TOWARDS THE GENERATION AND CHARACTERIZATION OF ARYLAMINE N-ACETYLTRANSFERASE DEFICIENT MICE

by

Hanan Abramovici

A Thesis submitted in conformity with the requirements for the Degree of Master of Science
Graduate Department of Pharmacology
University of Toronto

©Copyright by Hanan Abramovici 2000
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
ABSTRACT

TOWARDS THE GENERATION AND CHARACTERIZATION OF ARYLAMINE N-ACETYLTRANSFERASE DEFICIENT MICE

Hanan Abramovici
Degree of Master of Science, 2000
Graduate Department of Pharmacology
University of Toronto

N-acetyltransferases (NATs) are enzymes of phase II drug biotransformation that metabolize aromatic amine-containing compounds found in clinically useful drugs, in environmental pollutants and in cooked foods. Experimental studies have shown that NATs have both detoxifying and toxifying roles in the body. Epidemiological and animal studies have implicated NATs in drug-induced toxicological reactions and in carcinogenesis following exposure to aromatic amines. However, human epidemiological studies are limited by difficulties in accurately assessing the types, levels and lengths of exposure of individuals to aromatic amine compounds. Presently available animal models, while useful, have not yet truly revealed the overall or relative importance of mouse N-acetyltransferases in toxicology and disease. To avoid some of the inherent weaknesses in these studies and to more clearly understand the role of NATs in drug induced toxicity and carcinogenesis, the overall goal of this research was to create and characterize novel NAT deficient mouse strains.

Our first objective was to produce a NAT deficient mouse strain by gene targeting using R1 ES cells by: 1) introducing the mouse NAT1/ mouse NAT2 and mouse NAT3 gene targeting vectors into R1 ES cells, 2) designing PCR and Southern blotting screening methods to enable the detection of the mouse NAT targeted knockouts. Our second objective was to produce isozyme specific anti-mouse NAT antibodies for use in characterizing these strains by: 1) designing and constructing three mNAT-GST fusion protein antigens, 2) expressing the fusion proteins in bacteria and 3) purifying and injecting the antigens into rabbits to produce isozyme specific anti-mouse NAT antibodies.

We electroporated R1 ES cells with the mouse NAT gene targeting vectors, and selected and screened clones using the Southern blotting and PCR screening methods. We isolated several targeted mNAT3 ES cell clones which we will use to generate mouse NAT3 deficient mice. We have also successfully expressed the mouse NAT- GST fusion protein antigens in bacteria and isolated them for use in raising rabbit anti-mouse NAT antibodies.
The isolation of the heterozygously targeted \textit{mNAT3} ES cell clones represents a crucial first step in the production of a mouse NAT3 deficient strain and ultimately in the creation of completely NAT deficient mice. The successful production of three mouse NAT-GST antigens will be used to generate anti-mNAT antibodies which will find wide application in the characterization of the NAT deficient strains, in studies of mNAT isozyme tissue specific expression, and in measuring the specific activities of each isozyme. Furthermore, these antibodies will be useful in inhibition studies aimed at dissecting the contribution of each mouse NAT isozyme to aromatic amine biotransformation. Globally, the creation of mNAT deficient strains and their characterization will lead to a better understanding of the role of mNATs in toxicology and cancer.
ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Denis Grant for his valuable scientific input, guidance and support throughout the duration of my studies. I thank my advisor Dr. Richard Rozmahel for his valuable scientific input into this study.

I would like to especially thank Dr. Kim Sugamori who did the work in sections 2.2.2.1., 2.2.2.2., 2.2.2.6., and 2.2.2.10. of this thesis. Many thanks to Mrs. Jean Huang, Ms. Jennifer Skauag, Mrs. Xiaoli Lu, and Mrs. Violeta Yu- Plant for their help during the course of my studies. I would also like to thank the members of Dr. Denis Grant's lab, Geoff Goodfellow, Julie Green, and Kristi Durette for their helpful advice, humor and for making this an enjoyable experience.

I would like to warmly thank my parents for their love, guidance and support. This thesis is dedicated to them and to my grandparents who encouraged me to pursue a scientific education.

I would like to thank the National Cancer Institute of Canada (NCIC) for their financial support throughout the duration of these studies.
TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
TABLE OF CONTENTS iv
LIST OF FIGURES x
LIST OF TABLES xiv
LIST OF ABBREVIATIONS xv

PART ONE: INTRODUCTION 1

1.1. General introduction 1
1.2. Background 1

1.2.1. Genetic polymorphisms of DMEs 2

1.3. Arylamine N- Acetyltransferases (NATs) 8

1.3.1 Human NATs 8

1.3.1.1. Discovery and cloning of human NATs 8
1.3.1.2. Catalytic and kinetic NAT mechanisms 11
1.3.1.3. Localization of human NATs 13
1.3.1.4. The role of human NATs in biotransformation of xenobiotics 14
1.3.1.5. Genetic polymorphisms of human NATs and implications for carcinogenesis 18
1.3.1.6. Endogenous functions of human NATs 20

1.3.2 Mouse N- Acetyltransferases (mNATs) 20

1.3.2.1. Cloning of the mouse NATs 20
1.3.2.2. Mouse NAT nucleotide and amino acid homology 20
1.3.2.3. Mouse NAT kinetics 22
1.3.2.4. Tissue localization of mouse NATs 22
1.3.2.5. Genetic polymorphisms of mouse NATs 24
1.3.2.6. Modifying genes and NATs 25
1.3.2.7. Role of mouse NATs in carcinogenesis 25
1.3.2.8. Endogenous roles of mouse NATs 27

1.4. Gene targeting in the mouse 28
  1.4.1. Background 28
  1.4.2. Gene targeting vector design 31
  1.4.3. Transfection of ES cells and subsequent selection of targeted ES cells 33
  1.4.4. Screening of ES cells for targeted knockout of the gene 34
  1.4.5. Reconstitution of functional blastocysts 35

1.5. Overall goal, rationale and objectives 39
  1.5.1. Overall goal 39
  1.5.2. Rationale 39
  1.5.3. Specific objectives 40

PART TWO: MATERIALS AND METHODS 41

2.1. Materials 41
  2.1.1. Vectors, clones and bacterial strains
  2.1.2. Media and antibiotics for the propagation of bacterial strains 44
2.1.3. Isolation of DNA cloning and protein expression plasmids

2.1.4. Restriction enzymes, DNA modifying enzymes, DNA size ladders

2.1.5. Southern blots, Western blots, Sequencing, PCR, XL PCR

2.1.6. Embryonic Stem (ES) cell culture

2.1.7. mNAT- GST fusion protein expression

2.2. Methods

2.2.1. General methods

2.2.1.1. Bacterial growth conditions

2.2.1.2. Preparation of competent cells

2.2.1.3. Transformation of bacterial cells

2.2.1.4. Small scale isolation of plasmid DNA

2.2.1.5. DNA isolation from BAC clones

2.2.1.6. Isolation and purification of genomic DNA from 129 Sv/J mouse liver

2.2.1.7. Determination of DNA concentrations

2.2.1.8. Restriction enzyme digestions, electrophoresis, DNA fragment isolation and purification, dephosphorylations, and ligations

2.2.1.9. DNA sequencing

2.2.1.10. Polymerase Chain Reaction (PCR) amplification

2.2.1.11. Southern blotting

2.2.1.12. DNA probe labelling
2.2.2. Specific methods

2.2.2.1. ES cell culture

2.2.2.1.1. Standard growth conditions

2.2.2.1.2. Electroporation of ES cells with Vectors 1 and 2

2.2.2.1.3. Colony isolation and expansion

2.2.2.1.4. Freezing of ES cells

2.2.2.2. Preparation of Vectors 1 and 2 for electroporation

2.2.2.3. Production of the Vector 1 positive control for use in screening gene targeting events by PCR

2.2.2.4. Testing of the positive control for Vector 1

2.2.2.5. Cloning and sequencing of additional NAT3 3' UTR

2.2.2.6. Screening of the mouse NAT3 targeted knockout by Southern blotting

2.2.2.7. Production of the Vector 2 positive control for use in screening gene targeting events by XL PCR

2.2.2.8. Testing of the XL PCR reaction on the positive control and screening of the Vector 2 targeted events

2.2.2.9. Subcloning and sequencing of additional mouse NAT2 3' UTR

2.2.2.10. Screening of mouse NAT1 and 2 targeted knockout by Southern blotting

2.2.2.11. Subcloning and sequencing of additional mouse NAT1 5'UTR

2.2.2.12. Production of mouse NAT antigens

2.2.2.13. Conditions used in mNAT-GST fusion protein expression

2.2.2.13.1. Pilot experiment expression conditions
2.2.2.1.3.2. Bradford protein assay - standard curve preparation 69

2.2.2.14. Western blotting 70

2.2.2.15. Solubilization of inclusion bodies 72

PART THREE: RESULTS 74

3.1. Design and testing of Vector 1 positive control for PCR based screening of mouse NAT3 targeted knockout 74

3.2. XL PCR amplification and sequencing of additional mNAT3 3' UTR 74

3.3. Electroporation of ES cells with Vector 1 and selection 79

3.4. Design of the Vector 2 positive control for the XL PCR based screening method for Vector 2 targeted knockout of mNAT1 and mNAT2 84

3.5. Design of Southern blotting screening method to verify the XL PCR results of the targeted knockout of mNAT1 and mNAT2 84

3.6. Electroporation and selection of ES cells targeted with Vector 2 85

3.7. Design of Southern blotting screening assay for detection of targeted knockout of mNAT1 and mNAT2 85

3.8. Electroporation of ES cells with Vector 2 and selection 92

3.9. Subcloning and sequencing of additional mouse NAT1 5' UTR 92

3.10 Production of mNAT antigens 95

3.11. mNAT- GST expression studies 95

3.12. Isolation and solubilization of the mNAT1 and mNAT3 GST fusion proteins from inclusion bodies 101

PART FOUR DISCUSSION 104

4.1. Overview 104

4.2. Mouse NAT3 targeted knockout 104
4.3. Mouse NAT1/2 targeted knockout 106
4.4. Reconstruction of Vector 2 108
4.5. Production of mouse NAT antigens 109
4.6. Initial pilot expression studies of the mouse NAT- GST fusion proteins 111
4.7. Solubilization of the mouse NAT1 and 3 GST fusion proteins 112
4.8. Summary 115
4.9. Future studies and conclusion 116

BIBLIOGRAPHY 118
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.0</td>
<td>The fate of drugs and toxins in the body</td>
<td>3</td>
</tr>
<tr>
<td>Figure 1.1</td>
<td>The interaction of various factors, intrinsic and extrinsic to the body exerting changes on the genetic profile of an individual which in turn can affect drug response or biotransformation of xenobiotics</td>
<td>4</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Bimodal distribution of plasma concentration of isoniazid in a human population reflecting interindividual differences in acetylation capacity</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Pathways of arylamine metabolism</td>
<td>17</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Schematic depicting the process of homologous recombination</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Overall gene targeting scheme</td>
<td>30</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Design of a gene targeting vector</td>
<td>32</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Screening selected ES cells for the targeted gene knockout</td>
<td>36</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Diagram of generation of functional blastocysts and implantation into pseudopregnant surrogate mice</td>
<td>37</td>
</tr>
<tr>
<td>Figure 2.0</td>
<td>Construction of Vector 1</td>
<td>42</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Construction of Vector 2</td>
<td>43</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Cloning of the Vector 1 positive control</td>
<td>61</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>XL PCR amplification (using T7 and mNAT primers) and subcloning into the pCR®2.1 vector of a 3 kb fragment containing additional mNAT3 3' UTR sequence</td>
<td>64</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Cloning of the Vector 2 positive control</td>
<td>65</td>
</tr>
</tbody>
</table>
Figure 2.5. Cloning of the mNAT open reading frames from the mNAT pET3a constructs into the pGEX fusion protein expression vector

Figure 3.0. PCR based screening strategy for detection of mNAT3 knockout

Figure 3.1. Verification of the presence and orientation of the Vector 1 positive control insert

Figure 3.2. Results of the amplification of 1 ng of the Vector 1 positive control

Figure 3.3. Restriction map of the mouse NAT3 gene region

Figure 3.4. Results of the XL PCR amplification of additional mNAT3 3' UTR from Lambda clones containing the mNAT3 open reading frame (using mNAT and T7 or T3 primers)

Figure 3.5. Verification of the clones containing the 3 kb XL PCR product in pCR®2.1 by digestion with Kpn I

Figure 3.6. Southern blotting screening of Vector 1 targeted ES cells

Figure 3.7. Representative autoradiogram of targeted ES cell genomic DNA isolated from Vector 1 electroporated ES cells

Figure 3.8. Verification of the Vector 2 positive control clones by digestion with Sac I

Figure 3.9. Diagrams depicting the XL PCR amplification of the Vector 2 positive control and the Vector 2 targeted events

Figure 3.10. Agarose gel electrophoresis of the XL PCR amplification of the Vector 2 positive control, mNAT1 and mNAT2 Vector 2 and a clone containing the Vector 2 positive control insert in the reverse orientation
Figure 3.11. Agarose gel electrophoresis of the XL PCR amplification of 1 picogram of the Vector 2 positive control in 100, 200, and 300 ng of untargeted ES cell genomic DNA 89

Figure 3.12. Autoradiogram of the initial round of XL PCR screening of Vector 2 targeted ES cells 90

Figure 3.13. Restriction map of the mNAT1 and mNAT2 loci and the location of the 600 bp mNAT1 and 2 Southern probe, demarcated by the stars between the Sph I and Hpa I restriction sites 91

Figure 3.14. Design of untargeted and Vector 2 targeted ES cell genomic DNA with Bam HI and probing with the 600 bp Southern probe 93

Figure 3.15. Autoradiogram of a Southern blot of BAC clone 256A7 digested with restriction endonucleases and probed with the mNAT1 and mNAT3 ORF probes 94

Figure 3.16. Agarose gel electrophoresis of the PCR amplification of the 870 bp mNAT1, 2 and 3 open reading frames from the mNAT1, 2, and 3 pET3a constructs 96

Figure 3.17. Verification of the nature of the PCR products amplified from the mNAT1, 2 and 3 pET3a constructs 97

Figure 3.18. Agarose gel electrophoresis of Bam HI and Sal I digested clones containing the mNAT1, 2, and 3 open reading frames subcloned into the pGEX fusion protein expression vector 97

Figure 3.19. Coomassie stained SDS- PAGE gel showing the results of initial studies on the expression of the mNAT1-GST protein 98

Figure 3.20. Coomassie stained SDS- PAGE gel showing the results of initial studies on the expression of the mNAT2- GST protein 98

Figure 3.21. Coomassie stained SDS- PAGE gel showing the results of initial studies on the expression of the mNAT3- GST protein 99
Figure 3.22. Coomassie stained SDS-PAGE gel showing the results of attempts at shifting the expression of the mNAT1-GST protein from the pellet to the supernatant by lowering the IPTG induction from 1.0 mM to 0.1 mM and 0.01 mM.

Figure 3.23. Immunoblots of mNAT1,2 and 3 GST fusion proteins probed with the anti-GST antibody

Figure 3.24. Coomassie stained SDS-PAGE gel showing the isolation and solubilization of the mNAT1 and mNAT3-GST proteins from within inclusion bodies
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 1.0.</th>
<th>Selected examples of genetically polymorphic drug metabolizing enzymes</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1.</td>
<td>Human <em>NAT1</em> variant alleles</td>
<td>9</td>
</tr>
<tr>
<td>Table 1.2.</td>
<td>Human <em>NAT2</em> variant alleles</td>
<td>10</td>
</tr>
<tr>
<td>Table 1.3.</td>
<td>Affinities and activities of recombinant human NATs (rhNATs) expressed in COS-1 cells</td>
<td>12</td>
</tr>
<tr>
<td>Table 1.4.</td>
<td>Clinically useful drugs, containing aromatic amine functional groups, that are NAT2 substrates</td>
<td>15</td>
</tr>
<tr>
<td>Table 1.5.</td>
<td>Some examples of known heterocyclic amine carcinogens that are NAT substrates</td>
<td>16</td>
</tr>
<tr>
<td>Table 1.6.</td>
<td>Nucleotide and amino acid homologies between human and mouse NATs</td>
<td>21</td>
</tr>
<tr>
<td>Table 1.7.</td>
<td>Mouse and human NAT substrate selectivities and activities</td>
<td>23</td>
</tr>
<tr>
<td>Table 1.8.</td>
<td>Selected examples of genes that have been studied by the gene targeting method in mice</td>
<td>38</td>
</tr>
<tr>
<td>Table 2.0.</td>
<td>Oligonucleotide primers used for sequencing and PCR</td>
<td>54</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>2-AF</td>
<td>2-aminofluorene</td>
<td></td>
</tr>
<tr>
<td>Ac-CoA</td>
<td>acetyl-CoA</td>
<td></td>
</tr>
<tr>
<td>Ac-NAT</td>
<td>acetylated NAT</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>arginine</td>
<td></td>
</tr>
<tr>
<td>ArNH₂</td>
<td>arylamine</td>
<td></td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
<td></td>
</tr>
<tr>
<td>CIP</td>
<td>calf intestinal phosphatase</td>
<td></td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
<td></td>
</tr>
<tr>
<td>CoASAc</td>
<td>acetyl coenzyme A</td>
<td></td>
</tr>
<tr>
<td>CYP</td>
<td>cytochrome P450</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>cysteine</td>
<td></td>
</tr>
<tr>
<td>DAC</td>
<td>deacetylase</td>
<td></td>
</tr>
<tr>
<td>ddNTP</td>
<td>dideoxynucleotidetriphosphate</td>
<td></td>
</tr>
<tr>
<td>DME</td>
<td>drug metabolizing enzyme</td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
<td></td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleotide triphosphate</td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>embryonic stem</td>
<td></td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
<td></td>
</tr>
<tr>
<td>FMO</td>
<td>flavin monooxygenase</td>
<td></td>
</tr>
<tr>
<td>G418</td>
<td>geneticin 418</td>
<td></td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
<td></td>
</tr>
<tr>
<td>GSTM</td>
<td>glutathione-S-transferase MU</td>
<td></td>
</tr>
<tr>
<td>GSTTI</td>
<td>glutathione-S-transferase Tl</td>
<td></td>
</tr>
<tr>
<td>hNAT</td>
<td>human NAT</td>
<td></td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>2-amino-3-methylimidazo[4,5-f]quinoline</td>
<td></td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pair</td>
<td></td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
<td></td>
</tr>
<tr>
<td>Kₘ</td>
<td>apparent Michaelis constant</td>
<td></td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
<td></td>
</tr>
<tr>
<td>LIF</td>
<td>leukemia inhibitory factor</td>
<td></td>
</tr>
<tr>
<td>LOH</td>
<td>loss of heterozygosity</td>
<td></td>
</tr>
<tr>
<td>MelIQ</td>
<td>2-amino-3,4-dimethylimidazo[4,5-f]quinoline</td>
<td></td>
</tr>
<tr>
<td>MelIQx</td>
<td>2-amino-3,8-dimethylimidazo [4,5-f]quinoxaline</td>
<td></td>
</tr>
<tr>
<td>mNAT</td>
<td>mouse NAT</td>
<td></td>
</tr>
<tr>
<td>NAT</td>
<td>N-acetyltransferase</td>
<td></td>
</tr>
<tr>
<td>Neo</td>
<td>neomycin resistance gene</td>
<td></td>
</tr>
<tr>
<td>NOAT</td>
<td>N,O-acetyltransferase</td>
<td></td>
</tr>
<tr>
<td>NTD</td>
<td>neural tube defect</td>
<td></td>
</tr>
<tr>
<td>O.D.</td>
<td>optical density</td>
<td></td>
</tr>
<tr>
<td>OAT</td>
<td>O-acetyltransferase</td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
<td></td>
</tr>
<tr>
<td>p-ABG</td>
<td>para-aminobenzoylglutamate</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>procainamide</td>
<td></td>
</tr>
<tr>
<td>PABA</td>
<td>para-aminobenzoic acid</td>
<td></td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
<td></td>
</tr>
<tr>
<td>PAS</td>
<td>para-aminosalicylate</td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
<td></td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
<td></td>
</tr>
<tr>
<td>PGK</td>
<td>phosphoglycerate kinase</td>
<td></td>
</tr>
<tr>
<td>PhIP</td>
<td>2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine</td>
<td></td>
</tr>
<tr>
<td>PHLOP</td>
<td>polymerase halt-mediated linkage of primers</td>
<td></td>
</tr>
<tr>
<td>PHS</td>
<td>prostglandin-H-synthetase</td>
<td></td>
</tr>
<tr>
<td>RE</td>
<td>restriction endonuclease</td>
<td></td>
</tr>
<tr>
<td>rhNAT1</td>
<td>recombinant human NAT1</td>
<td></td>
</tr>
<tr>
<td>rhNAT2</td>
<td>recombinant human NAT2</td>
<td></td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecylsulfate</td>
<td></td>
</tr>
<tr>
<td>SMZ</td>
<td>sulfamethazine</td>
<td></td>
</tr>
<tr>
<td>SULT</td>
<td>sulfotransferase</td>
<td></td>
</tr>
<tr>
<td>TK</td>
<td>thymidine kinase</td>
<td></td>
</tr>
<tr>
<td>UGT</td>
<td>glucuronosyltransferase</td>
<td></td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
<td></td>
</tr>
<tr>
<td>V&lt;sub&gt;max&lt;/sub&gt;</td>
<td>apparent maximum velocity</td>
<td></td>
</tr>
<tr>
<td>XL PCR</td>
<td>extra long PCR</td>
<td></td>
</tr>
<tr>
<td>YT</td>
<td>yeast tryptone</td>
<td></td>
</tr>
</tbody>
</table>
PART ONE: INTRODUCTION

1.1. GENERAL INTRODUCTION

Epidemiological and animal studies have shown that N-acetyltransferases (NATs) play a role in drug induced toxicity and carcinogenesis following exposure to aromatic amines. However, controversy exists on the overall importance of NATs in toxicity and cancer because of the inherent weaknesses of epidemiological studies in accurately assessing the levels, lengths and types of exposure of individuals to aromatic amines. Animal studies have overcome some of these problems, but there are differences in the types of mutations producing the phenotypes, differences in substrate selectivity between the NATs, as well as differences in the tissue distribution and level of expression of these enzymes. To overcome these difficulties, the overall goal of this research is to create and characterize NAT deficient mouse strains. These strains will be used as a model to more clearly determine the role of these enzymes in the modulation of toxicity and carcinogenesis following exposure to aromatic amines.

1.2. BACKGROUND

Humans and animals are continuously exposed to various natural and man-made chemical substances, known as xenobiotics, that may have toxic potential. In order to survive, nature has endowed humans and animals with a unique set of biotransforming enzymes, or "drug metabolizing enzymes" (DMEs) that are designed to convert these xenobiotics to more readily excretable forms. There is evidence to suggest that DME genes initially evolved through the interaction between animals and plants and the need for each to either protect itself from poisoning and toxicity or to ensure survival respectively (Nebert 1997). For example, the mixed-function monooxygenases or cytochrome P450s (CYPs) comprise a family of DMEs which is thought to have descended from a common ancestor gene that arrived on the scene ~ 400 million years ago when animals started to invade dry land and came into contact with various new terrestrial plant forms (Nelson et al. 1996).

The ability of an organism to deal with an almost endless variety of xenobiotics relies mainly on the fact that DMEs exhibit both enzyme multiplicity and overlapping substrate specificity. DMEs can be segregated into two broad categories: the phase I DMEs and the phase II DMEs.
The phase I DMEs comprise the mixed-function monooxygenases (CYPs), flavin containing monooxygenases (FMOs), prostaglandin-synthases, esterases, hydrolases, amidases and peptidases (Riddick 1998). The phase I DMEs carry out oxidation, reduction and hydrolysis of xenobiotics and their metabolites. The phase I DMEs may either bioactivate and increase the toxic potential of substrates, inactivate or not alter the pharmacological activity of substrates (Riddick 1998).

The phase II DMEs comprise the UDP-glucuronyltransferases, glutathione-S-transferases, sulfotransferases, N-acetyltransferases and others (Riddick 1998). The phase II DMEs are preoccupied with the chemical conjugation of compounds that may be produced from phase I reactions, rendering them for the most part inactive and increasing their solubility for excretion in the urine (Riddick 1998). However, there are cases when phase II DMEs may bioactivate xenobiotics as well.

The phase I and phase II DMEs may either act in concert with one another, allowing the detoxification of substances or may antagonize each other, yielding toxic products. As such, these DMEs are like double-edged swords: although they are designed to protect the organism, in some cases they may inadvertently cause the body harm (Figure 1.0). The outcome of these reactions largely depends on the environment, the levels and lengths of exposure to these substances, the physiological state of the organism, its gender, its age, racial background, and the types, levels, affinities and activities of its DMEs as well as the expression of modifying genes (Figure 1.1). Because of these multiple interacting variables, it is often extremely difficult to predict the outcome of exposure of an organism to xenobiotics.

1.2.1. Genetic polymorphisms of DMEs

Although the fate of many xenobiotics relies on the multiplicity of factors mentioned above, much interest has also focused on the genetically based variation of DMEs as an important determinant of an individual's response to foreign chemicals. For example, from the administration of the tuberculostatic drug isoniazid in the early 1950s, Bonicke and Reif (1953) noticed that in some recipients the drug was excreted largely in the conjugated form in the urine, while in others the drug remained mostly unchanged. In 1957 Biehl noted a bimodal distribution in the elimination half-life of isoniazid (Figure 1.2). In the same year, Mitchell and Bell found the same bimodal distribution and categorized individuals as either rapid or slow metabolizers of isoniazid.
Figure 1.0.
The fate of drugs and toxins in the body.
Diagram A shows that a prodrug/protoxin may be biotransformed by DMEs to the active form of the drug allowing effective pharmacological action; alternatively the drug may be biotransformed to a non potent form without pharmacological effect. Diagram B shows the role of DMEs in drug induced toxicity, carcinogenesis or detoxification; again the double role of DMEs in bioactivation and bioinactivation of xenobiotics is emphasized (From Wormhoudt et al. 1999).
Figure 1.1.
The interaction of various factors, intrinsic and extrinsic to the body exerting changes on the genetic profile of an individual which in turn can affect drug response or biotransformation of xenobiotics (From Vesell 1991).
Though it was shown many years later that the variable metabolism of isoniazid was in fact caused by genetic polymorphisms in the N-acetyltransferase \textit{NAT2} gene (Blum et al. 1990), it also became clear that many other DMEs were also polymorphic (Table 1.0.). These genetic polymorphisms are important in drug efficacy, drug induced toxicity, and carcinogenesis. They are also important in the context of initial drug design, testing, and eventual drug administration to the population. For example, a genotypic polymorphism in the CYP2C19 gene (De Morais et al. 1994) produces a truncated protein with no activity that is phenotypically manifested as poor metabolism of clinically useful drugs such as omeprazole, diazepam, and propanolol. This phenotype is observed in 3-5\% of Caucasians, while 18-23\% of Asians display the defective phenotype (De Morais et al. 1994). A defective CYP2D6 protein, found in 6-10\% of Caucasian populations (Idle and Smith 1979) may lead to poor metabolism of some classes of drugs such as antiarrhythmics, antihypertensives, beta-blockers, and monoamine oxidase inhibitors among many others. In terms of toxicity and in terms of rational drug design, these percentages can come to represent significant segments of the population.

Genetic polymorphisms of DMEs are also important because of their role in carcinogenesis. For instance, polycyclic hydrocarbon epoxides (found in cigarette smoke) which are substrates for the glutathione-S-transferase MU (GSTM) class of DMEs have been shown to be highly carcinogenic, and a 2.3 fold increased risk of lung adenocarcinoma has been observed in heavy smoking individuals who lack the GSTM protein (Anttila et al 1994; Daly et al 1994). Not surprisingly, it is becoming increasingly clear that DMEs interact to produce variable phenotypes. For example, Hayashi et al. (1992) report that a combination of the \textit{CYP1A1*2B} allele and the null \textit{GSTM1*0} allele results in a 9- fold increase in squamous cell carcinoma of the lung.
Figure 1.2.
Bimodal distribution of plasma concentration of isoniazid in a human population reflecting interindividual differences in acetylation capacity. The distribution shows the segregation of rapid acetylators from slow acetylators (From Evans et al. 1960).
<table>
<thead>
<tr>
<th>ENZYMES</th>
<th>CYTOCHROMES P450</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESTERASES</td>
<td></td>
</tr>
<tr>
<td>butyrylcholinesterase</td>
<td>debrisoquine hydroxylase</td>
</tr>
<tr>
<td>paraoxonase/ arylesterase</td>
<td>S- mephenytoin hydroxylase</td>
</tr>
<tr>
<td></td>
<td>phenytoin-4-hydroxylase</td>
</tr>
<tr>
<td>TRANSFERASES</td>
<td></td>
</tr>
<tr>
<td>N-acetyltransferases</td>
<td></td>
</tr>
<tr>
<td>thiopurine methyltransferase</td>
<td></td>
</tr>
<tr>
<td>catechol-O-methyltransferase</td>
<td></td>
</tr>
<tr>
<td>sulfotransferases</td>
<td></td>
</tr>
<tr>
<td>glutathione- S- transferase</td>
<td></td>
</tr>
<tr>
<td>UDP-glucuronosyltransferases</td>
<td></td>
</tr>
<tr>
<td>glucosyltransferases</td>
<td></td>
</tr>
<tr>
<td>DEHYDROGENASES</td>
<td></td>
</tr>
<tr>
<td>alcohol dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>aldehyde dehydrogenase</td>
<td></td>
</tr>
<tr>
<td>dihydropyridine dehydrogenase</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.0.**

Selected examples of genetically polymorphic drug metabolizing enzymes (Adapted from Kalow and Grant 1995)
1.3. ARYLAMINE N-ACETYLTRANSFERASES (NATs)

1.3.1. Human NATs

1.3.1.1. Discovery and cloning of human NATs

By examining acetylation from livers of slow and fast acetylators, Jenne and Orser (1965) proposed that there were two different enzymes responsible for the fast and slow acetylator phenotypes although it was not until 25 years later that this observation was confirmed by molecular cloning techniques.

In 1990, Blum et al. cloned three human NAT genes: NAT1, NAT2 and NATP. It was found that the NAT2 gene was responsible for the original "rapid"/"slow" acetylation polymorphism seen with isoniazid. Although NAT1 was initially thought to be monomorphic because the NAT1 specific substrates PAS and PABA could not distinguish between fast and slow acetylators, it was later discovered that NAT1 is in fact also polymorphic. The third human NAT gene, NATP, is a non-functional pseudogene. A summary of the genetically variable NAT1 and NAT2 alleles is found in Table 1.1 and 1.2.

All three NAT genes have been localized to chromosome 8 p23.1-21.3 with the NAT1 and NAT2 genes between 170-360 kb apart and with the NATP gene within 140 kb of NAT2. The physical order of the human NAT genes is telomere-NAT1*-NATP-NAT2*-centromere (Blum et al. 1990, Hickman et al. 1994, Matas et al. 1997).

The human NAT1 and 2 genes each include an 870 bp intronless open reading frame encoding a 290 amino acid cytosolic protein that is 33-34 kDa in deduced size (Blum et al. 1990). The NAT1 gene produces its entire mRNA transcript from a single exon, while the NAT2 transcript is produced from a 5' noncoding exon located 8 kb upstream of the translation start site and a coding and 3' noncoding exon downstream of the translation start site (Blum et al. 1990, Ebisawa and Deguchi 1991).

The NAT1 and NAT2 protein coding regions are 87% identical at the nucleotide level and their encoded proteins are 81% identical. (Hickman et al. 1998, Stanley et al. 1996, Windmill et al. 1997). Forty to seventy percent of Caucasians in Europe and North America display a slow acetylator phenotype (Weber et al. 1985, Evans et al. 1989).
<table>
<thead>
<tr>
<th>Allele</th>
<th>Nucleotide change(s)</th>
<th>Amino acid change(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAT1*3</td>
<td>C1095A</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*4</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*5</td>
<td>G350,351C, G497-499C, A884G, Δ976, Δ1105</td>
<td>Arg117Thr, Arg166Thr, Glu167Gln</td>
</tr>
<tr>
<td>NAT1*10</td>
<td>T1088A, C1095A</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*14A</td>
<td>G560A, T1088A, C1095A</td>
<td>Arg187Gln</td>
</tr>
<tr>
<td>NAT1*14B</td>
<td>G560A</td>
<td>Arg187Gln</td>
</tr>
<tr>
<td>NAT1*15</td>
<td>C559T</td>
<td>Arg187Stop</td>
</tr>
<tr>
<td>NAT1*16</td>
<td>[AAA] immediately after 1091, C1095A</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*17</td>
<td>C190T</td>
<td>Arg64Trp</td>
</tr>
<tr>
<td>NAT1*18A</td>
<td>Δ3 between 1064-1087, T1088A, C1095A</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*18B</td>
<td>Δ3 between 1064-1091</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*19</td>
<td>C97T</td>
<td>Arg33Stop</td>
</tr>
<tr>
<td>NAT1*20</td>
<td>T402C</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*21</td>
<td>A613G</td>
<td>Met205Val</td>
</tr>
<tr>
<td>NAT1*22</td>
<td>A752T</td>
<td>Asp251Val</td>
</tr>
<tr>
<td>NAT1*23</td>
<td>T777C</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*24</td>
<td>G781A</td>
<td>Glu261Lys</td>
</tr>
<tr>
<td>NAT1*25</td>
<td>A787G</td>
<td>Ile263Val</td>
</tr>
<tr>
<td>NAT1*26A</td>
<td>[TAA] insertion between 1066 and 1091, C1095A</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*26B</td>
<td>[TAA] insertion between 1066 and 1091</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*27</td>
<td>T21G, T777C</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*28</td>
<td>[TAATAA] deletion between 1085-1090</td>
<td>None</td>
</tr>
<tr>
<td>NAT1*29</td>
<td>T1088A, C1095A, Δ1025</td>
<td>None</td>
</tr>
</tbody>
</table>

**Table 1.1.**  
Human NAT1 variant alleles  
(Adapted from the Arylamine N-Acetyltransferase (EC 2.3.1.5.) Nomenclature website http://www.louisville.edu/medschool/pharmacology/NAT.html). From a heterogenous Toronto population sample, Hughes et al. (1998) reports NAT1*4, NAT1*10 as common with frequencies of 0.715, and 0.222 (n=228). NAT1*11, NAT1*14, NAT1*15 are present at frequencies of 0.021, 0.028, 0.014 (n=228), while the other alleles presented in this table are rare. Note NAT1*4 is the "wild type" allele.
<table>
<thead>
<tr>
<th>Allele</th>
<th>Nucleotide change(s)</th>
<th>Amino acid change(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAT2*4</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>NAT2*5A</td>
<td>T₃⁴¹C, C₄⁸¹T</td>
<td>Ile¹¹⁴Thr, Lys²⁶⁸Arg</td>
</tr>
<tr>
<td>NAT2*5B</td>
<td>T₃⁴¹C, C₄⁸¹T, A⁸⁰³G</td>
<td>Ile¹¹⁴Thr, Lys²⁶⁸Arg</td>
</tr>
<tr>
<td>NAT2*5C</td>
<td>T₃⁴¹C, A⁸⁰³G</td>
<td>Ile¹¹⁴Thr, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*5D</td>
<td>T₃⁴¹C</td>
<td>Ile¹¹⁴Thr, Lys²⁶⁸Arg</td>
</tr>
<tr>
<td>NAT2*5E</td>
<td>T₃⁴¹C, G⁵⁹⁰A</td>
<td>Arg¹⁹⁷Gln, Lys²⁶⁸Arg</td>
</tr>
<tr>
<td>NAT2*5F</td>
<td>T₃⁴¹C, G⁵⁹⁰A</td>
<td>Arg¹⁹⁷Gln, Gly²⁸⁶Glu</td>
</tr>
<tr>
<td>NAT2*6A</td>
<td>C²⁸²T, G⁵⁹⁰A</td>
<td>Arg¹⁹⁷Gln, Gly²⁸⁶Glu</td>
</tr>
<tr>
<td>NAT2*6B</td>
<td>C²⁸²T, G⁵⁹⁰A</td>
<td>Gly²⁸⁶Glu, Lys²⁶⁸Arg</td>
</tr>
<tr>
<td>NAT2*6C</td>
<td>C²⁸²T, G⁵⁹⁰A, A⁸⁰³G</td>
<td>Lys²⁶⁸Arg, Arg⁶⁴Gln</td>
</tr>
<tr>
<td>NAT2*6D</td>
<td>T¹¹¹C, C²⁸²T, G⁵⁹⁰A</td>
<td>Lys²⁶⁸Arg, Arg⁶⁴Gln</td>
</tr>
<tr>
<td>NAT2*7A</td>
<td>G⁸⁵⁷A</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*7B</td>
<td>C²⁸²T, G⁸⁵⁷A</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*12A</td>
<td>A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*12B</td>
<td>C²⁸²T, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*12C</td>
<td>C⁴⁸¹T, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*13</td>
<td>C²⁸²T</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14A</td>
<td>G¹⁹¹A</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14B</td>
<td>G¹⁹¹A, C²⁸²T</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14C</td>
<td>G¹⁹¹A, T₃⁴¹C, C₄⁸¹T, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14D</td>
<td>G¹⁹¹A, C²⁸²T, G⁵⁹⁰A</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14E</td>
<td>G¹⁹¹A, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14F</td>
<td>G¹⁹¹A, T₃⁴¹C, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*14G</td>
<td>G¹⁹¹A, C²⁸²T, A⁸⁰³G</td>
<td>Arg⁶⁴Gln, Arg¹⁹⁷Gln</td>
</tr>
<tr>
<td>NAT2*17</td>
<td>A³⁴³C</td>
<td>Gin¹⁴⁵Pro</td>
</tr>
<tr>
<td>NAT2*18</td>
<td>A²⁴⁵C</td>
<td>Lys²⁸²Thr</td>
</tr>
</tbody>
</table>

**Table 1.2.**

Human NAT2 variant alleles

(Adapted from the Arylamine N-Acetyltransferase (EC 2.3.1.5.) Nomenclature website [http://www.lousiville.edu/medschool/pharmacology/NAT.html](http://www.lousiville.edu/medschool/pharmacology/NAT.html). In a sample of Caucasian-Canadians, the NAT2 rapid acetylator alleles (*NAT2*4 (wild type), *NAT2* *12, *NAT2* *13*) were found at a frequency of 0.240 (n=226). The *NAT2* *5* had a frequency of 0.470 (n=226). *NAT2* *6* had a frequency of 0.230 (n=226). The *NAT2* *7* had a frequency of 0.060 (From Grant et al. 1997).
However, only 5% of Canadian Eskimos are slow acetylators, while 87% of Egyptians and 90% of Moroccans display the slow acetylator phenotype (Blum et al. 1991).

1.3.1.2. Catalytic and kinetic NAT mechanisms

NAT proteins catalyze the acetylation of aromatic amine substrates, such as arylamines and hydrazines, by using a Ping-Pong Bi-Bi reaction mechanism where an acetate molecule from Acetyl-CoA is transferred to the substrate (Weber and Cohen 1967). According to Voet & Voet (1995), the Bi-Bi reaction mechanism requires two substrates (the Acetyl-CoA, and the arylamine or hydrazine) and produces two products (CoA and the acetylated arylamine or hydrazine). The Ping-Pong mechanism is one that requires that the first product is released from the enzyme before the second substrate has attached to the enzyme such that the Acetyl-CoA and the arylamine or hydrazine do not contact one another on the surface of the NAT enzyme (Voet and Voet 1995). The NAT acetylation mechanism for an arylamine is depicted below:

\[
\begin{align*}
(1) \quad \text{NAT} + \text{Ac-CoA} & \rightarrow \text{Ac-NAT} + \text{CoA} \\
(2) \quad \text{Ac-NAT} + \text{ArNH}_2 & \rightarrow \text{NAT} + \text{ArNH-Ac}
\end{align*}
\]

In the first reaction (1), the acetyl-CoA (Ac-CoA) donates its acetate group to the NAT enzyme, producing the activated form of the NAT enzyme (Ac-NAT) and the reduced CoA cofactor. In the second reaction (2), the arylamine contacts the acetylated NAT enzyme, and the acetate is transferred to the arylamine, forming an arylamide, while the NAT enzyme is regenerated.

While NAT1 and NAT2 are 81% identical at the amino acid level, enzyme kinetic studies have shown large differences between them with respect to their intrinsic stabilities and their kinetic selectivities for amine acceptor substrates (Dupret et al. 1994). Both recombinant and liver human NAT1 have high affinity (low $K_m$) for para-aminosalicylate (PAS) and para-aminobenzoic acid (PABA) while recombinant and liver human NAT2 prefer sulfamethazine (SMZ) and procainamide (PA) (Grant et al. 1990). From the same study, it was also found that NAT2 had a very high affinity for the known carcinogen aminofluorene (AF), while NAT1's affinity for AF was comparable to its affinity for PAS and PABA (Table 1.3).
<table>
<thead>
<tr>
<th>Substrate</th>
<th>rhNAT1</th>
<th>rhNAT2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\text{Km (\text{\mu M})})</td>
<td>(\text{Vmax (nmol/min/unit)})</td>
</tr>
<tr>
<td>SMZ</td>
<td>1,160</td>
<td>28</td>
</tr>
<tr>
<td>PA</td>
<td>29,700</td>
<td>510</td>
</tr>
<tr>
<td>PAS</td>
<td>11</td>
<td>1,280</td>
</tr>
<tr>
<td>PABA</td>
<td>15</td>
<td>1,250</td>
</tr>
<tr>
<td>AF</td>
<td>18</td>
<td>1,395</td>
</tr>
</tbody>
</table>

**Table 1.3.**

Affinities and activities of recombinant human NATs (rhNAT) expressed in COS-1 cells (From Grant et al. 1990). N/A: In this study, the substrate PABA showed apparent substrate activation with NAT2, but the enzyme degraded at higher PABA concentrations preventing determination of kinetic constants (Grant et al. 1990).
Chicken, rabbit, cat, hamster, mouse, rat and human NAT1 and NAT2 contain three conserved cysteine residues at positions 44, 68 and 223 (Delomenie et al. 1997, Trepanier et al. 1998). Studies in Salmonella typhimurium (Watanabe et al. 1992) have shown that Cys\textsuperscript{69}, equivalent to Cys\textsuperscript{68} in human NAT, is strongly conserved suggesting that this amino acid could be the active site of NAT. Dupret and Grant (1992) found that Cys\textsuperscript{68} is involved in the catalytic function of NATs allowing the transfer of the acetate from CoASAc to the acceptor amine substrate. Site-directed mutagenesis studies by Delornenie et al. (1997) showed that Arg\textsuperscript{9} and Arg\textsuperscript{64} were not directly involved in amine substrate binding, thus not affecting the $K_m$ of the enzyme for its substrate. However, changes in the $V_{\text{max}}$ as a result of amino acid substitutions at these two sites revealed that these two arginine residues are essential for the stability of the NAT enzyme, and possibly for maintaining the configuration of the active site.

1.3.1.3 Localization of human NATs

Human NAT1 appears to be more ubiquitously expressed in the body while NAT2 expression is more limited to the digestive system. For example, human NAT2 mRNA is more abundant in liver than NAT1 (Ohsako and Deguchi 1989) but both NAT1 and NAT2 mRNA expression has been found in the esophagus, stomach, small intestine, and in the colon (Windmill et al. 1997). Antibodies specific for NAT1 and NAT2 detected their expression in epithelial cells of the small intestine, along the whole length of the villus from the tip to the crypt of Lieberkühn (Hickman et al. 1998). In the colon the NATs were only expressed at the apical membrane of the epithelial cells (Hickman et al. 1998). Human NAT1 and NAT2 activity has also been found in the liver and intestinal epithelial tissue (Ohsako and Deguchi 1990, Ilett et al. 1994). However, only NAT1 activity is found in red blood cells (Ward et al. 1992 and Risch et al. 1996), and in lymphocytes (McQueen and Weber 1980, Coroneos et al. 1991, Cribb et al. 1991). Furthermore, while human NAT1 and NAT2 mRNA was found to be expressed in bladder and ureter (Windmill et al. 1997), Stanley et al. (1996) only report the presence of NAT1 protein in the bladder and not NAT2 using an anti-NAT polyclonal antisera. These results were confirmed by Badawi et al. (1995) and Janezic (1997) by studies of enzyme kinetics, although very low levels of NAT2 may still be present. Recently, NAT1 and NAT2 activity and NAT1 immunoreactivity were detected in human placenta with NAT1 activity 1000 fold greater than NAT2 activity (Smelt et al. 1998).
1.3.1.4. The role of human NATs in biotransformation of xenobiotics

NATs are known to be important in drug induced toxicity, and a summary of clinically useful drugs containing arylamine and hydrazine functional groups (which are substrates for NATs) is found in Table 1.4. Table 1.5 summarizes a few of the known heterocyclic amine carcinogens that are NAT substrates and includes the sources of these compounds.

As shown in Figure 1.3, NATs can participate at many points in the pathways of arylamine metabolism. As an example, an arylamine (1) is N-acetylated to an acetamide (2). The acetamide is then metabolized by CYP1A2 or prostaglandin-H-synthetase (PHS) to produce hydroxamic acid (3). This compound can then be excreted in the urine by conjugation with a glucuronosyl group catalyzed by UDP-glucuronosyltransferase (4). The fate of the arylamine can be altered if the arylamine at point 1 is first metabolized by CYP1A2 or PHS, making it an hydroxylamine (5). Metabolism by the O-acetyltransferase function of NAT transforms it into an acetoxy ester (6). The acetoxy ester can spontaneously decompose into a highly electrophilic arylnitrenium ion (7) which either binds to cellular protein or DNA. Binding to cellular proteins can lead to cytotoxicity, and binding to DNA can produce DNA adducts which may represent the initiating phases of cancer.

Figure 1.3 demonstrates that the metabolism of an arylamine is quite complex. The nature of the substrates, the relative affinities and activities of the DMEs for the substrates, in addition to their expression levels relative to each other and their tissue localization, will all act to determine the ultimate fate of the compound.
<table>
<thead>
<tr>
<th>Drug</th>
<th>Indication</th>
<th>Possible phenotype-related side effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>aminoglutethimide</td>
<td>breast cancer</td>
<td>dizziness, somnolence, nausea, vomiting, diarrhea, myelosuppression</td>
</tr>
<tr>
<td>amonafide</td>
<td>breast cancer</td>
<td>myelosuppression, dizziness, tinnitus, nausea, vomiting</td>
</tr>
<tr>
<td>amrinone</td>
<td>severe heart failure</td>
<td>arrhythmia, thrombocytopenia, gastro-intestinal symptoms</td>
</tr>
<tr>
<td>clonazepam</td>
<td>epilepsy</td>
<td>dizziness, somnolence, hangover</td>
</tr>
<tr>
<td>dapsone</td>
<td>leprosy</td>
<td>allergic skin reactions, methemoglobinemia, myelosuppression</td>
</tr>
<tr>
<td>isoniazid</td>
<td>tuberculosis</td>
<td>peripheral neuropathy, hepatotoxicity</td>
</tr>
<tr>
<td>hydralazine</td>
<td>hypertension</td>
<td>lupus erythematosus-like syndrome</td>
</tr>
<tr>
<td>nitrazepam</td>
<td>insomnia</td>
<td>dizziness, somnolence, hangover</td>
</tr>
<tr>
<td>procainamide</td>
<td>tachyarrhythmia</td>
<td>lupus erythematosus-like syndrome, gastro-intestinal symptoms, hypotension</td>
</tr>
<tr>
<td>sulfasalazine</td>
<td>Crohn's disease,</td>
<td>nausea, headache, gastro-intestinal symptoms</td>
</tr>
<tr>
<td></td>
<td>rheumatoid arthritis</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.4.
Clinically useful drugs, containing aromatic amine functional groups, that are NAT2 substrates. NAT2 genetic polymorphisms may lead to the listed side effects in susceptible individuals (From Cascorbi et al. 1999).
### Table 1.5.

Some examples of known heterocyclic amine carcinogens that are NAT substrates (Adapted from Sugimura 1997).

<table>
<thead>
<tr>
<th>Carcinogen</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQx)</td>
<td>Fried beef, salmon steak (baked)</td>
</tr>
<tr>
<td>2-Amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP)</td>
<td>Fried ground beef, salmon steak (baked)</td>
</tr>
</tbody>
</table>
Figure 1.3.
Pathways of arylamine metabolism.
The ultimate fate of an arylamine depends on which DME it first encounters, the relative affinities and turnover rates of DMEs for the substrate, and the types of interactions between the DMEs. UGT, Glucuronosyltransferase; NAT, N-Acetyltransferase; DAC, Deacetylase; SULT, Sulfotransferase; CYP1A2, Cytochrome P4501A2; PHS, Prostaglandin H-Synthase; OAT, O-Acetyltransferase; NOAT, N,O-Acetyltransferase (From Grant et al. 1997).
1.3.1.5. Genetic polymorphisms of human NATs and implications for carcinogenesis

Much interest has centred around genetic polymorphisms of DMEs as the basis for interindividual variations in susceptibility to cancer. For example, epidemiological studies by Risch et al. (1995) and Filiadis et al. (1999) showed that NAT2 slow acetylators who were either smokers or who had occupational exposure to arylamines had a higher risk of bladder cancer than controls. Risch et al. (1995) also found that there was a higher proportion of slow acetylators in bladder cancer patients who had no documented exposure to arylamine procarcinogens. This is a seemingly contradictory finding because it argues that exposure to arylamines is not necessary for bladder carcinogenesis in genetically susceptible individuals. However it should be kept in mind that exposure was very likely but that there were difficulties in accurately assessing exposure in these epidemiological studies. Hsieh et al. (1999) and Taylor et al. (1998) showed that patients carrying the NAT1*10 allele in combination with a NAT2 slow acetylator capacity were more susceptible to smoking-induced bladder cancer than patients who were NAT2 slow acetylators but lacked the NAT1*10 allele. They proposed that the NAT2 rapid acetylator can compete with the CYP1A2 pathway and rapidly detoxify arylamines in the liver such that very minute amounts of CYP1A2 hydroxylated arylamines are produced. However, if a person is a NAT2 slow acetylator, then the competing CYP1A2 reaction may predominate and the hydroxylated arylamines may be transported to the bladder epithelium where a NAT1*10 protein could bioactivate them to acetoxy esters (through the O-acetylating function of NAT1). Interestingly the activity of CYP1A2 is increased in smokers, through enzyme induction by chemicals in cigarette smoke, such as benzo[a]pyrene (Riddick 1998). Thus CYP1A2 biotransformation will predominate over the NAT pathway, such that larger amounts of hydroxylated arylamines become available for bioactivation in the bladder epithelium of a NAT1*10 individual (Taylor et al 1998). Accordingly, Vineis et al. (1994) found an increased amount of 4-aminobiphenyl-DNA adducts in urinary bladder epithelia in NAT2 slow acetylators while Badawi et al. (1995) found significantly higher arylamine DNA-adduct levels in bladder epithelium in individuals who were NAT2 slow and NAT1*10 acetylators.

A study by Brockmoller et al. (1996) showed that NAT2 slow acetylators had little risk of bladder cancer with no exposure to arylamines, but that the risk increased with increasing exposure to arylamines. In as yet unpublished and preliminary results, Brockmoller et al. have found that NAT1 has little or no role in the bioactivation of arylamines in bladder epithelium. Contrary to other studies, NAT2 acetylation capacity was not found to be an important factor in bladder cancer, as there was no significant difference in urinary mutagenicity between smokers
who were NAT2 fast and NAT2 slow acetylators (Hirvonen et al. 1994). Similarly, no association was found between 4-aminobiphenyl DNA adducts and NAT2 genetic polymorphisms (Martone et al. 1998).

Interestingly, two recent studies have shown that there was a detectable loss of heterozygosity (LOH) at the NAT2 allelic locus in tumors from bladder cancer patients (Matas et al. 1996, Schnakenberg et al. 1998). Furthermore, losses of chromosomal material on chromosome 8p23.2-8p22 and 8p21.3-8p11.22 detected in colon cancer patients suggest the loss of two tumor suppressor genes thought to be near the NAT2 gene locus 8p21.3-23.1 (Knowles et al. 1993, Farrington et al. 1996, Fujiwara et al. 1993).

NAT genetic polymorphisms have also been implicated in the etiology of colon cancer. Studies by Lang et al. (1986) and Ilett et al. (1987) have demonstrated that NAT2 rapid acetylators were at an increased risk of developing colorectal cancer. Bell et al. (1995a,b) showed that NAT1*10 acetylators were at greater risk of developing colon cancer especially when they were also NAT2 rapid acetylators. Chen et al. (1998) found that a subgroup of men who were 60 years of age or older, who had a high intake of red meat, and were either NAT1 or NAT2 rapid acetylators or both had a slightly higher incidence of colorectal cancer. However, other studies have shown the opposite (Kirlin et al. 1991, Ladero et al. 1991, Oda et al. 1994, Lin et al. 1998, Probst-Hensch et al. 1995, Probst-Hensch et al. 1996, Welfare et al. 1997, Hubbard et al. 1998).

Individuals who had hereditary nonpolyposis colon cancer with a mutation in the hMLH1 mismatch repair gene (an exon 16 deletion) and who were NAT1*10 acetylators had an earlier onset of colorectal cancer. Interestingly, Zhang et al. (1999) showed that mutations or loss of the hMLH1 gene may prevent cellular apoptosis, contributing to the immortalization of cancerous cells. If these patients also expressed GSTT1 as well, they were at an increased risk of developing distal colonic tumors (Moisio et al. 1998).

Other studies have indicated an increased risk of lung cancer in NAT2 rapid acetylators (Cascorbi et al. 1996), or a decreased risk of alcoholic liver cirrhosis in NAT2 slow acetylators (Rodrigo et al. 1999). Bandmann et al. (1997) and Ladero et al. (1989) showed that NAT2 slow acetylators were predisposed to developing Parkinson’s disease and a study by Agundez et al. (1998) showed that they also had an earlier onset of the disease. A conflicting report was issued by Harhangi et al. (1999) who found no association between NAT2 polymorphism and Parkinson’s disease. In sum, controversy exists over the exact role of NATs in disease. As mentioned at the outset, epidemiological studies provide limited information because of the
difficulties in accurately measuring exposure. In addition, cancer is a multi-variable disease where the role of NATs may be modulated by other DMEs or by other genetic factors such as modifying genes. The creation and characterization of NAT deficient mouse strains will help in further elucidating the role of NATs in disease.

1.3.1.6. Endogenous functions of human NATs

Although no clear endogenous roles have been found for human NATs, Minchin et al. (1995) have demonstrated the ability of human NAT1 to acetylate the folic acid metabolite p-aminobenzoylglutamate (p-ABG). Among its many important roles, folic acid is crucial for proper closure of the neural tube during embryogenesis, and deficiency of folic acid during pregnancy is a leading cause of spina bifida (Eskes 1998, Klein 1996, Butterworth and Bendich 1996). In addition, folic acid has been shown to have a protective effect against orofacial clefting (Shaw et al. 1995). Interestingly, Ward et al. (1995b) speculated that folate deficiency during pregnancy could be related to the level of NAT1 activity.

1.3.2. Mouse N-acetyltransferases (mNATs)

1.3.2.1. Cloning of the mouse NATs

Studies carried out by Martell et al. (1991, 1992) found two NAT genes in each of the A/J and C57BL/6 mouse strains, designated as NAT1 and NAT2 (in this thesis known as mNAT1 and mNAT2). Kelly and Sim (1994) subsequently discovered a third mouse NAT gene, which was named NAT3 (in this thesis known as mNAT3).

All three mNAT genes contain intronless open reading frames 870 bp in size, producing 290 amino acid proteins which have a deduced size of 33-34 kDa and are 31 kDa in size on SDS-PAGE gels (Kelly and Sim 1994, Martell et al. 1991, Martell et al. 1992). The mNAT1 and mNAT2 genes are found on chromosome 8 30.0-31.0 corresponding to the syntenic region of human chromosome 8 p21.3-23.1 (http://www.informatics.jax.org). Studies by Yu-Plant (1998) and Martell et al. (1991) found that the mNAT1 and mNAT2 genes are 9 kb apart with mNAT1 located upstream of mNAT2. The exact location of mNAT3 is still unknown, but studies carried out in this research thesis have revealed that it may be located within ~130 kb of mNAT1 and 2 (Section 4.4).
1.3.2.2. Mouse NAT nucleotide and amino acid sequence homology

Mouse *NAT3* is 74 % identical with *mNAT1* and 78 % identical with *mNAT2* at the nucleotide level, while *mNAT1* and *mNAT2* are 84 % identical at the nucleotide level (Kelly and Sim 1994). At the amino acid level, the *mNAT1* and *mNAT2* proteins are 82 % identical, while *mNAT3* and *mNAT1* are 68 % identical and *mNAT3* and *mNAT2* are 74 % identical (Kelly and Sim 1994). Table 1.6 summarizes nucleotide and amino acid homology between human and mouse NATs.

<table>
<thead>
<tr>
<th>Nucleotide</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human <em>NAT1</em></td>
<td>Human <em>NAT2</em></td>
</tr>
<tr>
<td><strong>Mouse <em>NAT1</em></strong></td>
<td>80 %</td>
</tr>
<tr>
<td><strong>Mouse <em>NAT2</em></strong></td>
<td>79 %</td>
</tr>
<tr>
<td><strong>Mouse <em>NAT3</em></strong></td>
<td>76%</td>
</tr>
</tbody>
</table>

**Table 1.6**
Nucleotide and amino acid homologies between human and mouse NATs (Adapted from Kelly and Sim 1994).
1.3.2.3. Mouse NAT kinetics

Table 1.7 shows the results of studies carried out on recombinant 129 Sv/J mNATs expressed in *E. coli* by Yu-Plant (1998). mNAT1 showed a higher affinity (lower Km) for 2-AF, than for PAS (346 fold) and SMZ (2762 fold). mNAT2 had a higher affinity for PAS, 10 fold less for 2-AF and 10 000 fold less for SMZ than PAS. mNAT3 only acetylated 2-AF (Yu-Plant 1998). In kinetic comparisons with human NATs, mouse NAT1 resembles human NAT; while mouse NAT2 resembles human NAT1.

1.3.2.4. Tissue localization of mouse NATs

Autoradiograms of Northern blots have shown no difference in the amount of mNAT1 and mNAT2 liver poly(A)^+ RNA from both genotypically and phenotypically different mice (De Leon et al. 1995). However, mNAT2 mRNA is more abundant than mNAT1 as judged by the differences in band intensities (De Leon et al. 1995). Further studies by this group have shown that the steady-state level of hepatic mNAT2 mRNA is 150 times greater than that of mNAT1, and that the messages are translated at similar rates (De Leon et al. 1995).

Mouse urinary bladder was shown to acetylate both PABA and AF, which indicates that it may contain either mNAT1 and/or mNAT2 (Mattano and Weber 1987). Ware et al. (1996) have shown greater activity of mNAT2 in the cecum (large intestine) of the mouse, and greater mNAT1 activity in the proximal small intestine, while Levy et al. (1992) support the view that mNAT1 plays a larger role in acetylation capacity in the colon. Studies by Chung et al. (1993) showed that mouse NAT was expressed in 33 tissues with absence in the blood plasma and seminal vesicles. However, the study did not look at specific mNAT1 and mNAT2 expression patterns because it failed to discriminate between the specific kinetic activities of mNAT1 and mNAT2. Furthermore, tissue immunoblots used a non-specific anti-mNAT2 antibody which could also pick up mNAT1 expression. Stanley et al. (1996) used a specific anti-mNAT2 antibody to show that mNAT2 was expressed in liver, lung, kidney, spleen, small intestine, urinary bladder, and skin. Stanley et al. (1998) detected the expression of mNAT2 in the cerebellum, and in the large motor neurons in the spinal cord. Expression of mNAT2 during embryogenesis was found in heart, intestine and neural tube (Stanley et al. 1998).
### Table 1.7.

Mouse and Human NAT substrate selectivities and activities. 
N/D = not detectable, N/A = not applicable (From Yu-Plant 1998).

<table>
<thead>
<tr>
<th></th>
<th>2- AF</th>
<th>PAS</th>
<th>SMZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse NAT1</td>
<td>Km (µM)</td>
<td>0.16± 0.08</td>
<td>55.4± 11.5</td>
</tr>
<tr>
<td></td>
<td>Vmax  (nmol/min/mg)</td>
<td>5.3± 0.37</td>
<td>1.1± 0.05</td>
</tr>
<tr>
<td></td>
<td>Vmax/ Km</td>
<td>33.4</td>
<td>0.02</td>
</tr>
<tr>
<td>Mouse NAT2</td>
<td>Km (µM)</td>
<td>15.6± 0.8</td>
<td>1.5± 0.6</td>
</tr>
<tr>
<td></td>
<td>Vmax  (nmol/min/mg)</td>
<td>17342.2± 211.7</td>
<td>6449.9± 644.9</td>
</tr>
<tr>
<td></td>
<td>Vmax/ Km</td>
<td>1111.7</td>
<td>4299.9</td>
</tr>
<tr>
<td>Mouse NAT3</td>
<td>Km (µM)</td>
<td>108.61± 20.3</td>
<td>N/D</td>
</tr>
<tr>
<td></td>
<td>Vmax  (nmol/min/mg)</td>
<td>14.6± 0.9</td>
<td>N/D</td>
</tr>
<tr>
<td></td>
<td>Vmax/ Km</td>
<td>0.13</td>
<td>N/A</td>
</tr>
<tr>
<td>Human NAT1</td>
<td>Km (µM)</td>
<td>18</td>
<td>12± 6</td>
</tr>
<tr>
<td></td>
<td>Vmax  (nmol/min/mg)</td>
<td>1395</td>
<td>4220± 480</td>
</tr>
<tr>
<td></td>
<td>Vmax/ Km</td>
<td>78</td>
<td>352</td>
</tr>
<tr>
<td>Human NAT2</td>
<td>Km (µM)</td>
<td>1.1</td>
<td>5300± 1060</td>
</tr>
<tr>
<td></td>
<td>Vmax  (nmol/min/mg)</td>
<td>36</td>
<td>0.91± 0.05</td>
</tr>
<tr>
<td></td>
<td>Vmax/ Km</td>
<td>33</td>
<td>0.00017</td>
</tr>
</tbody>
</table>
Taken together, preliminary data exists on the tissue localization of the mNATs. However, further studies identifying the localization of mNATs through the use of specific antibodies would be desirable in order to more completely understand the role of the mNATs in drug metabolism and carcinogenesis. Furthermore, the use of such specific antibodies would permit a more accurate determination of the specific activities (V_max) of the individual mNATs and allow for the verification of successful gene targeting to abolish NAT expression.

1.3.2.5. Genetic polymorphisms of mouse NATs

There are seventeen mouse strains classified as mNAT2 rapid acetylators, among them the C57BL/6J, BALB/cJ and the 129 SV strains (Glowinski and Weber 1982). Slow acetylator strains include the A/J and AHe/J strains (Glowinski and Weber 1982). The genetic basis of the interstrain acetylator polymorphism in mice is localized to the mNAT2 gene where an adenine at position 296 in the mNAT2*8 allele is replaced with a thymine, changing the amino acid at position 99 from asparagine to isoleucine in slow acetylating strains (mNAT2*9). This amino acid substitution changes the hydropathic profile of the peptide around amino acid 99 from hydrophilic to hydrophobic, perhaps affecting the activity or stability of the protein (Martell et al. 1991).

Martell et al. (1992) showed that there was no significant difference between the amount of hepatic mNAT1 mRNA in rapid and slow acetylating strains, or in mNAT2 mRNA between these strains, indicating that the point mutation in the ORF of mNAT2 does not affect the steady-state levels of mNAT2 mRNA. In addition, the translation efficiency of the mNAT2*8 mRNA was reported to be ~2 fold greater than that of the mNAT2*9 message and the stability of the mNAT2*8 protein was proposed to be 15 times greater than the *9 (Martell et al. 1992). These two factors together could account for the observed differences in mNAT2 V_max between rapid and slow strains, which have nearly identical mNAT2 K_m's. An alternate theory is proposed by De Leon et al. (1995). They report instead that the mNAT2 proteins are found in equal amounts in both slow and fast strains. The differences between the slow and fast strains are explained by differences in the affinities (K_m's) of the mNAT2s for their substrates as well as the unstable nature of the mNAT2*9 protein. De Leon found however that the protein in the slow strains was not subject to cellular proteases that could degrade the slow protein. Instead, the amino acid substitution made the protein more catalytically labile. As such, De Leon et al. (1995) proposed that the hypothesized similarities between the instabilities of the mNAT proteins from humans
and mice (as suggested by Martell et al. 1992) as the origin of the slow acetylator polymorphism phenotype is not valid.

1.3.2.6. Modifying genes and NATs

Studies in recombinant inbred mouse strains have suggested a possible role for modifier genes in modulating the expression of mNATs. Initial breeding studies by Glowinski and Weber (1982) found that there were some deviations from the expected phenotypes produced from crosses between rapid acetylating C57BL/6 and slow acetylating A/J mice. These deviations were proposed to be the result of modifier genes acting on the expression and/or activity of the mouse NATs. These findings were corroborated by Hultin and Weber (1984) and then by Hein et al. (1988).

1.3.2.7. Role of mouse NATs in carcinogenesis

In humans, genetic polymorphisms in DMEs have been implicated in modulating the toxicological and carcinogenic outcome from exposure to toxicants and carcinogens as discussed in section 1.2.1. Studies using homozygous rapid, homozygous slow and congenic mice have been carried out to model the human condition. Levy and Weber (1992), found that female slow acetylators had higher bladder 2-AF adduct levels than their rapid counterparts whereas male rapid acetylators had higher levels of 2-AF bladder DNA adducts than their slow counterparts. In addition, males had higher overall 2-AF bladder DNA adduct levels compared to females. When the levels of hepatic 2-AF DNA adducts were examined, rapid males and females showed higher DNA adduct levels than their slow counterparts with slow females showing greater levels of DNA adducts than slow males (Levy and Weber 1989). The authors suggest that for female rapid acetylators, increased amounts of activated arylamines in the liver go on to form hepatic DNA adducts, while the slow acetylators pass on more unmetabolized arylamine pro-carcinogen into the kidney and bladder where local activation of the arylamines produces the urinary DNA adducts. This proposed metabolic pathway is similar to the one proposed in the human. However, although this theory is interesting, it does not hold true for male mice where the reverse situation is seen.

Very interesting studies regarding the carcinogenicity of aromatic amines in the digestive system in mice have been carried out by Ohgaki et al. (1991), Nerurkar et al. (1995), Kristiansen (1996), Sasaki et al. (1998), Thorgeirsson et al. (1999), and Turteltaub et al. (1999). MeIQ, MeIQx, IQ, PhIP and other heterocyclic amines found in cured and cooked foods as well as in
cigarette smoke condensate were shown to be carcinogenic in mice with localization of tumors mostly to the liver, but also to the lung, the lymphatic system, blood, spleen, and forestomach (Ohgaki et al. 1991). One study suggested that hepatic cancer may be accelerated by a synergistic interaction of MeIQ induced mutations in cancer causing genes along with overexpression of the oncogene c-myc, known to be overexpressed in hepatic cancer (Thorgeirsson et al. 1999). Interestingly, Ohgaki's results correlate well with the presence and activity of the mNATs in these tissues (1991). In the same study, the authors noted differences between species tested, gender, and distribution of the tumors that are thought to be related to differences in the amount and inducibility of cytochrome P450s, the enzymes responsible for the initial bioactivation of heterocyclic amines (Degawa et al. 1987).

Nerurkar et al. (1995) found differences in IQ-DNA adducts in mice, which involved interactions between the CYP450s and NATs. It seems likely that CYP1A2 is involved in the N-hydroxylation of IQ to N-OH-IQ, which is the activated metabolite responsible for DNA adducts in the colon, kidney and bladder (Nerurkar et al. 1995). The hypothesis is that CYP1A2 in the liver produces N-OH-IQ from IQ. The latter is then transported to the kidney, bladder or colon where it undergoes local tissue biotransformation by NATs to the aryl nitrenium ion. This ion can bind to DNA and forms adducts or bind to proteins inducing cytotoxicity. The study found that adduct levels were significantly higher in the colon of mice that were NAT2 rapid and that expressed CYP1A2 but not CYP1A1 in the liver. The second highest adduct levels were found in mice that were NAT2 slow and expressed CYP1A2 and CYP1A1. These observations are interesting in light of their striking similarities to the human (Lang et al. 1994). Interestingly, induction of CYP1A1 in these mice showed a decreased level of DNA adducts in the colon in both slow and rapid acetylating mice, suggesting a possible role for CYP1A1 in the detoxification of IQ (Nerurkar et al. 1995).

Using the nuclear aberration assay, Kristiansen (1996) found that IQ administered to mice resulted in nucleotoxic effects in the crypts of colon epithelium, supporting the cytotoxic effects of IQ, or its metabolite N-OH-IQ. Using the sensitive single cell gel electrophoresis (SCG) assay to detect genotoxicity caused by heterocyclic amines, Sasaki et al. (1998) found significant DNA damage in the colon of mice but not in the small intestine. Lastly, Turteltaub et al. (1999) showed that administration of heterocyclic amines to mice at comparable levels to those found in humans produced bioactive metabolites that entered the circulation and reached target tissues. However, the same study also found that in general, at comparable levels of heterocyclic amines, humans had a greater tendency to form adducts than mice, especially in the blood and colon.
Despite some tantalizing similarities between humans and mice with regard to the biotransformation of aromatic amines, in light of the aforementioned factors, comparisons with the human condition should be interpreted with caution as is usually the case with animal models. Nevertheless, the production and characterization of NAT deficient mouse strains will help to further clarify the roles of NATs in carcinogenesis in the mouse and perhaps in the human.

1.3.2.8. Endogenous roles of mouse NATs

Although much is known about the roles of mNATs in metabolism of xenobiotics, relatively little is known about the endogenous functions of these enzymes. Recently, attention has been focused on para-aminobenzoylgutamate (p-ABG) as a potential substrate for mNATs. p-ABG is a folic acid catabolite and recent studies by Minchin (1995) and Ward et al. (1995b) have shown that it is a substrate for human NAT1 and mouse NAT2 (Estrada-Rogers et al. 1998).

Folic acid is an essential vitamin in humans and is believed to be important in preventing neural tube defects in developing embryos (MRC vitamin study research group 1991). During pregnancy in humans and in mice, there is an increased turnover of folic acid, and supplementation in humans has been recommended in order to prevent neural tube defects (NTDs) in children. Interestingly, it has been observed during embryogenesis that mNAT2 is highly expressed in the neural tube in mice during neural tube closure (Stanley et al. 1998). Recently, a hormone responsive element sensitive to glucocorticoids (GCs) has been found upstream of mNAT2 (Estrada-Rogers 1998). Strikingly, the incidence of NTDs in the mouse has been shown to increase after the administration of GCs (Liu and Erikson 1986). In addition to its protective role in NTDs, folic acid has been shown to prevent orofacial clefting in developing embryos, and studies have implicated the genetic region around the NAT genes in determining susceptibility to this malformation (Shaw et al. 1995, Karolyi et al. 1988, 1990). The significance of these findings is as yet unclear but may signal the involvement of many interacting genes or even possibly modifying genes in affecting phenotype. Evidently, the involvement of the NATs, if any, in these disorders remains to be elucidated and the use of NAT deficient mouse strains may help clarify the role of NATs in embryological development.
1.4. GENE TARGETING IN THE MOUSE

1.4.1. Background

Specific targeting and deletion of selected gene(s), known as gene targeting or knockout, is a powerful technique permitting the examination of the role of specific gene(s) while keeping the rest of the genetic background constant (Thomas and Capecchi 1986, Sedivy and Joyner 1992). As a first step, a gene targeting vector is introduced into embryonic stem (ES) cells derived from the inner cell mass of 3.5 day old blastocysts. Gene targeting allows a gene(s) to be knocked out at this early embryonic stage, when cellular differentiation has not yet occurred. Subsequently, when the cell divides during embryogenesis, the mutation can in theory be passed on to all the tissues.

This breakthrough in genetic technology came about from three different lines of work. First, Evans and Kaufman (1981) isolated embryonic stem (ES) cells from the inner cell mass of blastocyst stage embryos and managed to propagate them as in vitro cell lines. Second, the discovery of leukemia inhibitory factor (LIF) in 1988 by AG Smith et al. and Williams et al., as the agent necessary to keep ES cells in an undifferentiated state was an important milestone in enabling the process. Third, the principles and practice of gene knockout techniques like creating and using gene targeting vectors and selecting targeted clones by using selectable markers, were developed by Lin et al. (1985), Smithies et al. (1985), Doetschman et al. (1987) and Thomas and Capecchi (1986, 1987).

The technique of gene targeting relies on the process of homologous recombination whereby stretches of homologous DNA sequences can recombine with one another. This process is illustrated in Figure 1.4. In the first step, two homologous sequences find each other. In the second step, each strand breaks away from its parent strand. The third step involves the exchange of one homologous sequence for the other and the final step involves sealing the portions of the strands that were initially broken (Sedivy and Joyner 1992). This process of homologous recombination occurs naturally in living cells. In gene targeting, homologous recombination between the targeting (knockout) vector and the endogenous gene abolishes the function of the gene product. The overall gene targeting scheme is depicted in Figure 1.5.
Figure 1.4.
Schematic depicting the process of homologous recombination.
In step 1, two homologous sequences find each other and align. In step 2, strand breakage occurs. In step 3, strand exchange between homologous regions takes place. In step 4, breakpoint sealing (adapted from Sedivy and Joyner 1992).
Clone and map targeted gene
  \[\downarrow\]
Design targeting vector
  \[\downarrow\]
Electroporate ES cells with targeting vector
  \[\downarrow\]
Selection of ES cells on antibiotic media
  \[\downarrow\]
Screening of surviving ES cell clones by Southern blotting and/ or PCR
  \[\downarrow\]
Morula aggregation or blastocyst injection
  \[\downarrow\]
Blastocyst implantation into pseudopregnant mouse
  \[\downarrow\]
Breeding of chimeric pups

**Figure 1.5.**
Overall gene targeting scheme
1.4.2. Gene targeting vector design

To design a gene targeting vector, the gene locus must first be cloned and mapped. The protein coding region(s) of the gene and the untranslated regions will later be cloned into a targeting vector and form the regions of homology. Importantly, these two elements should be from the same strain of mouse as the ES cells into which the targeting vector will go. This demand for so called isogenic DNA is important because even the subtlest interstrain differences in sequence can drastically lower the success of gene targeting (Thomas and Capecchi 1987, Te Riele et al. 1992). Using the information gathered from cloning and mapping of the regions of homology, arms of homology are formed by cutting the cloned sequence with restriction endonucleases and subcloning the fragments into the gene targeting vector. Targeting vectors usually comprise a long arm of homology and a short arm of homology (which can be designated as a 5' arm or a 3' arm depending on their positions in the vector but independent of length). These arms flank the cloned target gene disrupted by a selectable marker. The most popular method of disrupting or knocking out gene function relies on the use of the neomycin resistance (Neo) gene (Silver 1995). In addition to knocking out the expression of the gene of interest, the Neo gene also acts as a selection marker (Figure 1.6).

The size of the arms of homology can be as crucial to the success of the knockout as the use of isogenic DNA to make the gene targeting vector. Hasty et al. (1991a) reported that for successful gene targeting in ES cells, there must be a minimum total amount of DNA homology of 1.9 kb between the targeting vector and the endogenous locus when using a replacement type gene targeting vector, the vector most commonly used to knockout genes (DeChiara et al. 1990, Koller et al. 1990, McMahon et al. 1990, Schwartzberg et al. 1989, Soriano et al. 1991, Thomas and Capecchi 1990). A modest increase of 200 bp from 1.7 kb to 1.9 kb increased the targeting efficiency 21 fold (Hasty et al. 1991a). The same authors report that the total length of homology is more important than the relative lengths of the arms of homology and they add that a short arm of homology as small as 472 bp in size is still sufficient for successful homologous recombination.
Figure 1.6.
Design of a gene targeting vector.
Diagram A shows the endogenous locus, with the grey areas representing the 5' and 3' untranslated regions (UTRs), and the black box representing the protein coding region of the gene to be knocked out. Diagram B shows the gene targeting vector, with corresponding 5' and 3' arms of homology. The gene is disrupted by the neomycin gene (Neo), and the TK gene lies outside the region of homology. The thick black line(s) represent the area of homology between the endogenous locus and the targeting vector.
1.4.3. Transfection of ES cells and subsequent selection of targeted ES cells

The gene targeting vector can be transfected into ES cells by a process known as electroporation (Chu et al. 1987). Prior to electroporation, the gene targeting vector is linearized by cutting with a restriction enzyme and then added to ES cells placed into specially designed cuvettes. The cuvettes are then subjected to a brief electrical pulse that temporarily disrupts the cell membrane, allowing the gene targeting vector to enter the cell.

Only about half of the ES cells survive the electroporation procedure. The survivors must then be screened for incorporation of the gene targeting vector in the ES cell genome and separated from the remaining cells. As such, after a brief recovery period the next step in the gene targeting experiment involves incubating (plating) the cells in ES cell medium that contains the mammalian cell antibiotic G418 (an analogue to the prokaryotic antibiotic neomycin). ES cells that express the exogenously provided Neo gene survive. This is known as positive selection (Mansour et al. 1988).

Once the gene targeting vector has entered the cell, it can either undergo homologous recombination, non-homologous recombination or it can be degraded. Non-homologous recombination is greatly favored over homologous recombination by a ratio of 30-40,000:1 (Hasty et al. 1991b). This means that the majority of the ES cells used in a gene targeting experiment have not had the desired gene knocked out and they have randomly incorporated the gene targeting vector in another region of their genome. Mansour et al. (1988) developed a technique that selects against ES cells that have undergone non-homologous recombination (negative selection). This relies on using a gene targeting vector equipped with a thymidine kinase (TK) gene which resides in a region of the targeting vector outside the region of homology shared between the targeting vector and the endogenous locus (Figure 1.6). In a homologous recombination event, the TK gene lying outside the region of homology will not be incorporated into the endogenous locus, while non-homologous events will incorporate some or all of the targeting vector with the possibility of incorporating the TK gene. Expression of the TK gene produces the enzyme thymidine kinase which phosphorylates thymidine and allows it to be incorporated into replicating DNA. ES cells that have survived electroporation, and have survived selection for neomycin expression (and hence uptake of the targeting vector), are then placed in cell culture medium containing the toxic thymidine analog gancyclovir, that when phosphorylated is incorporated into elongating DNA but causes the process to terminate. This is of course detrimental to the survival of the cells. ES cells expressing TK will phosphorylate gancyclovir and will not survive the selection procedure. Thus, after selection for Neomycin expression and
against TK expression, the ES cells remaining represent a pool of clones that could contain a correctly knocked out gene.

1.4.4. Screening of ES cells for targeted knockout of the gene

Screening can be done in two ways: by Southern blotting or by PCR. Screening by Southern blotting is performed by cutting targeted ES cell genomic DNA with restriction endonucleases (RE), electrophoresing the cut DNA through an agarose gel and then transferring the fractionated DNA to a nylon membrane. A radioactively labelled sequence of DNA referred to as a probe, is then hybridized to the membrane and used to distinguish between the untargeted and the targeted locus by relying on a noticeable difference in DNA fragmentation pattern between the loci.

The probe is designed to anneal to a region outside the region of homology but within a predetermined stretch of sequence (a region of interest) bordered by restriction endonuclease (RE) sites on either side. The key to the procedure lies in being able to gauge a difference in size between the untargeted locus and the targeted locus after digestion of targeted ES cell genomic DNA by restriction endonucleases corresponding to the RE sites flanking the region of interest. Differences in size between the targeted and untargeted loci may be caused by either: (1) the addition of a new RE site from the introduction of new genetic material from the targeting vector; (2) removal of an endogenous RE site through a replacement with new genetic material from the targeting vector; or (3) by keeping the same two RE sites in the targeted and untargeted loci and noting a change in size through incorporation of new genetic material from the targeting vector (Figure 1.7A). A heterozygous targeting event should produce one "wild type" band and a second "targeted band" of equal intensities. Once correctly targeted ES cells have been isolated, the clone(s) are expanded and re-screened for the correctly targeted event.

The PCR method is faster and less technically demanding but does not discriminate between targeting at one locus or both. The method relies on the amplification of a DNA sequence of known length, found in the targeted locus, by using one primer that anneals to the Neomycin gene with a second primer annealing to an untargeted region outside the region of homology (i.e. found exclusively in the endogenous locus and not in the targeting vector). This is depicted in Figure 1.7B. The amplification of this specified length of sequence can only occur if both the Neo gene and the untargeted region are in the correct orientation. Thus, this is an effective and rapid method of screening for targeted recombinants is used as an adjunctive and
independent method of confirming the results from Southern blotting (Zimmer and Gruss 1989, Joyner et al. 1989).

1.4.5. Reconstitution of functional blastocysts

The next step is the injection of targeted ES cells into host strain blastocysts or aggregation of the targeted ES cells with morulas derived from a host strain to produce a functional blastocyst (Figure 1.8). Once the blastocyst is reformed it is then implanted into a pseudopregnant host mother (i.e. CD1, MF1, or C57BL/6). The use of a morula strain different from the ES cell strain is important for two reasons (Schwartzberg et al. 1989). The first is that strain differences allow for easy identification of the chimeras by using coat color as a guide. The second is that some strain combinations produce more or less chimerism. This factor can of course also affect the extent to which the targeted mutation found in the 129 Sv component is transmitted into the germ line.

Following blastocyst injection or morula aggregation, functional blastocysts are implanted into the uteri of pseudopregnant female hosts and carried to term. Transmission of the targeted mutation is tested by breeding the chimeras with wild type mice from the same strain as the host. If chimeras contain ES cell derived sex cells, gene targeted mice that are heterozygous for the targeted mutation will be produced. Heterozygotes can then be intercrossed to produce homozygous knockouts.

The accessibility and power of the gene targeting technique have given scientists a new tool to use in understanding the roles of genes and have led to the production of a vast array of different knockouts (Table 1.8).
A. Screening by Southern blotting

Wild type fragment size

Restriction enzyme site A

Restriction enzyme site B

Probe

Targeted gene fragment size

B. Screening by PCR

Targeted region

Targeted region

Targeted locus

Neo primer

Primer annealing in untargeted region

PCR

PCR product

Figure 1.7.

Screening selected ES cells for the targeted gene knockout.

Diagram A depicts the screening of targeted ES cells by Southern blotting. Diagram B shows the screening of targeted ES cells by PCR. One primer anneals to the neomycin resistance gene (Neo), and the other to a region found only in the endogenous locus (untargeted region). Correct amplification signals successful gene knockout (see text for details).
Inject ES cells into blastocyst

Make single cell suspension

ES cells

Transfer blastocysts to pseudopregnant mother

Make small clumps of ES cells

Aggregate ES cells with 2 morulae

Figure 1.8.
Diagram of generation of functional blastocysts and implantation into pseudopregnant surrogate mice (From Sedivy and Joyner 1992).
<table>
<thead>
<tr>
<th>Targeted Gene</th>
<th>Use</th>
<th>Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-Hydroxytryptamine receptor (5-HT 1A, 5-HT 1B)</td>
<td>depression/anxiety</td>
<td>Zhuang et al. 1999</td>
</tr>
<tr>
<td>Beta adrenergic receptor (Beta 1, Beta 2)</td>
<td>cardiovascular function</td>
<td>Rohrer et al. 1999</td>
</tr>
<tr>
<td>Melanocortin-4 receptor (MC4-R)</td>
<td>obesity</td>
<td>Huszar et al. 1997</td>
</tr>
<tr>
<td>Cystic fibrosis transmembrane regulator (CFTR)</td>
<td>cystic fibrosis</td>
<td>Rozmahel et al. 1996</td>
</tr>
<tr>
<td>Dopamine receptor 2 (D2R)</td>
<td>learning/memory, disorders of mood, motivation</td>
<td>Aiba et al. 1996</td>
</tr>
<tr>
<td>Histamine receptor 1 (H1R)</td>
<td>allergy, inflammation, gastric secretion</td>
<td>Inoue et al. 1996</td>
</tr>
<tr>
<td>Cytochrome P450 2E1 (CYP 2E1)</td>
<td>xenobiotic metabolism</td>
<td>Lee et al. 1996</td>
</tr>
</tbody>
</table>

Table 1.8.
Selected examples of genes that have been studied by the gene targeting method in mice.
1.5. OVERALL GOAL, RATIONALE, OBJECTIVES

1.5.1. Overall goal

The overall goal of this research is to create and characterize N-acetyltransferase (NAT) deficient mouse strains as a model to determine the role of these enzymes in the modulation of toxicity and carcinogenesis following exposure to arylamines and hydrazines.

1.5.2. Rationale

While both animal studies and human epidemiological evidence suggest an important role for NATs in drug toxicity and in carcinogenesis, these studies have many limitations. First, there are difficulties in accurately assessing the levels, lengths and types of human exposure to aromatic amines. Second, many epidemiological reports have limited themselves to examination of only one DME at a time, when interaction between the DMEs probably forms the basis for many of the observed toxicological and carcinogenic reactions. Furthermore, although both natural animal models, and congenic models of the NAT polymorphism exist, there are differences in the types of mutations producing the phenotypes, differences in substrate selectivity between the enzymes, the tissue distribution and level of expression of these enzymes, and as yet unidentified factors (modifying genes) that influence the phenotype produced. An animal model that controls for some or all of these confounding factors is needed.

Production of a completely deficient NAT mouse strain by gene targeting and characterization of the phenotype would serve to further the understanding of the involvement of NATs in drug metabolism, toxicology, and carcinogenesis while avoiding the potential pitfalls of the other animal models mentioned above. In addition, targeted knockout of the NATs in mice would permit the elucidation of the endogenous physiological roles of NATs, since involvement of these enzymes in metabolism of one endogenous compound, as well as their presence during embryonic development has been noted. In the future, examination of the role of modifying genes in modulating acetylator phenotype may be more clearly elucidated by elimination of the NAT genes and examination of interstrain differences in acetylation capacity. With this comprehensive knowledge in hand, it may be possible to more accurately predict an individual's susceptibility to drug toxicity or cancer and implement effective preventive strategies.
1.5.3. Specific objectives

To overcome the limitations of epidemiological studies and current NAT animal models aimed at understanding the role of NATs in drug induced toxicity and carcinogenesis following exposure to aromatic amines we have undertaken studies to:

Produce a NAT deficient mouse strain by gene targeting using R1 ES cells

1) Introduce the mNAT1/ mNAT2 targeting vector and the mNAT3 targeting vector into R1 ES cells

2) Design PCR and Southern blotting screening methods to enable the detection of the mNAT1/ mNAT2 and mNAT3 targeted knockouts

Produce anti-mouse NAT antibodies for use in characterizing the NAT deficient mouse strain

1) Design and construct an mNAT- GST fusion protein to act as antigen

2) Express the mNAT- GST fusion protein

3) Purify the fusion protein and inject in rabbits to produce anti-mNAT antibodies
PART TWO: MATERIALS AND METHODS

2.1. MATERIALS

2.1.1. Vectors, clones, and bacterial strains

Mp5, a pBluescript® (pBSK-) vector containing ~5 kb of mouse NAT3 5' UTR, the 870 bp mouse NAT3 open reading frame, and ~1.5 kb 3' UTR from the 129/SvJ mouse strain was obtained from Dr. Nicola Hughes (Hospital for Sick Children, Toronto, Ontario). Bacteriophage Lambda DASH II vectors containing either the 129/SvJ mouse NAT1 gene (lambda clone 10) or the 129/SvJ mouse NAT2 gene (lambda clones 1, 2, 3, 7, 9, 14, 17, 18, 19, 22, and 23) or the 129/SvJ mouse NAT3 gene (lambda clones 1, 9, 14, 23) were obtained from Dr. Andrea Gaedigk (Hospital for Sick Children). NAT1-pBCMV, a construct containing ~0.8 Kb 5' UTR, the mNAT1 870 bp open reading frame, and ~4 Kb of 3' UTR and NAT2 pBCMV, a construct containing ~5.4 Kb 5' UTR, the mNAT2 870 bp open reading frame and ~3 Kb of 3' UTR were obtained from Violeta Yu Plant. Construct NAT2 3'UTR/ pCR2.1, containing the 129Sv/J mouse NAT2 ORF and ~ 8 Kb of 3' UTR, was obtained from Violeta Yu-Plant. Vector 1, a gene targeting vector used to knock out mNAT3, containing both the neomycin and thymidine kinase genes in addition to the mouse NAT3 open reading frame and selected regions of homology was obtained from Violeta Yu-Plant (Figure 2.0). Vector 2, a gene targeting vector containing both neomycin and thymidine kinase genes in addition to selected regions of homology of mNAT1 and mNAT2 and used to knock out mNAT1 and mNAT2 simultaneously, was obtained from Violeta Yu-Plant (Figure 2.1). pPNT, a gene targeting vector containing the neomycin and thymidine kinase genes, was a gift from Dr. Roger Gaedigk (Hospital for Sick Children, Toronto, Ontario).

The pCR®2.1 cloning vector, designed to facilitate the cloning of PCR amplified DNA fragments, was from Invitrogen (Carlsbad, CA). The DNA cloning vector pBluescript® (pBSK-) was from Stratagene (LaJolla, CA). The bacterial artificial chromosome plasmid (pBACE3.6) vectors 256A7, 277N15, 286K12, each containing approximately 80-130 Kb of cloned 129 Sv/J genomic mouse DNA were gifts from Dr. Richard Rozmahel (Hospital for Sick Children, Toronto, Ontario).

DNA fragments subcloned into cloning vectors were grown in E.coli strains DH5α (GIBCO BRL, Life Technologies Inc., Burlington, Ontario), INVα F' (Invitrogen) or XL1 MRF' blue (Stratagene). Protein expression vectors were grown in strain XA 90 provided as a gift by Dr. G. Verdine (Harvard University, Boston, MA).
Figure 2.0.
Construction of Vector 1.
The upper diagram represents the endogenous locus. The lower diagram, the targeting vector. In the upper schematic, the Kpn I site downstream of the open reading frame was artificially created by PCR for subcloning into the pPNT knockout vector and does not represent an endogenous site. The thin single straight lines, with numbers below them indicating size of fragments, represent the areas of homology. Vector 1 contains a 6.5 kb (blunted) 5' long arm of homology, a 1.8 kb neomycin resistance gene (Neo) inserted into the mNAT3 open reading frame, and a 1.0 kb (Kpn I- Kpn I) 3' short arm of homology. The thymidine kinase gene is denoted by TK (Figure adapted from Yu-Plant 1998).
Figure 2.1.
Construction of Vector 2.
The upper diagram represents the endogenous locus and the lower diagram the targeting vector. The knockout vector contains a 0.6 kb (blunted) 5' short arm of homology, a 1.8 kb neomycin resistance gene (Neo) which replaces ~9 kb of sequence between mNAT1 and mNAT2, and a 3.8 kb (blunted) 3' long arm of homology. The thymidine kinase gene is denoted by TK (Figure adapted from Yu-Plant 1998).
The prokaryotic protein expression vectors NAT1 pET3a, NAT2 pET3a, and NAT3 pET3a were obtained from Violeta Yu-Plant. The prokaryotic fusion protein expression vector pGEX was from AmershamPharmaciaBiotech [Pharmacia] (Baie D’Urfe, Quebec).

2.1.2. Media and antibiotics for the propagation of bacterial strains

LB (Luria-Bertani) agar (GIBCO BRL) was used as a solid media to grow bacterial transformations and streaks. LB broth base (GIBCO BRL) was used as a liquid media to grow bacterial cultures. To select bacteria carrying the various DNA cloning and protein expression vectors, ampicillin, kanamycin, or chloramphenicol were added as required to solid growth media and liquid culture media (Sigma Chemical Co., St. Louis, MO).

2.1.3. Isolation of DNA cloning and protein expression plasmids

The QIAGEN Spin Miniprep Kit (QIAGEN Inc. Mississauga, Ont.), QIAGEN Maxi Kit, Pharmacia Flexiprep kit as well as lab-made miniprep solutions were used to extract and purify plasmids. To isolate size separated DNA fragments from agarose gels, the QIAEX DNA purification kit from QIAGEN was used.

2.1.4. Restriction enzymes, DNA modifying enzymes, DNA size ladders

Restriction enzymes were from Pharmacia, GIBCO BRL, and New England Biolabs Inc. (Mississauga, Ont.). Calf Intestinal Phosphatase (CIP) was from Pharmacia. T4 DNA ligase was from GIBCO BRL. Bacteriophage Lambda DNA used to generate a DNA size ladder was from Pharmacia. The 1 kb ladder was from GIBCO BRL. The 100 bp ladder was from Pharmacia.

2.1.5. Southern Blots, Western Blots, Sequencing, PCR, XL PCR

DNA probes were labelled using the T7 Quickprime® Kit and gamma $^{32}$P dCTP both items purchased from Amersham. The hybridization solution ExpressHyb™ was from Clontech Inc. (Mississauga, Ont.). The Southern blotting vacuum transfer apparatus was from Pharmacia.
Nylon membranes for DNA immobilization (Hybond N+) were from Amersham. Hybridization ovens Hybaid Micro-4 were from Interscience (Markham, Ont). The UV DNA cross linking machine (Stratalinker 1800) was from Stratagene. The film used to detect a DNA hybridization signal, BIOMAX, was from Kodak (Rochester, NY).

Protein concentrations were determined using a colored protein binding reagent from Bio-Rad Laboratories Inc. (Mississauga, Ont.). The Mighty Small II SE 250 Western Blotting apparatus was from Hoefer (San Francisco, CA). The NovaBlot Multiphor II protein transfer apparatus was from Pharmacia. The nylon membranes for protein immobilization (Hybond C+) were from Amersham. The Enhanced Chemiluminescence (ECL) kit was from Pharmacia. The film used to detect an antibody hybridization signal, Hyperfilm ECL, was from Pharmacia.

DNA sequencing was in part performed using the T7 Sequencing Kit™ from Pharmacia. $\alpha$-$^{35}$S dATP was from Amersham. The sequencing apparatus (Base Ace™ vertical sequencing apparatus) was from Stratagene. Other DNA sequencing was performed at the Hospital for Sick Children Sequencing Facility (Toronto, Ont).

Oligonucleotide primers were ordered from Dalton Chemical Labs Inc. (North York, Ont), and ACGT Corp. (Toronto, Ont). The Polymerase Chain Reaction (PCR) buffers, deoxynucleotide triphosphates (dNTPs), and AmpliTaq® DNA polymerase were purchased from Perkin-Elmer (Applied Biosystems Canada Inc., Miss. Ont). The eXtra Long (XL) PCR kit was purchased from Perkin-Elmer. All PCR and XL PCR reactions were carried out in a Perkin Elmer GeneAmp® PCR System 9600 machine.

2.1.6. Embryonic Stem (ES) cell culture

The R1 pluripotent ES cell line was originally derived from Dr. Janet Rossant’s lab (Mount Sinai Hospital, Toronto, Ont). A batch of these cells was kindly provided by Dr. Richard Rozmahel (Hospital for Sick Children). Complete ES cell culture medium contained Dulbecco’s Modified Eagle Medium (DMEM), ESGRO™ (LIF) to inhibit ES cell differentiation, glutamine, MEM sodium pyruvate, and MEM non-essential amino acids from GIBCO BRL, fetal bovine serum (FBS) from Hyclone Laboratories Inc. (Logan, Utah) and 2-mercaptoethanol (tissue culture grade) from Sigma Inc. Gene targeted cells were selected with the mammalian neomycin analog G418 from GIBCO BRL and gancyclovir from Sigma Inc. Trypsin powder was from DIFCO-Bacto (Detroit, MI). Magnesium/Calcium-free Phosphate Buffered Saline (PBS) was from GIBCO BRL. ES cells were electroporated with a Gene-Pulser (0.4 cm cuvette) apparatus from Bio-Rad.
2.1.7. mNAT- GST fusion protein expression

Lysis of bacterial cells to release cellular proteins was performed by using a Vibra-Cell sonicator from Sonics and Materials Inc. (Danbury, CT).

All other reagents and solutions were purchased from local suppliers and were of analytical grade.

2.2. METHODS

2.2.1. General methods

2.2.1.1. Bacterial growth conditions

Individual Petri dishes containing LB agar and appropriate antibiotic (100 μg/ml ampicillin, 50 μg/ml kanamycin or 20 μg/ml chloramphenicol) were streaked with glycerol stocks of bacteria containing the desired cloning or expression plasmids or were used to plate transformed bacteria. Plates were incubated at 37°C for 16-18 hours.

Liquid bacterial cultures were grown by inoculating 3 ml of LB broth base (containing appropriate antibiotic) with single bacterial colonies picked from LB agar plates. Cultures were grown for 16-18 hours at 37°C, with shaking at 250-280 rpm.

2.2.1.2. Preparation of competent cells

To make standard competent E.coli cells, a 2 ml volume of SOB (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄) was inoculated with a single colony of either DH5α or XA 90. The culture was grown for 16-18 hours at 37°C with shaking at 210 rpm, and then 500 μl of this culture was used to restart a 50 ml culture of SOB. The culture was grown at 37°C with shaking at 210 rpm until the optical density (O.D.) at 600 nm reached 0.3-0.4. The culture was then placed on ice for 10 minutes. The culture was centrifuged at 15,000 rpm, 4°C, 5 minutes, the supernatant removed and the pellet resuspended in 1/3rd the original volume with transformation buffer (80mM CaCl₂·2H₂O, 20mM MnCl₂·4 H₂O, 9mM MgCl₂·6H₂O, 10mM potassium acetate, 10% glycerol, pH 6.4). The resuspended solution was chilled on ice for 20-30 minutes. It was then spun down at 15,000 rpm, 4°C, 5 minutes and the supernatant discarded. The pellet was resuspended in 1/12th the
original volume with transformation buffer. The solution was chilled on ice 10 minutes and cells were aliquoted into eppendorf tubes in 200 μl volumes. This was followed by shock freezing in liquid nitrogen and storage at -80° C. The transformation efficiency of these cells was around 1 X 10^6 colonies/ μg of pBSK- plasmid.

2.2.1.3. Transformation of bacterial cells

Competent cells were thawed on ice and 80-90 μl were aliquoted in a 1.5 ml eppendorf tube and mixed with either 10 μl of ligation mix or 1 ng (10μl) of plasmid. The mixture was chilled on ice for 30 minutes, heat shocked at 42° C for 90 seconds, and then chilled on ice for 2 minutes. 1 ml of SOC (SOB supplemented with 12 mM glucose) was added and the cells were incubated at 37° C for 1 hour with shaking at 250 rpm. For transformations of bacteria with ligations, the cells were spun down at 13,000 rpm for 1 minute, and the supernatant was removed. The cells were concentrated by resuspension in 100-300 μl of the remaining SOC. 50-100 μl of the mixture was then plated onto Petri dishes containing LB agar supplemented with the appropriate antibiotic. The plates were then incubated at 37° C for 16-18 hours. Individual colonies were picked with a sterile metal loop and were then immediately used to inoculate a 3 ml volume of LB broth (with antibiotic) to grow overnight cultures.

For transformations of plasmid vectors not directly derived from fresh ligations, 50-100 μl of the solution was plated directly, without prior concentration, onto plates containing appropriate antibiotic and incubated 16-18 hours at 37° C.

2.2.1.4. Small scale isolation of plasmid DNA

Three independent methods were used to isolate plasmid DNA from bacterial cells. However, all three methods are based on the alkaline lysis method (Sambrook et al. 1989). For standard minipreps, 1.5 ml of an overnight 3 ml bacterial culture was aliquoted to a 1.5 ml sterile Eppendorf tube. The tube was centrifuged at room temperature, 13,000 rpm for 1 minute to pellet the bacterial cells. The supernatant was removed by aspiration, and the pellet was resuspended in 150 μl of ice cold P1 solution (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.4 mg/ml RNAse A) by vortexing. 150 μl of P2 solution (200 mM NaOH, 1% SDS) was added and the solution was mixed by gentle inversion and incubated for 5 minutes at room temperature. 150 μl of P3 solution (5M KOAc, 11.5% (w/v) glacial acetic acid) was then added and the solution was
mixed by gentle inversion. The tube was incubated on ice for 10 minutes and then centrifuged at 18,000 rpm, 4°C, for 10 minutes. The supernatant was transferred to a new sterile Eppendorf tube. 900 µl of 100 % ethanol was added and the solution was vortexed for a few seconds (if the plasmid was smaller than 10 kb). The tube was again centrifuged at 18,000 rpm, 4°C, 10 minutes to pellet the plasmid DNA. The supernatant was aspirated with a fine tipped Pasteur pipette to avoid losing the plasmid pellet. The pellet was then washed with 100-200 µl of ice cold 70 % ethanol to remove residual salts. The solution was spun down at 18,000 rpm, 4°C, 5 minutes. The supernatant was removed, and the pellet was left to dry at 37°C for 5-10 minutes or until it became translucent. Depending on the pellet size, 20-100 µl of sterile TE 8.0 (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) was used for resuspension.

When higher quality plasmid DNA preparations were needed for sequencing or PCR, small scale isolation of plasmid DNA using the flexiprep method (Pharmacia) was carried out according to the manufacturer's instructions. For small scale preparations of ultra pure plasmid DNA to be sequenced at the Hospital for Sick Children Sequencing Facility, the QIAprep spin miniprep protocol was carried out according to the manufacturer's instructions.

2.2.1.5. DNA isolation from BAC clones

A single bacterial colony was seeded into 3 ml of LB media supplemented with 20 µg/ml of chloramphenicol in a 15 ml snap-cap polypropylene tube. The culture was grown for 16 hours, with shaking at 280 rpm, 37°C. The culture was then centrifuged at 3,000 rpm for 10 minutes, at room temperature to collect the cells. All solutions were added to the side of the tube, and gently mixed to avoid shearing the DNA. The supernatant was discarded and the pellet was resuspended in 300 µl of P1 solution (15 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ ml RNAse A). 300 µl of P2 solution (0.2 N NaOH, 1 % SDS) was added and the solution gently shaken. The tube was incubated at room temperature for 5 minutes. 300 µl of P3 solution (3 M KOAc, pH 5.5) was then added to the tube while gently shaking and incubated on ice for 5 minutes. The tube was centrifuged at 3000 rpm for 10 minutes, the supernatant transferred to a sterile 1.5 ml Eppendorf tube and further centrifuged at 14,000 rpm for 10 minutes. The supernatant was then transferred to a new 1.5 ml Eppendorf tube that contained 800 µl of ice cold isopropanol. The solution was mixed by inversion, and placed on ice for at least 5 minutes. The tube was centrifuged at 14,000 rpm for 15 minutes at 4°C. The supernatant was removed and 500 µl of 70 % ethanol was added. The tube was inverted several times to wash to DNA pellet and centrifuged for 5 minutes at 14,000 rpm, 4°C. The supernatant was removed and the
pellet air dried at room temperature. When the pellet turned white, 30 μl of sterile distilled water was added. The tube was incubated at room temperature for 10 minutes and occasionally tapped to promote the resuspension of the pellet. 5-10 μl were used for restriction digests.

2.2.1.6. Isolation and purification of genomic DNA from 129 Sv/J mouse liver

129 Sv/J mouse livers were kindly donated by Dr. Richard Rozmahel's laboratory (Hospital for Sick Children, Toronto, Ont.). Genomic DNA was isolated according to the methods of Sambrook et al. (1989). The livers were cut into very small fragments with a sharp razor blade and then placed into a mortar that had been pre-chilled with liquid nitrogen. A pestle was then used to grind the frozen liver tissue into a very fine powder. The powder was placed into a 15 ml polystyrene tube and 10 volumes of extraction buffer (10 mM Tris-HCl pH 8.0, 0.1 M EDTA, 20 μg/ml RNase A, 0.5 % SDS) was added. The solution was then incubated at 37° C for 1 hour. 100 μg/ml of Proteinase K was added and mixed in with the solution. The tube was placed in a water bath at 50° C for 3 hours, with periodic swirling. The solution was cooled to room temperature, and equal volumes of phenol and chloroform were added and the solution was inverted to mix. The tube was centrifuged at 13,000 rpm for 5 minutes and the top aqueous layer was removed and placed into a sterile tube. An equal volume of chloroform was added to remove excess phenol and inverted to mix. The tube was centrifuged at 13,000 rpm for 5 minutes and the top aqueous layer was removed and placed into a fresh tube. 2 volumes of 100 % ethanol and 0.2 volumes of 10 M ammonium acetate were added and the tube was inverted to mix. The tube was centrifuged at 13,000 rpm for 10 minutes at 4° C. The supernatant was aspirated and the pellet was washed with 70 % ethanol. The tube was centrifuged again at 13,000 rpm for 5 minutes and the supernatant removed. The pellet was air dried for 5 minutes and resuspended in TE 8. To solubilize the pellet, the tube was warmed at 37° C for 10-20 minutes.

2.2.1.7. Determination of DNA concentrations

A 100 μl aliquot of DNA diluted in sterile water was placed in a 100 μl quartz cuvette and the optical density (O.D.) at 280 and 260 nm (UV spectrum) was measured. The O.D. ratio of 260 nm/280 nm estimated the purity of the DNA sample (a ratio greater than 1.8 was deemed satisfactory). The O.D. at 260 nm was used to calculate the concentration of the DNA in the
sample. For double stranded DNA an O.D. of 1 unit = 50 µg/ml of DNA, while for single stranded DNA, an O.D. of 1 at 260 nm = 33.3 µg/ml DNA. The following calculation was used:

\[
\text{DNA O.D. constant (either 50 or 33.3 µg/ml depending on the sample)} \times \text{O.D. at 260 nm} \times \text{dilution factor} \times 1 \text{ ml/1000 µl} = \text{amount of DNA in sample}
\]

2.2.1.8. Restriction endonuclease digestions, electrophoresis, DNA fragment isolation and purification, dephosphorylations, and ligations

Restriction digests

Restriction endonuclease digestions were carried out under optimal conditions as specified by each manufacturer. For BAC clone digests, 10 units of restriction enzyme were used for each µl of BAC clone DNA. The time allowed for restriction endonuclease digestion was between 1 and 24 hours depending on the enzyme used and the extent of the digestion as monitored by electrophoresing the digested DNA along with uncut DNA on an agarose gel. Enzymes were inactivated either by incubation at 65°C for 20 minutes or by the addition of loading buffer dye (0.25 % bromophenol blue, 0.25 % xylene cyanol, 30 % glycerol in 1 X Tris-acetate EDTA).

Electrophoresis

DNA samples mixed with loading buffer dye were size separated on agarose gels (1 X Tris-acetate-EDTA) of varying agarose concentration and stained with 0.5 µg/ml ethidium bromide for visualisation of DNA under ultraviolet light. Visualization of DNA fragments and photographic records of the gels were carried out using the GDS 7500 gel documentation system from Diamed Lab Supplies Inc. (Mississauga, Ont.).

DNA fragment isolation and purification

The QIAEX gel purification kit was used to extract and purify DNA fragments from agarose gels. DNA cut with restriction endonucleases was size separated on an agarose gel, and the desired DNA fragment was cut out with a sharp, clean, razor blade. The gel slice was placed in a sterile 1.5 ml Eppendorf tube, weighed, and the amount of gel solubilization buffer (QX1) was determined based on a 3:1 buffer to gel slice weight ratio. The corresponding amount of solubilization buffer (3 M NaI, 4 M NaClO₄, 10 mM Tris-HCl pH 7.0, and 10 mM sodium thiosulphate) was added to the gel slice and the tube was vortexed briefly to promote dissolution
of the gel slice. DNA binding QIAEX beads were reconstituted by vortexing for 60 seconds. 10 μl of the beads were added to the tube, and the tube was vortexed (or flick mixed for plasmids > 10 kb) for a few seconds to promote binding of the DNA to the beads. The tube was then incubated at 50° C for 10 minutes with vortexing every 2 minutes. The mixture was centrifuged at 13,000 rpm, room temperature, for 1 minute. The supernatant was aspirated and the beads were washed with 500 μl of QX2 solution (8M NaClO₄, 10 mM Tris-HCl pH 7.0). The previous two steps were repeated again. The beads were then washed with 500 μl QX3 (70 % ethanol, 100 mM NaCl, 10 mM Tri-HCl pH 7.5) and centrifuged. This was repeated once more. The supernatant was removed, and the pellet was allowed to air dry until it became white. The pellet was then resuspended in 20 μl TE 8 and incubated at 50° C for 5 minutes with occasional vortexing or flick mixing. The beads were then pelleted by centrifugation at 13,000 rpm, room temperature for 1 minute and the supernatant containing the DNA was transferred to a sterile 1.5 ml Eppendorf tube.

**Dephosphorylation**

To prevent religation of cut DNA cloning vectors and to maximize ligation of inserts into these vectors, dephosphorylation of 5’ phosphate groups on the vectors was performed. After restriction endonuclease digestion of the plasmid vector in a 50 μl volume, the solution was heated to the appropriate temperature necessary to inactivate the enzyme. The solution was then cooled to room temperature for 20 minutes. 1 unit of Calf Intestinal Phosphatase (CIP) was added to the solution along with a final concentration of 1X One-Phor-All (OPA) enzyme buffer (10 mM Tris-acetate pH 7.5, 10 mM magnesium acetate, 50 mM potassium acetate ) from Pharmacia. The solution was incubated at 37° C for 30 minutes to 1 hour. Loading buffer dye was added and the mixed solution was electrophoresed through an agarose gel and gel purified as described above.

**Ligations**

A 1:1, 3:1, or 5:1 insert to vector ratio of molar ends was used depending on the sizes of the inserts and vectors. In a volume of 20μl, 1 unit T4 DNA ligase (GIBCO BRL) was added to a solution of plasmid and insert along with a final concentration of 1X T4 DNA ligation buffer (GIBCO BRL). A self- ligation control consisted of all the same reagents (without the insert) and CIP treated plasmid. The solution was incubated at room temperature for 4 hours to overnight and then stored at 4° C or -20° C if not used right away. In general, 10 μl of the ligation and self-ligation control were used for bacterial transformations.
2.2.1.9. DNA sequencing

DNA inserts were sequenced either in lab or at the Hospital for Sick Children Sequencing Facility. For in lab sequencing, the Pharmacia T7 Quickprime® kit was used. For sequencing using the T7 Quickprime® kit, DNA templates were prepared by using the Pharmacia Flexiprep kit. The concentration of the preparation was determined and approximately 1-2 µg (in a volume of up to 16 µl) of double stranded plasmid DNA was used for sequencing along with the addition of 4 µl of 2 M NaOH to denature the DNA. The mixture was incubated at room temperature for 10 minutes, and 6.66 µl of 7.5 M NH₄Ac and 60 µl of 100 % ethanol was added to precipitate the DNA. The mixture was vortexed and placed at -20°C overnight or -80°C for 10-15 minutes. The mixture was then spun down at 18,000 rpm for 10 minutes at 4°C. The supernatant was removed and the pellet was washed with ~ 100 µl of 70 % ethanol. The mixture was centrifuged for 2-3 minutes at 18,000 rpm and the supernatant was removed. The pellet was air-dried at 37°C for 5 minutes and resuspended in 4 µl sterile distilled water. 1 µl of annealing buffer (1 M Tris-HCl pH 7.5, 5 mM DTT, 100 µg BSA/ml, 5 % glycerol) and 2 µl of sequencing primer (~ 10 ng/µl) were added to the 4 µl of denatured DNA template. The mixture was then vortexed, briefly spun down, incubated at 37°C for 20 minutes and then cooled to room temperature for 5 minutes. 1-3 sets of 1.5 ml Eppendorf tubes labelled either A,C,G, or T were placed aside and 1.25 µl of A,C,G, or T “short mix” (dideoxynucleotides terminating the polymerase catalyzed DNA elongation along with non limiting concentrations of all four dNTPs) were aliquoted to the corresponding tubes.

A master mix was then prepared. The standard master mix protocol is designed for 2 reactions. In a sterile 1.5 ml Eppendorf tube, 0.5 µl of sterile water, 3 µl of labeling mix A (1.375 µM each of dCTP, dGTP, dTTP), (3n/ 8 µl) µl (where n = set of 2 reactions) of T7 polymerase, (2n- x µl T7 DNA polymerase) µl enzyme dilution buffer (20 mM Tris-HCl pH 7.5, 5 mM DTT, 100 µg BSA/ml, 5 % glycerol) and 0.5 µl of 35S dATP. 3 µl of the master mix was added to the annealed template and primer at room temperature to initiate T7 DNA polymerase catalyzed DNA elongation. This mixture was incubated for 5 minutes while the tubes with A,C,G,T dideoxynucleotides were placed in a tabletop heater set at 37°C. 2.3 µl of the labeling reaction was added in sequence (every 15 seconds) to each of the tubes containing the dideoxynucleotides. After a 5 minute incubation at 37°C, each reaction was terminated with the addition of 2.5 µl of stop solution (0.3% each of bromophenol blue and xylene blue FF, 10 mM EDTA pH 7.5, 97.5% deionized formamide) at 15 second staggered intervals. 1.5 –2 µl of each of the stopped DNA elongation reactions (one set of A,C,G,T reactions) were heated at 95-100°C for 5 minutes and then loaded side by side in adjacent wells on a prewarmed (45 min-1hr at 80
Watts) 6% polyacrylamide sequencing gel for size separation by electrophoresis. After complete electrophoresis, the gels were placed onto Whatman paper along with an overlying cover of Saran wrap and dried in a Bio-Rad Model 583 gel drier for 1-2 hours at 80°C. The gels were then exposed to BIOMAX autoradiographic film overnight at room temperature.

2.2.1.10. Polymerase Chain Reaction (PCR) amplification

PCR was used used in several instances. It was used to produce DNA fragments of interest for subcloning into either DNA cloning vectors such as pBSK-, pCR2.1, or in expression vectors such as the pGEX fusion protein vector. It was also used in screening for the presence of the NAT gene knockout (correctly targeted event), as well as in making DNA probes for Southern blotting and BAC clone screening. All of the primers used for PCR are listed in Table 2.0.

For standard PCR, each reaction contained 1 X GeneAmp PCR buffer with MgCl₂ (final concentrations of 1.5 mM MgCl₂, 50 mM KCl, 10 mM Tris-HCl pH 8.3 and 0.001 % (w/v) gelatin), 100 µM dNTPs, 8 pmol of each primer, 1 U Amplitaq® DNA polymerase, and 1 ng of lambda or bacterial plasmid vector DNA in a final reaction volume of 20 µl. The components were mixed together in a master mix, with the exception of the DNA template, and kept on ice. The master mix was aliquoted to 200 µl PCR tubes placed on ice. To minimize non-specific amplifications, the tubes were then placed in the GeneAmp 9600 PCR machine set at 85°C, and the addition of the DNA template followed shortly after (this is referred to as a “Hot Start”). Primer hybridization temperatures were determined by establishing the temperature at which 50% of the primer-template DNA duplex is single stranded (this is known as Tm). The simplified formula used to calculate the Tm was 2(A+T) + 4(G+C) (Wallace et al. 1979). However, the hybridization temperatures of both primers did not always match, and as such the Tm in those cases was used as a starting point with optimal hybridization temperatures being determined empirically. Cycle parameters were usually 94°C for 1 minute followed by 30-35 cycles of 94°C for 10 seconds (denaturation), Tm - 5°C for 10 seconds (primer hybridization), 72°C for 1 sec/40 nucleotides of expected PCR product (DNA elongation). A final extension time of 5 min at 72°C was added after the last cycle followed by a final cooling stage to 4°C.
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer sequence (5'-3')</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGK/PCR-probe/5'</td>
<td>ACGAGATCCAGCAGCCCTCTGTTTC</td>
<td>68.0</td>
</tr>
<tr>
<td>N3/ PCR-probe/3'</td>
<td>GCCTGGAGGACAAACATCCTTTTC</td>
<td>70.0</td>
</tr>
<tr>
<td>PGK/Neo 2</td>
<td>ATAGCCCTGAAGAAAGAGTTC</td>
<td>53.4</td>
</tr>
<tr>
<td>S/E Reverse</td>
<td>TCGAGCCTTTCAAAATTTTCAC</td>
<td>51.9</td>
</tr>
<tr>
<td>5' mNAT3 probe Jn99</td>
<td>CTTCCATTGGTGTATCATGCTAA</td>
<td>54.4</td>
</tr>
<tr>
<td>3' mNAT3 probe Jn99</td>
<td>TGGGACTCCATGGAGCAGAAAG</td>
<td>59.8</td>
</tr>
<tr>
<td>N2 3'UTR-PCR1</td>
<td>CTGTATGCTTGTCTCTGATCTTTAC</td>
<td>57.0</td>
</tr>
<tr>
<td>N2 3'UTR-PCR2</td>
<td>GCATGCTCAGCTGCCCTTTCTAC</td>
<td>60.0</td>
</tr>
<tr>
<td>N2 3'UTR SEQ P.1</td>
<td>TGAGGCAATCTTTGCTAGT</td>
<td>53.4</td>
</tr>
<tr>
<td>N2 3'UTR SEQ P.3</td>
<td>AATTTCTATGTAATCTAGTT</td>
<td>43.1</td>
</tr>
<tr>
<td>N2 3'UTR SEQ P.4</td>
<td>AACTTCTCTCAGCATTTTCT</td>
<td>49.2</td>
</tr>
<tr>
<td>N2 3'UTR primer5</td>
<td>AGACATATTACTATAATCA</td>
<td>45.1</td>
</tr>
<tr>
<td>N2 3'UTR primer6</td>
<td>GCATCAGAATAAAATCTCTGAC</td>
<td>50.0</td>
</tr>
<tr>
<td>N2 3'UTR SP7</td>
<td>GGTATCAGATGGAGTACGTCAGCG</td>
<td>58.1</td>
</tr>
<tr>
<td>N2 3'UTR SP8</td>
<td>ACTTCCTCAAGCTAAATATG</td>
<td>49.2</td>
</tr>
<tr>
<td>mNAT</td>
<td>TTGCCTTACGGAACCATGGGAC</td>
<td>58.4</td>
</tr>
<tr>
<td>Ha.mNAT1-BamHI</td>
<td>CGCGCG/GATCCCTGGACATCGAAGCATACTTTTGAAAG</td>
<td>69.7</td>
</tr>
<tr>
<td>Ha.mNAT1-Sal I</td>
<td>CGCGCG/TGAGCTAAAAATAGTAAAAACTAGTTCCATTG</td>
<td>65.3</td>
</tr>
<tr>
<td>Ha.mNAT2-BamHI</td>
<td>CGCGCG/GATCCCTGGACATCGAAGCTACTTTTGAAAG</td>
<td>70.8</td>
</tr>
<tr>
<td>Ha.mNAT2-Sal I</td>
<td>CGCGCG/TGAGCTAAAAATGGTAAAAATCGATCCATTG</td>
<td>66.5</td>
</tr>
<tr>
<td>Ha.mNAT3-BamHI</td>
<td>CGCGCG/GATCCCTGGACATTGAGACATCTTTTGAAAG</td>
<td>68.6</td>
</tr>
<tr>
<td>Ha.mNAT3-Sal I</td>
<td>CGCGCG/TGAGCTAAAAATGAAAAAGGACATTGCCAC</td>
<td>65.3</td>
</tr>
</tbody>
</table>

Table 2.0. Oligonucleotide primers used for sequencing and PCR
Restriction endonuclease recognition sites are underlined, / represents the actual cleavage site. The start (ATG) and stop (CTA) mNAT ORF translation codons are in bold. The Tm is the temperature at which half the primer can hybridize to its target sequence. Primers were from Dalton Chemical Labs Inc., ACGT Corp., General Synthesis & Diagnostics.
2.2.1.11. Southern blotting

Southern blotting (Southern 1975) was performed using the vacuum transfer method. DNA digested with restriction enzymes or DNA amplified by PCR was electrophoresed in agarose gels (usually about 0.7%). A nylon membrane (Hybond N+, Amersham) was then cut slightly smaller than the size of the gel, soaked in sterile distilled water, and placed in the vacuum transfer apparatus. The gel was then gently placed onto the membrane, covering it on all sides. Warm, semi-liquified agarose was added into the wells of the gel and the sides of the gel and when solidified, created an air tight seal. The apparatus was then sealed and the gel was covered with 0.2 N HCl (for depurination). The vacuum pump was then turned on to a pressure setting of 40 cm H₂O. After 20-25 minutes, the HCl was removed, the gel washed twice with distilled water, and covered with denaturing solution (1.5M NaCl, 0.5M NaOH). After 20-25 minutes, the denaturing solution was removed, the gel washed twice with distilled water, and covered with neutralizing solution (1M Tris-HCl pH 7.4, 1.5M NaCl). Following another 20-25 minute incubation period, the neutralizing solution was removed, the gel washed with distilled water and then covered with 400 ml of 6 X SSC (1.1 M NaCl, 0.09 M NaCitrate). This step was allowed to proceed for 1 hour. When complete, the gel was lifted to break the suction, the vacuum turned off, and the positions of the individual wells of the gel marked with pencil on the nylon transfer membrane. The membrane was then washed with 6 X SSC to remove residual agarose, and briefly dried on Whatman paper. The DNA on the membrane was immobilized by UV crosslinking. The agarose gel was placed in a water bath containing several drops of ethidium bromide to permit restaining of the gel and determination of the extent of DNA transfer to the nylon membrane.

2.2.1.12. DNA probe labeling

DNA probes used in Southern blotting were either produced by PCR amplification or by restriction enzyme digestion from inserts contained in DNA cloning vectors. The probes were labelled with 50μCi of ³²P dCTP (Amersham) by the random priming method using the Pharmacia Quickprime® Kit. After restriction enzyme digestion or PCR amplification, the DNA probe was run through an agarose gel and gel purified. 50 ng of DNA was then used in the labeling reaction in a total initial volume of 34 μl. The mixture was boiled at 100° C for 3 minutes to denature the DNA, and then briefly centrifuged. 10 μl of reagent mix (dATP, dGTP, dTTP and random deoxyribonucleotide 9 mers) were added along with 5 μl of ³²P dCTP (50μCi), and 1 μl of T7 DNA polymerase (10.6 units/μl) in a total final volume of 50 μl. The solution was vortexed to mix the contents and incubated at 37 °C for 30-45 minutes. The labeled DNA
solution along with 200 µl of TE 8 was filtered through a Sephadex G-50 matrix placed in a 1 cc hypodermic syringe (Sambrook et al. 1989). The syringe was placed upside down in a 12 ml polypropylene tube, with the nozzle facing a 1.5 ml Eppendorf tube. The syringe was centrifuged at 2500 rpm, 4° C, for 2 minutes to remove unincorporated label, with the flow through collected in the Eppendorf tube. The extent of probe labeling (specific activity) was determined by taking 1 µl of the labeled, filtered probe and placing it into a scintillation vial filled with 4 ml of scintillation fluid. The vial was placed in a Beckman LS6500 scintillation counter and the specific activity determined. For each membrane 1 X 10^6 counts to 2 X 10^6 counts/ml of hybridization solution was used.

The membrane was placed inside a hybridization bottle and prehybridized for 1 hour at 60° C in a Hybaid MICRO-4 rotisserie-type hybridization oven with 5-15 ml of ExpressHyb™ solution depending on the size of the membrane. After 1 hour, the prehybridization solution was removed and fresh ExpressHyb solution was added. The probe was boiled at 100° C for 3 minutes to denature the DNA and then chilled quickly on ice. The probe was then added to the hybridization solution and the bottle placed at 60° C in the hybridization oven for 1 hour-overnight.

After hybridization, the solution was discarded, and the membrane was washed with 5-20 ml of 2 X SSC, 0.05 % SDS three times at room temperature for 40 minutes. This was followed by two washes with 0.1 X SSC, 0.1 % SDS at 50° C for 30 minutes. The membrane was then removed from the bottle, quickly blotted dry (to remove excess moisture, but not allowed to dry completely) and placed in Saran wrap. The wrapped membrane was placed in a film cassette and exposed to BIOMAX film with 2 enhancing screens (Fisher Biotech L-Plus) for 1-72 hours at room temperature or -80° C.

2.2.2. Specific methods

2.2.2.1 ES cell culture

2.2.2.1.1. Standard growth conditions

Culturing, harvesting, electroporation, and selection of R1 ES cells were performed by Mrs. Jean Huang (Dr. Richard Rozmahel's laboratory, Hospital for Sick Children) and by Dr. K. Sugamori (Dr. Denis Grant's laboratory, Hospital for Sick Children). R1 ES cells were cultured on gelatinized plates in either 10 ml (for 10 cm culture dishes) or 1.5 ml (for 24-well plates) of
complete ES cell culture medium. The medium contained DMEM supplemented with 20 % Fetal Bovine Serum (FBS), 2 mM glutamine, 0.1 mM 2-mercaptoethanol, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, and 500 units/ ml LIF (a cytokine polypeptide inhibiting the differentiation of ES cells). Cells were incubated in an incubation atmosphere of 5 % CO₂ and 95 % air at 37°C.

To make gelatinized plates, the surface of each plate was coated with 2 ml of 0.1 % gelatin solution. The gelatin solution was removed after 5 minutes by aspiration and the plates were allowed to air dry for 15 minutes.

To harvest ES cells, the cell media was first aspirated and the plates or wells rinsed with calcium/magnesium-free PBS. The media was aspirated and the plates or wells were covered with a thin film of 0.25 % trypsin/ 0.04 % EDTA, citrate-saline solution. After a 3 minute incubation, complete ES cell medium was added to stop the trypsinization and cells were then pipetted up and down to achieve a homogenous single-cell suspension.

2.2.2.1.2. Electroporation of ES cells with Vectors 1 or 2

Media was aspirated from 80 % confluent, exponentially growing, early passaged R1 ES cells in 10 cm culture dishes. The cells were then washed with 5 ml calcium/magnesium-free PBS. Cells were trypsinized with 1 ml trypsin and suspended in 8 ml of complete ES cell culture medium. Single-cell suspensions were pooled into 50 ml capped conical tubes and cells were pelleted by centrifugation at 1000 rpm for 10 minutes. The pellets were then washed with calcium/magnesium-free PBS and resuspended at a density of approximately 7 X 10⁶ cells /ml or 5 X 10⁶ cells in ~800 μl in ice cold calcium/magnesium-free PBS. The cell suspensions were then placed into 0.4 cm wide sterile electroporation cuvettes and 50 μg (50 μl) of the targeting vector were added and pipetted up and down to mix. The cuvettes were then immediately placed on ice. The ES cells were electroporated for 5-6 msec at 500 μF, 0.24 kV to allow the vector to enter the cells. After electroporation, the cuvettes were placed on ice for 20 minutes to allow the ES cells to recover. The contents of all the cuvettes were pipetted into a sterile tissue culture flask containing 100 ml of complete ES cell culture medium. The contents were well mixed and 10 ml of the suspension were plated into 10 cm gelatinized plates. The cells were allowed to recover for 2 days, with daily changing of the media in the absence of any selection.
2.2.1.3. Colony isolation and expansion

Cells that had both taken up the vector (and express the neomycin gene) and that had lost the TK gene were selected by plating the targeted ES cells in complete ES cell culture media supplemented with 200 μg/ ml G418 and 2 X 10⁶ M gancyclovir. To determine the proportion of cells that had undergone non homologous recombination, one plate was set aside which contained the targeted ES cells in complete ES cell medium and G418 alone. The media was changed every day to remove dead cells and wastes and selection was continued for a total of 7 days.

At the end of the selection period, the ES cell medium was removed and the ES cells were washed with calcium/ magnesium-free PBS. 1 ml of calcium/ magnesium-free PBS was added to each plate and individual colonies were picked through sterile pipette tips using a stereo microscope (placed within a laminar flow tissue culture hood) to visualize the cells. Each colony was placed into 1.5 ml complete ES cell culture medium (without antibiotics) in a single well within a 24 well tissue culture plate. ES cells were allowed to recover for 2 days and then trypsinized and allowed to grow for 4 days. The media was changed every 2 days.

Cells were trypsinized and suspended in 1 ml of complete ES cell culture medium. 900 μl of cell suspension was transferred from each well into separate 1.5 ml Eppendorf tubes. 1.5 ml of complete ES cell culture medium was added to the remaining 100 μl of the cell suspension in the wells. These remaining cells were grown for 4 days and then frozen down. The cells transferred to the 1.5 ml Eppendorfs were pelleted by centrifugation at 5000 rpm for 5 minutes. Pellets were resuspended in 500 μl of PK buffer (100 mM Tris-HCl pH 7.8-8.0, 5 mM EDTA, 0.2 % SDS, 200 mM NaCl), and 400 μg/ml proteinase K and incubated overnight at 42° C with shaking at 225 rpm. The next day, 1 ml of 100 % ethanol was added to each tube. The tubes were then mixed by inversion to precipitate the DNA. The DNA was pelleted by centrifuging at 12 000 rpm for 10 minutes. The ethanol was removed and the DNA pellet was air dried for 5 minutes by placing the tubes with caps open in a heating block set at 37° C. The pellets were then resuspended in 80 μl sterile distilled water. Clones that tested positive for the targeted event were transferred into 12 well plates, allowed to grow for several days (with change of media), then transferred into 6 cm plates, expanded and then transferred into 10 cm plates.

When clones reached 80 % confluency in the 10 cm plates, ES cells were harvested. The media was removed by aspiration and the cells washed with 4 ml Calcium/ Magnesium-free PBS. The media was suctioned off and cells were trypsinized for 5 minutes. 4 ml of DMEM was added to stop the trypsinization reaction. Detached cells were pipetted up and down to produce
a homogenous single cell suspension. The cells were then collected into a plastic 10 ml tube. The tube was centrifuged for 8 minutes at 5000 rpm. The supernatant was removed and the pellet was resuspended in 200 μl PK buffer and 20 μg/μl proteinase K and transferred into a 1.5 ml Eppendorf tube. The tube was incubated overnight in a water bath set at 37 °C. The DNA was precipitated the next morning by adding an equal volume of isopropanol to the solution followed by inversion of the tube to mix. The DNA was collected by twisting it around a disposable sterile pipette tip and then it was placed in a new, sterile 1.5 ml Eppendorf tube containing 50 μl sterile distilled water.

2.2.2.1.4. Freezing of ES cells

To preserve the ES cells in their plates for extended periods of time, ES cell culture medium was replaced with freezing medium consisting of complete ES cell culture medium with 25 % FBS and 10 % DMSO. The plates were pre-cooled on ice for 30 minutes prior to freezing. The plates were put into the ~80 °C freezer in covered styrofoam boxes which were sealed with packing tape.

2.2.2.2. Preparation of Vectors 1 and 2 for electroporation

To prepare gene targeting vectors for electroporation into R1 ES cells, the QIAGEN Maxi kit was used and the procedure was carried out according to the manufacturer's instructions. In the final step, the pellets were air dried for 10 minutes and resuspended in 500 μl of sterile distilled water.

Before electroporation of the ES cells, the vectors were linearized with the Not I restriction endonuclease. Small aliquots of the digested and undigested vector were electrophoresed through an agarose gel to determine the success of linearization. The vector was then precipitated out of solution with 2 volumes of 100 % ethanol, centrifuged, and the pellet was resuspended in sterile distilled water to a final DNA concentration of 1 μg/μl.
2.2.2.3. Production of the Vector 1 positive control for use in screening gene targeting events by PCR

Construct Mp5 was cut with Kpn I, releasing a 1.7 kb fragment spanning 520 bp of the open reading frame all the way to the end of the known 3' UTR sequence (1.2 Kb) corresponding to mouse NAT3. The digest was run on an agarose gel, the 1.7 kb fragment isolated and gel purified. An empty pPNT gene targeting vector was linearized with Kpn I, run on an agarose gel, gel purified and treated with Calf Intestinal Phosphatase to prevent religation of the vector. The 1.7 kb fragment was then subcloned by ligation into the linearized, CIP treated vector to produce the positive control template for Vector 1 (mouse NAT3 knockout) targeted events (Figure 2.2). The ligated product was then transformed into DH5α cells, and plated. Colonies were picked, indexed, streaked on a fresh LB agar plate for later use and also used to inoculate overnight cultures of LB broth. Minipreps of each of the overnight cultures were made and the plasmids were then digested with Kpn I and Bam HI restriction enzymes to determine the presence of the insert and its orientation.

Colonies corresponding to the digests showing the correctly subcloned fragments were then grown up in overnight cultures and the plasmid isolated by the flexiprep method.

2.2.2.4. Testing of the positive control for Vector 1

PCR was used to amplify the 1.7 kb piece from the Vector 1 positive control. Initial experiments used 1 ng of the positive control. The sensitivity of the PCR reaction was optimized by diluting the positive control from 1 ng to 1 picogram. 100-200 ng of genomic DNA isolated from untargeted R1 ES cells were spiked with 1 picogram of the positive control. A 1.7 kb sequence in a 3 X 10^9 bp mouse genome would represent 0.00006 % of the genome. 1 picogram of the positive control spiked into 100 ng of genomic DNA would make up 0.001 % of the DNA solution, while in 200 ng it would make up 0.0005 % of the genomic DNA solution. The mixture was subjected to the standard PCR method with a cycle denaturation time of 20 sec at 94 °C, annealing for 20 sec at 55 °C, and an extension time of 1 min 20 sec at 72 °C.
Figure 2.2.
Cloning of the Vector 1 positive control.
2.2.2.5. Cloning and sequencing of additional NAT3 3' UTR

To design a Southern blotting screening assay for the mouse NAT3 knockout, additional mouse NAT3 3' UTR sequence was needed. Lambda clones 1, 9, 14, and 23 contained all or parts of the mouse NAT3 open reading frame and varying lengths of 3' UTR. XL PCR on these selected lambda clones was used to amplify more of the mNAT3 3'UTR using the mNAT and T3 or T7 primers.

For XL PCR, the GeneAmp® XL PCR kit was used. The XL PCR method uses the rTth DNA polymerase in a Hot Start PCR reaction facilitated by using the AmpliWax® PCR Gem (a wax bead) to separate the rTth polymerase from the primers until the PCR reaction reaches 85° C. This method reduces or eliminates PCR mis-priming and primer oligomerization side reactions. In addition, the rTth polymerase also has a 3'-5' exonuclease proofreading function which corrects nucleotide misincorporations.

Two individual master mixes were prepared. The lower master mix was made first. It consisted of a 1 X final concentration of 3.3 X XL Buffer (Tricine, potassium acetate, glycerol, and DMSO), a final concentration of 2.75 mM magnesium acetate solution, 2mM final concentration of dNTPs, and 12 pmol final concentration of each primer. The master mix was vortexed, spun down and placed on ice. 20 µl of the lower master mix were aliquoted into ice chilled 200 µl PCR reaction tubes. One AmpliWax® PCR gem was added to each tube containing the lower reaction mix. The tubes were then placed in the GeneAmp® 9600 PCR machine and the temperature set at 80° C for 5 minutes to melt the PCR gem. This was followed by cooling to 25° C for 5 minutes to solidify the melted PCR gem. This solidification step served to separate the upper mix from the lower mix.

The upper master mix was then prepared. It consisted of a final concentration of 1 X 3.3 X XL Buffer, and 2 U of the rTth XL DNA polymerase. The PCR cycling parameters were: 94° C for 1 minute at the start followed by 32 cycles of 94° C for 15 seconds, 51° C for 30 seconds, and 70° C for 10 minutes. A final extension time of 5 minutes at 70° C followed the cycles, ending with a cooling stage to 4° C.

The reactions were run on an agarose gel and the largest PCR amplified fragment (3 kb) was excised, gel purified and subcloned into the pCR 2.1 vector using the TA Cloning® Kit according to the manufacturer's instructions (Figure 2.3). The 3 kb fragment was digested with Kpn I to confirm the identity of the fragment. The TA Cloning® kit facilitates the subcloning of
PCR amplified products by taking advantage of the non-specific addition of adenine nucleotides to the 3' ends of PCR products by DNA polymerases such as Taq and rTth. Bacterial clones were picked, plasmids isolated and then digested with Eco RI and Kpn I to isolate those plasmids that had insert and to determine the orientation of the insert. The plasmid containing the correctly oriented insert was then sequenced and overlapping sequences were compiled into a full-length sequence using MacVector™ software (Eastman Kodak Company, Rochester NY).

The probe was verified by cutting 10 µg of 129 Sv/J genomic DNA with 100 units Xba I, followed by Southern blotting and probing with an 860 bp probe designed from the new sequence.

2.2.2.6. Screening of the mouse NAT3 targeted knockout by Southern blotting

After electroporation of Vector 1 into R1 ES cells and double selection, 10 µg of genomic DNA was isolated from surviving ES cells, cut with 100 units of Xba I and electrophoresed on a 0.9 % agarose gel, Southern blotted and probed with the 860 bp probe.

2.2.2.7. Production of the Vector 2 positive control for use in screening gene targeting events by XL PCR

Figure 2.4 outlines the production of the Vector 2 positive control. The NAT2 3'UTR/pCR2.1 vector was cut with Eco RI, yielding a 5.0 kb piece (containing the entire mNAT2 ORF and the Pst I-Sac I 3' arm of homology). This fragment was subcloned downstream of the neomycin resistance gene into an Eco RI cut, CIP treated empty pPNT gene targeting vector. The ligation was transformed into DH5α cells and plated. Individual colonies were picked, indexed and streaked on a fresh LB agar Petri plate as usual. Overnight cultures were grown from the colonies, minipreps made, and each digested with Eco RI to first eliminate the colonies not bearing any insert. Colonies with insert were digested separately with Sac I alone to determine which clones had the insert oriented in the correct direction.
Figure 2.3.
XL PCR amplification (using T7 and mNAT primers) and subcloning into the pCR® 2.1 vector of a 3 kb fragment containing additional mNAT3 3' UTR sequence.
NAT2 3’ UTR/ pCR 2.1
~9.4 kb

Cut with Eco RI

5.0 kb fragment

Cut pPNT with Eco RI

pPNT
7.2 kb

Ligate 5.0 kb fragment into pPNT

Vector 2 positive control
~12.2 kb

Figure 2.4.
Cloning of the Vector 2 positive control.
2.2.2.8. Testing of the XL PCR reaction on the positive control and screening of the Vector 2 targeted events

Once a positive clone was found, an overnight culture of it was grown from a streaked plate, and the plasmid isolated by the flexiprep method. XL PCR was used and as mentioned above, two individual master mixes were prepared. The lower master mix was made first. It consisted of a 1X final concentration of 3.3X XL Buffer (Tricine, potassium acetate, glycerol, and DMSO), a final concentration of 2.75 mM magnesium acetate solution, 2mM final concentration of dNTPs, and 12 pmol final concentration of primer PGK-NeoII and 12 pmol concentration of primer S/E reverse. The master mix was vortexed, spun down and placed on ice. 20 μl of the lower master mix were aliquoted into ice chilled 200 μl PCR reaction tubes. One AmpliWax® PCR gem was added to each tube containing the lower reaction mix. The tubes were then placed in the GeneAmp® 9600 PCR machine and the temperature set at 80°C for 5 minutes to melt the PCR gem. This was followed by cooling to 25°C for 5 minutes to solidify the melted PCR gem. This solidification step served to separate the upper mix from the lower mix.

The upper master mix was then prepared. It consisted of a final concentration of 1X 3.3 X XL Buffer, and 2 U of the rTth XL DNA polymerase. The upper master mix was vortexed, spun down, and placed immediately on ice and cooled for about 5 minutes. 15 μl of this master mix were aliquoted to the tubes. 13 μl of sterile distilled water were added to the tubes and then 2 μl of genomic DNA isolated from selected ES cells electroporated with vector 2, were added. The 3.8 kb target sequence makes up about 0.00013% of the mouse genome (3 X 10^9 bp). 15 μl sterile, distilled water was added to one control tube. 10 μl (~ 200 ng) of genomic DNA from untargeted ES cells were aliquoted to two separate tubes and then 1 pg of the positive control template was added to one tube and 0.01 pg of the positive control template was added to the other tube. The positive control target sequence of 4.3 kb would make up 0.00014% of the mouse genome (3 X 10^9 bp). 1 picogram in 200 ng represents 0.0005% of the genomic DNA solution and 0.01 picograms in 200 ng represents 0.000005% of the genomic DNA solution. The abundance of the endogenous sequence in targeted ES cell genomic DNA is thus much greater than that of the spiked controls, and if a signal could be detected with the positive control, a similar signal should also be detectable in the targeted ES cell genomic DNA.

The tubes were then placed in the PCR machine, and the cycle parameters were: 94°C for 1 minute; 30 cycles each of 94°C for 15 seconds; 55°C for 30 seconds; 70°C for 5 minutes. A final extension time of 7 minutes at 72°C was followed by 4°C until removal of the samples from the PCR machine. The reaction contents of each tube were mixed with 10 X loading buffer.
dye and then 20 μl of the mix was loaded onto a 0.8 % agarose gel. The gel was electrophoresed and then Southern blotted. The membrane was probed with a 1.0 kb DNA fragment isolated from digestion of NAT2 3'UTR/ pCR2.1 with Hind III.

2.2.2.9. Subcloning and sequencing of additional mouse NAT2 3' UTR

2.4 kb of mouse NAT2 3'UTR was subcloned into the pBSK- DNA cloning vector by cutting the NAT2 3'UTR/ pCR2.1 construct and pBSK- with Sac I. Ligations were transformed into DH5α cells. Individual colonies were picked, indexed and streaked on fresh LB agar plates. LB broth was inoculated with individual colonies, cultures grown and plasmids isolated by miniprep. Plasmids were cut with Not I to linearize and with EcoRI to determine orientation. Correctly oriented inserts were then sequenced. The resulting sequence was analyzed for stretches of repetitive DNA sequences by comparing the sequence to known sequences on the GenBank BLAST sequence similarity search function (National Center for Biotechnology Information (NCBI), Bethesda, Maryland). Standard PCR with the N2 3'UTR PCR1 and N2 3'UTR PCR2 primers was used to amplify a 600 bp probe for Southern blotting screening of the mNAT1 and 2 targeted knockout. To confirm that the 600 bp probe would anneal to a unique region (the region of interest) and not to any repetitive sequences, 10 μg of genomic DNA (isolated from the 129 Sv/J mouse liver), digested with 100 units Bam HI, Southern blotted and probed with the 600 bp probe.

2.2.2.10. Screening of mouse NAT1 and 2 targeted knockout by Southern blotting

After electroporation of ES cells with Vector 2 and double selection with G418 and gancyclovir, 10 μg of genomic DNA from ES cells was digested with 100 units Bam HI, electrophoresed on an 0.6 % agarose gel, Southern blotted and probed with the 600 bp probe.

2.2.2.11. Subcloning and sequencing of additional mouse NAT1 5' UTR

Bacterial Artificial Chromosome pBAC clones containing on average 80-130 kb of the 129 Sv/J mouse genomic DNA sequence were used to provide more 5' UTR corresponding to mouse NAT1. A 129 Sv/J pBAC mouse library was screened using mouse NAT1 and mouse NAT3 probes. Mouse NAT1 and mouse NAT3 probes were generated by cutting mouse NAT1 and
mouse NAT3 open reading frames out from pET3a constructs by digestion with Nde I. The digests were run on agarose gels, the 870 bp fragment corresponding to the mouse NAT open reading frame cut out, and gel purified. This was repeated once more to ensure the absence of any part of the pET3a plasmid which could hybridize to the pBAC vector or DNA sequences during the screening of the BAC library. Screening of the pBAC library courtesy of Ms. Jennifer Skauag (laboratory of Dr. L.C. Tsui, Hospital for Sick Children, Toronto, Ont). BAC clones containing the mouse NAT1, 2 and 3 ORFs and corresponding sequences were grown and pBAC DNA isolated as described. The clones were then digested with Bam HI, Hind III, Kpn I, Sac I, and Xba I restriction endonucleases, the cut DNA electrophoresed on a 0.7 % agarose gel, transferred to a nylon membrane by Southern blotting, and probed with the mouse NAT1 and mouse NAT3 ORF probes. Subcloning and sequencing of an 8 kb band containing 1 kb of previously unmapped NAT1 5' UTR is now underway.

2.2.2.12. Production of mouse NAT antigens

Figure 2.5 outlines the PCR amplification and subcloning of the mouse NAT 1, 2 and 3 open reading frames into the pGEX vector to generate the mNAT-pGEX fusion protein construct. In addition to amplifying the entire mouse NAT 870 bp open reading frame, the primers (Ha. mNAT1- Bam HI, Ha. mNAT1- Sal I, Ha. mNAT2- Bam HI, Ha. mNAT2- Sal I, Ha. mNAT3- Bam HI, Ha. mNAT3- Sal I) also added Bam HI and Sal I restriction enzyme sites to allow easy directional cloning of the open reading frame into the pGEX fusion protein expression vector. The primers were designed to allow the insert to be in frame with the start and stop codons of the pGEX vector. The identities of the mNATs amplified by PCR were first verified by running the PCR products on agarose gels along with markers, and confirming that the products migrated close to the 900 bp band. The bands were excised from the gel, purified and cut with selected restriction enzymes and results compared to the expected fragments predicted from established restriction enzyme maps of the mouse NATs. The band corresponding to mNAT1 was cut with BstE II, mNAT2 was cut with Pst I and mNAT3 was cut with Kpn I.

The pGEX vector was prepared by cutting with Bam HI and Sal I to generate cohesive termini. The linearized vector was run on an agarose gel to remove the stuffer fragment, gel purified and ligated to the mNAT1, 2 and 3 PCR amplified products that had also been cut with Bam HI and Sal I. A self ligation control containing the Bam HI, Sal I cut pGEX vector and all other components except the insert was performed alongside the mNAT-pGEX ligation to test the extent of religation of the vector. The ligations were then transformed into DH5α cells, plated
onto LB amp plates, and incubated at 37°C overnight. Individual colonies were picked, indexed and streaked on fresh LB amp plates. A single colony was then used to seed a 3 ml overnight culture in LB broth with ampicillin. Minipreps of the cultures were then prepared and isolated plasmids were cut with Bam HI alone. Those plasmids that linearized and ran 870 bp higher than the other plasmids were isolated and cut with Sal I to release the 870 bp band. Plasmids that had inserts were then prepared by the flexiprep method and sequenced. Correctly cloned plasmids were then transformed into XA 90 bacterial cells, plated and grown overnight at 37°C.

2.2.2.13. Conditions used in mNAT- GST fusion protein expression

2.2.2.13.1. Pilot experiment expression conditions

In 15 ml polypropylene test tubes individually picked XA 90 colonies, containing the mNAT-GST fusion protein plasmid, were seeded into 3 ml of LB broth supplemented with ampicillin and grown overnight at 37°C, 280 rpm. In the morning, 150 µl of the overnight cultures were used to restart 5 ml of LB broth supplemented with ampicillin. The restarted cultures were incubated at room temperature with shaking at 280 rpm until the O.D. at 600 nm was between 0.4-0.5. IPTG at a final concentration of 1 mM was then added to induce the expression of the mNAT- GST plasmid. The cultures were grown overnight at room temperature with shaking at 250 rpm. The culture was sonicated 5 X 15 seconds using a 50 % duty cycle and an output of 4 until the solution appeared less viscous and more yellowish. The tubes were then spun down at 3000 X g, 10 min at 4°C. The supernatants were removed, aliquoted to 1.5 ml Eppendorf tubes, and kept. The pellets were resuspended in 1 ml 1 X PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄ pH 7.3), aliquoted to 1.5 ml Eppendorf tubes and kept. If not used right away, the tubes were flash frozen with liquid nitrogen.

2.2.2.13.2. Bradford Protein Assay-Standard curve preparation

To be able to determine the amount of expressed protein, a standard curve was plotted with O.D. at 595 nm on the Y axis and with amount of protein in micrograms on the X axis. Varying concentrations of BSA dissolved in 1 X PBS acted as protein standards. Enough Bradford dye reagent to allow three measurements of each protein standard was diluted 1/5 into distilled water and mixed well. 20 µl of each protein amount standard was added to 980 µl of
diluted reagent in 1.5 ml Eppendorf tubes, vortexed and allowed to incubate at room temperature for 10 minutes. The O.D. at 595 nm was averaged for each of the BSA protein standards and plotted against amount of protein expressed in micrograms. A best fit straight line was then calculated and used as a standard curve.

2.2.2.14. Western blotting

To visualize the mNAT- GST fusion protein, determine its relative molecular size, and estimate its abundance relative to other cellular proteins, Western blotting was performed (Laemmli, 1970). The samples were boiled at 100° C for 5 minutes and loaded onto 10 % SDS PAGE gels. The gels were electrophoresed at 32-33 milliamps/gel until the dye ran to the bottom of each gel.

For Coomassie staining, the gels were removed from the electrophoresis apparatus, and soaked in Coomassie solution (45 % (v/v) methanol, 45 % (v/v) distilled water, 10 % (v/v) acetic acid, 2.5 % (w/v) Coomassie Brilliant Blue) for 1 hour with gentle rocking. The Coomassie solution was decanted, and the gels were destained (45 % (v/v) methanol, 45 % (v/v) distilled water, 10 % (v/v) acetic acid) for half an hour with gentle rocking followed by further destaining overnight. The protein bands were then visualized under visible light using the gel documentation system.

For immunoblotting, the gels were removed from the electrophoresis apparatus. The Hybond- C super nylon membrane was soaked in distilled water and then in transfer buffer (0.04 M Glycine, 0.05 M Tris-HCl, 1.3 mM SDS, 20 % (v/v) methanol) for 1 minute while the blotting papers were soaked only in transfer buffer. The nylon membrane was placed beneath the gel with blotting papers above and below. After transfer, the membrane was rinsed with distilled water 2-3 times to remove traces of the gel, and then stained for 10 minutes with Ponceau Red dye until the bands were visible. The location of the marker bands on the membrane were marked in pencil and the membrane was rinsed with distilled water and blocked overnight with milk- TNT solution (4 % (w/v) skim milk powder, 0.01 % (w/v) thimersol, 0.2 % (w/v) Tween 20, 10 mM Tris-HCl pH 8.0, 150 mM NaCl) with gentle shaking. The blocking solution was removed, the membrane rinsed with distilled water, and 20 ml of anti-GST antibody added (1: 4000 dilution). Prior to use, the anti-GST antibody sera was diluted 1: 2000 into 1 X TNT and incubated overnight with 4 mg/ml boiled and native XA 90 lysates to remove antibodies capable of recognizing endogenous E.coli proteins.
Cloning of the mNAT open reading frames from the mNAT pET3a constructs into the pGEX fusion protein expression vector. B = Bam HI, S = Sal I.
The antibody solution was decanted from the membrane the blot rinsed with 1 X TNT six times for 5 minutes each and a 1:20,000 dilution of secondary antibody was added (horseradish peroxidase labelled goat anti-rabbit IgG). After an hour, the solution was decanted and kept. The membrane was rinsed with 1 X TNT six times for 5 minutes each and an enhanced chemiluminescence (ECL) mixture was made by combining ECL reagent 1 and ECL reagent 2. The membranes were each soaked in ECL mixture for 1 minute, briefly dried and then placed in Saran wrap within a film cassette.

2.2.2.15. Solubilization of inclusion bodies

The method of Williams et al. (1995) was used for both the small scale and large scale solubilizations. For small scale solubilizations, a single bacterial colony was used to seed a 3 ml volume of LB broth supplemented with ampicillin. The culture was grown overnight, at 37° C, with shaking at 280 rpm. In the morning, a 12 ml volume of LB ampicillin was inoculated with 600 μl of the overnight culture. Expression of the mNAT-GST fusion protein was induced with 1.0 mM IPTG and grown for 3 hours at 37° C, 280 rpm with shaking. The culture was centrifuged at 3000 rpm, 4° C, 10 minutes. The pellet was resuspended in 1 ml of inclusion body sonication buffer (25 mM HEPES pH 7.7, 100 mM KCl, 12.5 mM MgCl₂, 20 % glycerol, 0.1 % (v/v) Nonidet P-40, 1 mM DTT). After sonication, 10 μl of 10 mg/ml lysozyme were added to the tube followed by a 30 minute incubation period on ice. Next, the suspension was placed at -70° C for 1 hour and then it was thawed rapidly in a water bath at room temperature, and cooled on ice. The suspension was sonicated 5 times 15 seconds at a 50 % duty cycle, and an output control of 4 until the solution appeared less viscous. The lysed, sonicated protein suspension was then centrifuged at 8000 X g for 10 min, 4° C to pellet the inclusion bodies. The supernatant was aspirated and the pellet was washed twice (with vortexing) with ice cold RIPA buffer (0.1 % (v/v) SDS, 1 % (v/v) Triton X-100, 1 % (w/v) sodium deoxycholate in TBS [25 mM Tris-HCl pH 7.5, 150 mM NaCl]). The tube was centrifuged after each wash and the supernatant removed. The pellet was dried and transferred to a 1.5 ml Eppendorf tube. The volume of the pellet was estimated and 4 volumes of 10 % SDS were added to the pellet. The pellet was solubilized by intermittently pipetting up and down, vortexing, and heating at 95° C for 1 hour.

The sample was then diluted with 9 volumes of 1 X PBS followed by overnight dialysis of the sample against a 100 fold volume of PBS containing 0.05 % SDS. The dialysis was performed at room temperature to prevent the SDS from crystallizing. Dialysis tubing with a molecular weight cutoff of 12-14,000 kDa was used. The sample containing the solubilized
inclusion bodies was then dialyzed again against a 100 fold volume of 1X PBS with 0.01 % SDS for four hours at room temperature. The concentration of protein in the solution was then determined as stated above.

For large scale solubilizations, single colonies were seeded into 10 ml of LB broth (supplemented with ampicillin). The culture was grown overnight at 37°C, with shaking at 280 rpm. In the morning, 3 ml of the overnight culture was used to inoculate 500 ml of LB broth (supplemented with ampicillin). The culture was incubated at 37°C, 150 rpm, until the O.D. at 600 nm (vis) reached 0.4-0.5. The culture was then induced with 1.0 mM IPTG and grown under the same conditions as above for 3 hours. 250 ml of culture were then aliquoted to individual centrifuge tubes, and centrifuged at 3000 X g for 10 min at 4°C to pellet the cells. Each pellet (corresponding to 250 ml of culture) was resuspended in 10 ml of inclusion body sonication buffer. 500 µl of 10 mg/ml lysozyme were added to each 10 ml of resuspended pellet and incubated on ice for 30 minutes. The suspension was then placed at -70°C for 1 hour followed by rapid thawing in a water bath at room temperature, and then placed on ice. The suspension was sonicated 15 times for 15 seconds at a 50 % duty cycle, and an output control of 4 until the solution appeared less viscous and more yellowish. The lysed, sonicated protein suspension was then centrifuged at 8000 X g for 10 min, 4°C to pellet the inclusion bodies. The supernatant was aspirated and the pellet was washed twice (with vortexing) with ice cold RIPA buffer. The tubes were centrifuged after each wash and the supernatants removed. The pellets were dried and transferred to 1.5 ml Eppendorf tubes. The volume of the pellet was estimated (did not exceed 500 µl) and 4 volumes of 10 % SDS were added to each pellet. The pellets were solubilized and dialysis was performed exactly as stated above.
PART THREE: RESULTS

3.1. Design and testing of Vector 1 positive control for PCR based screening of mouse NAT3 targeted knockout

To provide a rapid method of screening and identifying ES cells that have had the mNAT3 gene disrupted by the neomycin resistance gene cassette from Vector 1, a PCR based screening assay was designed. The assay relied on the PGK/PCR-probe/5' primer and the N3/PCR-probe/3' primer which amplify a 1.7 kb DNA fragment spanning the neomycin gene phosphoglycerate kinase (PGK) promoter (terminator) sequence, 480 bp of mNAT3 coding sequence and through into 1.32 kb of mNAT3 3' UTR (Figure 3.1). A positive control for this assay was needed to first test its viability, and as a means of ensuring that the assay would work when using it to routinely screen the targeted ES cells. The 1.7 kb DNA fragment was first subcloned into pBSK- as described in Materials and Methods. Bacterial colonies putatively containing the Vector 1 positive control plasmid were isolated and the plasmid extracted by the miniprep method. Restriction endonuclease digests to determine the presence of the 1.7 kb insert and its relative orientation were performed using Bam HI (Figure 3.1).

The PCR assay was first tested by amplifying 1 ng of the Vector 1 positive control (Figure 3.2). A band at 1.7 kb corresponding to the expected fragment size was seen. Serial dilutions down to 1 pg of the Vector 1 positive control template were used to test the sensitivity of the PCR assay. 100 and 200 ng of genomic DNA from untargeted ES cells were spiked with 1 pg of the Vector 1 positive control to mimic the biological environment in which the targeted mutation would have taken place (i.e. within the genomic DNA). The results (not shown) of these initial experiments were not consistently reproducible, which would preclude the use of this positive control in routine screening of Vector 1 targeted ES cells.

3.2. XL PCR amplification and sequencing of additional mNAT3 3' UTR

To design a Southern blotting screening method for ES cells targeted with the mNAT3 knockout vector (Vector 1), additional mNAT3 3' UTR sequence was obtained (Figure 3.3). With this additional information, a Southern probe could be produced and a suitable restriction enzyme site could be found. XL PCR on bacteriophage Lambda clones containing the mNAT3 open reading frame, using the T7 or T3 universal primers with the mNAT primer (which anneals to the
Normal gene

Targeted gene

PCR based screening strategy for detection of mNAT3 knockout.

A forward primer (a, PGK/Neo) anneals to the phosphoglycerate kinase terminator sequence of the Neomycin resistance gene (Neo). A reverse primer (b, mNAT3 PCR probe 3') anneals to a region just outside the 3' arm of homology. The result is amplification of a 1.7 kb PCR product indicative of a successful targeting event.
Figure 3.1.
Verification of the presence and orientation of the Vector 1 positive control insert. The upper diagram shows that digestion with Bam HI will yield a 1.7 kb piece (insert) and a 7.2 kb piece (dashed line, pPNT vector). The lower diagram shows the results of clones digested with Bam HI. Correctly oriented inserts (lanes 1,2,3). M, marker lane.
Figure 3.2.
Results of the amplification of 1 ng of the Vector 1 positive control. Lane 1, blank; lanes 2 and 3, correctly amplified PCR product 1.7 kb in size using PGK/Neo as the forward primer and N3/PCR-probe/3' as the reverse primer; lane 4, negative control (water); M, marker lane.
Figure 3.3.
Restriction map of the mouse NAT3 gene region. The area enclosed by the double arrows shows previously cloned and mapped sequence. The asterisks show the newly cloned and sequenced region (Figure adapted from Yu-Plant 1998).
start of the \textit{mNAT3} ORF), was used to obtain approximately 3 kb of additional sequence (Figure 3.4). The identity of the PCR amplified product was confirmed by digestion of the fragment with \textit{Kpn} I and comparison with previously mapped sequence. The 3 kb fragment was then subcloned into \textit{pCR} 2.1 as described in Materials and Methods. Clones were verified by digestion with \textit{EcoR} I and \textit{Kpn} I (Figure 3.5). The clone (\textit{Lambda} 23 3:1 more 3' UTR) containing the 3 kb fragment was then sequenced. Overlapping sequences were compiled into a full-length sequence using MacVector software.

Analysis of the 3 kb sequence revealed 1 kb of previously unmapped sequence. An 860 bp fragment within this new sequence was found that could be used as a probe in conjunction with \textit{Xba} I restriction endonuclease digestion using a new \textit{Xba} I restriction site in the newly sequenced 3' UTR. Digestion of wild type genomic DNA with \textit{Xba} I and probing would show a 3.1 kb piece whereas in the targeted event, a new \textit{Xba} I site 700 bp downstream of the wild type \textit{Xba} I would be created from the insertion of the \textit{Neo} gene. Genomic DNA of heterozygously targeted ES cells cut with \textit{Xba} I would reveal two bands: one wild type band at 3.1 kb and a new mutant band at 2.4 kb (Figure 3.6).

The putative probe sequence was matched against known sequences in GenBank to ensure the absence of repetitive sequences. PCR primers (5' \textit{mNAT3} probe and 3' \textit{mNAT3} probe) were synthesized allowing for convenient amplification of the probe. Isolation of genomic DNA from 129 Sw/J liver followed by cutting with \textit{Xba} I, electrophoresis on a 0.7 % agarose gel, Southern blotting and probing with the 860 bp probe ensured that the probe did not hybridize to any repetitive sequences.

3.3. Electroporation of ES cells with Vector 1 and selection

Following verification of the screening method, ES cells were electroporated with Vector 1 and underwent positive-negative selection to isolate targeted clones. Out of 5 X 10^6 ES cells electroporated, 290 colonies survived single selection with G418 alone (as counted on one plate only). 112 colonies in a total of 11 plates (about 10 colonies per plate) survived double selection with G418 and gancyclovir. The ratio of 290 G418 resistant colonies to 10 G418 and gancyclovir resistant colonies represents a roughly 30 fold enrichment factor. All colonies having survived double selection were then screened by Southern blotting for the targeted knockout of \textit{mNAT3}. Out of 112 colonies, 5 showed a correctly targeted allele (Figure 3.7).
Figure 3.4.
Results of the XL PCR amplification of additional *mNAT3* 3' UTR from Lambda clones containing the *mNAT3* open reading frame (using mNAT and T7 or T3 primers). Bands of different size were generated (lanes 1 through 4). The largest band at 3 kb in lane 4 was isolated and cloned into the pCR® 2.1 cloning vector. Lane 5, negative control (water).
Verification of the clones containing the 3 kb XL PCR product in pCR® 2.1 by digestion with *Kpn* I. The upper diagram a shows that digestion of a correctly oriented clone with *Kpn* I will yield a 480 bp fragment and a 5.4 kb fragment (dashed line). Diagram b shows digestion of an incorrectly oriented clone with *Kpn* I, yielding a 2.5 kb fragment and a 4.4 kb fragment (dashed line). Lane 1 in the lower diagram shows a 2.5 kb band that indicates the presence of the insert (incorrect orientation) and a 4 kb band (vector). Lane 3 shows a 5.4 kb band and a 480 bp band (correct orientation). Lanes 2 and 4, empty pCR® 2.1 vector.
Southern blotting screening of Vector 1 targeted ES cells.

Upper diagram a. shows the wild type mNat3 allele and upper diagram b. shows the targeted mNat3 allele. The Neo gene adds an Xba I site 700 bp downstream of the endogenous Xba I site. Digestion of the wild type DNA with Xba I and probing would show a 3.1 kb band whereas in heterozygous targeted DNA, under the same conditions a 3.1 kb and 2.4 kb band would be seen. The lower diagram shows the results expected from an autoradiogram of the wild type (lane 1) and heterozygously targeted alleles (lane 2) after Xba I digestion, Southern blotting and probing with the 860 bp probe.
Figure 3.7.
Representative autoradiogram of targeted ES cell genomic DNA isolated from Vector 1 electroporated ES cells. Digestion of the DNA with Xba I and probing with the 860 bp Southern probe shows a wild type allele at 3.1 kb and a targeted allele at 2.4 kb.
3.4. Design of the Vector 2 positive control for the XL PCR based screening method for vector 2 targeted knockout of mNAT1 and mNAT2

To provide a rapid method of screening and identifying ES cells that had had the mNAT1 and 2 genes disrupted and 9 kb of intervening sequence removed by the neomycin resistance gene cassette from Vector 2, an XL PCR based screening assay was used. The assay relied on the amplification of a 3.8 kb piece spanning the PGK terminator sequence of the neomycin resistance gene, ~ 3.7 kb of the Pst I- Sac I arm of homology, and ~ 90 bp of sequence outside the region of homology. However, because the region downstream of the 3' arm of homology was found to contain repetitive sequences, a reverse PCR primer designed to anneal in this region had to be carefully designed. 80 bp of DNA sequence within this area that were not part of the stretch of repetitive sequences were isolated. From those 80 bp, a 21 mer oligonucleotide (S/ E Reverse) having the same Tm as the Neo/PGK II primer was designed and synthesized.

The XL PCR assay required the amplification of a rather large piece of DNA (3.8 kb), and it was necessary to ensure that this size fragment could be successfully amplified. As such, a positive control was designed by subcloning a 5.0 kb Eco RI cut DNA fragment from the NAT2 3' UTR/ pCR 2.1 construct into an Eco RI cut empty pPNT plasmid as described in Materials and Methods. Bacterial colonies putatively containing the Vector 2 positive control plasmid were isolated and the plasmid extracted by the miniprep method. Plasmids were then digested with Eco RI to determine which had inserts, followed by digestion of those selected clones with Sac I to determine the orientation of the insert (Figure 3.8). The XL PCR assay on the positive control would amplify a 4.3 kb piece, 500 bp larger than the fragment expected from the true XL PCR assay (Figures 3.9a and b and Figure 3.10). The assay on the positive control was tested by adding 1 picogram of the positive control to 100, 200 and 300 ng of untargeted ES cell genomic DNA followed by XL PCR, and then agarose gel electrophoresis of the XL PCR reaction (Figure 3.11). All three lanes on the gel in this figure show the expected amplification of a 4.3 kb PCR product in 100, 200 and 300 ng of untargeted ES cell genomic DNA.

3.5. Design of Southern blotting screening method to verify the XL PCR results of the targeted knockout of mNAT1 and mNAT2

Because visualization of the XL PCR amplified product on an agarose gel is not always possible due to its extremely low abundance, Southern blotting of the gel followed by probing allows more definite confirmation of the nature of the XL PCR product produced. The probe
used was generated by digestion of the NAT2 3’ UTR/ pCR 2.1 construct with Hind III followed by isolation and gel purification of the 1 kb fragment. This 1 kb fragment probe lies within the amplified region but does not share any of the primer sequences used in the XL PCR reaction (Figure 3.9b).

3.6. Electroporation and selection of ES cells targeted with Vector 2

Out of 5 X 10^6 ES cells electroporated with Vector 2, 500 colonies survived single selection with G418 (as counted on one plate). 224 colonies in total (in 10 plates for an average of 22.4 colonies per plate) survived double selection with G418 and gancyclovir. The ratio of 500 G418 resistant colonies to 22.4 G418 and gancyclovir resistant colonies represents a 22 fold enrichment factor. Subsequent screening of the surviving clones by XL PCR and Southern blotting of the PCR product revealed 18 clones that showed a targeted allele. Figure 3.12 shows an autoradiogram of the XL PCR screening of a sample of 18 Vector 2 targeted clones. Only 3 out of these 18 display a hybridization signal at 3.8 kb (indicating a targeted allele).

From an overall cumulative total of 18 putatively positive clones, 4 which showed the strongest signal were selected for expansion. However, screening of the expanded clones failed to reproduce the 3.8 kb signal that indicates a targeted allele.

3.7. Design of Southern blotting screening assay for detection of targeted knockout of mNAT1 and mNAT2.

To design a DNA probe that could be used in Southern blotting screening for mNAT1 and mNAT2 targeted events, a 2.4 kb stretch of DNA sequence downstream of the Pst I- Sac I 3' arm of homology of Vector 2 was sequenced. The NAT2 3' UTR/ pCR 2.1 construct was digested with Sac I to release a 2.4 kb piece that was then subcloned into pBSK+. Presence of insert and orientation was determined by cutting the plasmids with Not I and Eco RI. Correctly oriented inserts were sequenced, and only 600 bp between an Sph I and Hpa I site in the 3' UTR of mNAT2 (Figure 3.13) did not have repetitive sequences as determined by matching with known sequences in GenBank using the BLAST sequence similarity search program. A Southern blot of genomic DNA cut with Bam HI and probed with the 600 bp probe revealed a single band at 23 kb (data not shown).
Figure 3.8.
Verification of the Vector 2 positive control clones by digestion with Sac I. The upper diagram shows a correctly oriented clone digested with Sac I yielding a 4.7 kb band, a 4.4 kb band, a 1.9 kb band, and a 1.4 kb band. Lanes 1 and 3 show the correct pattern (4.7 kb, 4.4 kb, 1.9 kb, 1.4 kb) indicating the presence of the insert in the correct orientation. Lane 2 shows an incorrectly oriented clone. M, marker lane.
Diagram a depicts the XL PCR amplification of the Vector 2 positive control yielding a 4.3 kb product. Diagram b depicts the XL PCR amplification of a correctly targeted allele and its result, a 3.8 kb product. The identity of the PCR product is verified by Southern blotting and probing with the 1.0 kb probe lying within the PCR product but not sharing any of the primer sequences.

1, Forward primer PGK/Neo 2; 2, Reverse primer S/E reverse; Neo, Neomycin resistance gene.
Figure 3.10.
Agarose gel electrophoresis of the XL PCR amplification of the Vector 2 positive control, mNAT1 and mNAT2 Vector 2 and a clone containing the Vector 2 positive control insert in the reverse orientation. Lane 2 shows the correctly amplified band at 4.3 kb, 500 bp larger than a true targeting event (3.8 kb). The size of the band is verified by DNA markers as well as digestions of DNA constructs yielding fragments of predetermined sizes (lanes 1 and 3). The mNAT1 and mNAT2 targeting Vector 2 does not show a visible band in lane 4. The clone containing the Vector 2 positive control insert in the reverse orientation does not show a band either (lane 5). M, marker lane.
Figure 3.11.
Agarose gel electrophoresis of the XL PCR amplification of 1 picogram of the Vector 2 positive control in 100, 200, and 300 ng of untargeted ES cell genomic DNA (lanes 1, 2 and 3 respectively). Lane 4, negative control (water). M, marker lane.
Figure 3.12.
Autoradiogram of the initial round of XL PCR screening of Vector 2 targeted ES cells. Leftmost brackets delineate signals from three putative positives. Rightmost brackets delineate signals from the 1 picogram positive control and 0.01 picogram positive control.
Figure 3.13.

Restriction map of the *mNAT1* and *mNAT2* loci and the location of the 600 bp *mNAT1* and 2 Southern probe, demarcated by the stars between the *Sph I* and *Hpa I* restriction sites (Figure adapted from Yu-Plant 1998)
Targeted knockout of mouse NAT1 and 2 creates a new Bam HI site derived from the Neo gene (Figure 3.14). Digestion of heterozygous targeted ES cell genomic DNA with Bam HI, Southern blotting and probing with the 600 bp probe would produce a 14 kb band representing the targeted allele, and a 23 kb band representing the wild type allele shown in Figure 3.14.

3.8. Electroporation of ES cells with Vector 2 and selection

Following electroporation of 5 X 10^6 ES cells (per plate) with Vector 2, 132 colonies survived selection with G418 alone (as counted in one plate). A total of 84 colonies from an overall total of 10 plates (which averages to 8.4 colonies per plate) survived double selection with G418 and gancyclovir. The ratio of 132 G418 resistant colonies to 8.4 G418 and gancyclovir resistant colonies represents about a 16 fold enrichment factor. Screening of the 84 surviving colonies with the Southern blotting method described above only displayed a 23 kb wild type allele without the 14 kb targeted allele.

3.9. Subcloning and sequencing of additional mouse NAT1 5' UTR

Bacterial Artificial Chromosome “pBAC” clones containing on average 80-130 kb of the 129 Sv/J mouse genomic DNA sequence were used to provide more 5' UTR corresponding to mouse NAT1. A 129 Sv/J pBAC mouse library was screened using mouse NAT1 and mouse NAT3 probes. Two clones, 256A7 and 277N15 were cut with Eco RI, Southern blotted and separately probed with mouse NAT1 and mouse NAT3 ORF probes. Autoradiograms revealed a signal at ~ 2kb with either probe which corresponded roughly to the expected pattern from mouse NAT3 (data not shown). BAC clone 256A7 was digested with Bam HI, Hind III, Kpn I, and Xba I restriction endonucleases. The cut DNA was electrophoresed on a 0.7 % agarose gel, Southern blotted, and simultaneously probed with the mouse NAT1 and mouse NAT3 ORF probes (Figure 3.15). Results were compared to the restriction maps of the mouse NATs. Digestion with Bam HI revealed two bands close together at 20- 23 kb that could correspond to mouse NAT1 and 2 (as previously documented from work on Vector 2 Southern screening) and a mouse NAT3. Hind III revealed an 8.9 kb band corresponding to mouse NAT2, a 6.5 kb band corresponding to mouse NAT1, and a 2.6 kb band corresponding to mouse NAT3. Digestion with Kpn I showed 4 bands out of 4 expected bands, but no size comparisons could be made with the present map as Kpn I mapping is incomplete. Digestion with Xba I revealed one 9 kb fragment corresponding to mouse NAT2, one 3 kb fragment corresponding to mouse NAT3, and one 8 kb
Figure 3.14.

Digestion of untargeted and Vector 2 targeted ES cell genomic DNA with *Bam* HI and probing with the 600 bp Southern probe. Upper diagram a. shows that *Bam* HI digestion of the wild type (untargeted) allele and probing would yield a 23 kb band. Upper diagram b. shows the configuration of the targeted allele. Digestion with *Bam* HI and probing with the 600 bp probe would yield a 14 kb band (a 9 kb reduction in size) because the Neo gene adds a new *Bam* HI site. This size difference corresponds to the amount of sequence replaced by the Neo gene. The lower diagram depicts the expected results from an autoradiogram of wild type (lane 1) and heterozygously targeted alleles (lane 2).
Figure 3.15.
Autoradiogram of a Southern blot of BAC clone 256A7 digested with restriction enzymes and probed with the mNAT1 and mNAT3 ORF probes.
Lane 1, digestion with Xba I yields 9.0 kb, 8.0 kb, 3.0 kb fragments. Lane 2, digestion with Kpn I yields four fragments (unmapped). Lane 3, digestion with Hind III yields 8.9 kb, 6.5 kb, 2.6 kb fragments. Lane 4, digestion with Bam HI yields 23 kb, 20 kb and 9.4 kb fragments.
fragment that could correspond to \textit{mNAT1}. Digestion of clone 277N15 with \textit{Hind} III showed a band at 5 kb which corresponds to the expected size determined from the \textit{mNAT3} restriction map (not shown).

3.10. Production of mNAT antigens

Mouse NAT protein coding regions were amplified by PCR, generating an 870 bp fragment (Figure 3.16). The identity of the PCR products corresponding to \textit{mNAT1}, \textit{mNAT2} and \textit{mNAT3} was verified by restriction enzyme digestion of each PCR product (Figure 3.17). The 870 bp fragment was subcloned into the pGEX vector as described in materials and methods. Following subcloning and transformation into DH5\(\alpha\) cells, colonies were picked, miniprepped and the isolated plasmid digested with \textit{Bam} HI. Plasmids which displayed an 870 bp shift were then digested with \textit{Sal} I (Figure 3.18). The clones which released an 870 bp fragment corresponding to the mouse NAT protein coding region were then sequenced. For mouse \textit{NAT1}, a nucleotide substitution at position 244 resulted in replacement of an adenine in Balb/c and C57BL/6 mice for a guanine in 129 Sv/J. This results in a change from methionine to valine at amino acid position 82. There were no differences between 129 Sv/J mouse \textit{NAT2} sequence and Balb/c and C57BL/6. For mouse \textit{NAT3}, a nucleotide substitution at position 25 resulted in replacement of an adenine in Balb/c and C57BL/6 mice for a guanine in 129 Sv/J. This results in a change from arginine to glycine at amino acid position 9.

3.11. mNAT- GST expression studies

Initial pilot experiments were carried out to determine the expression profiles of the mouse NAT- GST fusion proteins. Growth at room temperature in either LB or 2X YT media, with shaking at 280 rpm and a 1.0 mM IPTG concentration showed that the mNAT1- GST and mNAT3- GST fusion proteins were expressed mainly in the pellet fraction (most likely within inclusion bodies) with apparent absence in the supernatant (Figures 3.19 and 3.21 respectively). The mNAT2- GST fusion protein was expressed in both the supernatant and in the pellet (Figure 3.20). Lowering the IPTG concentration to 0.1 mM and 0.01 mM did not shift the expression of the mNAT1- GST fusion protein to the supernatant (Figure 3.22).

The localization of the mNAT- GST fusion protein expression as well as the identity of each of the proteins were detailed by performing immunoblots of the fusion proteins with
Figure 3.16.
Agarose gel electrophoresis of the PCR amplification of the 870 bp mNAT1 (lanes 1 to 3), 2 (lanes 4 to 6), and 3 (lanes 7 to 9) open reading frames from the mNAT 1, 2 and 3 pET3a constructs. Lane 10, negative control (water). M, marker lane.
Verification of the nature of the PCR products amplified from the mNAT 1, 2 and 3 pET3a constructs. The upper diagram shows the three mNATs and the positions of the restriction enzyme sites. Lower diagram: Lane 1, digestion of the PCR products from lanes 1 to 3 in figure 3.16 with Bst E II produce one band at ~530 and another at ~330 bp indicating mNAT1. In lane 2, digestion of the PCR products from lanes 4 to 6 in figure 3.16 with Pst I produce one band at ~670 and another at ~190 bp indicating mNAT2. In lane 3, digestion of the PCR products from lanes 7 to 9 in figure 3.16 with Kpn I produce one band at ~480 and another at ~390 bp indicating mNAT3. M, marker lane.

Agarose gel electrophoresis of Bam HI and Sal I digested clones containing the mNAT 1, 2 and 3 open reading frames subcloned into the pGEX fusion protein expression vector. An 870 bp band corresponding to the size of the mNAT open reading frame and an upper band at 4.9 kb (pGEX vector) are released. Lane 1, mNAT1; lane 2, mNAT2; lane 3, mNAT3. M, marker lane.
Figure 3.19.
Coomassie stained SDS-PAGE gel showing the results of initial studies on the expression of the mNAT1-GST protein. Lane 1, GST (IPTG induced) supernatant; lane 2, GST (IPTG induced) pellet; lane 3, mNAT1-GST (IPTG uninduced) supernatant; lane 4, mNAT1-GST (IPTG uninduced) pellet; lane 5, mNAT1-GST (IPTG induced) supernatant; lane 6, mNAT1-GST (IPTG induced) pellet. Note the band in lane 6 which corresponds to the expected size of the protein at 59 kDa.

Figure 3.20.
Coomassie stained SDS-PAGE gel showing the results of initial studies on the expression of the mNAT2-GST protein. Lane 1, GST (IPTG induced) supernatant; lane 2, GST (IPTG induced) pellet; lane 3, mNAT2-GST (IPTG uninduced) supernatant; lane 4, mNAT2-GST (IPTG uninduced) pellet; lane 5, mNAT2-GST (IPTG induced) supernatant; lane 6, mNAT2-GST (IPTG induced) pellet. Note the bands in lanes 5 and 6 which correspond to the expected size of the protein at 59 kDa.
Figure 3.21.
Coomassie stained SDS-PAGE gel showing the results of initial studies on the expression of the mNAT3-GST protein. Lane 1, GST (IPTG induced) supernatant; lane 2, GST (IPTG induced) pellet; lane 3, mNAT3-GST (IPTG uninduced) supernatant; lane 4, mNAT3-GST (IPTG uninduced) pellet; lane 5, mNAT3-GST (IPTG induced) supernatant; lane 6, mNAT3-GST(IPTG induced) pellet. Note the band in lane 6 which corresponds to the expected size of the protein at 59 kDa.
Figure 3.22.
Coomassie stained SDS-PAGE gel showing the results of attempts at shifting the expression of the mNAT 1- GST protein from the pellet to the supernatant by lowering the level of IPTG induction from 1.0 mM to 0.1 and 0.01 mM. Lane 1, GST (IPTG induced) supernatant; lane 2, GST (IPTG induced) pellet; lane 3, mNAT 1- GST (IPTG uninduced) supernatant; lane 4, mNAT 1- GST (IPTG uninduced) pellet; lane 5, mNAT 1- GST (0.1 mM IPTG induced) supernatant; lane 6, mNAT 1- GST (0.1 mM IPTG induced) pellet; lane 7, mNAT 1- GST (0.01 mM IPTG induced) supernatant; lane 8, mNAT 1- GST (0.01 mM IPTG induced) pellet.
probing by the anti- GST antibody (Figure 3.23). The upper and lower blots respectively show the expression of the 59 kDa mNAT1- GST and mNAT3- GST fusion proteins in the pellet. The middle blot shows the dual expression of the 59 kDa mNAT2- GST fusion protein both in the supernatant and in the pellet.

3.12. Isolation and solubilization of the mNAT1 and mNAT3 GST fusion proteins from inclusion bodies

Following the method of Willams et al. 1995, inclusion bodies containing the mNAT1 and mNAT3 fusion proteins were isolated and solubilized to release the fusion proteins (Figure 3.24). The top gel shows the solubilization and enrichment of the mNAT1- GST fusion protein. In lane 1, a thin band at 59 kDa (the fusion protein) is visible in the total protein lysate. Lane 5 shows the enrichment and solubilization of the mNAT1- GST fusion protein. The bottom gel shows the isolation and solubilization of the mNAT3- GST fusion protein. Lane 1 in the lower gel shows the total protein lysate. Lane 5 indicates enrichment and solubilization of the mNAT3- GST fusion protein.
Immunoblots of mNAT1, 2 and 3 GST fusion proteins probed with the anti-GST antibody. **Top blot:** lane 1, GST (IPTG induced) supernatant; lane 2, mNAT1- GST (IPTG uninduced) supernatant; lane 3, mNAT1- GST (IPTG uninduced) pellet; lane 4, mNAT1- GST (IPTG induced) supernatant; lane 5, mNAT1- GST (IPTG induced) pellet. Note the expression of GST in the supernatant at 29 kDa and the expression of the mNAT1- GST in the pellet. **Middle blot:** lane 1, GST (IPTG induced) supernatant; lane 2, mNAT2- GST (IPTG uninduced) supernatant; lane 3, mNAT2- GST (IPTG uninduced) pellet; lane 4, mNAT2- GST (IPTG induced) supernatant; lane 5, mNAT2- GST (IPTG induced) pellet. Note the expression of GST in the supernatant at 29 kDa and similar expression levels of mNAT2- GST in the supernatant and pellet. **Bottom blot:** lane 1, GST (IPTG induced) supernatant; lane 2, mNAT3- GST (IPTG uninduced) supernatant; lane 3, mNAT3- GST (IPTG uninduced) pellet; lane 4, mNAT3- GST (IPTG induced) supernatant; lane 5, mNAT3- GST (IPTG induced) pellet. Note the expression of GST in the supernatant at 29 kDa and the expression of the mNAT3- GST in the pellet.
Coomassie stained SDS-PAGE gel showing the isolation and solubilization of the mNAT1 and mNAT3-GST proteins from within inclusion bodies. **Top gel:** lane 1, total protein lysate from mNAT1-GST expression; lane 2, supernatant from the 1st spin; lane 3, supernatant after washing with RIPA; lane 4, pellet after washing with RIPA; lane 5, supernatant post dialysis; lane 6, pellet post dialysis. Note in lanes 5 and 6, the significantly higher enrichment of the mNAT1-GST protein in the supernatant compared to the pellet indicating that the protein was released from the inclusion bodies. **Bottom gel:** lane 1, total protein lysate from mNAT3-GST expression; lane 2, supernatant from the 1st spin; lane 3, supernatant after washing with RIPA; lane 4, supernatant after second washing with RIPA; lane 5, supernatant post dialysis; lane 6, pellet post dialysis. Note the significantly higher enrichment of the mNAT3-GST protein in the supernatant compared to the pellet indicating that the protein was released from the inclusion bodies.
PART FOUR: DISCUSSION

4.1. Overview

Epidemiological and animal studies have suggested that NATs may play an important role in aromatic amine drug induced toxicity and in carcinogenesis following exposure to aromatic amines. However because of inherent weaknesses in these studies and models, the goal of our research is to clarify the role of NATs in the modulation of aromatic amine induced toxicity and carcinogenesis by creating and characterizing novel NAT deficient mouse strains. As first steps towards this goal, we have electroporated ES cells with both Vectors 1 and 2 which knock out mouse NAT3 and mouse NAT1/2 respectively. We have developed both Southern and PCR screening methods to select ES cell clones that have been successfully targeted. In preparation for the characterization of these NAT deficient strains, we have produced mouse NAT antigens to be used in raising anti-mouse NAT antibodies. These antibodies will be used to verify the absence of mouse NATs in the knockout strains, to detail the tissue expression of mouse NATs in wild type mice, and to permit more accurate determination of mouse NAT activities in enzyme kinetic studies. The results gleaned from such studies will help to identify the importance of NATs in pharmacology and disease.

4.2. Mouse NAT3 targeted knockout

Cloning and sequencing of XL PCR amplified mNAT3 3' UTR revealed an extra 1 kb of new sequence that contained a previously unmapped Xba I restriction enzyme site. The identification of this site allowed the design of a Southern blotting screening procedure for the detection of the mouse NAT3 knockout. The Xba I site in the targeting vector's Neo gene introduces itself in the targeted locus 700 bp downstream of the endogenous Xba I site. When targeted ES cell genomic DNA is then cut with Xba I and probed, the targeted allele is 700 bp smaller than the wild type allele.

R1 ES cells were electroporated with Vector 1 to knockout the expression of mouse NAT3. Double selection resulted in a 30 fold enrichment factor which falls within the reported range. However, the level of enrichment achieved varies widely and may even vary from experiment to experiment under the same conditions. As such it is often difficult to accurately compare enrichment factors with those listed in the literature (Wurst and Joyner 1993, Hasty et
al. 1991a and b, Zhang et al. 1994, Snouwaert et al. 1992, Te Riele et al. 1992). Nonetheless, such comparisons can serve as rough guidelines. On autoradiograms of Southern blots containing \textit{Xba I} digested targeted ES cell genomic DNA, five clones showed a correctly targeted allele and equal band intensity indicating heterozygous targeting.

A PCR screening method was also developed using specific primers designed to anneal to the neomycin resistance gene and to a region located outside the 3' arm of homology amplifying a 1.7 kb fragment representing the targeted allele. A positive control for the Vector 1 targeted event was designed to verify that a fragment of this size could be reliably amplified using these primers and under similar conditions to those in the actual screening process. The assay was able detect a 1.7 kb product from starting concentrations as low as 1 picogram in a mixture containing up to 200 ng of genomic DNA. There was some difficulty in generating a consistent signal from this positive control which could preclude its use in ES cell clone screening. In addition, difficulties could arise from using this positive control because it generates a PCR fragment of the same size as the targeted allele. To avoid this pitfall, one could use a different downstream reverse primer for the positive control and for the ES cell clone screening while maintaining the same upstream forward primer. This approach would however introduce another confounding factor (the use of two different primers). This problem could be rectified by redesigning the Vector 1 positive control in the same manner as the Vector 2 positive control (section 4.3). A size difference between the positive control and ES cell clone DNA would be created by subcloning additional mouse \textit{NAT3} sequence (480 bp) upstream of the \textit{Kpn I} site located in the \textit{mNAT3} ORF. The upstream and downstream primers would be identical for the positive control and for the ES cell clones, and a size difference could be easily detected.

In general, the advantages of this PCR screening assay are that it can be used to rapidly screen a large number of targeted ES cell colonies that have survived double selection and narrow down the number of colonies that must be subsequently screened by Southern blotting. In addition, the assay also provides a means of independently confirming the results of the Southern blotting screening. One of the disadvantages of screening by PCR is that the assay does not reveal the zygosity of targeting and does not ensure homologous recombination has occurred throughout the targeted locus. As such, investigators tend to rely on the Southern blotting screening method as the "gold standard" for ensuring that complete targeting has taken place.
4.3. Mouse NAT1/2 targeted knockout

Due to the original design of the genomic library there were many difficulties in creating a Southern blotting strategy for screening the mouse NAT1 and 2 targeted knockout. There was no sequence information available upstream of the mouse NAT1 region that could be used to identify potential restriction enzyme sites and a suitable Southern blotting probe. As such, the focus fell on using the region downstream of the mouse NAT2 region.

Initial designs for a probe in the area just downstream of the 3’ arm of homology (Sac I- Eco RI 300 bp fragment) were unsatisfactory, since as the probe contained repetitive elements homologous to the 3’ UTR of the alpha-4 integrin subunit gene, as well as to the 5’ enhancer region of the immunoglobulin J chain gene (checked against documented sequences in GenBank using the BLAST program). This was simultaneously verified by Southern blotting of restriction endonuclease cleaved 129 Sv/J genomic DNA and probing. Further sequencing downstream would have been necessary, and to temporarily circumvent this time consuming procedure, a PCR screening method was implemented. An 80 bp stretch of DNA within the Sac I- Eco RI fragment that did not contain any repetitive motifs was isolated. A reverse PCR primer was designed within this 80 bp stretch that when used with the Neo primer, would amplify a 3.8 kb DNA fragment indicating a targeted allele.

A positive control was created to ensure that such a large fragment of DNA could be successfully amplified. A 500 bp size difference between the positive control and the actual targeted allele product was detected by size fractionation on an agarose gel, Southern blotting and probing. Although it was possible to visually detect a signal from the XL PCR amplified positive control from a starting concentration of 1 picogram, no PCR signal from targeted ES cells could be seen on an agarose gel.

R1 ES cells were electroporated with Vector 2 to knockout the expression of mouse NAT1 and NAT2. A 22 fold enrichment factor was achieved with double selection which falls within the range of reported enrichment (Wurst and Joyner 1993, Hasty et al. 1991a and b, Zhang et al. 1994, Snouwaert et al. 1992, Te Riele et al. 1992). Confirmation of mNAT1 and 2 knockout by expansion and rescreening of clones initially showing a targeted allele, was unsuccessful.

It is difficult to account for such an observation but there can be several hypothetical explanations for the loss of the targeted mutation. The first is that the original signal indicating
the presence of a targeted clone was an artefact or contamination from the positive control. Alternatively, Frohman and Martin (1990) describe a situation where a false positive could have been formed by a process known as polymerase halt-mediated linkage of primers (PHLOP). If during the extension of one of the strands the polymerase enzyme halts, it produces a truncated product which can then anneal (like a primer) to a complementary strand and then be extended to include the sequence of the other primer. This new product would be identical in every way to a true homologous recombinant event. As such, subsequent PCR amplification of this product would signal a false targeted event. Another explanation could be that the majority of the clones in a pool were untargeted but had survived selection. Within this pool, there was a smaller proportion of true targeted clones. Initial PCR screening produced a true signal but prolonged storage of the cells followed by prolonged passaging might have compromised the viability of these targeted clones. A last explanation could be that the targeting vector picked up the region just outside the 3' area of homology from the endogenous locus by incomplete strand exchange and then inserted itself in a non-homologous fashion elsewhere in the genome (Soriano et al. 1991). If in addition, the TK cassette was also mutated the ES cells would not express TK and would survive the gancyclovir selection (Hasty and Bradley 1993). Screening by PCR would produce a false positive signal of the correct size.

Before attempting a second round of ES cell electroporations with Vector 2, additional 3' UTR downstream of the 3' arm of homology was sequenced to develop a Southern blotting screening assay. Only a 600 bp DNA fragment contained between the Sph I and Hpa I sites in the 2.4 kb Sac I- Sac I region downstream of the mouse NAT2 ORF was non-repetitive and could be used as a probe. The only restriction endonuclease site that could be used with the probe was the Bam HI site lying downstream of the probe. Although we did not have sequence information upstream of the mouse NAT1 region, we hypothesized that a Bam HI site would certainly be found there. Since digestion of untargeted ES cell genomic DNA with Bam HI and probing with the 600 bp probe showed a 23 kb band. Importantly, the Neo gene contains a Bam HI site and replacement of the mNAT1 and mNAT2 ORFs and 9 kb of intervening sequence during targeting by the Neo gene would create a new Bam HI site ~ 9 kb downstream of the endogenous Bam HI site. This size difference corresponds to a 14 kb band representing the targeted allele.

132 colonies (counted in one plate) survived selection with G418 alone after the second electroporation of ES cells with Vector 2. Of those, 84 in total survived double selection which translates into a 16 fold enrichment factor. This level of enrichment again falls within reported ranges. Unfortunately, screening of the surviving clones by Southern blotting again did not reveal any targeted alleles (0 positives).
Failure to successfully target this locus likely means that Vector 2 will have to be reconstructed. Although a study by Hasty et al. (1991a) showed that a 472 bp short arm of homology was sufficient to produce a targeted knockout of the \textit{hppt} gene, and Vector 2 had a 600 bp short arm, targeting of the mouse \textit{NAT1} and 2 was not achieved. There can be three explanations for this. First, the construct used by Hasty et al. (1991a) shared \textasciitilde 6 kb of total homology with the endogenous locus, while Vector 2 only had \textasciitilde 4.3 kb of homology. Notably, successful homologous recombination occurred with Vector 1 which shared 7.5 kb of total homology with the endogenous locus. In his article, Hasty clearly delineates that one of the most important factors affecting targeting success is the total amount of homology between the targeting vector and the endogenous locus, not necessarily the size of the short arm of homology. Alternatively the construction of Vector 2 itself may have prevented successful targeting. The 3' arm of homology in Vector 2 had originally been cloned by XL PCR from Lambda clones containing the mouse \textit{NAT2} ORF. Changes in the original sequence could have been introduced during the PCR process. Because of these artificially introduced "polymorphisms", the success of homologous recombination could have been compromised. Alternatively, the arms of homology in Vector 2 did not incorporate a recombinational hot spot that would greatly increase the success of targeting (Sedivy and Joyner 1992). According to Ponticelli et al. (1992) these so-called hot spots are thought to be recognition sites for recombination-promoting enzymes which act at high frequency at or near the recombinational hot spot. Either way, all the explanations argue in favor of reconstruction of Vector 2.

\subsection*{4.4. Reconstruction of Vector 2}

To reconstruct vector 2, additional 5' UTR upstream of mouse \textit{NAT1} will have to be cloned and sequenced to generate a longer 5' arm of homology. To do this, a 129 \textit{Sw/J} BAC library was screened with the mouse \textit{NAT1} ORF probe to isolate clones containing mouse \textit{NAT1} and 2 ORFs as well as additional sequence upstream of \textit{NAT1}.

Multiple restriction enzyme digestion, Southern blotting and probing of BAC clone 256A7 with \textit{mNAT1} and \textit{mNAT3} probes revealed an extra 1 kb of previously unmapped sequence upstream of the \textit{mNAT1} ORF that could be used to generate a new 5' arm of homology, more diagnostic restriction sites, and a Southern probe for knockout screening. Efforts are now underway to subclone and sequence this novel region.
 Although all three mouse NATs have been cloned and the localization of mouse NAT1 and 2 revealed (Martell et al. 1992), the location of mouse NAT3 is still unknown. Results from the BAC clone work have shown that mouse NAT3 probably lies within ~135 kb of mouse NAT1 and 2.

4.5. Production of mouse NAT antigens

In preparation for the characterization of the NAT deficient mouse strains, mouse NAT antigens were produced to raise polyclonal anti-mouse NAT antibodies. These antibodies will be used to confirm the absence of mouse NAT expression in the NAT deficient strains. In addition, to more clearly understand the roles of mouse NATs in aromatic amine induced toxicity and carcinogenesis, it is necessary to know the tissue localization and expression patterns of the mouse NATs. With this knowledge in hand, comparisons can then be made between wild type mice and knockouts regarding differences in tissue specific histopathology following exposure to aromatic amine procarcinogens and a role for NATs in this scheme more clearly defined.

In the past investigators have either looked at enzyme activity with non-specific mouse NAT substrates as an indicator of mouse NAT expression, or only examined the presence of mouse NAT2 while neglecting mouse NAT1 and 3. In general then, there is a need for more extensive studies on the tissue selective expression of mouse NATs and as mentioned above, in addition to detailing the expression of mouse NATs, it will also be important for our studies, to confirm the absence of the mouse NATs. As such, preparatory work was undertaken to produce mouse NAT antigens in order to raise specific rabbit anti-mouse NAT polyclonal antibodies. While monoclonal antibodies have many advantages such as specificity of binding and homogeneity, and can be produced in unlimited quantities, production of polyclonal antibodies was preferred because while being well suited to answer our experimental questions, they are also cheaper and much easier to make than monoclonal antibodies (Harlow and Lane 1988).

In general two methods can be used to generate antigens: coupling an antigenic peptide to a carrier or using a bacterially expressed fusion protein containing the desired antigenic region coupled to a fusion tag. Generating anti-mouse NAT antibodies by using a peptide antigen coupled to a carrier is technically more challenging and the antibodies may not recognize the native antigen (Harlow and Lane 1988). Furthermore, this method is more expensive than bacterial expression of a fusion protein antigen. Therefore, we opted for the latter strategy using the entire mouse NAT ORF as an antigen to raise the antibodies to be followed by subtractive
absorption of the cross reacting antibodies to generate polyclonal mouse NAT isozyme specific antibodies. Due to the extensive amino acid identity between the three mouse NATs (68-82%) the subtractive absorption step may yield a low titer isozyme specific serum. We realize that this purification procedure will be laborious and time consuming, however the resulting antibodies will be highly isozyme specific.

Previous work by Yu-Plant (1998) involving heterologous expression of mouse NATs in the pET3a expression system in bacteria has shown that mouse NAT1 and mouse NAT3 proteins are expressed in the insoluble fraction (likely within inclusion bodies), while mouse NAT2 is expressed in the soluble fraction. To be able to accurately characterize mouse NAT1 and 3 in terms of kinetics and to generate antibodies, it is desirable to have expression of a soluble protein and also to maximize the level of expression of this soluble protein.

One popular method of producing solubilized proteins is by cloning the protein of interest into a fusion protein plasmid vector (pGEX) expressing glutathione-S-transferase (GST) followed by its expression in E. coli bacterial strains (Xu et al. 1996, Cachot et al. 1998, Basu et al. 1997, Komori et al. 1997, Pakdel et al. 1994, Sundaram et al. 1996, Su et al. 1996, Zhou et al. 1999, Fernando et al. 1999, Tas et al. 1997, Williams et al. 1995). The pGEX plasmid contains the lacI8 gene which confers tight control of the Tac promoter and thus GST expression. To achieve high levels of fusion protein expression, the Tac promoter is induced by addition of IPTG to the culture media. Because GST is water soluble, its expression is directed to the soluble fraction. Therefore insoluble proteins are cloned into the pGEX fusion protein vector to increase their solubility. Besides affecting solubility, the GST moiety serves as a handle allowing the purification of the GST fusion protein from the rest of the cellular proteins by affinity chromatography on Glutathione Sepharose 4B. In addition, the GST portion may then be cleaved from the fusion protein by specific proteases such as thrombin or Factor Xa, allowing the isolation of the protein of interest and its subsequent purification. The expressed GST protein is also small in size compared to other fusion protein systems and does not interfere with antibody-antigen interactions as seen with protein A fusions (Williams et al. 1995). Lastly, as opposed to conventional means of producing antigens, subcloning of the entire mouse NAT ORF into the pGEX vector followed by its expression in E. coli was simple and inexpensive. Some disadvantages of using the GST system are that some of the fusions may still be insoluble, precluding the use of affinity chromatography for purification. Some fusions may even disrupt the ability of GST to bind glutathione, further complicating the purification procedure (Williams et al. 1995).
Although the pGEX system is designed to allow expression of insoluble proteins in the soluble fraction, this does not always happen (Soler et al. 1995, Bettadapura et al. 1998, Yuan et al. 1997, Fernando et al. 1999, Sundaram et al. 1996, Pakdel et al. 1994, Komori et al. 1997). This inability of the GST moiety to effectively solubilize the cloned protein can be due to the aggregation of the incompletely folded fusion polypeptides (Grisshammer and Nagai 1995). These aggregated proteins are usually found in association with bacterially produced insoluble inclusion bodies. Inclusion bodies can contain ribosomes, RNA polymerase and DNA that co-precipitate with the nascent polypeptide chains (Hartley and Kane 1988). The nature of the inclusion bodies varies widely, with some being more easily soluble than others. Factors affecting inclusion body formation are the type of culture media used, the growth temperature, the type of expression vector used and/or the sequence of the gene to be expressed (Grisshammer and Nagai 1995). In addition, the use of strong promoters such as Tac can result in the formation of inclusion bodies as large amounts of protein accumulate rapidly in the bacterial cell and the hastily produced polypeptides readily aggregate to one another (Grisshammer and Nagai 1995).

However, the fusion proteins may be extracted and solubilized from the insoluble inclusion bodies by treatment with strong denaturants such as urea, guanidine-HCl, strong base, acetonitrile, or SDS and high heat (Grisshammer and Nagai 1995, Williams et al. 1995). The binding of the GST to the Glutathione- Sepharose 4B relies on a correctly folded and active GST moiety (Grisshammer and Nagai 1995), and under the strongly denaturing conditions used to solubilize the fusion proteins from inclusion bodies, the GST folding and activity would be compromised. As such, after solubilization of the fusion proteins, a process of dialysis and renaturation is needed to restore GST conformation and activity. Refolding and renaturation of proteins is a difficult and not always successful task with the major difficulty being to avoid re-aggregation of the proteins during the renaturation process (Grisshammer and Nagai 1995).

4.6. Initial pilot expression studies of the mouse NAT- GST fusion proteins

Initial small-scale expression studies on the mouse NAT1 and 3 indicated that expression was slightly better in LB medium than in 2 X YT. Importantly, in either medium, the expression of the fusion proteins was limited to the insoluble fraction as seen by Coomassie staining and immunoblots. This situation differed from the soluble expression of mouse NAT2. These variations in expression profiles are identical to the differences seen in the expression of the mouse NATs in the pET3a expression system (Yu-Plant 1998) and in the pET5a expression system (Kelly and Sim 1994). Since the expression conditions were identical for all three mouse
NATs within their respective experiments, these results suggest that some of the differences in amino acid sequence between mouse NAT1/3 and 2 determine the solubility profile of the protein, rather than the strength of the promoter.

Examination of the immunoblots of the mouse NAT- GST fusion proteins probed with the anti-GST antibody revealed the expression profile of the mouse NATs. Expression of mouse NAT1- GST and mouse NAT3- GST was limited to the insoluble pellet fraction, while there was greater expression of mouse NAT2 in the soluble fraction, with less expression in the insoluble pellet. It is interesting to note that the anti-GST antibody picks up three other bands below the mouse NAT1- GST fusion protein and 2-3 other bands below the mouse NAT3- GST fusion protein. It is tempting to speculate that these extra bands may be proteolytically degraded fusion proteins or prematurely terminated and/or improperly folded fusions. In addition, there is a band at or near the same level as the GST protein in the mouse NAT3- GST blot which could perhaps indicate cleavage of the GST portion of the fusion from the mouse NAT protein.

To try to shift the expression of mouse NAT1 and 3 from the insoluble to the soluble fraction both the incubation temperature and the IPTG concentration were varied as suggested by GST gene fusion system manual (PharmaciaBiotech). However, neither condition had the desired effect. As such, a reverse approach was used in which expression of the inclusion bodies would be maximized followed by the solubilization of the fusion protein from the pelleted fraction.

4.7. Solubilization of the mouse NAT1 and 3 GST fusion proteins

In order to allow the binding of the mouse NAT1 and 3 fusions to the glutathione Sepharose 4 B matrix for affinity purification, it was necessary to solubilize the mouse NATs from the inclusion bodies, and the method of Williams et al. (1995) was used. However, there are many other solubilization protocols available. Solubilization with urea is a very commonly used method (Komori et al. 1997, Pakdel et al. 1994, Fernando et al. 1999, Yuan et al. 1997, Bettadapura et al. 1998, Grisshammer and Nagai 1995). Another method is the use of laurysarcosine (Sarkosyl) which is reported to be a good detergent for non-denaturing solubilization of inclusion body proteins (Frankel et al. 1991). However, Sarkosyl may prevent the binding of the GST moiety to the glutathione affinity matrix (Soler et al. 1995), and has had limited success with solubilization of human NAT- GST fusion proteins in our lab. Other methods of solubilization include Zwittergen (palmityl sulfobetaine) which does not interfere with binding to one type of glutathione (GSH) affinity matrix (Soler et al. 1995). Guanidine-HCl,
strong base and acetonitrile are also commonly used (Grisshammer and Nagai 1995). Williams et al. (1995) report the use of 10% sodium-dodecylsulfate (SDS) and high heat as being superior to using urea or Guanidine-HCl to generally solubilize inclusion bodies, and as one of the most reliable methods of solubilizing GST fusion proteins for use in antibody production. Furthermore, Williams reports that unlike proteins solubilized in urea or Guanidine-HCl, SDS solubilized proteins tend to remain in solution after removal of the solubilizing agent by dialysis. After solubilization with 10% SDS and heat, the solution is then dialyzed step wise to 0.05 % and then to 0.01 % SDS. Interestingly, Smith and Johnson (1988) report that SDS concentrations below 0.03 % do not significantly affect binding of the GST moiety to its affinity matrix.

Small scale studies were carried out to solubilize the mouse NAT1 and 3 fusion proteins from the inclusion bodies. Results showed that the inclusion bodies were successfully solubilized, with almost 100 % transfer of the fusion protein from the insoluble to the soluble fraction. However, the now soluble fusion protein fraction still contained many other contaminating proteins, and the next step was to purify the fusion protein from the rest of the cellular proteins by adsorption onto the Glutathione Sepharose 4B beads. Unfortunately, binding of the fusion to the beads was very poor and most of the protein remained unbound (data not shown). The apparent inability of the GST moiety to bind the matrix can have several explanations, although all of them involve the loss of GST activity. The first explanation is that the sonication itself or the heat and oxidation generated may have contributed to the loss of GST activity (Grisshammer and Nagai 1995). Alternatively, the SDS and high heat used in the solubilization permanently denatured the GST moiety making it impossible for the GST to bind the matrix. Even with the extensive dialysis, proper refolding and renaturation of the fusion protein may not have been achieved. Lastly, the fusion partner may have altered the conformation of the GST moiety, preventing binding to the affinity matrix (Pharmacia Biotech GST manual).

Cleavage of the GST moiety from the fusion protein with thrombin to release the native mouse NAT protein was attempted but with no success. The inability of thrombin to cleave the peptide between the mouse NAT protein and the GST partner could have been caused by improper refolding of the fusion protein and/ or by the compromising of the GST protein structure by the mouse NAT protein fusion partner. Nevertheless, the solubilization protocol was successful in solubilizing the mouse NAT1 and 3 GST fusions.

Although it would have been preferable to use the native mouse NAT proteins as antigens to avoid producing anti- GST antibodies as well, we decided that the antibodies raised against GST could be removed at a later stage. Several studies have used the full GST fusion protein to
generate polyclonal antibodies against the subcloned protein (Pakdel et al. 1994, Komori et al. 1997, Zhou et al. 1999, Su et al. 1996, Tas et al. 1997, Basu et al. 1997). Subsequent purification of the antiserum can either be performed by filtration of the antiserum through an affinity matrix coated with \textit{E. coli} whole cell lysates and GST to remove non-\textit{mNAT} specific antibodies and antibodies raised against GST (Williams et al. 1995). Alternatively, an affinity matrix coated with the fusion protein made by linking the GST part to the matrix with only the subcloned protein moiety available for antibody binding, could also be used. An easier method would involve preadsorption of the antiserum with Western blots containing GST alone to remove the anti-GST antibodies, while incubation with whole cell lysates would remove the remaining non \textit{mNAT} specific antibodies.

Since the entire mouse NAT open reading frames were used for each fusion protein and all three mouse NAT proteins likely share many similar antigenic sites, the mouse NAT1, 2 and 3 antibodies would exhibit some level of cross-reactivity. This would seem at first to preclude using them to specifically detect each mouse NAT isozyme. However, the purity and specificity of each antiserum for its parent protein can be optimized by removal of the cross reacting antibodies by preincubation of the specific antisera with their fusion protein counterparts. For example, the mouse NAT1 antiserum can be incubated with the mouse NAT2- GST fusion protein and the mouse NAT3- GST fusion protein. The remaining antiserum would only contain anti-mNAT1 antibodies.

The antibodies will be used to confirm the absence of mouse NAT expression in the knockouts by probing to tissue cytosols that have been subjected to denaturing environments such as SDS- PAGE. They will also be extensively used in the quantification of the amount of immunoreactive mouse NAT isozyme in kinetic assays of mouse NAT activity and ultimately in the determination of the activities of each specific mouse NAT. This information when also compared with the NAT deficient strains can for example be used to identify the role of NATs in local tissue bioactivation of aromatic amine procarcinogens. Lastly, these antibodies can be used in detailing the tissue expression patterns of the mouse NAT isozymes in wild type mice through immunohistochemistry. The results from these latter studies in association with the enzyme kinetic data will help elucidate the role of NATs in aromatic amine induced toxicity and carcinogenesis.

Because the anti-mouse NAT antibodies will be generated by injection of the denatured antigen into rabbits it is likely that many of the antibodies raised will be specific for epitopes that were only unmasked as a consequence of denaturation, while another proportion of the
antibodies will recognize epitopes that were resistant to denaturation and presented in the physiological conformation of the enzyme. As such, caution should be exercised in using these antibodies in co-immunoprecipitation experiments or in inhibition studies which require the recognition of antigenic sites presented in the physiological conformation of the enzyme.

4.8. Summary

The N-acetyltransferases are thought to play an important role in xenobiotic biotransformation by detoxifying and bioactivating aromatic amine containing compounds found in clinically useful drugs, foods, and environmental pollutants. Epidemiological and animal studies have pointed to genetic polymorphisms in the N-acetyltransferase genes as factors influencing acetylation capacity, a characteristic thought to be important in aromatic amine induced toxicity and carcinogenesis.

Epidemiological studies, although implicating the N-acetyltransferases in toxicity and cancer, are limited by difficulties in accurately assessing the types, levels and lengths of exposure of individuals to aromatic amine containing compounds. In addition, there are differences in other drug metabolizing enzymes amongst populations tested that are not always controlled for, as well as in other, as yet unidentified modifier genes which may differentially influence different populations to toxicity and cancer.

Data from natural mouse models of the acetylation polymorphism (actually an interstrain difference) and congeneric models have shown that differences in acetylation capacity may affect susceptibility to aromatic amine induced carcinogenesis. However, these studies are hampered by the fact that the NAT2 defect still permits low levels of acetylation and that the mouse NAT1 and NAT3 proteins may compensate for this because of their overlapping substrate selectivities. Furthermore, these studies have also shown the effects of unidentified modifier genes and gender which influence the disease phenotype. The role of NATs may also be difficult to assess with congeneric strains because these strains may contain very small chromosomal segments from the donor strain flanking the donor gene of interest (the NAT gene) that may play a role in modulating acetylation capacity.

No studies have yet detailed the isozyme specific tissue expression of mouse NATs. Such data is crucial in gaining a better understanding of the role of NATs in toxicity and
carcinogenesis, especially with regard to tissue specific damage caused by suspected aromatic amine procarcinogens.

To remedy some of these problems, the goal of this research is to create mouse strains deficient in N-acetyltransferase expression. These NAT deficient strains will help determine if the NATs play an important role in toxicity and disease and what that role is. In addition, these strains will help determine if there are any endogenous roles for the NATs and if they have any function in embryological development.

We have undertaken important methodological steps to address our research goals. The mouse NAT gene targeting vectors were introduced into ES cells and heterozygously targeted mNAT3 clones were selected. We have designed and used both Southern blotting and PCR assays to detect the presence of NAT gene targeting. With the aid of newly cloned untranslated sequences upstream of mouse NAT1, a second targeting vector knocking out mouse NAT1 and 2 is currently being developed. The production of mouse NAT antigens will now allow polyclonal anti-NAT antibodies to be raised in rabbits and used in the characterization of the NAT deficient mouse strains. Furthermore, these antibodies will also serve to detail the tissue localization and activities of each of the mouse NATs to gain deeper understanding of their roles in drug biotransformation and disease.

4.9. Future studies and conclusion

In the near future mNAT3 targeted ES cell clones will be used in morula aggregations to regenerate functional blastocysts and will be implanted into a pseudopregnant mouse host stain. Heterozygous chimeric pups will be intercrossed to produce homozygous mNAT3 knockout mice. With the reconstruction of Vector 2, heterozygous mouse NAT1/2 knockouts will be produced, and will be bred to homozygosity. These two homozygous knockout strains will then be bred to generate entirely NAT deficient mouse strains. These strains will be compared to wild type strains for failure to thrive and any physiological changes. Absence of mouse NAT mRNA and protein expression will be confirmed using mouse NAT specific riboprobes and the anti-mouse NAT antibodies. We also plan to examine the in vivo pharmacokinetics of aromatic amines and their metabolites in knockouts and compare the results to wild type strains.

To more clearly understand the role of NATs in arylamine bioactivation, tissue cytosols from NAT deficient strains will be compared to wild types in tests to measure the rates of carcinogen-DNA adduct formation in different tissues. Knockouts and wild type mice will also
be used in *in vivo* arylamine sensitivity tests using the $^{32}$P-postlabelling technique. After administration of known procarcinogens to knockouts and wild type mice, organs will be removed, tissues will be collected and DNA adduct levels and patterns compared. We will also compare the knockouts and the wild type mice in terms of tissue histopathology after exposure to aromatic amine procarcinogens. With detailed knowledge of the tissue specific NAT expression patterns we will then be in a better position to understand the role of NATs in cancer. Lastly, these strains will be particularly well suited to long-term carcinogenicity studies. These studies are particularly important as cancer is not only a multi-variable disease but also one which progresses over an individual's lifetime. Administration of procarcinogens to wild type mice and NAT deficient strains over prolonged periods will be followed by detailed histological comparisons to determine the role of NATs in cancer.

Overall, the engineering of mouse strains lacking the N-acetyltransferase genes will help to further clarify the role of these genes in drug induced toxicity and carcinogenesis. It is also hoped that these mouse strains will help in the development and testing of clinically useful drugs in the future.
BIBLIOGRAPHY


Chung JG, Levy GN and Weber WW. Distribution of 2-Aminofluorene and p-

Coroneos E, Gordon JW, Kelley SL, Wang PD, Sim E. Drug metabolising N-


Yu Plant V. *Creation and characterization of N-acetyltransferase deficient strains*. University of Toronto 1998.


