0026-895X/91/020218-10$03.00/0

MOLECULAR PHARMACOLOGY, 40:218-227

Molecular Genetic Basis of Rapid and Slow Acetylation in Mice

KAREN J. MARTELL, KOSTAS P. VATSIS, and WENDELL W. WEBER

Department of Pharmacology, The University of Michigan Medical School, Ann Arbor, Michigan 48109-0626

Received February 18, 1991; Accepted May 13, 1991

SUMMARY

The molecular genetic basis of *N*-acetylation polymorphism has been investigated in inbred mouse models of the human acetylation polymorphism. Two genomic clones, *Nat1* and *Nat2*, were isolated from a C57BL/6J (B6) mouse (rapid acetylator) genomic library. The *Nat1* and *Nat2* genes both have intronless coding regions of 870 nucleotides and display >47% deduced amino acid similarity with human, rabbit, and chicken *N*-acetyltransferases. Amplification of *Nat1* and *Nat2* from A/J (A) mouse (slow acetylator) genomic DNA by the polymerase chain reaction and subsequent sequencing revealed that *Nat1* was identical in B6 and A mice, whereas *Nat2* contained a single nucleotide change from adenine in B6 to thymine in A mice. This nucleotide substitution changes the deduced amino acid at position 99 from asparagine in B6 to isoleucine in A mice. Hydropathy analysis revealed that this amino acid change alters the hydropathy of the flanking peptide segment in NAT2 from hydrophilic in the B6 mouse to hydrophobic in the A mouse. The amino acid change occurs in a region of the gene where no polymorphism has yet been reported in human or rabbit *NAT2* and may represent an important structural domain for *N*-acetyltransferase activity. *Nat1* and *Nat2* have the same 5′ to 3′ orientation in the B6 mouse; the two genes are separated by approximately 9 kilobases, with *Nat1* located 5′ of *Nat2*.

Hepatic NAT (EC 2.3.1.5) metabolizes a wide spectrum of amine and hydrazine drugs and arylamine carcinogens (reviewed in Ref. 1). This activity varies severalfold in humans, and the difference in drug-acetylating capacity is transmitted as a single gene trait (2, 3), which has come to be known as the acetylation polymorphism. Individuals can be classified phenotypically, on the basis of this polymorphism, as rapid or slow acetylators. Differences in acetylator status are implicated in idiosyncratic reactions to amine and hydrazine drugs such as isoniazid, procainamide, hydralazine, and sulfasalazine (4) and may also modulate the susceptibility of individuals to arylamine-induced carcinogenesis after occupational exposure to benzidine and β-napthylamine (5). Thus, the human acetylation polymorphism is an important factor in drug therapy and occupational exposure to hazardous environmental chemicals.

Numerous reports have appeared during the past 25 years that present biochemical evidence for more than one hepatic NAT in humans (6, 7) and in several animal models for the human polymorphism (8–10). Our laboratory has investigated acetylation polymorphism in rabbits (11), hamsters (12, 13), and mice (14–16), which have proven valuable in biochemical and toxicological investigations (17). For the past 15 years, we have concentrated on developing inbred mouse models for the human NAT polymorphism and determining the molecular basis of NAT polymorphism in mice (16, 18). We have demonstrated that the hepatic acetylation polymorphism in A (slow acetylator) and B6 (rapid acetylator) inbred mouse strains is determined by the *Nat*[1] locus on mouse chromosome 8 (19). This polymorphism is due, as in humans, to a difference in the activity of a single hepatic NAT (16). The slow and rapid hepatic NATs of A and B6 mice have very similar biochemical and physicochemical properties, but they do appear to differ qualitatively, because they are differentially inhibited by dimethylsulfoxide and exhibit K_m differences for the arylamine carcinogen 2-aminofluorene (16, 20). We have concluded that the rapid and slow mouse NATs are isoforms with minimal structural differences (16). In addition, we have recently reported biochemical evidence for the existence of a second hepatic NAT in B6 and A mice that did not exhibit an activity difference between the two strains (14, 21).

Recently, recombinant DNA studies of humans and rabbits have provided new perspectives on the drug-metabolizing acetylator genes that may contribute to an improved understanding of the variation in human acetylating capacity (22, 23). In this report, the first on the molecular genetic basis for acetylation polymorphism in B6 and A mice, we describe two Nat genes, *Nat1* and *Nat2*, that were obtained by library screening and PCR amplification. The relative orientation, the proximity,

This work was partially supported by United States Public Health Service Grants GM 27028, GM 44965, 5T32 GM 07544, and M01 RR 00042.

[1]*Nat* refers to the gene encoding mouse NAT protein, according to the convention established in Ref. 19. According to current usage, in contrast to that for the mouse, NAT genes of the human and rabbit are both referred to as *NAT* (38, 39).

ABBREVIATIONS: NAT, acetyl coenzyme A:arylamine *N*-acetyltransferase; A, A/J; B6, C57BL/6J; PCR, polymerase chain reaction; kb, kilobase(s); RFLP, restriction fragment length polymorphism; SDS, sodium dodecyl sulfate; SSC, standard saline citrate; *Taq*, *Thermus aquaticus*.

the nucleotide sequences, and the deduced amino acid sequences of these two genes are presented.

Materials and Methods

Animals. Inbred B6 and A mice were purchased from The Jackson Laboratory (Bar Harbor, ME). Mice were housed two to six/standard shoebox cage and were maintained at 25° on a 12-hr light/dark cycle. Animals were given water and Purina Mouse Chow 9F 5020 *ad libitum*. Mature mice of both sexes were used.

Materials. Restriction endonucleases and other DNA-modifying enzymes were purchased from Boehringer Mannheim Biochemicals (Indianapolis, IN) or BRL (Bethesda, MD). All chemicals and buffers were molecular biology grade and were obtained from Sigma (St. Louis, MO) or IBI (New Haven, CT). Nitrocellulose paper, [γ-^{32}P]ATP (6000 Ci/mmol), and [α-^{32}P]dCTP (3000 Ci/mmol) were purchased from Amersham (Arlington Heights, IL). A Sequenase Version II kit was purchased from USB (Cleveland, OH). Plasmids and phage DNA were obtained from the following sources: pGEM vectors were from Promega (Madison, WI), λ DASH II was from Stratagene (La Jolla, CA), and M13mp18 and M13mp19 were from Boehringer Mannheim Biochemicals. Positive displacement pipets and tips used for PCR experiments were purchased from Rainin (Woburn, MA). The oligonucleotide probes used for PCR amplification, sequencing, and mapping analysis were synthesized by the University of Michigan DNA Synthesis Facility.

PCR amplification protocols. The following protocols were used to amplify the partial mouse NAT clone B6Nat-1, the human NAT2 gene, herein designated HNat-1, and the A mouse Nat1 and Nat2 genes. DNA amplifications were carried out by the method of Saiki *et al.* (24), with minor modifications. Genomic DNA for PCR was isolated from liver using a phenol-chloroform extraction procedure (25), as modified in our laboratory (26). Amplification primers were constructed with approximately 50% G/C content and with similar T_m values, according to established guidelines (27). Primer pairs were designed so that they would not hybridize to each other. Amplification reaction mixtures (100 μl) contained 2 μg of genomic DNA, 1 μg of each amplification primer, 250 μM each of dATP, dCTP, dGTP, and dTTP, 10 mM Tris·HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 3 mM dithiothreitol, and 5 units of *Taq* DNA polymerase (Beckman); mixtures were layered with 100 μl of light mineral oil. Procedures to avoid false positives were used (28). The reaction mixtures were amplified for 30 cycles, with 1.5 min at 94° for denaturation, 1 min at 55° for annealing, and 3 min at 72° for extension; one final extension step for 10 min at 72° was also carried out. Amplification reactions were carried out in an Ericomp TwinBlock thermal cycler (San Diego, California). PCR products were visualized after separation on a 1% agarose gel, in 1× TAE (40 mM Tris-acetate buffer, pH 8.0, 1 mM EDTA), and staining with ethidium bromide.

Construction of pB6Nat-1 and the 3′-specific probe (pHNat-1). Two probes were generated by PCR from B6 mouse and human DNA and were designated B6Nat-1 and HNat-1, respectively. B6Nat-1 corresponds to the 5′ end of the coding region of the mouse Nat1 gene, and HNat-1 corresponds to the human NAT2 gene, as reported (29). The 3′-specific probe was created from the 3′ part of the coding region of HNat-1. B6Nat-1 and the 3′-specific probe, which hybridized to nonoverlapping regions of the genes, were used for library screening, Southern blot analysis, and mapping of restriction sites, as indicated in Results.

The construction of pB6Nat-1, reported previously (30), was performed as follows: two amplification primers corresponding to the rabbit NAT cDNA, rnat (31), were used in conjunction with B6 genomic DNA to amplify the target sequence by PCR. The sequence of the 5′ amplification primer was GC/GGATCC/ATGGACATTGAAGCATATT (GC/linker/amplification primer). This primer corresponded to nucleotides 1–19 (where A of the initiating ATG codon is nucleotide 1) and contained a *Bam*HI linker. The 3′ amplification primer corresponded to nucleotides 409–435 and contained *Eco*RI linkers followed by the bases G and C (GC/GAATTC/TTGAGGCTGATCTTTCCCAGAAATTAA). The G and C bases in both primers were inserted to aid in endonuclease digestion (27). For amplification, the annealing temperature was lowered to 50° to allow sufficient annealing of the nonhomologous sequences. The PCR product was subcloned into M13mp18 and M13mp19, and 10 clones were sequenced. A single M13mp18 clone was determined to have an errorless sequence by consensus of all the sequences. The sequence corresponds to nucleotides 1–435 of *Nat1* (Fig. 2). The M13mp18 insert was subsequently ligated into the *Bam*HI and *Eco*RI sites of pGEM3zf(+) (Promega) to make the recombinant plasmid pB6Nat-1.

The 3′-specific probe spanning from nucleotide 480 (*Kpn*I site) to nucleotide 856 (*Bam*HI site) was prepared from the sequence of the human *NAT2* gene (29). The 1.9-kb segment of human *NAT2* from nucleotide 721 upstream of the initiating methionine to nucleotide 1170 was amplified with primers corresponding to the terminal 5′ and 3′ ends of the reported sequence. The 5′ primer (GC/AAGCTT/GAATTCAGTGCTCTCCCTGTGCACCCAC) consisted of G and C bases and a *Hin*dIII linker and spanned from nucleotide 721 to nucleotide 694 upstream from the initiating methionine. The 3′ primer (GC/GAGCTC/GGAATTCAACAATAAATCAATATTAAAT) consisted of the G and C bases and an *Sst*I linker and was complementary to nucleotides 1143–1170. The PCR product was subcloned into M13 vectors. Of the 10 M13 sequences analyzed, a single M13mp19 clone contained a correct sequence, in comparison with that reported in the literature (29). The 1.9-kb insert from this clone was subcloned into pGEM7zf(−) to make the recombinant plasmid pHNat-1. The 3′-specific probe was prepared by digesting pHNat-1 with *Kpn*I and *Bam*HI and then purifying the resulting fragments from agarose gel using the Geneclean kit from Bio 101 (La Jolla, CA).

Construction and screening of the B6 genomic library. A B6 genomic library was constructed in the cloning vector λ DASH II (Stratagene). Genomic DNA was isolated from liver as described above (see PCR amplification protocols). The DNA was partially digested with *Sau*3AI and size-fractionated by centrifugation on a continuous sucrose density gradient (10–40%). DNA fragments of 15–20 kb were ligated into λ DASH II arms, which had been sequentially digested with *Bam*HI and *Hin*dIII to inactivate the stuffer fragment, and were then separated from the stuffer fragment on a sucrose density gradient (10–40%). The DNA was packaged *in vitro* with Gigapack Gold (Stratagene) and plated on the restrictive host P2(PLK-17). The number of independent recombinants before amplification was determined to be 7.5×10^5. The library was screened with the insert of pB6Nat1, which was ^{32}P-labeled with a multiprime labeling system (Amersham) to a specific activity of 1.5×10^9 cpm/μg of DNA. Nitrocellulose filters were prehybridized at 37° for 4 hr in a solution containing 5× SSC (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.4), 5× Denhardt's (1× Denhardts = 0.02% bovine serum albumin, 0.02% polypyrrolidine), 50% formamide, 0.5% sarcosyl, 1 mM EDTA, 50 mM sodium phosphate buffer, pH 6.5, and 50 μg/ml yeast tRNA. Probe was added and hybridized at 37° for 18 hr. Filters were washed sequentially with 2× SSC, 0.5% SDS, for 5, 10, and 15 min at room temperature and then washed under stringent conditions at 65° in 2× SSC, 0.1% SDS, for 15 and 30 min.

DNA sequencing. Positive clone 5 from the B6 genomic library was subcloned into pGEM7zf(−) (Promega) and sequenced according to a protocol modified for sequencing of double-stranded plasmids by the dideoxy chain-termination method (32) outlined in the Sequenase Version II kit (USB). Direct sequencing of PCR-amplified *Nat1* and *Nat2* from A mice was carried out by the method of Engelke *et al.* (33), as modified by McGuire *et al.* (34). For ^{32}P-end-labeling of oligonucleotide primers, we modified the procedure in only one respect; unincorporated [^{32}P]ATP was separated from the labeled oligonucleotide on D-25 disposable columns (IBI). Sequencing reaction mixtures were subjected to electrophoresis through 6% polyacrylamide gels containing 8 M urea, in 1× TBE (0.135 M Tris-borate buffer, pH 8.8, 2.6 mM

EDTA). Gels were dried under vacuum, and autoradiography was performed with KODAK X-OMAT film for 2–6 hr at −70°.

Southern blot analysis of mouse Nat genes. High molecular weight genomic DNA was prepared from liver (25) and digested with *Bam*HI, *Eco*RI, and *Hin*dIII. The digested DNAs were separated on an 1× TAE-0.8% agarose gel and transferred to nitrocellulose by standard Southern blot procedures (25). The blots were hybridized with the insert of pB6Nat-1. Prehybridization, hybridization, and washing conditions were the same as those for the library screening described above. Blots were exposed to KODAK X-OMAT film with intensifying screens for 2 weeks at −70°.

Results

Amplification of a partial mouse Nat clone. A segment of the coding region of a B6 mouse Nat1 gene was amplified by PCR with primers based on the rabbit cDNA sequence, rnat (31). Several amplification primers were designed to anneal to putative conserved segments of the coding regions of rnat, which had >85% nucleotide similarity to human *NAT* (29). One pair of primers amplified a fragment that was 435 nucleotides long. Sequencing of this gene fragment revealed that it started with the initiating methionine codon and it had 45%, 66%, and 74% similarity with the deduced amino acid sequence of the 5′ half of the coding region of chicken liver NAT cDNA (35), rabbit NAT cDNA (31), and the human NAT gene (29), respectively. The 435-base pair product was originally referred to as B6Nat-1. This gene fragment was subcloned to make the construct pB6Nat-1, which was used as a probe for library screening.

Cloning and sequencing of two B6 Nat genes. Twelve positive plaques were identified on duplicate nitrocellulose lifts from the initial screening of 450,000 plaques from the B6 genomic library. Southern blot analysis with restriction digests of positive clones indicated that one of these positives plaques, clone 5, contained two Nat genes. Digestion of clone 5 with *Hin*dIII alone yielded 9-kb and 1-kb bands (Fig. 1A). Digestion with *Hin*dIII and *Eco*RI, which cuts within the polylinker region, released an additional band of 5.5 kb. The 9-kb and 5.5-kb bands both hybridized to pB6Nat-1 (Fig. 1B), indicating

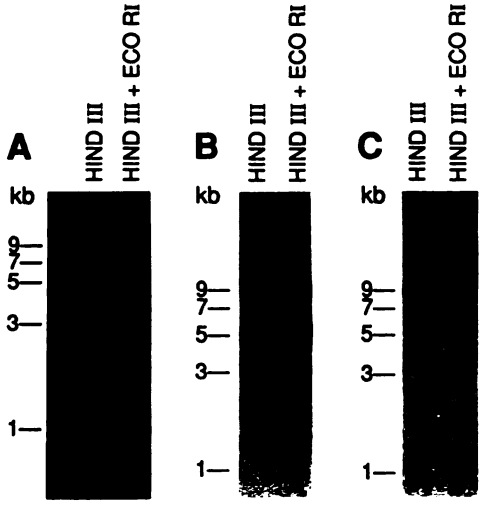

Fig. 1. Restriction digests and Southern blots of mouse clone 5. A, 1% 1× TAE agarose gel of genomic clone 5 digested with *Hin*dIII and *Hin*dIII plus *Eco*RI. Band sizes are indicated by a 1-kb ladder to the *left* of the gel. B, Autoradiogram of a Southern blot of the gel shown in A probed with the 5′-specific probe. C, Autoradiogram of a Southern blot of the gel shown in A probed with the 3′-specific probe.

that both bands contained at least part of the 5′ coding region of the Nat gene. In addition, both bands hybridized to the 3′-specific probe (Fig. 1C), which indicated that there were two Nat genes within clone 5.

For sequencing and mapping analysis, the two fragments of clone 5 were subcloned into pGEM7zf(−) to make the recombinant plasmids pB6Nat-5H (the 9-kb *Hin*dIII fragment from clone 5) and pB6Nat-5HE (the 5.5-kb *Hin*dIII/*Eco*RI fragment from clone 5). The inserts of these recombinant plasmids are referred to as 5H and 5HE, respectively. We began sequencing in the middle of the coding regions of 5H and 5HE with internal primers based on the sequence of pB6Nat-1. Additional sequencing primers were constructed as more sequence data were obtained, until the entire coding regions and 200 nucleotides extending into both the 5′ and 3′ untranslated regions were sequenced. The initial 435 nucleotides of the coding region of 5HE (Fig. 2) were identical to those identified in pB6Nat-1, indicating that the partial sequence obtained from the PCR was correct. The sequences of 5HE (Fig. 2) and 5H (Fig. 3) indicated that they were related gene products with intronless open reading frames of 870 nucleotides. The coding regions had 84% nucleotide similarity and 81% deduced amino acid similarity. The 5′ noncoding regions were 49% identical, and the 3′ noncoding regions were 42% identical.

PCR amplification and direct sequencing of Nat genes from A mice. In order to investigate the molecular basis of *Nat* polymorphism in B6 and A mice, the two homologous Nat genes were amplified from A mouse genomic DNA. Because the 5′ and 3′ noncoding regions of 5HE and 5H were dissimilar, amplification primers could be designed to anneal specifically to either the 5′ or 3′ noncoding region of 5H or 5HE, so that each gene could be amplified separately. For amplification of 5HE, primers were used that corresponded to nucleotides −155 to −132 and to nucleotides 962 to 968 (Fig. 2); for amplification of 5H, the primers used corresponded to nucleotides −183 to −157 and to nucleotides 953 to 979 (Fig. 3). The expected sizes of the amplified products were 1.2 kb for 5HE and 1.1 kb for 5H. There was no DNA contamination or any primer-dimer artifact, because no band was observed after PCR amplification of primers without genomic DNA (data not shown).

The A Nat genes were sequenced from pooled aliquots of five independently amplified reaction mixtures. Direct sequencing of the pooled sample was carried out to minimize possible errors introduced by the *Taq* DNA polymerase (24, 36). The sequence of the 5HE PCR product from A genomic DNA was identical to the sequence of 5HE from the B6 genomic clone. However, 5H contained a change from adenine in B6 mice to thymine in A mice, at nucleotide 296 (Fig. 3). This mutation resulted in a change from asparagine in B6 to isoleucine in A mice, at position 99 in the deduced amino acid sequence. According to Kyte and Doolittle (37), asparagine is a hydrophilic amino acid residue with a hydropathy index of −3.5, whereas isoleucine is the most hydrophobic amino acid residue, with a hydropathy index of 4.5. Using the Kyte-Doolittle moving segment approach to determine the average hydropathy of a peptide segment within a protein (Fig. 4), it is evident that the presence of a nonpolar isoleucine residue in the NAT from A mice causes a regional change toward greater hydrophobicity, and this may affect protein secondary structure.

Henceforth, we shall refer to the invariant 5HE gene as *Nat1* and to the variant 5H gene as *Nat2*, in analogy to other reports

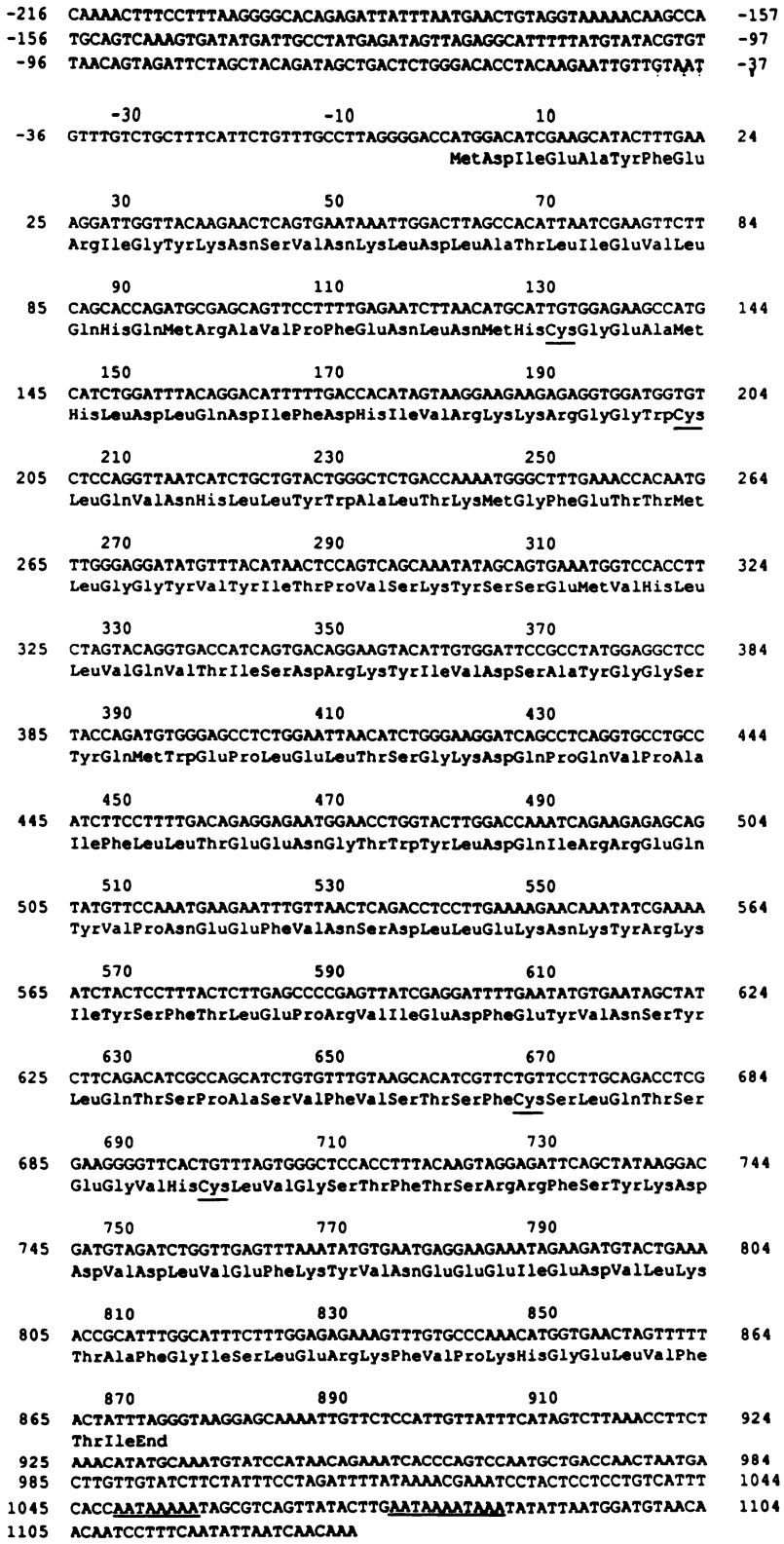

Fig. 2. Nucleotide and deduced amino acid sequences of the mouse 5HE (*Nat1*) genomic clone. Nucleotide residues are numbered in the 5′ to 3′ direction, where +1 represents the first nucleotide of the coding region. Putative poly(A)$^+$ signals and four cysteines conserved in mouse, rabbit, and human are *underlined*.

in which genetically variant and invariant acetylator genes have been characterized (23, 38, 39).

Linkage of *Nat1* and *Nat2* in B6 mice. The locations of the Nat1 and Nat2 genes in the genomic insert of clone 5 were mapped to determine the relative orientation and the distance between the two genes. Initially, *Nat1* and *Nat2* were mapped with restriction endonucleases that digested infrequently within their respective constructs, pB6Nat-5HE and pB6Nat-5H. The enzymes *Bgl*I, *Nco*I, *Nde*I, and *Ssp*I were used for *Nat1; Bgl*I, *Nco*I, and *Pst*I were used for *Nat2*. Each construct was also digested with a combination of the above mentioned enzymes and enzymes that cleave within the polylinker of the plasmid vector. Co-digestion in this manner was done to separate the insert from the vector and to determine the orientation of the insert within the plasmid. Southern blots of these digests were prepared and then probed with ^{32}P-labeled oligonucleotide

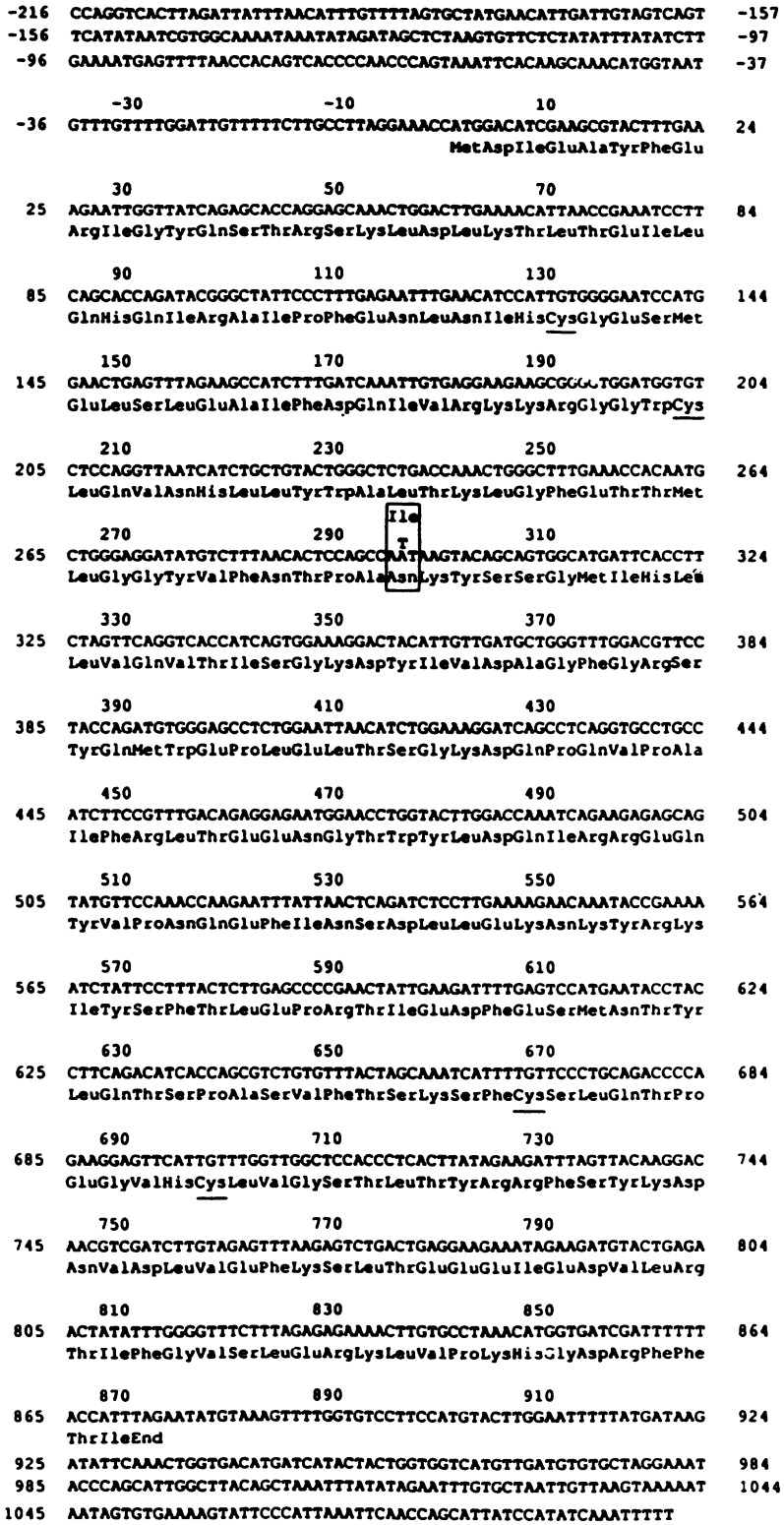

Fig. 3. Nucleotide and deduced amino acid sequences of the mouse genomic clone 5H (*Nat2*). Nucleotide residues are numbered in the 5′ to 3′ direction, where +1 represents the first nucleotide of the coding region. The four cysteine residues conserved in mouse, rabbit, and human are *underlined*. The nucleotide and deduced amino acid residue in A mice that differ from those in B6 mice are shown in the *upper lines* in the *open box*.

probes that hybridized to known locations within the sequence, to map the restriction sites. The maps of pB6Nat-5H and pB6Nat-5HE were deduced, and this information was used to map the Nat genes in the original genomic clone 5, using additional restriction enzymes (see Fig. 5). *Nat1* is 5′ to *Nat2*, and the genes are separated by approximately 9 kb.

Southern blot analysis of Nat mouse genes. Southern blots of digested genomic DNA from A and B6 mice were probed with pB6Nat-1, to estimate the number of Nat genes (Fig. 6). High molecular weight DNA isolated from B6 and A mice was digested with *Bam*HI, *Eco*RI, and *Hin*dIII. These enzymes do not cleave within the sequenced regions of *Nat1* and *Nat2*. Separate *Bam*HI and *Eco*RI digestions gave a single band of >23 kb in both strains. *Eco*RI digestion also revealed a RFLP, indicated by an additional band at 2 kb in the A mouse DNA. A second RFLP was observed with *Hin*dIII-

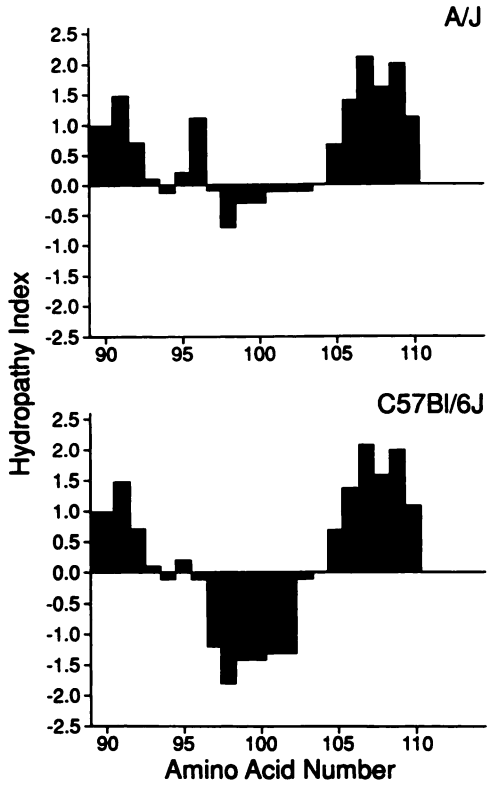

Fig. 4. Hydropathy plots of the deduced amino acid sequences of NAT2 from A and B6 mice, using the moving segment approach of Kyte and Doolittle (37). The moving segment approach was carried out as follows. Hydropathy values assigned to each amino acid in a group of seven consecutive amino acids were averaged, and this average hydropathy value was assigned to the central amino acid in the group. The seven-amino acid segment was continuously shifted by one amino acid from residue 90 to residue 110, to observe the effect of the change in amino acid 99 from asparagine in B6 to isoleucine in A mice. The *x*-axis indicates the number of the central amino acid in the window of seven amino acids. The *y*-axis shows the average hydropathy value assigned to the central amino acid; the extent of hydrophobicity is indicated by increasing positive values, and the extent of hydrophilicity is indicated by negative values.

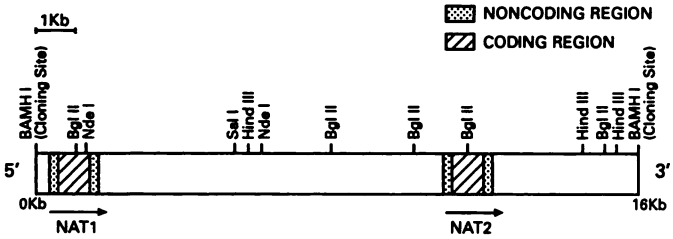

Fig. 5. Linkage of B6 mouse Nat1 and Nat2 genes. The schematic representation shows the restriction sites used to map the genes. Distance in kb is shown by the scale at the *upper left*. The 5′ to 3′ orientation of the two genes is indicated by *directional arrows*. *Nat1* is located on the 5.5-kb segment between the cloning site and a *Hin*dIII site. *Nat1* was referred to earlier as 5HE because it was within a *Hin*dIII/*Eco*RI fragment of clone 5. The *Eco*RI site of *Nat1* came from the polylinker region of the cloning vector; consequently, no *Eco*RI site is shown in the map. The Nat2 gene is located between two *Hin*dIII sites that are 9 kb apart. The map also accounts for the 1-kb *Hin*dIII fragment of clone 5 (see Fig. 1A).

digested DNA. *Hin*dIII digestion resulted in three bands of approximately 9.4, 6.6, and 4.4 kb in B6 mice and gave bands at 9.4, 6.6, and 2.8 kb in A mice. Because *Hin*dIII does not cut within the sequenced regions of the two cloned Nat genes, the appearance of three bands in *Hin*dIII-digested B6 DNA indicates the existence of at least three Nat genes.

Discussion

We report the cloning and sequencing of two genes, *Nat1* and *Nat2*, from the rapid acetylator B6 and slow acetylator A inbred mouse strains. *Nat1* is identical in B6 and A mice, whereas *Nat2* differs by a single nucleotide that results in a change in the deduced amino acid sequence from asparagine in B6 mice to isoleucine in A mice. This change from a polar residue in B6 mice to a nonpolar residue in A mice is accompanied by a change of 8 hydropathy units, according to the Kyte-Doolittle hydropathy index (37). As a consequence, the hydropathy profile of the peptide surrounding amino acid 99 changes from hydrophilic in the B6 to hydrophobic in the A mouse (Fig. 4). Such a change could affect the activity or stability of the protein. This proposal is consistent with earlier reports from our laboratory (16), in which it was proposed that hepatic NAT enzymes in A and B6 mice had minimal structural differences, because their physicochemical and biochemical properties were so similar, but were distinctly different because they exhibited different K_m and V_{max} values for the arylamine carcinogen 2-aminofluorene (20). However, evidence was also obtained for a difference in the amount of NAT protein between A and B6 (16), which could contribute to the V_{max} difference. Hence, differences in both protein affinity and quantity could be involved in the mouse NAT polymorphism (16).

The deduced amino acid sequence of multiple NATs has been reported for humans (23), rabbits (38, 39), and chickens (35, 40, 41). Fig. 7 shows many conserved regions between these four species. Cysteine residues, important for catalysis (42), are conserved at positions 44, 68, and 223. Amino acids in the vicinity of these cysteines are also highly conserved. Additionally, several basic amino acid residues are conserved near cysteine 68, which is consistent with our suggestion (43) that basic amino acids surround the catalytic site of NAT. The cysteine residue at position 233 is conserved in the mammalian species but not in the chicken. Certain other segments of conserved amino acids are discernible elsewhere in the sequence, including the amino-terminal sequence and amino acids 140–146. These conserved regions may point to functional domains of the NAT protein. Table 1 summarizes the overall percentage of identity of the mouse NAT coding region with the other species, at the nucleotide and amino acid levels. Generally, the mouse NATs are more similar to human than to rabbit NATs and are least similar to the chicken NATs.

Southern blots of *Hin*dIII digests of B6 and A DNA (Fig. 6) show three bands. This finding suggests the presence of an Nat gene in addition to *Nat1* and *Nat2* in mice. Evidence for three Nat genes has also been reported in chickens and humans. In the chicken, three distinct functional NAT cDNAs have been reported (35, 40), whereas in humans two functional genes and a pseudogene have been reported (23). Bands of 9.4 kb and 6.6 kb are observed in the *Hin*dIII digests of DNA from both A and B6 mice (Fig. 6), but the third band is clearly different between the two species, i.e., 2.8 kb in A versus 4.4 kb in B6. Still further, *Eco*RI digests show that two bands are present in the A mouse DNA, whereas only one band is present in B6 DNA. The observed nucleotide change in *Nat2* between B6 and A mice does not create a *Hin*dIII or *Eco*RI site and, therefore, cannot explain the RFLPs between the two mice strains. At

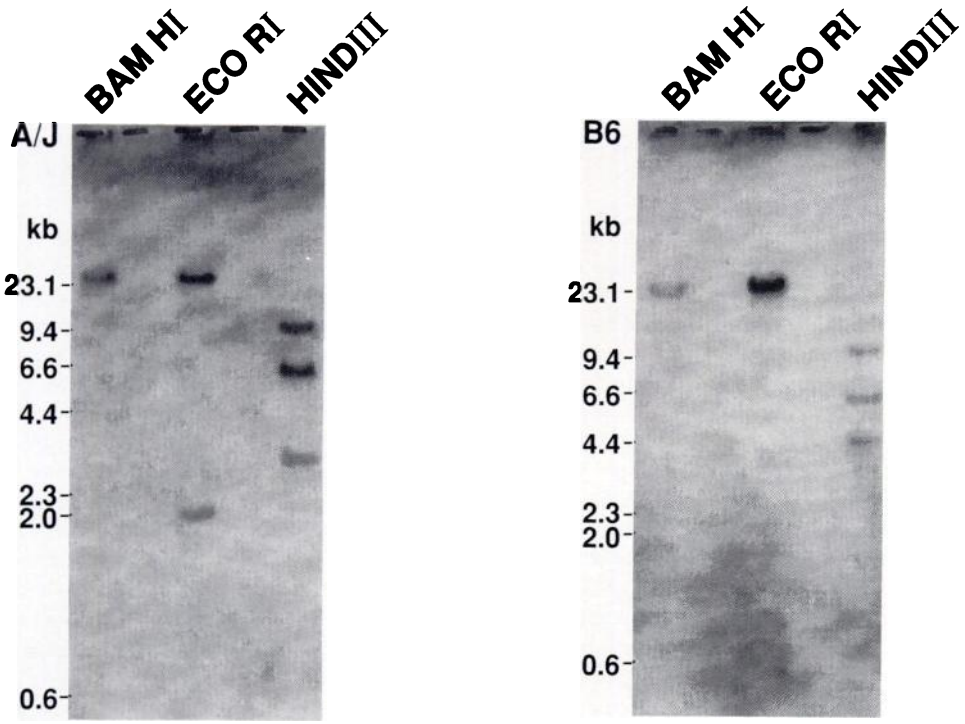

Fig. 6. Southern blot analysis. High molecular weight A and B6 DNA was digested with *Bam*HI, *Eco*RI, and *Hin*dIII. Wild-type λ phage DNA digested with *Hin*dIII was used for size markers and is shown, in kb, to the *left* of the autoradiograms. The Southern blots were probed with pB6Nat-1.

present we do not know the extent to which this polymorphism in a third gene is involved in the mouse NAT polymorphism or whether this is a functional Nat gene or a pseudogene. Although the structural basis and significance of the *Hin*dIII and *Eco*RI RFLPs are unknown at this time, these RFLPs should be useful in testing whether the same mutation that accompanies the NAT polymorphism in B6 and A mice also occurs in other inbred strains whose acetylator status has been determined (15).

By comparing the sizes of the bands obtained in the *Hin*dIII digests of B6 DNA in the Southern blot in Fig. 6 with the map in Fig. 5, we can putatively identify the 6.6-kb band with *Nat1* and the 9.4-kb band with *Nat2*. *Nat1* may be contained in the 6.6-kb fragment (Fig. 6), because *Nat1* is located between the 5′ cloning site and the *Hin*dIII site approximately 5.5 kb away (Fig. 5). The 4.4-kb band of *Hin*dIII-digested B6 DNA in Fig. 6 is too small to account for the presence of *Nat1* on the 5.5-kb segment. The presence of *Nat1* in the larger 6.6-kb fragment (Fig. 6) can be accounted for if we assume that there is a *Hin*dIII site approximately 1 kb upstream from the 5′ cloning site (Fig. 5). Similarly, we propose that *Nat2* is contained in the 9.4-kb fragment (Fig. 6), because *Nat2* is located between two *Hin*dIII sites approximately 9 kb apart in the map (Fig. 5). The 4.4-kb band may contain the third gene mentioned in the previous paragraph, although this remains to be established. The fact that neither *Bam*HI nor *Eco*RI cleaves between *Nat1* and *Nat2* of B6 DNA (Fig. 6) is consistent with the mapping analysis presented in Fig. 5.

Nat1 and *Nat2* have an intronless coding region of 870 nucleotides, which is of interest because intronless genes with more than 800 coding-region nucleotides appear relatively infrequently in vertebrate genomes (44, 45). The mouse Nat genes are not unique in this respect, because the human (23) and rabbit (38, 39) NAT genes also have been reported to have intronless open reading frames of 870 nucleotides. The lack of introns within the coding region of *Nat1* and *Nat2* raises the possibility that they may be pseudogenes. This seems unlikely, because *Nat1* and *Nat2* lack several features that are generally associated with processed pseudogenes, e.g., pseudogenes are usually found on different chromosomes than the parent genes, their 5′ and 3′ untranslated regions often contain direct repeats that flank the site of integration of the retroposon, and they often contain terminal repeat codons that interrupt their open reading frames (46). More importantly, the human NAT1 and NAT2 genes, which have intronless coding regions, have been transfected into COS-1 cells, and both have expressed functional proteins (23).

Close inspection of Figs. 2 and 3 reveals the presence of an AG pair six nucleotides upstream from the initiating ATG codon in both mouse genes. This AG pair is preceded by 13 nucleotides that display 85% (*Nat1*) and 92% (*Nat2*) similarity to the consensus sequence for a 3′ splice acceptor site (46). In agreement with the findings of Blum *et al.* (23), we have also observed identical segments at the corresponding region of human *NAT1* and *NAT2* (47). A comparison of the human NAT1 and NAT2 cDNAs (48) with their respective genes (23) reveals that these splice acceptor sites are utilized by the NAT2 cDNA but not the NAT1 cDNA. This observation suggests that human *NAT2* may contain another exon encoding the 5′ untranslated region of the cDNA. The organization of the corresponding mouse Nat1 and Nat2 genes is not known, but the high degree of similarity to the human NAT genes within the potential splice acceptor sites suggests that other exons 5′ to the open reading frame may exist.

Efficient expression of both Nat genes may be compromised by the presence of several ATGs present 5′ to that designated as the translation initiator in the mouse genes, unless the aforementioned splice acceptor sites are utilized. *Nat1* has six upstream ATGs, five of which terminate before the coding region begins. The remaining ATG (at position −36) terminates within codon 17 in the coding region. *Nat2* has four upstream ATGs, one of which terminates before the coding region begins.

```
                1                                                  *    50
  Mouse NAT1    MDIEAYFERI GYKNSVNKLD LATLTEVLQH QMRAVPFENL NMHCGEAMHL
  Mouse NAT2    ---------- --QSTRS--- -K----I--- -I--I----- -I------E-
  Human NAT1    ------L--- ---K-R---- -E---DI--- -I-------- -I---D--D-
  Human NAT2    ---------- -----R---- -E---DI-E- -I-------- -----Q--E-
  Rabbit NAT1   ------YQ-- ----PR---- -ES--DIF-- -I-T--Y--- SI----S-E-
  Rabbit NAT2   ------YQ-- ----PR---- -ES--DIF-- -I-T--Y--- SI----S-E-
  Chicken NAT3  ---KE--A-- S-GG-YE-P- -E----IF-- HIQ------- SI----TIE-
  ChickenNat10  -NL-E--A-T ---G-LENQ- -E---DIF-- HI-------- SI----KIT-
  Chicken NATL  -N-QE--S-- SFDG-HKDA- -Q---AIF-- HIQ-I----- S-----TID-

                51                  *                              100
  Mouse NAT1    DLQDIFDHIV RKKRGGWCLQ VNHLLYWALT KMGFETTMLG GYVYITPVSK
  Mouse NAT2    S-EA---Q-- ---------- ---------- -L-------- ---FN--AN-
  Human NAT1    G-EA---QV- -RN------- ---------- TI-------- --G-S--AK-
  Human NAT2    G-EA------ -RN------- --Q------- TI--Q----- --F--P--N-
  Rabbit NAT1   --EA---Q-- -RN------- --Y------- TT-------- -F-CGSHTD-
  Rabbit NAT2   --EA---Q-- -RN------- --Y------- TT-------- -F--GSNND-
  Chicken NAT3  --AATY-K-- --------ME N----S---K TL-YNV-L-- AK---PEHDA
  ChickenNAT10  E-EHVYNK-- H-------ME N-Q--G-V-K CL-YD-SF-- A--FNPHENA
  Chicken NATL  ---ATYNK-- K-------ME T-Y--F---K E--YDICV-- -NS-EPAKKA

                101                                                150
  Mouse NAT1    YSSEMVHLLV QVTISDRKYI VDSAYGGS.Y QMWEPLELTS GKDQPQVPAI
  Mouse NAT2    ---G-I---- -----GKD-- --AGF-R-.- ---------- ----------
  Human NAT1    --TG-I---L ----DGTN-- --AGF-R-.- ---Q----I- --------CV
  Human NAT2    --TG-----L ----DG-N-- --AGS-S-S. ---Q----I- --------C-
  Rabbit NAT1   --TG-I--I- ----NG-N-- --AGF-R-.- ---Q-V--I- --------S-
  Rabbit NAT2   --TG-I--I- ----NG-N-- --AGF-R-.- ---Q-V--I- --------S-
  Chicken NAT3  -ADDID---L K-VLH--S-- --GGF-MA.- -L-Q-M--I- ------T-G-
  ChickenNAT10  -ATI-T---- K-V-EGKA-- --AGF-V-.- ---Q-M--V- ------A-G-
  Chicken NATL  -TDEIN-I-L K-V-KGSS-- --AGF--GP- ---L-ML-I- ------I-G-

                151                                                200
  Mouse NAT1    FLLTEENGTW YLDQIRREQY VPNEEFVNSD LLEKNKYRKI YSFTLEPRVI
  Mouse NAT2    -R-------- ---------- ---Q--I--- ---------- --------T-
  Human NAT1    -R------F- ------TQ-- I-----LH-- ---DS----- -----K--T-
  Human NAT2    -C----R-I- ---------- IT-K--L--H --P-K-HQ-- -L------T-
  Rabbit NAT1   -R-R--GE-- -------Q-H --DQ--L--E ---RKTH--L -C---Q--T-
  Rabbit NAT2   -R-R--GE-- -------Q-H --DQ--L--E ----KI-Q-L -C---Q--T-
  Chicken NAT3  -RFV------ --EKVK-K-- ---HSDSAPH NVD-EVC-RV -L---Q--D-
  ChickenNAT10  -RF--K-AI- --EKM--K-- I--QN-S--- ----KDC--V -M-S----TV
  Chicken NATL  -RFI-D--I- --EKVK-KH- --EGSVPLT- NP-MGNI--L --------KH-

                201                  *                             250
  Mouse NAT1    EDFEYVNSYL QTSPASVFVS TSFCSLQTSE GVHCLVGSTF TSRRFSYKDD
  Mouse NAT2    ----SM-T-- --------T- K-------P- ---------L -Y-------N
  Human NAT1    ----SM-T-- ----S---T- K--------PD -------F-L -H---N---N
  Human NAT2    ----SM-T-- ----T-S-IT --------P- --Y----FIL -Y-K-N---N
  Rabbit NAT1   -E--SA-T-- -I--S-P-LD K-I-----P- -------LIL -F-TYN--EN
  Rabbit NAT2   -E--SA-T-- -E--S---LD K-I-----P- -------L-L -S-TYN--EN
  Chicken NAT3  -E-RAR-LH- --A-D-L--T K-I-----PD --RA---WKL -EIKYN---N
  ChickenNAT10  ---CFQCT-- ----D-L-TK K-I-T---TD -FRA-I-W-L -ETKYN--EN
  Chicken NATL  D--QEL-A-- -VA-DTILQK K-I-----TD -FYA---W-- SEMKYK--E-

                251                                  291
  Mouse NAT1    VDLVEFKYVN EEEIEDVLKT AFGISLERKF VPKHGELVFT I
  Mouse NAT2    -------SLT --------R- I--V-----L -----DRF-- -
  Human NAT1    T--I---TLS -----K---N I-N---Q--L -----DTF-- -
  Human NAT2    T-------TLT ---V-E---N I-K---G-NL ---P-DGSL- -
  Rabbit NAT1   T------VLT ---V-E---- I-N---GK-L -S-N-N-F-- -
  Rabbit NAT2   T------VLT ---V-G---- I-N---GK-L -S-N-H-S-- -
  Chicken NAT3  M----IRILA D--M-KT--E K-N-T-DK-- --INTSRLSL F
  ChickenNAT10  M-----ITLK D--V-KT--D K-N-T----L --INVKGFY- -
  Chicken NATL  A--LQTTTLT D--V-KT--D K-N-V--N-L I-VNVRG... .
```

Fig. 7. Comparison of the deduced amino acid sequences of mouse, human, rabbit, and chicken arylamine NATs. –, Residues that are identical to those in mouse NAT1; those that differ are shown by their single-letter amino acid codes. *, Conserved cysteines, ·, Gap introduced to optimize the alignment. Deduced amino acid sequences were obtained from the following sources: mouse NAT1 and NAT2, Figs. 2 and 3; human NAT1 and NAT2 (23), rabbit NAT1 (38) and NAT2 (39), and chicken NATL (where L indicates liver), Ref. 35; and chicken NAT3 [p-NAT-3 (40)] and chicken NAT10 [p-NAT-10 (40)], Ref. 40.

TABLE 1

Percentage of identity between mouse and human, rabbit, or chicken *NAT* coding regions

	Identity						
	Murine		Human[a]		Rabbit		Chicken[d]
	Nat1	*Nat2*	*Nat1*	*Nat2*	*Nat1*[b]	*Nat2*[c]	
	%						
Nucleotide							
Murine NAT1	100	84	80	82	77	77	61
Murine NAT2	84	100	79	79	76	76	58
Amino acid							
Murine NAT1	100	81	73	72	66	67	47
Murine NAT2	81	100	80	74	72	73	47

[a] From Ref. 23.
[b] From Ref. 38.
[c] From Ref. 39.
[d] From Ref. 35.

The other three ATGs occur only 100 nucleotides upstream of the initiating methionine, and all terminate within codon 22. Utilization of the putative splice acceptor sites would keep the upstream ATGs from interfering with translation in the intact organism, but the presence of these ATGs could interfere with production of the NAT protein in a mammalian expression system. Furthermore, none of the upstream ATGs observed in *Nat2* is in frame, and none of these ATGs in *Nat1* or *Nat2* fits the consensus sequence CCA/GCCATGG identified as a favorable context for initiation by eucaryotic ribosomes (49, 50). The start codons of mouse *Nat1* and *Nat2*, identified by comparison with other species (23, 31, 35), most closely approximate the consensus sequence for initiation of translation, i.e., the sequence encompassing the initiating codon of *Nat1* is GGACCATGG, whereas that for *Nat2* is AAACCATGG.

Currently, we are attempting to express the cloned Nat1 gene and Nat2 genes from the B6 and A mice, using a mammalian expression system. We plan to investigate the differential substrate specificities of the NATs, their kinetic properties, and other possible differences between them, such as message and protein abundance and stability. The activity profiles we obtain from the expressed mouse Nat genes will lend further insight into the functional nature of NAT polymorphism in mice and how its molecular basis compares with that of human NAT polymorphism. Such information will be useful in understanding more completely the role of acetylation polymorphism in mediating toxic and carcinogenic effects of xenobiotics.

Acknowledgments

We wish to thank Michael F. Clarke for his technical expertise and advice. We also thank Susan S. Mattano, Gerald N. Levy, Timothy P. Angelotti, and Michael Collard for valuable discussion and critical evaluation of the manuscript. The preprint of the nucleotide sequence of rnat obtained from Urs Meyer (Basel) is gratefully acknowledged.

References

1. Weber, W. W. *The Acetylator Genes and Drug Response*. Oxford University Press, New York (1987).
2. Evans, D. A. P., K. A. Manley, and V. A. McKusick. Genetic control of isoniazid metabolism in man. *Br. Med. J.* **2**:485–491 (1960).
3. Evans, D. A. P., and T. A. White. Human acetylation polymorphism. *J. Lab. Clin. Med.* **63**:394–403 (1964).
4. Weber, W. W., and D. W. Hein. Clinical pharmacokinetics of isoniazid. *Clin. Pharmacokinet.* **4**:401–422 (1979).
5. Hein, D. W. Acetylator genotype and arylamine-induced carcinogenesis. *Biochim. Biophys. Acta* **948**:37–66 (1988).
6. Jenne, J. W. Partial purification and properties of the isoniazid transacetylase in human liver: its relationship to the acetylation of *p*-aminosalicylic acid. *J. Clin. Invest.* **44**:1992–2002 (1965).
7. Grant, D. M., F. Lottspeich, and U. A. Meyer. Evidence for two closely related isozymes of arylamine *N*-acetyltransferase in human liver. *FEBS. Lett.* **244**:203–207 (1989).
8. Weber, W. W., G. Drummond, J. N. Miceli, and R. Szabadi. Genetic control and isozymic composition of drug acetylating enzymes, in *Third International Conference on Isozymes* (C. Market, ed.). Yale University, New Haven, CT, 813–828 (1975).
9. Hein, D. W., J. G. Omichinski, J. A. Brewer, and W. W. Weber. A unique pharmacogenetic expression of the *N*-acetylation polymorphism in the inbred hamster. *J. Pharmacol. Exp. Ther.* **220**:8–15 (1982).
10. Weber, W. W. The molecular basis of hereditary acetylation polymorphisms. *Drug Metab. Dispos.* **14**:377–381 (1986).
11. Hein, D. W., T. N. Smolen, R. R. Fox, and W. W. Weber. Identification of genetically homozygous rapid and slow acetylators of drugs and environmental carcinogens among established inbred rabbit strains. *J. Pharmacol. Exp. Ther.* **223**:40–44 (1982).
12. Hein, D. W., W. G. Kirlin, R. J. Ferguson, and W. W. Weber. Inheritance of liver *N*-acetyltransferase activity in the inbred hamster. *J. Pharmacol. Exp. Ther.* **233**:584–587 (1985).
13. Hein, D. W., W. G. Kirlin, R. J. Ferguson, and W. W. Weber. Biochemical investigation of the basis for the genetic *N*-acetylation polymorphism in the inbred hamster. *J. Pharmacol. Exp. Ther.* **234**:358–364 (1985).
14. Tannen, R. H., and W. W. Weber. Inheritance of acetylator phenotype in mice. *J. Pharmacol. Exp. Ther.* **213**:480–484 (1980).
15. Glowinski, I. B., and W. W. Weber. Genetic regulation of aromatic amine *N*-acetylation in inbred mice. *J. Biol. Chem.* **257**:1424–1430 (1982).
16. Mattano, S. S., S. Land, C. M. King, and W. W. Weber. Purification and biochemical characterization of hepatic arylamine *N*-acetyltransferase from rapid and slow acetylator mice: identity with arylhydroxamic acid *N,O*-acyltransferase and *N*-hydroxyarylamine *O*-acetyltransferase. *Mol. Pharmacol.* **35**:599–609 (1989).
17. Weber, W. W., and D. W. Hein. *N*-Acetylation pharmacogenetics. *Pharmacol. Rev.* **37**:25–79 (1985).
18. Glowinski, I. B., and W. W. Weber. Biochemical characterization of genetically variant aromatic amine *N*-acetyltransferases in A/J and C57BL/6J mice. *J. Biol. Chem.* **257**:1431–1437 (1982).
19. Mattano, S. S., R. P. Erickson, M. N. Nesbitt, and W. W. Weber. Linkage of Nat and Es-1 in the mouse and development of strains congenic for *N*-acetyltransferase. *J. Hered.* **79**:430–433 (1988).
20. Mattano, S. S., and W. W. Weber. Kinetics of arylamine *N*-acetyltransferase in tissues from rapid and slow acetylator mice. *Carcinogenesis (Lond.)* **8**:133–137 (1987).
21. Hein, D. W., A. Trinidad, T. Yerokun, R. J. Ferguson, W. G. Kirlin, and W. W. Weber. Genetic control of acetyl coenzyme A-dependent arylamine *N*-acetyltransferase, hydrazine *N*-acetyltransferase, and *N*-hydroxy-arylamine *O*-acetyltransferase enzymes in C57BL/6J, A/J, AC57F1, and the rapid and slow acetylator A.B6 and B6.A congenic inbred mouse. *Drug Metab. Dispos.* **16**:341–347 (1988).
22. Blum, M., D. M. Grant, A. Demierre, and U. A. Meyer. *N*-Acetylation pharmacogenetics: a gene deletion causes absence of arylamine *N*-acetyltransferase in liver of slow acetylator rabbits. *Proc. Natl. Acad. Sci. USA* **86**:9554–9557 (1989).
23. Blum, M., D. M. Grant, W. McBride, M. Heim, and U. A. Meyer. Human arylamine *N*-acetyltransferase genes: isolation, chromosomal localization, chromosomal localization, and functional expression. *DNA Cell Biol.* **9**:193–203 (1990).
24. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Washington D. C.)* **239**:487–491 (1988).
25. Maniatis, T., E. F. Fritsch, and J. Sambrook. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
26. Levy, G. N., and W. W. Weber. High-performance liquid chromatographic analysis of ^{32}P-postlabeled DNA-aromatic carcinogen adducts. *Anal. Biochem.* **174**:381–392 (1988).
27. Innis, M. A., D. H. Gelfland, J. J. Sninsky, and T. J. White. *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York (1990).
28. Kwok, S., and R. Higuchi. Avoiding false positives with PCR. *Nature (Lond.)* **339**:237–238 (1989).
29. Grant, D. M., M. Blum, A. Demierre, and U. A. Meyer. Nucleotide sequence of an intronless gene for a human arylamine *N*-acetyltransferase related to polymorphic drug acetylation. *Nucleic Acids Res.* **17**:3978 (1989).
30. Martell, K. J., K. P. Vatsis, and W. W. Weber. Cloning of polymorphic mouse NAT. *FASEB J.* **4**:753A (1990).
31. Blum, M., D. M. Grant, A. Demierre, and U. A. Meyer. Nucleotide sequence of a full-length cDNA for arylamine *N*-acetyltransferase from rabbit liver. *Nucleic Acids Res.* **17**:3589 (1989).
32. Sanger, F., S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467 (1977).
33. Engelke, D. R., P. A. Hoener, and F. A. Collins. Direct sequencing of enzymatically amplified human genomic DNA. *Proc. Natl. Acad. Sci. USA* **85**:544–548 (1988).
34. McGuire, M. C., C. P. Nogueira, C. F. Bartels, H. Lightstone, A. Hajra, A. F. L. Van Der Spek, O. Lockridge, and B. N. La Du. Identification of the structural mutation responsible for the dibucaine-resistant (atypical) variant form of human serum cholinesterase. *Proc. Natl. Acad. Sci. USA* **86**:953–957 (1989).
35. Ohtomi, M., Y. Sakamoto, K. Uyemura, and T. Deguchi. Arylamine *N*-acetyltransferase from chicken liver. II. Cloning of cDNA and expression in Chinese hamster ovary cells. *J. Biol. Chem.* **263**:7534–7538 (1988).
36. Ennis, P. D., J. Zemmour, R. D. Salter, and P. Parham. Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. *Proc. Natl. Acad. Sci. USA* **87**:2833–2837 (1990).
37. Kyte, J., and R. F. Doolittle. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105–132 (1982).
38. Blum, M., M. Heim, and U. A. Meyer. Nucleotide sequence of rabbit *Nat1* encoding monomorphic arylamine *N*-acetyltransferase. *Nucleic Acids Res.* **18**:5287 (1990).
39. Blum, M., M. Heim, and U. A. Meyer. Nucleotide sequence of rabbit *Nat2* encoding polymorphic liver arylamine *N*-acetyltransferase (NAT). *Nucleic Acids Res.* **18**:5295 (1990).
40. Ohtomi, M., M. Sasaki, and T. Deguchi. Two arylamine *N*-acetyltransferases from chicken pineal gland as identified by cDNA cloning. *Eur. J. Biochem.* **185**:253–261 (1989).
41. Ohtomi, M., T. Ichikawa, and T. Deguchi. Arylamine *N*-acetyltransferase from the chicken kidney: characterization of the enzyme and localization of the mRNA by *in situ* hybridization. *Biomed. Res.* **11**:263–267 (1990).

42. Andres, H. H., A. J. Klem, L. M. Schopfer, J. K. Harrison, and W. W. Weber. On the active site of liver acetyl-CoA: arylamine *N*-acetyltransferase from rapid acetylator rabbits (III/J). *J. Biol. Chem.* **263**:7521–7527 (1988).
43. Andres, H. H., H. J. Kolb, R. J. Schreiber, and L. Weiss. Characterization of the active site, substrate specificity and kinetic properties of acetyl-CoA:arylamine *N*-acetyltransferase from pigeon liver. *Biochim. Biophys. Acta* **746**:193–201 (1983).
44. Hawkins, J. D. A survey on intron and exon lengths. *Nucleic Acids Res.* **16**:9893–9908 (1988).
45. Dorit, R. L., L. Schoenbach, and W. Gilbert. How big is the universe of exons? *Science (Washington D. C.)* **250**:1377–1382 (1990).
46. Rogers, J. H. The origin and evolution of retroposons. *Int. Rev. Cytol.* **93**:187–279 (1985).
47. Vatsis, K. P., K. J. Martell, and W. W. Weber. Diverse point mutations in the human gene for polymorphic *N*-acetyltransferase. *Proc. Natl. Acad. Sci. USA*, 1991 in press.
48. Ohsako, S., and T. Deguchi. Cloning and expression of cDNAs for polymorphic and monomorphic arylamine *N*-acetyltransferases from human liver. *J. Biol. Chem.* **265**:4630–4634 (1990).
49. Kozak, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**:857–872 (1984).
50. Kozak, M. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**:5233–5262 (1981).

Send reprint requests to: Dr. Wendell W. Weber, Department of Pharmacology, Medical Science Building I, Room 6322, University of Michigan Medical School, Ann Arbor, MI 48109-0626.