPont Cronex Lightning Plus intensifying screen.

PCR

Oligonucleotides were synthesized by Synthetic Genetics (San Diego) and were purified by gel electrophoresis. The sequence of the downstream primer was CATAT ATCAA ATATA CTAC, and the sequence of the upstream primer was TCCTG ATTTT ATTTT TGATG. The downstream primer was end-labeled using the reagents in the BRL 5' terminus-labeling kit (Bethesda Research Laboratories, Bethesda, MD). The primer was labeled with ^32P-γ-ATP (6,000 Ci/mmol; New England Nuclear, Boston) to a final specific activity of ~3.6 x 10^7 cpm/ pmol. Approximately 140 pmol of radiolabeled primer and 160 pmol of nonradiolabeled primer were added to the PCR reaction mix to give a final primer concentration of 3 μM.

dNTPs were obtained in powder form (Sigma, St. Louis), dissolved in water, and adjusted to pH 7.5 with NaOH. A buffer containing 100 mM Tris (pH 8.0), 50 mM MgCl₂ was obtained from New England Biolabs. Ta DNA polymerase, Lot 21, 1 U/μl, was from New England Biolabs.

The 100-μl reaction mixture contained 5 μg genomic DNA, 3 μM of each primer, 1.5 mM of each dNTP, 10 mM Tris (pH 8.0), 5 mM MgCl₂. The denaturation time for the first three cycles was 3 min; the remainder of the denaturations were performed for 1 min. The cycle was (1) denaturation in a boiling water bath, (2) hybridization of primers to template for 45 s at room temperature, followed by a 15-s incubation in a 37 C water bath, (3) addition of 1 U Ta DNA polymerase followed by brief vortexing, and (4) DNA polymerization for 4 min in a 37 C water

bath. The cycle of denaturation, hybridization, and polymerization was repeated 30 times.

After 30 cycles of amplification the reaction mix was extracted with phenol, ethanol precipitated, dissolved in water, and loaded onto an 8% polyacrylamide gel. The desired 224-bp amplified product was located by autoradiography, electroeluted from the gel, and sequenced by the Maxam-Gilbert method (Maxam and Gilbert 1980).

Results and Discussion

DGGE can detect mutations only in the lower-temperature melting domains, so it is important to determine what portion of the molecule becomes single stranded when the DNA melts. This can be determined by use of a computer algorithm, developed by L. Lerman, that predicts DNA melting behavior solely from the base-pair sequence. The calculated melting map of the wild-type homoduplex (see fig. 2) shows a low-temperature melting domain of ~100 bp, from base-pair position 300 to base-pair position 400. Thus, we expect to be able to resolve mutations in ~100 bp of the 169-bp probe; these 100 bp represent 15% of the coding region of HPRT. Mutations
Figure 3  Autoradiograms of the denaturing gradient gels. Twenty micrograms of genomic DNA was digested with TaqI and MboII and then hybridized to an HPRT exon 3, wild-type, ss, 164-bp probe. A. 10%–33% denaturant. Lane 1: 5S, 107 cpm ss probe; lane 2: MUNICH, HPRTtsuh, genomic DNA (lymphoblast line); lane 3: 107 cpm ss probe hybridized to HPRT cDNA propagated in a dam− strain R. 5%–80% denaturant. Lane 1: 5S, 107 cpm ss probe; lane 2: DAM− 107 cpm ss probe hybridized to HPRT cDNA propagated in a dam− strain, lane 3: DAM− 107 cpm ss probe; lane 4: Δ, RJ853 genomic DNA; HPRT deletion lymphoblast line; lane 5: WT, TK- genomic DNA (lymphoblast line with four X chromosomes); The behaviour of the melted state at 15%–33% and 10%–35%, denaturing gradient gels. The DNA undergoes an abrupt transition to the partially denatured state at 23% denaturant (data not shown). The electrophoresis time for the 5%–10% gel must be extended to allow the DNA to reach concentrations of 23% denaturant. This results in a greater separation of the unbound ss probe from the wild-type sequence than the wild-type and mutant samples are on the gels as occurring from base-pair position 234 to base-pair position 300 will lie in the high-temperature melting domain of the molecule and should not be separated from the wild-type sequence by DGGE.

The calculated melting map of the heteroduplex formed by hybridization of the wild-type probe to HPRTtsuh is given in figure 2. The effect of a mismatch can be approximated by a modification of the melting algorithm (L. Lerman, personal communication; see legend to fig. 2). The melting algorithm predicts that the mismatched heteroduplex will melt at a lower temperature than the perfectly base-paired homoduplex.

Autoradiograms of the denaturing gradient gels are shown in figure 3A and 3B. The wild-type DNA shows a band at 23% denaturant (WT; fig. 3A, lane 4, and fig. 3B, lane 5). The cell line deleted for HPRT (Δ; fig. 3B, lane 4) shows no signal, while the cell line with four X chromosomes (4X; fig. 3B, lane 6) shows a strong band; this cell line is expected to show a more intense signal than does the wild-type because HPRT is X linked. HPRTtsuh is displaced −1.2 cm from the wild-type band (MUNICH; fig. 3A, lane 3). Later sequencing of HPRTtsuh revealed that the DGGE shift was due to a single-base-pair substitution.

Denaturing gradient gels are sensitive to the methylation state of the molecule. Therefore, a methylated molecule could appear to be a mutant. In mammalian DNA, the only methylated base is 5-methyl cytosine.
which is found almost exclusively at the sequence Cpg (Wyatt 1951; Vanyushin et al. 1970). However, the region of exon 3 that we are examining has no CpG sequences, and thus no 5-methyl cytosine residues are expected to confound our results.

To investigate the effect of methylation, we have propagated the cDNA for HPRT in both dam+ and dam- bacterial strains, and the effect of bacterial methylation can be seen in the denaturing gradient gel. A single dam site exists at base-pair position 357 in exon 3; dam methylase will produce an N5-methyl adenine residue in the sequence GATC (Hattman et al. 1978). The presence of a single 6-methyl adenine destabilizes the molecule (DAM+; fig. 3A, lane 2, and fig. 3B, lane 3) so that it melts at a lower temperature than does the unmethylated sequence (DAM-; fig. 3A, lane 5, and fig. 3B, lane 2). The molecule is actually hemimethylated, since unmethylated ss probe was hybridized to fully methylated cDNA.

Although four HPRT pseudogenes exist (Patel et al. 1984), there is no pseudogene interference in this system. Pseudogenes might have created undesirable background problems in the denaturing gradient gels. A cell line that is totally deleted for HPRT shows no signal when hybridized to ss exon 3 probe and resolved on a denaturing gradient gel (Δ; fig. 3B, lane 4); the HPRT deletion line also shows no signal at the wild-type position when the DNA is run on a conventional polyacrylamide gel (data not shown). No pseudogene cross-hybridizing species can be seen. Apparently these pseudogenes do not have restriction sites such that a fragment of ~169 bp is produced when genomic DNA is digested with MboII and TaqI.

The 169-bp probe appears to be well suited for work with DGGGE because it has a distinct, small, high-temperature melting domain. A run of six consecutive guanines near the 5′ end of the molecule helps form the high-temperature melting domain, in effect forming a small GC-rich clamp (for discussion of GC clamps, see Myers et al. 1985a, 1985b). We expect to be able to detect nearly all base-pair substitution mutations that occur in the 100-bp low-temperature melting domain of the molecule. These 100 bp represent ~15% of the coding region of HPRT. DGGGE may be well suited for analyzing other loci possessing low-temperature melting domains adjacent to a short higher-temperature melting domain.

To sequence HPRT$_{Munch}$, we used PCR to generate a sufficient quantity of template for direct sequencing. We have made one important modification of the PCR, namely, the substitution of Taq DNA polymerase (Keohavong et al. 1988) for the Klenow fragment that was employed in the original technique. The error rate for the Klenow fragment is reported to be $10^{-10}$~$10^{-9}$ error/bp, while the error rate for Taq DNA polymerase is reported to be $10^{-9}$~$10^{-8}$ error/bp (Loeb and Kunkel 1982). DNA amplified with Taq DNA polymerase may contain fewer polymerase-induced errors than does DNA amplified with Klenow fragment. DNA amplified
with the Klenow fragment has been cloned and sequenced (Scharf et al. 1986); in one case the frequency of nucleotide misincorporation was $<1.5 \times 10^{-3}$ errors/bp, and in another case the error rate was $<1.7 \times 10^{-3}$ errors/bp.

After 30 cycles of amplification, an ethidium bromide-stained gel (Fig. 4) shows a distinct band $\sim 224$ bp in length, which is the expected size of the amplified exon 3 product; the exon itself is 184 bp and each primer is 20 bp. After 30 cycles of amplification, $\sim 150$ ng of exon 3 was obtained, which represents $\sim 6 \times 10^{10}$ copies. Five micrograms of genomic DNA contains $\sim 7 \times 10^9$ copies of HPRT, so 30 cycles of amplification produced an $8.7 \times 10^3$-fold increase; this represents an average efficiency of $\sim 58\%$/cycle ($1 + X)^n = Y; X = \text{efficiency}, n = \text{cycles}, Y = \text{amplification}$; see Saiki et al. 1985 for discussion of efficiency).

A single radiolabeled primer was used during the amplification so that the amplified product could be directly sequenced by the Maxam-Gilbert method. Sequence analysis of the amplified product showed that HPRT<sub>Munch</sub> contained a single-base-pair substitution, a CG-to-AT transversion at base-pair position 397 (see fig. 5). This base-pair substitution is consistent with the amino acid substitution reported for HPRT<sub>Munch</sub>, a serine-to-arginine substitution at amino acid 103 (Wilson and Kelley 1984).

DGGE provides a method to localize base-pair mutations to a region of DNA, and the PCR can be used to generate sufficient DNA for direct sequencing. We are applying both techniques to the study of human mutagenesis both in vivo and in vitro (Thilly 1985; Cariello and Thilly 1986). DGGE and PCR may have

---

**Figure 5** Direct sequencing of HPRT<sub>Munch</sub> and wild-type DNA amplified by PCR. DNA amplification was performed for 30 cycles by using one end-labeled primer. After electrophoresis of the amplification mixture on an 8% polyacrylamide gel, the 224-bp band was located by autoradiography, excised, and electrophoresed in the gel. The DNA was chemically cleaved by the method of Maxam and Gilbert (1980) and loaded onto an 8% polyacrylamide, 8 M urea sequencing gel; the gel was dried and exposed to Kodak XAR-5 film with a DuPont Cronex Lightning Plus intensifying screen at $-70$ C for 48 h. The first four lanes show the sequence of the (+) strand of HPRT<sub>Munch</sub>, and the next four lanes correspond to the sequence of the wild type; the sequence of the wild-type (+) strand is given. A single-base-pair substitution in HPRT<sub>Munch</sub> was found, a G-to-T transversion that corresponds to a C-to-A transversion at base-pair position 397 of the (+) strand. The arrow indicates the position of the mutation in the sequencing gel, and the asterisk indicates the location of the mutation in the wild-type sequence.
many applications in the analysis of mutation and human disease.

Acknowledgements

The authors wish to thank L. Lerman for the use of his melting algorithm and for helpful discussions. This work was supported by Department of Energy grant DE-FG02-86ER60449 and National Institute of Environmental Health Sciences grants NIH-S-P01-ES00597, NIH-S-P30-ES02109, NIH-S-P01-ES01640, and NIH-S-P01-ES03926.

References


Cariello et al.


6.1.2 DNA damage produced by ethidium bromide staining and exposure to ultraviolet light

We have examined routine DNA manipulations, namely, ethidium bromide (EB) staining and exposure to ultraviolet (UV) light, in order to determine the extent of DNA damage produced by these procedures. Many of the lesions produced by EB and UV light should affect the ability of the DNA helix to properly base pair and therefore should alter the melting behavior of the DNA. Denaturing gradient gel electrophoresis (DGGE) is extremely sensitive to the primary and secondary structure of the DNA molecule. This technique can resolve base pair substitutions in perfectly base-paired homoduplex DNA (1,2), single base pair mismatches in heteroduplex DNA (3), and DNA methylation states (4). DGGE conditions and melting characteristics of the 169 bp TaqI/MboII exon 3 fragment of human HPRT are given elsewhere (5). DNA was end-labelled, run on an 8% PAGE, and the gel was stained for 10 min with 1 μg/ml EB. Bands were located by autoradiography, cut from the gel, and a gel slice was exposed to 254 nm UV, 300 nm UV, or no exposure. The 254 nm light source was a Mineralight Lamp Model UVGL-58 (Ultra-Violet Products, San Gabriel, CA) placed 24 cm from the gel slice. The 300 nm UV light source was a Fotodyne Transilluminator Model 3-1000 (New Berlin, WI); the gel slice was placed on the exciter filter. To determine the damage produced by UV light alone, the procedure was repeated without ethidium bromide. Each sample was analyzed on three different gels: a PAGE, a 7 M urea polyacrylamide gel, and a denaturing gradient gel. The PAGE and 7 M urea gel showed no differences between the samples, only the denaturing gradient gel revealed that significant DNA damage occurred. Even a 10 s exposure of EB-stained DNA to 300 nm UV light produced detectable damage. A scanning densitometer trace showed that only about 50% of the DNA exposed to 180 s of 300 nm UV light remained in the major band, as compared to the zero control. DNA without EB suffered considerably more damage than did EB-stained DNA, in fact, all of this DNA was damaged after a three minute exposure to 300 nm UV. This is consistent with reports which showed that dyes bound to DNA during irradiation significantly reduced the thymine dimer yield. UV-induced lesions melt at a lower temperature than does wild-type DNA; any adducts, including various pyrimidine dimers and purine photoproducts, which disrupts DNA hydrogen bonding could produce such a shift in melting behavior. Several discrete bands can be seen on the denaturing gradient gel, it would be possible to isolate and characterize the DNA in a particular band. Any DNA lesion that alters the melting behavior of DNA has the potential to be detected and isolated using denaturing gradient gel electrophoresis. The pattern of adducts produced by a wide range of compounds could be examined in this way, including alkylating agents which methylate the DNA, and compounds such as benz(a)pyrene which produce a bulky DNA adduct.

References
2. Lerman et al. (1994).
6.1.3 DNA amplification in vitro using T4 DNA polymerase

LABORATORY METHODS
DNA Amplification In Vitro Using T4 DNA Polymerase

PHOUTHONE KEOHAVONG, ALEXANDRA G. KAT, NEAL F. CARIELLO, and WILLIAM G. THILLY

ABSTRACT
We have evaluated in vitro DNA amplification by polymerase chain reaction using either T4 DNA polymerase or Klenow fragment of Escherichia coli DNA polymerase I. Both polymerases under optimal salt conditions permit efficient amplification of exon 3 of the hypoxanthine guanine phosphoribosyltransferase (HPRT) gene from human genomic DNA and from plasmid containing the HPRT cDNA. DNA sequences amplified from human genomic DNA, using two 20-nucleotide primers flanking the ends of the exon, showed a marked difference between the two polymerases. T4 DNA polymerase yielded only the expected amplified DNA fragment, whereas Klenow fragment produced many lower-molecular-weight bands in addition to the expected DNA fragment. On the basis of the reported fidelity of in vitro DNA synthesis using Klenow fragment and T4 DNA polymerase, it is expected that the latter will create substantially fewer errors during the amplification process. For these reasons, T4 DNA polymerase should be particularly valuable for amplification of sequences present at a very low frequency requiring many cycles of amplification to be detected.

INTRODUCTION
The recent development of a method allowing exponential synthesis of DNA segments from genomic DNA represents an important addition to the techniques of genetic analysis. The technique involves repeating a DNA chain-extension reaction from two primers, each complementary to the 3' end of the antiparallel strands of a desired segment of the genomic DNA. After 20 cycles of the polymerase chain reaction process, more than a 10^5-fold amplification of a 110-bp β-globin sequence from human genomic DNA was obtained (Saiki et al., 1985). However, during the process of DNA synthesis, it is expected that the polymerase will introduce errors in the newly synthesized DNA. Any error introduced by the polymerase will be subsequently amplified, resulting in an accumulation of mutated DNA sequences, the fraction of which will increase with the number of amplification cycles.

The fidelity of the polymerase typically used in these studies, Klenow fragment of Escherichia coli DNA polymerase I, is reported to be from 10^{-5} to 10^{-6} errors per base pair (for review, see Loeb and Kunkel, 1982). By sequencing clones containing the HLA DQα locus, Scharf et al. (1986) found an error frequency of about 1/600 after 27 cycles of amplification. We wish to analyze complex DNA populations containing small fractions of different mutations (Thilly, 1983; Cariello and Thilly, 1986), and error rates of 10^{-6} per base pair would create severe background problems. For these reasons, we explored the possibility of increasing the fidelity of this new molecular tool. Of particular interest is the previous demonstration that T4 DNA polymerase has a high fidelity and makes less than 10^{-7} errors per base pair (Kunkel et al., 1981). Consequently, the expected error frequency using T4 DNA polymerase would be at least 10-1000 times less than that expected for Klenow fragment.

In this study we established conditions that permit T4 DNA polymerase to amplify exon 3 of the HPRT gene from human genomic DNA using 20-nucleotide primers that immediately flank the 5' and 3' ends of the exon. We
used polyacrylamide gel electrophoresis to compare the pattern of products amplified by this polymerase with the pattern of products amplified by Klenow fragment. Our results are discussed in relation to those previously reported for Klenow fragment (Saiki et al., 1985; Scharf et al., 1986; Falcón and Mullis, 1987). The data with T4 DNA polymerase suggest that the use of this enzyme is a valuable modification of the polymerase chain reaction procedure, and we assess the feasibility of using T4 DNA polymerase under these conditions.

**MATERIALS AND METHODS**

**Enzymes and chemicals**

T4 DNA polymerase was purchased from New England Biolabs (Beverly, MA) and Klenow fragment of *E. coli* DNA polymerase I from Bethesda Research Laboratories (Gaithersburg, MD). 2’-Deoxynucleoside-5’-triphosphates were obtained either from Sigma (St. Louis, MO) in powder form or from Pharmacia (Piscataway, NJ) in 100 mM minimal salt solutions at pH 7.0 (this solution was used in experiments to determine optimum salt concentrations for T4 DNA polymerase). Powdered dNTPs were dissolved in water and adjusted to pH 7 with NaOH. The buffer (10×) for T4 DNA polymerase (100 mM Tris pH 8.0, 50 mM MgCl₂) was supplied by New England Biolabs. 20-Mer oligonucleotides were synthesized by Synthetich Genetics (San Diego, CA):

- **Primer 1**: 5’-CATATATATATAATATCACC-3’
- **Primer 2**: 5’-CTGATATTTTATGATG-3’
- **Primer 3**: 5’-ACAATACTCCTTTACG-3’
- **Primer 4**: 5’-GACTGAACTCCTTCGCTCAG-3’

To enable sequencing of the amplified DNA by the method of Maxam and Gilbert (1980) and also to follow the accumulation of product DNA, one of the primers was 5’-end-labeled with [γ-³²P]ATP (7000 Ci/mmol, New England Nuclear, Boston, MA) using T4 polynucleotide kinase and the reagents in the 5’ DNA terminus labeling system (Bethesda Research Laboratories). The average specific activity of the primers was 1.3-3.6 × 10⁷ cpm/pmol (53-160 Ci/mmol).

**Plasmid and genomic DNA**

Plasmid pHPT30, containing the entire human CDNA for HPRT, was kindly provided by Dr. C.T. Caskey (Baylor College of Medicine in Houston, TX). The 1-kb Pst I fragment containing the entire coding region was subcloned in pSP64 (Promega Biotech, Madison, WI). Plasmid DNA was isolated by alkaline lysis (Maniatis et al., 1982) and purified by CsCl centrifugation.

DNA was isolated from TK6, an HPRT⁺ diploid human male lymphoblastoid cell line (Skopek et al., 1978), as described by Blin and Stafford (1976) with the following modifications: (i) 10⁶ cells frozen as a washed 1-ml pellet were thawed and resuspended in 7.5 ml of 50 mM EDTA/10 mM Tris pH 7.5, then 180 µl of 8% Sarkosyl and 38 µl of 20 mg/ml proteinase K were added; (ii) after incubation at 50°C for 2 hr in a shaking water bath, the cell suspension was dialyzed against 50 mM EDTA/10 mM Tris pH 7.5 for 12-16 hr at 4°C; (iii) after dialysis, 100 µl of 10 mg/ml heat-treated RNase was added and incubated at 37°C for 2 hr; (iv) the cell suspension was extracted twice each with phenol, phenol/chloroform, chloroform; (v) the DNA was precipitated with sodium acetate/ethanol, resuspended in water, then precipitated again with ammonium acetate/isopropanol and resuspended in water.

**Polymerase chain reaction procedure**

The polymerase chain reaction procedure used for T4 DNA polymerase was a modification of the technique using Klenow fragment (Saiki et al., 1985). The procedure consists of repeating the following steps: (i) denaturation of the DNA template, (ii) annealing of the oligonucleotide primers to the denatured DNA, (iii) DNA synthesis. DNA was denatured in a volume of 100 µl by heating to 100°C in boiling water for 3 min for the first three cycles, then boiling for 1 min in subsequent cycles. The reaction mixture was then placed at room temperature for 45 sec, followed by a 15-sec incubation in a 37°C water bath, allowing the primers to hybridize to the template DNA. The polymerase was added and the mixture placed in a 37°C water bath for 3-4 min, during which time DNA synthesis occurred. These steps were repeated for n cycles. For 5 µg of genomic DNA, 0.8 units of T4 DNA polymerase or 1 unit of Klenow fragment were added per cycle; for 1 nM plasmid DNA, 1 unit of T4 or 5 units of Klenow fragment were added.

The 100-µl reaction mixture for T4 DNA polymerase contained 10 mM Tris pH 8.0, 5 mM MgCl₂, 33 mM Na⁺ (from neutralizing the dNTP with NaOH); 1.5 mM of each of the four dNTPs, and 3 µM of each primer. For genomic DNA a concentration of 2.5 mM of each dNTP was used. To analyze the effects of sodium salts on the efficiency of the amplification from plasmid template, we used minimal salt dNTPs (20 mM Na⁺) and added NaCl at 0, 25, 50, or 100 mM concentrations. The conditions for experiments with Klenow fragment were as described by Falcón and Mullis (1987), summarized in the legend of Fig. 2 (below). The cycle time was as described above for T4 DNA polymerase.

**Acrylamide gel analysis**

After amplification for n cycles as indicated in the legends of Figs. 2-6 (below), a 10-µl aliquot was taken from the reaction mixture and analyzed by electrophoresis on 6 or 8% polyacrylamide gels (37.5/1 = acrylamide/bisacrylamide) in TBE buffer. The gel was then stained with 0.5 mg/ml ethidium bromide and photographed under ultraviolet light.
DNA AMPLIFICATION USING T4 DNA POLYMERASE

Purification of the amplified DNA

The reaction mixture was extracted once with phenol/chloroform (1:1), twice with chloroform/isooamyl alcohol (24:1), and precipitated with ethanol. The DNA was then separated by polyacrylamide gel electrophoresis. Radioactive bands were located by autoradiography, cut out, and electroeluted onto dialysis membranes.

RESULTS

Figure 1 shows the schematic positions of the two types of 20-mer primers used in this study for the amplification of exon 3 of the HPRT gene: the "intron" primers (primers 1 and 2), complementary to the intron sequences immediately adjacent to exon 3, and the "exon" primers (primers 3 and 4), complementary to the extreme 3' and 5' ends of exon 3. It is expected that the DNA synthesized by the polymerase chain reaction procedure will give rise to a 224-bp fragment containing the intron primers and a 184-bp fragment with the exon primers. Both fragments should be blunt-ended and delimited by the 5' end of each primer.

In this study, we used the intron primers for the amplification from genomic DNA to increase the likelihood that the amplified exon 3 sequence came from the single copy of the HPRT gene and not from any of the four known cDNA cross-hybridizing sequences (Patel et al., 1984). Analysis of HPRT pseudogene sequences in our laboratory suggests that they contain exons without the large intervening sequences (I. Gennett and A. Beck, personal communication). Thus, these pseudogenes were not expected to possess sequences complementary to the exon primers and would therefore not be a substrate for amplification. The efficiency of the amplification can be estimated by assuming that the intron primers hybridize only to the single X-linked HPRT gene in male cells.

In our initial work to amplify exon 3 from genomic DNA, we used the Klenow fragment of E. coli DNA polymerase I under the conditions previously reported for this enzyme (Sakai et al., 1985) with intron primers (see Fig. 1) and 5 μg (1 × 10⁶ pmol) of genomic DNA. Analysis on an ethidium bromide-stained polyacrylamide gel (Fig. 2) showed that the products from 30 amplification cycles contained a series of sequences, including a fragment with the expected size of 224 bp. The 224-bp fragment was purified from the gel, and sequencing by the method of Maxam and Gilbert (1980) verified that it contained exon 3 of the HPRT gene (see Fig. 7, below, section b). By comparison of the ethidium bromide fluorescence with that of the molecular-weight markers DNA, we estimate that 30 amplification cycles produced approximately a 4 × 10⁶-fold increase in the expected 224-bp fragment. Smaller-molecular-weight fragments amplified by Klenow fragment can also be seen.

FIG. 1. Arrangement of template and primers used in the polymerase chain reaction for exon 3 of the HPRT gene. Two pairs of primers were used: (i) intron primers (primers 1 and 2), complementary to the intron sequences directly adjacent to exon 3, and (ii) exon primers (primers 3 and 4), complementary to the extreme 3' end of each strand of exon 3. Shown also are the expected DNA fragments from the amplification process using these primers.

FIG. 2. Electrophoretic analysis on a 6% polyacrylamide gel (acrylamide/bis = 37.5/1) of the DNA amplified with Klenow fragment from genomic DNA using intron primers. Genomic DNA (5 μg) was subjected to polymerase chain reaction for 30 cycles using the reaction mixture described by Falcoons and Mullis (1987) [1 μM of each primer, 1.5 mM of each dNTP, 30 mM Tris-acetate pH 7.9, 10 mM MgCl₂, 10 mM DTT, and 60 mM NaCl (TAE buffer)], with cycle time as described in Materials and Methods for T4 DNA polymerase. Lane KL contains 10 μl of amplified DNA sample, and lane M, 0.5 μg of pBR322/ Msp I marker DNA.
in Fig. 2, but these sequences have not yet been characterized. These smaller-molecular-weight bands could represent either (i) sequences unrelated to exon 3 that were amplified by the pairing of both primers to other segments of the genomic DNA, as has been reported for the amplification of the α-globin segment from genomic DNA with Klenow fragment (Scharf et al., 1986), or (ii) by-products of the amplification process arising from mispriming inside the accumulated 224-bp sequence, which has been observed for the amplification of β-globin gene (Faloona and Mullis, 1987).

To study the behavior of T4 DNA polymerase, we first optimized conditions for this enzyme by using the exon primers (see Fig. 1) and a plasmid containing the HPRT cDNA as a template. The polymerase chain reaction was carried out for 10 cycles as described in Materials and Methods. To follow the yield of the amplified DNA, a 10-μl aliquot was taken at the fourth cycle and at each successive cycle for analysis on a polyacrylamide gel. As shown in Fig. 3a, a greater number of amplification cycles produced an increase in the ethidium bromide-stained 184-bp fragment, resulting in a fragment concentration of approximately 125 nM after 10 cycles (starting concentration 1 nM). Since one of the primers used was end-labeled, we could follow the accumulation of the 184-bp product by radioactive counting. When the radioactivity versus cycle number is plotted on a semilogarithmic scale, the points are reasonably represented by a straight line, indicating that the formation of the 184-bp fragment follows an exponential function (Fig. 3b). By fitting the curve to the function:

$$\text{CPM}_N = \text{CPM}_0 [1 + y]^N - y^N - 1$$

where CPM$_N$ is the starting number of counts, CPM$_0$ is the number of counts after $N$ cycles (both are directly related to copy number, as described in the legend for Fig. 3b), and $y$ is the efficiency per cycle (Faloona and Mullis, 1987), the efficiency per cycle can be estimated as 67%, which is comparable to the estimates of Klenow fragment efficiency under similar conditions (Faloona and Mullis, 1987).

Figure 4 shows the influence of pH on T4 DNA polymerase performance. Optimal yields of 184-bp fragment were obtained in the range of pH 7.6 (lane B) to 8.0 (lane C). Either lowering the pH to 7.2 (lane A) or increasing the pH to 8.3 (lane D) caused the total yield of 184-bp fragment to drop by approximately 40% as judged by radioactive counting of the 184-bp fragments excised from the gel.

In addition, T4 DNA polymerase functioned optimally at low salt concentrations. As shown in Fig. 5, lane A, the best yield was obtained when no salt was added to supplement the sodium already present in the reaction mixture from the addition of a minimal salt dNTP solution (20 mM Na⁺ from a total of 6 mM dNTP). Addition of 25 mM (lane B), 50 mM (lane C), or 100 mM (lane D) NaCl reduced the total yield of 184-bp fragment by approximately 10, 50, and 90%, respectively, of the yield obtained with no added NaCl (lane A). Inclusion of DTT or β-mercaptoethanol in the reaction mixture was ineffective in improv-

FIG. 3. a. Electrophoretic analysis of exon 3 amplified from plasmid template containing the HPRT cDNA, using T4 DNA polymerase and the exon primers. Enzymatic amplification was performed for 10 cycles, starting with a 1 nM plasmid template concentration. After cycles 4-10, a 10-μl (1/10) aliquot was withdrawn and run on a 6% polyacrylamide gel (acylamide/bis = 37.5:1). After staining with ethidium bromide, the gel was photographed under UV light. Lanes marked cycles 4-10 show the appearance of the 184-bp fragment in each aliquot. Lane M contains 0.25 μg of pBR322/Msp I molecular weight marker. b. Formation of 184-bp fragments monitored by the incorporation of $^{32}$P 5'-end-labeled primer. One of the primers in the reaction mixture was 5'-end-labeled with $^{32}$P (about $1.3 \times 10^6$ cpm/pmol). The ratio of incorporation of labeled versus unlabeled primer was assumed to remain constant throughout the procedure. The bands corresponding to 184 bp, described in Fig. 3a, were excised and the radioactivity of the gel slices was counted. Since the ratio of labeled to unlabeled primers remains constant throughout the amplification reaction, there is a direct correlation between the number of counts per minute detected and the number of fragments.
FIG. 4. Electrophoretic analysis of the effect of pH on exon 3 amplification by T4 DNA polymerase. The conditions for T4 DNA polymerase were as described in Materials and Methods, except that the pH of the Tris buffer was adjusted to different values and one of the primers was 5'-endlabeled as described in the legend for Fig. 3b. In this way the 184-bp band could be excised and counted to determine the relative yield. The polymerase chain reaction was carried out for 10 cycles and a 10-μl aliquot (1/10 of the reaction mixture) was loaded on a 5% polyacrylamide gel (acylamide/bis = 37.5/1). After staining with ethidium bromide, the gel was photographed under ultraviolet light. The pH values of the Tris buffer used were pH 7.2 (lane A), pH 7.6 (lane B), pH 8.0 (lane C), and pH 8.3 (lane D). Lane M is 0.5 μg of pBR322/Msp I marker DNA.

FIG. 5. Electrophoretic analysis of the effect of NaCl concentration on exon 3 amplification by T4 DNA polymerase (lanes A–D) and Klenow fragment (lanes E–H). The conditions for T4 DNA polymerase were as described in Materials and Methods, apart from a varying NaCl concentration. The conditions for Klenow fragment were as described in the legend to Fig. 2, with the exception that the template was 1 nM plasmid and the NaCl concentration was varied. After 10 cycles, a 10-μl aliquot was run on a 5% polyacrylamide gel (acylamide/bis = 37.5/1), stained with ethidium bromide, and photographed under ultraviolet light. Lanes A and E, 0 mM NaCl; lanes B and F, 25 mM NaCl; lanes C and G, 50 mM NaCl; lanes D and H, 100 mM NaCl. Lane 1 shows the amplification of exon 3 with Klenow fragment under the same buffer conditions as used with T4 DNA polymerase in lane A, i.e., no added NaCl. Lane 3 shows the performance of Klenow fragment in T4 buffer but with the MgCl₂ concentration increased to 10 mM and DTT added to 10 mM. The 184-bp band was excised and the radioactivity was counted to determine the relative yield as in Fig. 3b.

The results indicate that T4 DNA polymerase can be used to amplify exon 3 from plasmid template. Therefore, as a second step, we amplified exon 3 from genomic DNA with T4 DNA polymerase, using the intron primers (Fig. 1). After 30 cycles, starting with 5 ng of genomic DNA (Fig. 6), a 224-bp fragment could be detected easily by ethidium bromide staining when one-tenth of the reaction mixture was analyzed on a polyacrylamide gel. We verified that this fragment contained exon 3 of the HPRT gene by sequencing according to the method of Maxam and Gilbert (1980) (Fig. 7, section a). Based on the intensity of the ethidium bromide fluorescence, T4 DNA polymerase produced approximately the same increase as that observed with Klenow fragment after 30 cycles, a 4 × 10-fold increase. Thus, from an estimated concentration of 25 fmol
FIG. 6. Electrophoretic analysis of the DNA amplified with T4 DNA polymerase from genomic DNA, using intron primers. The amplification was performed for 30 cycles, starting with 5 ng of genomic DNA. A total of 0.2 μl (1/10 of the reaction) was analyzed on an 8% polyacrylamide gel (acrylamide/bis = 19/1) as shown in lane T4; lane M contains 0.5 μg of pBR322/Msp I-size marker DNA.

for a single-copy sequence such as HPRT exon 3, we achieved a final yield of 10 nM (150 ng in 100 μl). This represents an efficiency per cycle of about 54% according to the equation:

\[(1 + Y)^N = 4 \times 10^6\]

where \(Y\) is the efficiency per cycle and \(N\) is the number of cycles performed (Saiki et al., 1985). These results show that T4 DNA polymerase can be used in the polymerase chain reaction procedure to amplify exon 3 from genomic DNA. Moreover, as can be seen in Fig. 6, there are no detectable co-amplified products when T4 polymerase is used for 30 cycles of amplification and only the expected 224-bp fragment is visible. This is in marked contrast to the results seen with Klenow fragment, in which many smaller-molecular-weight fragments are generated in addition to the desired fragment.

DISCUSSION

The polymerase chain reaction procedure (Saiki et al., 1985; Faloons and Mullis, 1987) represents a new cool per-
mitting rapid production of a desired fragment from genomic DNA in quantities sufficient for molecular manipulations, provided that some sequence information is known. This technique has been applied successfully to prenatal DNA-diagnostic procedures (Saitki et al., 1985) and to molecular cloning from genomic DNA (Scharf et al., 1986; Wrischnik et al., 1987). These studies, however, have all used Klenow fragment for enzymatic amplification.

We have demonstrated that amplification of a desired fragment from genomic DNA can also be accomplished using T4 DNA polymerase in the polymerase chain reaction technique. Based on the fidelity of DNA synthesis in vitro using the two polymerases (reviewed by Loeb and Kunkel, 1982), we expect the DNA fragments amplified with T4 DNA polymerase to contain a significantly lower proportion of inaccurate copies of the original template. Furthermore, DNA amplified with T4 DNA polymerase shows fewer co-amplified sequences than does DNA amplified with Klenow fragment. This improvement in the purity of the fragments expedites subsequent DNA recovery by reducing the risk of contamination from unwanted sequences.

Several explanations exist for the different patterns of amplification of the smaller-molecular-weight fragments produced by the two enzymes. This effect could result from the ability of Klenow fragment to begin synthesis from primers imperfectly paired to template, either in genomic DNA from sequences other than that expected or from within the amplified 224-bp fragment, giving rise to by-products. Alternatively, the hybridization specificity of the primers to the template could be greater due to the lower NaCl concentration in the buffer used for T4 DNA polymerase. The present study cannot differentiate among these possibilities, and it may be that both are involved.

The concentration of salt during amplification may influence the yield as well as the primer specificity. Addition of 25 mM NaCl or more in the reaction mixture diminishes the yield of amplified DNA with T4 DNA polymerase (Fig. 5). This may be due to the fact that an increase in ionic strength decreases the binding of T4 DNA polymerase to single-strand DNA (Newport et al., 1985). This decreased binding would be expected to lower the efficiency of synthesis of the 184-bp fragment, as observed. Alternatively, an increased salt concentration could favor increased secondary structure within the single-strand (heat-denatured) DNA template (Dory et al., 1959; Studier, 1969). These intrastrand secondary structures constitute pause sites for T4 DNA polymerase (Challberg and Englund, 1979; Huang et al., 1981; Roth et al., 1982) and could consequently result in a decreased yield of the amplified product. To obtain the best yield, it is important that the salt content of all components (primers, DNA template, and dNTPs) be strictly controlled. The optimal salt concentrations however would be expected to vary with the base composition of the primers, i.e., with the relative proportions of A to GC. Preliminary experiments indicate that the amplification reaction from a plasmid template or even reamplification from a purified fragment is significantly different from the amplification from genomic DNA.

Findings of an efficiency of 67% per cycle with a plasmid template and 54% per cycle with a genomic sequence may indicate that there is a discrepancy between the two conditions. In addition, analyses of optimal dNTP concentrations as well as of the concentration of primer to be used also indicate a difference in the requirements for genomic amplification.

The use of one 5'-end-labeled primer in the reaction mixture permits the direct sequence analysis of the amplified DNA by the method of Maxam and Gilbert (1980). Amplification with either polymerase for 30 cycles, starting from 5 µg of genomic DNA, produced sufficient 32P-end-labeled DNA for sequencing. After purification, even a small amount of the amplified DNA could be used as a stock of template for reamplification to a larger quantity, if necessary.

ACKNOWLEDGMENTS

We thank Dr. T. Featherstone and Dr. B.J. Sanderson for helpful suggestions and K.T. Dreher for help in preparing the manuscript. This work was supported by the Office of Health and Environmental Research, U.S. Department of Energy, Grant DE-FG02-86ER60448, and National Institute of Environmental Health Sciences Program Project Grants 5-P01-ES01640 and 5-P01-ES00597.

REFERENCES


KUNKEL, T.A., ECKSTEIN, F., MILDVAN, A.S., KOPPLITZ,

Address reprint requests to:
Dr. Phothonie Kehayong
Room E18-666
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

NOTE ADDED IN PROOF
The results reported here have involved T4 DNA polymerase lot #2 from New England Biolabs. Subsequent experiments with lots #22 and 23 from the same supplier have consistently shown lower efficiencies during amplification of HPRT exon 3 from genomic DNA. While the efficiency of lot #21 was about 54%, the efficiency of lots #22 and 23 was about 43%. As a result, a higher number of cycles was necessary to see a band in an ethidium bromide stained gel. At 35 cycles a 224 bp band was visualized, but the presence of at least one other co-amplified band was also noted.

Received for publication April 24, 1987, and in revised form October 12, 1987.
6.1.4 Use of denaturing gradient gel electrophoresis to determine mutational spectrum in human cells


USE OF GRADIENT DENATURED GELS TO DETERMINE MUTATIONAL SPECTRUM IN HUMAN CELLS

Neal F. Cariello and William G. Thilly
Massachusetts Institute of Technology
Department of Applied Biology
Cambridge, Massachusetts 02139

ABSTRACT

Based on the fact that mutagens induce specific patterns of gene mutations, this paper outlines a method to allow discrimination among mutagen-treated populations. The technique should allow direct screening of human tissue for genetic change, using human peripheral blood lymphocytes deficient in the enzyme hypoxanthine guanine phosphoribosyl transferase. The method is based on gradient denaturing gel electrophoresis, which separates short DNA molecules according to their melting properties. The melting behavior of DNA fragments is extremely sequence-dependent, and DNAs with single basepair substitutions often migrate differently. Even DNA fragments with the same basepair substitutions at different locations in the molecule have been resolved. Gradient-denaturing gel electrophoresis has the capacity to separate mutant DNA on the basis of the nature and position of the mutation.

INTRODUCTION

Genetic changes occur in human offspring, but the specific causes of the changes remain unknown. Nearly 1500 specific locus mutations have been identified in the human population1, and it has been estimated that 1-2% of all liveborns suffer a specific locus mutation associated with significant pathology. Now, at the other end of the spectrum there exists a plethora of tests showing the mutagenicity of various compounds in bacteriophage, bacteria, yeast, and Drosophila. Many of these compounds have been detected in the environment; however, no causal link has been made between environmental mutagens and human genetic changes. If such a link exists, the development of a sensitive yet rapid assay for DNA alteration may permit the tracing of a particular genetic change to a specific environmental exposure.

The problem of discerning the relationship between environmental mutagens and human genetic change is clearly difficult given the minute concentrations of mutagens found in the environment and the low frequency of any specific genetic change. However, one biological phenomenon, the appearance of a specific set of mutations correlated with a given mutagen, offers a solution to this complex problem. The key to any such analysis of genetic change is the concept of mutational spectrum2. Mutagens do not act

439
randomly on DNA, rather they produce unique alterations at certain locations. In short, mutagens show specificity with regard to the kind and position of mutation; this specificity can be thought of as a chemical fingerprint on the DNA, this fingerprint can identify the specific mutagen to which a population of cells has been exposed. Mutational analysis of cells sampled from the human population could reveal which segments of the population have suffered genetic damage from a particular mutagen.

**MUTATIONAL SPECTRUM**

The formation of a chemically-induced mutation is a complex process. In some instances, the compound reacts directly with DNA, forming a premutagenic lesion. This premutagenic lesion can then be acted upon by repair processes and finally acted upon by the replicative machinery. The opportunity for mutagenic specificity exists at every step. The chemical may: 1) react only with certain nucleotides 2) react preferentially with nucleotides in a specific sequence, or 3) react preferentially with DNA in a particular chromatin formation. The repair enzymes may correct certain lesions at a greater frequency, and finally, the replicative machinery may differentially misincorporate nucleotides opposite the lesion, depending on the surrounding sequence. The result is a pattern of genetic change induced by a particular chemical. The pattern of genetic change, with regard to the kind and position of mutation, is termed the mutational spectrum. Importantly, the pattern of induced genetic changes is non-random, and varies from chemical-to-chemical. Three important experimental studies will now be examined.

Pioneering work was done by Benzner and Freese. They mapped spontaneous and bromouracil-induced mutations in the rII region of the bacteriophage T4. 126 spontaneous and 67 bromouracil-induced mutations were mapped, using recombination with a known set of deletion mutants. In this way, the mutations could be localized to a certain portion of the genome. Hotspots for both spontaneous and bromouracil-induced mutations were found, and, importantly, the hotspots were found at different locations in the rII region. Therefore, if an experimenter was blindly given two sets of rII mutants, one spontaneous and the other bromouracil-induced, it would be possible for him to determine which population had been exposed to bromouracil. It should be stressed that the mutational spectrum has another layer of complexity superimposed, namely the selective conditions. Not only must the premutagenic lesion be processed as described above, but in this case the mutation must produce an altered plaque morphology. Any mutation that does not score positively under the selective conditions will be missed. Benzner and Freese were therefore examining a subset of the total genetic changes induced by bromouracil, yet this subset possessed sufficient specificity to allow discrimination among spontaneous and bromouracil-induced mutations.

The site-specificity of mutagens has been extensively studied in the lacI gene of E. coli by Miller and co-workers. Supressible amber and ochre mutants in the lacI gene were selected and their location determined by crosses with deletion mutants. Since the sequence of the lacI gene was known, the precise location of each mutation and the particular basepair change could be determined. The pattern of mutation was determined for spontaneous change and for N-nitroquinoline-1-oxide (NQO), ethylmethanesulfonate (EMS), 2-aminopurine (AP), N'-methyl-N'-nitro-N-nitrosoguanidine (MNNG) and ultraviolet light (UV). These results are summarized in Figure 1. Benzner's earlier observations were confirmed, namely, that mutagens act specifically with regard to kind and position of mutation.
Figure 1. Distribution of 3738 independent E. coli mutations arising from the GC→AT transition. The number of independent occurrences at each site is indicated by bar height.

From reference 5.

In these studies, 72 amber and ochre sites which can arise from a single basepair change were utilized; only the AT→GC transition was not detected. This system essentially examined the mutational spectrum at 72 noncontiguous nucleotides, representing about 7% of the target, the lacI gene. The selection system involved two steps: first screening for lacI mutants, then examining these mutants for amber and ochre suppressibility.
The nonsense mutants were then mapped to a particular codon and the exact nature of the mutation deduced.

Importantly, the mutagens were shown to have different hotspots. Only those basepair substitutions which produced nonsense codons in the 72 bp "window" could be detected, yet even with these limitations of target size and mutation type it was possible to discriminate among mutagens. The foundation for molecular epidemiology has thus been laid. Although a great number of mutants must still be analyzed to determine the mutational spectrum for a particular compound, once the spectrum is known, a smaller number of mutants could reveal the "fingerprint" of a particular compound.

Another study made use of certain forward selection systems which monitor mutations at relatively few nucleotide pairs. These systems are all similar, in that the selective agent used seems to be toxic to wild-type cells because it binds to, and inactivates, essential proteins. Resistance presumably occurs from a specific modification of the target protein which renders it functional but refractory to selective agent binding and/or inactivation.

This approach was developed and used by Phaik-Mooi Leong who examined mutational spectra in a human lymphoblast line. She examined the ability of different mutagens to induce resistance to the drugs 6-thioguanine (TG), ouabain (OUA), dichloro-bis[uranosylbenzimidazole (DRB), and podophyllotoxin (PPT).

Only restricted kinds of mutations, missense or small addition/deletions that do not shift the reading frame, would be expected to confer resistance to such toxins. Nonsense and frameshift mutations should not be detected in this assay since these mutations would presumably destroy the protein's functionality. The mutations producing toxin resistance would be restricted in kind to certain missense mutations and restricted in place to certain positions. The actual kinds and positions would be defined by the structure of each essential protein and defined by the binding sites within it for an existing toxin. OUA, PPT and DRB appear to act via this type of mechanism.

TG, however, is not mediated by the same mechanism. Loss of functional hypoxanthine guanine phosphoribosyl transferase (HPRT) will result in the TG phenotype. Since the HPRT gene product is nonessential for cell survival, it is expected that this locus will be sensitive to all forms of gene mutation, including deletions and frameshifts.

In this assay, the data were normalized by expressing the ratio of DRUG to TG. The TG fraction is a reflection of the frequencies of all types of mutation, while the DRUG fraction only reveals that portion of cells that have suffered a missense mutation at the drug binding site. The TG fraction is taken as a general indication of the frequency of mutation, and if a mutational hotspot exists for a particular drug-binding site, then an elevated DRUG/TG ratio is expected. In this system, the "window" for mutational spectrum is the different drug-binding sites. The data are shown in Figure 2; the results are normalized with respect to TG.

The results show that the spontaneous mutant fraction is different from the EMS-, MNNG- or NQO-induced mutant fraction. Furthermore, the system can discriminate between EMS- and MNNG-induced mutants and between EMS- and NQO-induced mutants. The system cannot discriminate among spontaneous and ICR-191-induced mutants. This latter finding is particularly noteworthy, since it indicates that the system is behaving as expected. ICR-191 generates essentially only frameshift mutations in bacteria. But the "small marker" assay used here is not expected to be
sensitive to frameshift mutations since these mutations would destroy the essential gene product. The results suggest that ICR-191 is not causing basepair substitution mutations in the human lymphoblast line; extrapolating from the prokaryotic system, we can predict that ICR-191 is causing frameshift mutations. The results also suggest that EMS, MNNG and NQO are causing basepair substitution mutations. Furthermore, it appears that there is specificity of mutation induced by EMS versus MNNG, and EMS versus NQO. This work is the first report of mutational spectrum in human cells, and this work shows that human cell populations can be used to derive mutational spectra of sufficient precision to resolve differences among populations mutated spontaneously and those populations mutated by certain mutagens.

What are the possible approaches for determining mutational spectrum in human samples? Several criteria must be met before a method would be
deemed appropriate for human sampling:

1) the sampling method should be reasonable; that is, it should not cause the donor undue pain or discomfort.
2) a sufficient number of genetically detectable events must be present in a sample.
3) the events detected must be of sufficient diversity so that a mutational spectrum can be determined.
4) the protocol, when extended to human sampling, should not require excessive laboratory time or cost.

A relatively new approach to studying DNA double-stranded structure, called gradient-denaturing gel electrophoresis, may allow the resolution of mutational spectrum from human cells. This system, developed by Fischer and Lerman, should allow the detection and enumeration of point mutations in specific sequences within total human genomic DNA.

GRADIENT-DEНАTURING GEL ELECTROPHORESIS

The use of gradient denaturing gels is based on the fact that the mobility of DNA in polyacrylamide gels is sensitive to the secondary structure of the molecule. As the concentration of a denaturant is increased, DNA in solution passes through three discrete states: 1) totally helical 2) partially helical—partially random coil (melted), and 3) strand separation. As the concentration of denaturant is increased, certain sections of the DNA will make the transition from helix to random coil. Therefore, under certain denaturing conditions, a section of the DNA molecule can be expected to be helical while another section will be melted. This type of DNA molecule will show a markedly reduced migration rate on a polyacrylamide gel. In practice, one often finds several isomer melting regions of a few hundred basepairs, separated by very steep temperature gradients. The mobility of the molecule will be severely retarded when the lowest-temperature isomer melting domain becomes random coil.

Fischer and Lerman have developed a gel system to examine the melting behavior of DNA. The system employs an increasing gradient of urea and formamide. When the DNA enters the gel, it is completely helical, and it will migrate to that concentration of urea/formamide where its lowest-melting domain becomes random coil; when this occurs the DNA's migration rate is reduced.

The melting of a domain is extremely sequence-dependent. The melting of a domain has been described by Fischer and Lerman as a cooperative phenomenon, in which a single basepair substitution can affect the temperature at which the entire domain melts. Fischer and Lerman have examined the effect of known point mutations on the mobility of a 536 bp lambda fragment. The results of this experiment are shown in Figure 3. To understand the power of these gels, it is necessary to examine this figure in some detail. In some cases, a single basepair substitution can shift the fragment about one centimeter; for example, mutation cy3019, which is a GC to TA substitution at position 129, results in retardation of the fragment about 0.9 cm above wild-type. Even two mutants identical in overall DNA composition can be resolved; for example, a TA-to-GC or a TA-to-CG at the same lambda position can be separated. cy2001 is a TA-to-GC at position 136, while cy3071 is a TA to CG at the same position, yet these fragments are separated by about 2 mm on a gradient denaturing gel. It should be noted that the four mutants with substitutions above basepair 153 migrate as wild-type. The basepair substitutions for the lambda fragments are also given in Figure 3.
Fischer and Lerman have also developed an algorithm which predicts the melting behavior of DNA solely from its sequence. They have used the Fixman and Freire\textsuperscript{9} modification of Poland's\textsuperscript{10} algorithm for calculation of the equilibrium melting transition as a function of sequence. Fischer and Lerman have replaced the two-valued stability parameters for basepairs by the set of 16 values for nearest-neighbor doublets as suggested by Gotoh and Tagashira\textsuperscript{11}. The resulting algorithm can predict DNA melting solely from the DNA sequence. The melting map of the wild-type 536 bp lambda fragment is shown in Figure 4. Note that the DNA is organized into

![Melting Map](image)

**Figure 4.** Calculated melting map of wild-type lambda fragment.

*From reference\textsuperscript{8}.*
isomelting domains of a few hundred basepairs. Only mutations in the
lowest-melting domain (1-153 bp) affected the mobility in the gel.
Mutations affecting a higher-melting domain were not seen because this
higher melting-domain did not get the opportunity to denature.

Up to now, gradient-denaturing gels have been used only to study the
behavior of homoduplex DNA (perfectly paired). How would heteroduplex
(mismatched) DNA run on a gradient-denaturing gel? It is expected that the
mismatch would severely destabilize an isomelting domain, causing the domain
to melt at reduced temperature. As is shown in Figure 5, wild-type and
mutant DNAs containing a single basepair substitution were allowed to
anneal, forming four DNA species: wild-type, mutant, AC mismatch, and GT
mismatch. Four discrete bands representing the different DNA duplexes are
resolved; it is assumed that the least stable species is the AC mismatch,
followed by the GT mismatch DNA.

This type of heteroduplex analysis has recently been extended to
detect single basepair changes in human genomic DNA. A 32p wild-type
single-stranded cloned probe was hybridized to total DNA with known point
mutations in IVS1 of the human beta globin gene (See Figure 6). Four
mutations were examined; all four species were displaced from the wild-type
and, more importantly, all four mutants were resolved from one another, as
is shown in Figure 7. The specific mutations are:

IVS1, 1 is a G->A at basepair position 1
IVS1, 6 is a C->T at position 6
IVS1, 110 is G->A at position 110
IVS1, 5 is G->T at position 5.

Basepair position 1 refers to the 5' nucleotide of IVS-1. Note that
the same substitution, G->A in mutants 1 and 110, produces mismatched DNA
which runs to discrete locations on a gradient-denaturing gel. An
individual heterozygous for the beta-thalassemia mutation is shown in lane
2; both alleles are resolved. A homozygous patient is given in lane 3; as
expected a single band is seen.

Gradient-denaturing gel electrophoresis seems to be a reasonable
method for detection of mutational spectra. The spectra would indicate the
relative position and frequency of the mutants on a gradient-denaturing

denaturing gel. Left to right: 1) wild-type, 2) G to A
substitution, 3 and 4) wild-type and mutant heated and
reannealed separately, 5) wild-type and mutant heated and
reannealed together, 6 and 7) other samples.

Figure 5. Effect of mispairing, using cloned DNA in gradient-

From reference 12.
Figure 6. Diagram showing the location of the four beta thalassemia mutations in the human beta globin gene. Basepair substitutions are shown.

From reference 13.
gel. A number of mutants could be examined by the method developed by Myers, et al. For example, if a GC->AT transition hotspot for MNNG existed at position 1 of IVS1 of the beta-globin gene and a GC->AT EMS hotspot existed at position 110, these species could be resolved (See Figure 6). By analyzing a sufficient number of the EMS- and MNNG-induced mutants, the mutagenic specificity of these compounds could be discerned.

APPLICATION OF GRADIENT-DENATURING GEL ELECTROPHORESIS TO EXAMINE MUTATIONAL SPECTRUM IN HPRT

A gene locus must meet several criteria if it is to be useful for mutational analysis, including:

1) mutants must be present in sufficient numbers in the sample for spectral analysis,
2) a nonessential locus is desired, allowing accumulation of mutations without killing the cell,
3) in vivo and in vitro selection systems must exist.

The nonessential 5-linked enzyme hypoxanthine guanine phosphoribosyl transferase (HPRT) is a likely candidate for further investigation by gradient-denaturing gel electrophoresis. HPRT mutants can be selected and cloned from human peripheral blood, and the mutant fraction appears to be on the order of 10^{-5} to 10^{-6} \textsuperscript{14,15,16}. Therefore, a 100 ml blood sample would yield between 100 and 1000 mutants for further spectral analysis. To devise a concrete method for obtaining spectra from human HPRT we will have to examine the structure of the gene in some detail.

The cDNA coding for human HPRT has been cloned and sequenced\textsuperscript{17}. The organization of the human HPRT genome is now known; the entire gene spans about 35 kb, and the locus is split into nine exons. The length of the intervening sequences (IVS) and exons are shown in Figure 8 (P.L. Patel, personal communication). Many of the exons are small; the largest coding portion of a single exon is contained in exon 3.

Gradient-denaturing gel analysis requires contiguous genomic DNA. The most straightforward approach is to make restriction enzyme cuts within exons of individual mutants and to probe this section of the exon with the wild-type radiolabeled probe. In this way, a mismatch at the site of the mutation would be formed. The heteroduplex may run differently than wild-type on a gradient-denaturing gel. However, many of the exons are small and do not contain the appropriate restriction enzyme sites. Exon three lends itself to this direct approach; it is the largest exon, at 184 bp,
comprising 28% of the reading frame. Restriction sites exist near the 5' and 3' ends of the exon, such that a 161-bp fragment can be freed; this fragment represents 25% of the coding region.

Another possible approach would be to make the cuts in the mutant's DNA 5' and 3' of each exon and then hybridize this fragment to a wild-type fragment of the same size. In this way, all of the coding portion for a particular exon and some flanking sequence would be available for analysis. The exons are closely spaced at the 3' end of the gene. Here the proper 400 bp fragment would contain exons 7 and 8 with IVS.

Using the melting algorithm of Lerman, we generated the melting profile of the fragment of exon 3 that can be cleaved by restriction enzymes; this is given in Figure 9. The easily-cleavable 161 bp fragment shows a lower-melting domain from 300 to 385 bp; this represents 13% of the coding region. It is assumed that mutations in this section will all be displaced from wild-type using heteroduplex formation. Perhaps even mutations in the highest-melting domain can be resolved. A frameshift mutation in the high-melting domain should destabilize the heteroduplex even more than a simple point mutation, and it may be possible to resolve a frameshift in the high-melting domain, whereas a point mutation may not be seen. To recapitulate, the 161-bp section of exon 3 appears to have, at minimum, about 85 bp that may be useful for spectrum; these 85 bp comprise 13% of the entire coding region of HPRT.

At this point, two related approaches to mutational spectrum could be tried, a clone-by-clone approach and an approach that polls the mutants. For the clone-by-clone approach, the DNA from each mutant would be individually isolated, and one mutant per lane would be run on a gel. If 100 total mutants were isolated, perhaps 10 would run differently than wild-type. If a mutational hotspot exists, several of the mutants should run to the same spot on the gel. In the mass approach, the DNA from some number of mutants would be pooled and run in a single lane. A densitometer trace should reveal any hotspots.

Clearly the pooled approach is the more desirable. The DNA from a large number of mutants would be pooled, cut with the appropriate restriction enzymes, hybridized to single-stranded radioactive probe and run on a gradient denaturing gel. The majority of the DNA should run as wild-type but mutational hotspots should produce a non-wild-type signal. The intensity of a non-wild-type band on an autoradiogram would be proportional to the frequency of its occurrence. In this way, the mutational spectrum could be generated by a scanning densitometer trace of the film. The distance from wild-type would represent the specificity of

Figure 8. Organization of the human HPRT Organization. Total size is about 35 kb. Exon and intron sizes are given.
Figure 9. Melting profile of human HPRT exon 3 fragment. 161-bp fragment can be generated by restriction enzyme digestion. This 161-bp fragment represents about 25% of the coding region.

The mutation, while the intensity of the band would indicate the frequency of occurrence. This is summarized in Figure 10. If a group of high-risk individuals -- for example, chemical workers -- showed identical mutational spectrum, it could be argued that something common to the experience of these workers is causing the pattern of genetic change.

A central problem is the generation of a spontaneous spectrum for a given individual. For example, a person with a defect in a DNA repair enzyme may produce a pattern of mutation suggesting that he or she has been heavily exposed to environmental mutagens; such a conclusion would be in error. It is possible that each person can be his own control. A relatively small number of wild-type T-lymphocytes, say $10^4$, could be expanded in vitro to $10^8$; during this time spontaneous mutagenesis will occur. Very few previously existing mutants would be present in the starting population of $10^8$ lymphocytes. HPRT clones from the expanded population would then be analyzed; this pattern would comprise the spontaneous mutational spectrum. This assumes that spontaneous in vitro mutagenesis closely resembles spontaneous in vivo mutagenesis.

The spontaneous spectrum would then be compared to the spectrum found in a large sample of the person's blood, say $10^8$ lymphocytes; this is summarized in Figure 11. For the latter spectrum, selective pressure would be applied immediately, selecting HPRT mutants generated in vivo. If the spontaneous spectrum is not identical to the spectrum found in the large-
Figure 10. Schematic representation of mutational spectra derived from two mutagens, using gradient-denaturing gels.

Figure 11. Schematic representation of spontaneous and non-spontaneous spectra from a human donor. Spontaneous spectrum is generated from the in vitro expansion of a small number of cells. The non-spontaneous spectrum is generated by selection of existing in vivo mutants.
scale sample, it can be argued that this individual has suffered induced mutations. The final step would be to match the spectrum of an environmental mutagen with that of the non-spontaneous spectrum found in an individual; the individual might have been exposed to this mutagen.

REFERENCES


6.2 Melting programs

6.2.1 MELT program

Three files are necessary to run the MELT program, (1) a file which calculates the GTV values, (2) a short file which provides input parameters for the MELT program, and (3) the MELT program. The programs are presented below.

This is the file that generates the GTV values.

```
cNOTE--- the file contains (bp-1) entries. If you asked the program
to analyze 200bp the FOR016.DAT file has 199 entries. The exact
number of entries in the Gotoh-Tagashira file will be need to get
things working later.

c
PARAMETER NB=1000,ISEQ=1
DIMENSION GTM(4,4)
DIMENSION GTSEQ(NB,ISEQ)
INTEGER SEQ(NB,ISEQ),CHAR(4)/'A','T','G','C'/
DATA ((GTM(MU,MH),MU=1,4),MH=1,4)/54.50, 36.73, 86.44, 54.71,
+57.02, 54.50, 97.73,58.42, 58.42, 54.71, 85.97, 72.55,
+97.73, 86.44, 136.12, 85.97/
10 FORMAT(5(10A1,1X)5X)
300 DO 301 IS=1,ISEQ
 WRITE(6,1000)
1000 FORMAT(1X,'ENTER NUMBER OF BASES YOU WISH TO ANALYZE')
 READ(5,*)NBMAX
 READ(IS+10,10)(SEQ(MN,IS),MN=1,NBMAX)
99 DO 100 MN=1,NBMAX
109 DO 110 J=1,4
 IF(SEQ(MN,IS),NE.,CHAR(J))GO TO 110
 SEQU(MN,IS)=J
 GO TO 100
110 CONTINUE
C UNKNOWN BASE CONDE IF GOT TO HERE
 WRITE(6,11)MN
11 FORMAT(1X,'ERROR: ',215)
100 CONTINUE
C DETERMINE GOTOH TM FOR EACH SEQUENCE POSITION
 MMAX=NBMAX-1
199 DO 200 MN=1,MMAX
```
MU=SEQ(MN,IS)
MD=SEQ(MN+1,IS)
GTSEQ(MN,IS)=GTM(MU,MD)

200 CONTINUE
WRITE(15+IS,12) (GTSEQ(MN,IS),MN=1,MMAX)

301 CONTINUE
12 FORMAT(10F7.3)
END
This is the setup file for MELT.

1, 188, 2.0E-5, 60.0, 0.2, 12.25, 160
This is the MELT program

TO USE: You should have already generated the GT values from
the program GTV and edited the B4MELT program to contain the
exact number of entries in the output of the GTV program (re-
member that the output of the GTV program is [input-1]).

If there are no OBJ or EXE files for MELT you must generate
them. See GTV program if you can’t remember how to do this.

Assign the file with the GT values to Fortran block 011. The
output of the GTV program was FOR016.DAT. To do this type:
ASSIGN FOR016.DAT FOR011

Assign the file B4MELT to Fortran block 003. Type in:
ASSIGN B4MELT.TXT FOR003

To run type: RUN MELT

Output with melting temperatures is FOR021.DAT

What follows is Lerman’s stuff, for those of you who may be
slightly curious.

G.MELT22 4*25*84

ENTRY FROM TERMINAL AFTER PROMPT PROVIDES EITHER MELTING MAP OR
HELIX PROBABILITY AT CONSTANT TEMPERATURE. PROBABILITY OUTPUT
IS IN FILE 20.

TO USE: BE SURE THAT THE FIRST LINE OF THE PROGRAM, 26, CONTAINS
DIMENSIONS AT LEAST AS LONG AS THE SEQUENCE TO BE ENTERED AS
MLENTH AND AT LEAST AS MANY TEMPERATURES AS ARE TO BE CALCULATED
AS NHMMX. NHMMX=(HIGHEST TEMP - TINIT)/TSTEP OR MORE. ENTER THE
NAME OF THE FILE CONTAINING THE GT VALUES OF THE BASIC SEQUENCE
AS LINE 16 OF GELS.FIRSTM20. LINE 9 MUST SPECIFY THE POSITIONS
OF THE FIRST AND LAST BASES IN THE CALC. LINE 11 CONTAINS THE
LOWEST TEMPERATURE, TINIT. LINE 12 HAS THE TEMPERATURE INCREMENT
TSTEP, AND LINE 14 HAS THE TOTAL NUMBER OF TEMPERATURES, NNM, NOT
MORE THAN THE VALUE OF NHMMX. AS LONG AS MLENTH AND NHMMX ARE LARGE
ENOUGH, THEY DO NOT HAVE TO BE CHANGED. START WITH
8ADD GELS.FIRSTM20, THEN 8ADD GELS.MELT22, WHEN PROMPTED ENTER
8XT AB22. ON COMPLETION OF THE CALC, COPY THE MELTMAP FROM 21.,
THE INVERSE TRANSITION WIDTH (AS ENTHALPY) FROM 22., AND THE
TRANSITION DIFFERENCES BETWEEN ADJACENT BASES FROM 23.

PARAMETER MLENTH=800, NHMMX=200
DIMENSION ST(MLENTH),B(MLENTH),T(MLENTH),P(MLENTH),A(10),B(10)
DIMENSION TM(MLENTH),E(10),D(10),BEXP(10),PDAR(NHMMX)
DIMENSION PI(MLENTH),P2(MLENTH),DTH(MLENTH),COHER(MLENTH)
DATA SIG,TSTEP,NNM,MMIN,MMAX/3.3E-05,0.2,250,1,MLENTH/
DATA TINIT/50.0/, DEL/12.25/
DATA A/4.5658083E+01, 8.4808384E+00, 1.2355326E+00,
+8.1231263E-01, 0.2553350E-02, 0.3549439E+03,
+4.5607812E-04, 0.9830478E-05, 0.1361430E-05,
+1.3474846E-06/
DATA B/2.5257303E+01, 8.4211493E+00, 2.9293422E+00,
+1.0287631E+00, 0.3619084E-01, 0.1270586E-01,
+4.4247335E-02, 0.1989989E-02, 0.4635086E-03,
+1.0067630E-03/
DATA DSKL1,DSKLG2/2*1,E-10/, TDKB /0.02/
READ(3,*)MMIN,MMAX,SIG,TINIT,TSTEP,DEL,NNM
DO 12 I=1,10
BEXP(I)=EXP(-B(I))
A(I)=A(I)*SIG
12 CONTINUE
100 FORMAT(10F7.3)
READ(11,100)ST(M),N=1,MMAX)
MPEN=MMAX-1
LENGTH=MMAX-MMIN+1
C 2000 FORMAT()
WRITE(6,2200)
2200 FORMAT(IX,'FOR HELIX PROBABILITY AT CONST TEMPERATURE, ENTER
-1, FOR MELTING MAP, ENTER 2')
READ(5,*)II1
IF(II1.GT.1)GO TO 36
15 WRITE(6,2300)
2300 FORMAT(IX,'ENTER TEMPERATURE, DEG CELSIUS')
READ(5,*)AT
K=1
GO TO 98
36 DO 1700 K=1,NMN
AT=TINIT+K*TSTEP
98 DO 8 M=MMIN,MMAX
X=((ST(M)+273.1)/(AT+273.1))*DEL
R(M)=EXP(DDEL-X)
8 CONTINUE
BETA=R(MMAX)
Q=0
92 DO 2 I=1,10
E(I)=0.
2 CONTINUE
T(MPEN)=R(MPEN)/(1+R(MMAX))
64 DO 4 M=MPEN,MMIN,-1
Q=0
93 DO 3 I=1,10
E(I)=T(M)*BEXP(I)*(1+E(I))
Q=Q+A(I)*E(I)
3 CONTINUE
BETA=T(M)*BETA
T(M-1)=R(M-1)/(1+Q+BETA)
4 CONTINUE
GAM=1
GSUM=G0
W=0
95 DO 5 M=MMIN,MPEN
GAM=T(M)*GAM
GSUM=GSUM+GAM
5 CONTINUE
GAM=1
91 DO 7 I=1,10
D(I)=0.
1 CONTINUE
P(MMIN)=1/(1+GSUM)
PSUM=0.
96 DO 6 M=MMIN,MPEN
PSUM=PSUM+P(M)
PCON=1/(1+PSUM)
6 P(M)=PCON
96 P(M)=P(M)
7 CONTINUE
GAM=T(M)*GAM
P(M)=R(MMIN)*GAM+PP+W
IF(TM(M).GT.0.0) GO TO 6
P(M)=P2(M)
P2(M)=P(M)
IF(P1(M).LE.5) AND (P2(M).LE.5) THEN
CALL VNTREF(P1(M),P2(M),AT,TO,TSTEP,EN)
TM(M)=TO
COHER(M)=1.9865*EN/((TO+273.1)*DEL)
ELSE
GO TO 6
END IF
6 CONTINUE
25 PBAR(K)=MMAX-MMIN-PSUM
   ATHT=-3.0-PBAR(K)/LENTH
   BDIS=5.E+03*LENTH**2ATHT
   DISK=R(MMAX)*P(MMIN)*GAM/BDIS
   IF(I21.EQ.2) GO TO 300
   WRITE(20,150)(P(M),M=MMIN,MMAX)
   WRITE(20,*)AT, MMIN, MMAX, DISK
   WRITE(6,2100)
2100 FORMAT(1X,'TO REPEAT CALC, NEW TEMPERATURE, ENTER 9')
   READ(5,*)22
   IF(Z22.EQ.9) GO TO 15
300 IF(DISK.LT.1.E-20) DISK=1.E-20
   DSKLG=ALOG10(DISK)
   DSKLG1=DSKLG2
   DSKLG2=DSKLG
   IF((DSKLG1.LT.-8.) .AND. (DSKLG2 GE.-8.)) THEN
      TDK8=AT-((DSKLG2+8.)/(DSKLG-DSKLG1))*TSTEP
   ELSE
      GO TO 1700
   END IF
1700 CONTINUE
   DTM(MMIN)=0.
200 DO 201 M=MMIN+1,MMAX
   DTM(M)=TM(M)-TM(M-1)
201 CONTINUE
150 FORMAT(10F7.4)
47 WRITE(21,100)(TM(M),M=MMIN,MMAX)
   WRITE(22,2002)(COHER(M),M=MMIN,MMAX)
2002 FORMAT(10F7.2)
   WRITE(21,*)TDK8
71 CONTINUE
120 FORMAT(10F8.4)
END
C---------------------------------------------------------------
SUBROUTINE VNTRP(Y1,Y2,AT,X,TSTEP,EN)
   F1=LOG(Y1/(1.-Y1))
   F2=LOG(Y2/(1.-Y2))
   V1=1./(AT-TSTEP)
   V2=1./AT
   EN=1.9865*(F2-F1)/(V2-V1)
   V0=V1-1.9865*F1/EN
   X=1./V0
RETURN
END
6.2.2 SQHT program

This is the setup file for SQHT.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5e-05</td>
<td>1 ' sigma'</td>
</tr>
<tr>
<td>60.0</td>
<td>2 ' TINIT'</td>
</tr>
<tr>
<td>0.3</td>
<td>3 ' Tstep'</td>
</tr>
<tr>
<td>12.25, 8, 10.0</td>
<td>4 ' (deltaS)/R, LFEND, ENDB'</td>
</tr>
<tr>
<td>75.0</td>
<td>5 ' retardation length'</td>
</tr>
<tr>
<td>'224bp.seq'</td>
<td>6 ' input sequence file'</td>
</tr>
<tr>
<td>'(5(10A1,1X),5X)'</td>
<td>7 ' sequence format'</td>
</tr>
<tr>
<td>224</td>
<td>8 ' length of molecule'</td>
</tr>
<tr>
<td>1</td>
<td>9 ' section for destab series, first doublet'</td>
</tr>
<tr>
<td>222</td>
<td>10 ' last doublet in series'</td>
</tr>
<tr>
<td>85.0</td>
<td>11 ' upper temperature limit'</td>
</tr>
<tr>
<td>1.00</td>
<td>12 ' CLOCK'</td>
</tr>
<tr>
<td>0.00</td>
<td>13 ' running time before time series starts'</td>
</tr>
<tr>
<td>0.0</td>
<td>14 ' POWER'</td>
</tr>
<tr>
<td>2.0</td>
<td>15 ' initial mobility'</td>
</tr>
<tr>
<td>50.0</td>
<td>16 ' decrease in Tm of doublets with mismatch'</td>
</tr>
<tr>
<td>'224.dat_sqht'</td>
<td>17 ' output file for results'</td>
</tr>
</tbody>
</table>
This is the SHQT program.

```
PROGRAM SHQT
sqt2x 2/8/88
The calculation gives the difference in depth in
the gradient, YDISP(site,tie), between the wild type molecule and
a similar sequence in which the stability temperatures of the two
adjacent doublets that enclose the site are each reduced by the
amount, DESTAB. The length of the molecule, the first and last
sites in the section in which destabilization at each site is to
be examined, the value of DESTAB, the temperature at the top of the
gradient, the name and format of the sequence file, the output
file, and other parameters are read from the set-up file, sqt2x.

Time values are determined from time=start+clock*t
Temperatures for dissociation constants of E-12,-10,-8,-6,-4 are calc.

All the print*9 eg statements removed by COMMENT to streamline this
thing. 20JUN88

PARAMETER (MBNX=1000, MMNX=200)
INTEGER ISQ(MBNO), NUM(4), ITEM, MFIRST, MLA, NMN, IND
CHARACTER*4 1 SEQ(MBNO), CHARUC(4), CHARLC(4)
CHARACTER*20 SQFILE, YDPSFL, SQRTNT
CHARACTER SPEC=40
REAL GT(MBNO), A(10), B(10), BEXP(10), PBAR(MMNX), PBAR1(MMNX), GTM(4,4)
REAL P(MBNO), R(MBNO), T(MBNO), FSTAB(MBNO), ST(MBNO), RETL
REAL Y0(10), YSUB(10), YDISP(MBNO,10), SIG, TSTEP, DEL
REAL START, CLOCK, POWN, TIME(10), TINIT, THAX, TDX(5), TDX0(5)
DATA SIG,TSTEP,MAX/3.3E-05,0.2,536/
DATA TSTEP/50.0/, DEL/12.25/, TMAX/95.0/
DATA A/65680B3E+01, 8408834E+00, 1255532E+00, +1871128E-01, .2854330E-02, .4354949E-03, +.6607812E-04, .9304079E-05, .1361403E-05, +.1374848E-06/
DATA B/.2525730E+01, .8421140E+00, .2929342E+00, +.10287619E+00, .3619084E-01, .1270588E-01, +.4423732E-02, .1498889E-02, .4635086E-03, +.1006763E-03/
DATA CHARUC 'A', 'T', 'G', 'C' /
DATA CHARLC /'a', 't', 'g', 'c' /
DATA (GTM(NUM,ND),NU=1.4,ND=1.4)/54.50, 36.73, 86.44, 54.71,
+57.02, 54.50, 97.73, 58.42, 58.42, 54.71, 85.97, 72.55,
+97.73, 86.44, 136.12, 85.97/
DATA IFILE/15/, TDX/5.0/, TDX0/5.0/
OPEN(3, FILE='sqt2x', status='OLD')
READ(3,*),SIG,ITEM,SPEC
READ(3,*),TSTEP,ITEM,SPEC
READ(3,*),TSTEP,ITEM,SPEC
READ(3,*),DEL, LFEND, ENDSTAB, ITEM, SPEC
READ(3,*),RETL,ITEM,SPEC
READ(3,*),SQFILE,ITEM,SPEC, SQRTNT, ITEM, SPEC
READ(3,*),SQFILE,ITEM,SPEC, MFIRST,ITEM,SPEC, MLA,ITEM,SPEC,
+MAX,ITEM, SPEC, CLOCK,ITEM, SPEC, START,ITEM, SPEC,
+POWR,ITEM, SPEC, VEL0,ITEM, SPEC, DESTAB,ITEM, SPEC
READ(3,*),YDPSFL,ITEM,SPEC
print *,nb,mfirst,mlast, clock, start, POWN, DESTAB
2110 FORMAT('SITE',1X,10F7.2)
2111 FORMAT(I4,1X,10F7.3)
2112 FORMAT(A)
2113 FORMAT(A)
2114 FORMAT(6I7)
2115 FORMAT(8F8.3)
2116 FORMAT(1X,3I5,2F8.2)
2117 FORMAT(10F7.3)
11 DO 12 I=1,10
2118 BEXP(I)=EXP(-B(I))
A(I)=A(I)*SIG
12 CONTINUE
OPEN(10, FILE='SQFILE, STATUS='OLD')```

READ(10, SQFRMT) (SEQ (MN), MN=1,NB)
99 DO 100 MN=1,NB
100 DO 110 JB=1,4
  IF((SEQ(MN),NE.CHARUC(JB)),AND.(SEQ(MN),NE.CHARLC(JB))) GO TO 108
  SEQ(MN)=JB
  GO TO 100
108 IF(JB.EQ.4) THEN
  PRINT*, 'sequence error at ',MN
  GO TO 999
ENDIF
110 CONTINUE
100 CONTINUE
print*, '100',
NMAX=NB-1
120 DO 121 MN=1,NMAX
  IBU=ISEQ(MN)
  IBD=ISEQ(MN+1)
  GT(MN)=GTM(IBU,IBD)
121 CONTINUE
870 DO 871 MN=1,NMAX
  ENDS=REAL(M)
  END3=REAL(MMAX-M+1)
  FSTAB(M)=ENDSTB*(EXP((-END5/LFEND)+EXP(-(END3)/LFEND))
871 CONTINUE
NHN=INT((TMAX-TINIT)/TSTEP)+1
INDEX=0
CALL MELTFR (MNBX,MMAX,NNN,TINIT,TSTEP,GT,A,BEXP,DEL,PBARI,P,R,T,
+FSTAB,ST,TDX,INDX)
C print*, (PBARI(K),K=1,NNN,20)
CALL TRAVEL(PBARI,TINIT,TSTEP,NNN,VELO,RETL,CLOCK,START,POWER,
+TIME,YO,T,MMAX)
C print*, (Y(J),J=1,NNN,20)
WRITE(*,2111) NN,J=1,10)
C DETERMINE GOTOH TM FOR EACH TRIPLET INCLUDING A SUBSTITUTION
IF(MFIRST.LT.2) MFIRST=2
IF(MLAST.GT.(NB-2)) MLAST=NB-2
239 DO 240 MN=MHFIRST,MLAST
  NNN=GT(MN-1)
  GT(MN)=GT(MN-1)-DESTAB
  GT(MN)=GT(MN)+DESTAB
C print *, NN,' ',SEQ(MN),SEQ(MN+1),',',GT(MN-1),GT(MN)
INDEX=NN
CALL MELTFR (MNBX,MMAX,NNN,TINIT,TSTEP,GT,A,BEXP,DEL, PBARI,
+P,R,T,FSTAB,ST,TDX,INDX)
C print*, (PBARI(K),K=1,NNN,20)
CALL TRAVEL(PBARI,TINIT,TSTEP,NNN,VELO,RETL,CLOCK,START,POWER,
+TIME,YSUB,T,MMAX)
WRITE(*,2111) NN,J=1,10)
C print*, (TDXQ(ID), ID=1,5)
350 DO 351 J=1,10
  YDSP(MN,J)=YSUB(J)-YO(J)
351 CONTINUE
GT(MN)=YDSP
GT(MN)=NN
240 CONTINUE
C PRINT *, YDSPFL
OPEN(iFILE, FILE=YDSPFL, STATUS='NEW')
WRITE(iFILE,2110) 'TIME(J),J=1,10'
WRITE(iFILE,2113) 'YDSP(MN,J),J=1,10)
WRITE(iFILE,2111)NN,J=1,10)
WRITE(iFILE,2113) 'LENGTH=NB,' RETRANTATION LENGTH', RETL
WRITE(iFILE,2113) 'INITIAL TEMPERATURE='TINIT,' VELZERO='VELO
210 DO 211 ID=1,5
  XD=3* (6-ID)
WRITE(IFILE,*)'T for logKd=',XD,'   = ',TDX0(ID)
CONTINUE
WRITE(IFILE,*)'START, CLOCK, POWR=',START,CLOCK,POWR
WRITE(IFILE,*)'YDSPL=',destab=','DESTAB,' from ',' SQFILE
WRITE(IFILE,2113)
END

SUBROUTINE MELTPR(MBXM,MMAX,MAXTPT,TSTART,TSTEP,ST,A,BEXP,DEL,
PBAR,PR,T,FSTAB,ST,TDX,INDEX)
INTEGER I,M,R,K,ID
REAL GT(MBXM),R(MBXM),T(MBXM),P(MBXM),PBAR(MAXTPT)
REAL X(10), BEXP (10), FSTAB (MBXM), ST (MBXM)
REAL E(10),D(10),TDX(5), XD
DATA DSKLO2/1.0E-18/, KZ/0/
DATA E/10**0.0/, D/10**0.0/

C print *,mbxm,mmmax,maxtpt,tstart,tstep
DO 51 M=1,MMAX
      ST(M)=GT(M)+FSTAB(M)
51 CONTINUE
DO 36 K=1,MAXTPT
      AT=TSTART+K*TSTEP
98 DO 8 M=1,MMAX
      X=((ST(M)+273.1))/((AT+273.1))*DEL
      R(M)=EXP (DEL-X)
8 CONTINUE
C print*,'/8'
      BETAR=1/(MMAX)
      Q=0
      MPEN=MMAX-1
      T(MPEN)=R(MPEN)/(1+R(MMAX))
64 DO 4 M=MPEN,2,-1
      Q=0
93 DO 3 I=1,10
      E(I)=T(M)*BEXP (I)*(1+EXP(I))
      Q=Q+A(I)*E(I)
3 CONTINUE
C print*,'/3'
      BETAT=T(M)+BETA
      T(TM+1)=R(TM+1)/(1+Q+BETAT)
4 CONTINUE
C GAM=1
      GSUM=0
5 CONTINUE
C print*,'/5'
      GAM=1
      P(1)=3/(1+GSUM)
      FSUM=0.
96 DO 6 M=1,MPEN
      FSUM+PSUM=P(M)
      PCON =T(M)/R(M)
      PP=P(M)*PCON
      N=0
97 DO 7 I=1,10
      W=W+A(I)*D(I)
      D(I)=T(M+1)*BEXP (I)*(PP+D(I))
    7 CONTINUE
      GAM=T(M)*GAM
      P(M+1)=P(1)*GAM+PP+W
6 CONTINUE
C print*,'/6'
      PBAR(K)=MMAX-PSUM
      ATHAT=2.8-3.2*PBAR(K)/MMAX
      BDIS=5.E+03*MMAX**ATHAT
DISK = R(MMAX) * P(1) * GAM/BDIS
IF (INDX.GT.0) GO TO 1700
IF (DISK.LT.1.E-16) GO TO 1700
IF (DISK.GT.1.E-03) GO TO 1700
C
print *,DISK
DSKLG = ALOG10(DISK)
DSKLG1 = DSKLG2
DSKLG2 = DSKLG
C print*,DSKLG1,DSKLG2
110 DO 111 ID = 1,5
   XD = 9.-ID
   IF ((DSKLG1.LT.XD).AND.(DSKLG2.GE.XD)) THEN
      KT = KT + 1
      TDX(KT) = AT - ((DSKLG2-XD)/(DSKLG-DSKLG1)) * TSTEP
   END IF
111 CONTINUE
1700 CONTINUE
RETURN
END

SUBROUTINE TRAVEL(PBAR,TINIT,TSTEP,MAXTPT,VEL0,RETL,CLOCK,START,
PBAR,TMAXTPT), Y(10), T(MAXTPT), CLOCK, TIME(10)
TSM=0.
30 DO 31 K = 1, MAXTPT
   CORBL=(PBAR(K)/RETL) * (1-EXP(-(MMAX-PBAR(K))/150.))
   RECVEL=EXP(PBAR(K)/RETL)
C RECVEL=EXP(CORBL)
   TSM=TSUM+RECVEL
   T(K) = TSTEP * TSM / VEL0
C PRINT*,CORBL,RECVEL,TSUM,T(K)
31 CONTINUE
400 DO 401 J = 1, 10
   TIME(J) = START + CLOCK * J
20 DO 21 K = 1, MAXTPT
   IF (TIME(J), LE, T(K)) THEN
      Y(K) = TIME(J) / T(K)
      Y(J) = Y(K) * TSTEP + TINIT
      goto 401
   ELSE IF (TIME(J), GT, T(K)) AND (TIME(J), LE, T(K+1)) THEN
      Y(K) = (TIME(J) - T(K)) / (T(K+1) - T(K)) + K
      Y(J) = Y(K) * TSTEP + TINIT
      goto 401
   END IF
21 CONTINUE
401 CONTINUE
RETURN
END
References


Goth, R. and Rajewsky, M.F. 1974. Molecular and cellular mechanisms associated with pulse-carcinogenesis in the rat nervous system by ethyl nitrosourea: ethylation of nucleic acids and elimination rates of ethylated bases from the
DNA of different tissues. Z. Krebsforsch., 82: 37-64.


Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis,


\(N\)-nitroso-\(N\)-methylurea during initiation