Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 6333–6337, July 1991
Medical Sciences

Diverse point mutations in the human gene for polymorphic *N*-acetyltransferase

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

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

Communicated by J. Lawrence Oncley, March 25, 1991 (received for review December 10, 1990)

ABSTRACT **Classification of humans as rapid or slow acetylators is based on hereditary differences in rates of N-acetylation of therapeutic and carcinogenic agents, but N-acetylation of certain arylamine drugs displays no genetic variation. Two highly homologous human genes for *N*-acetyltransferase (NAT; arylamine acetyltransferase, acetyl CoA:arylamine *N*-acetyltransferase, EC 2.3.1.5), *NAT1* and *NAT2*, presumably code for the genetically invariant and variant NAT proteins, respectively. In the present investigation, 1.9-kilobase human genomic *Eco*RI fragments encoding NAT2 were generated by the polymerase chain reaction with liver and leukocyte DNA from seven subjects phenotyped as homozygous and heterozygous acetylators. Direct sequencing revealed multiple point mutations in the coding region of two distinct *NAT2* variants. One of these was derived from leukocytes of a slow acetylator and was distinguished by a silent mutation (codon 94) and a separate G → A transition (position 590) leading to replacement of Arg-197 by Gln; the mutated guanine was part of a CpG dinucleotide and a *Taq* I site. The second *NAT2* variant originated from liver with low N-acetylation activity. It was characterized by three nucleotide transitions giving rise to a silent mutation (codon 161), accompanied by obliteration of the sole *Kpn* I site, and two amino acid substitutions: Thr for Ile (codon 114) and Arg for Lys (codon 268). Heterozygosity was detected in three *NAT2* samples: two were heterozygous for the rapid and one of the allelic variants, and the third was a compound heterozygote of both mutant alleles. The results show conclusively that the genetically variant NAT is encoded by *NAT2*.**

Hepatic cytosolic *N*-acetyltransferase (NAT; arylamine acetyltransferase, acetyl CoA:arylamine *N*-acetyltransferase, EC 2.3.1.5) participates in the detoxification of a plethora of hydrazine and arylamine drugs, as well as in activation pathways of occupational carcinogens (1, 2). A reactive cysteinyl thiolate is part of the active site of NAT (3), which displays a ping-pong reaction mechanism with a covalent acetyl-cysteinyl-NAT species as the catalytic intermediate (3, 4).

Hereditary differences in N-acetylation activity among individuals and in populations of diverse raciogeographic origin have led to a phenotypic classification of humans as rapid or slow acetylators. This genetic heterogeneity in NAT activity, widely referred to as the N-acetylation polymorphism, is determined in people and several animal species by a single autosomal gene with two major alleles expressed codominantly (2, 5). On the other hand, rates of elimination *in vivo* and acetylation *in vitro* of certain drugs (e.g., *p*-aminosalicylic acid) do not differ appreciably among rapid and slow acetylators, indicating that metabolism of these substrates takes place without genetic variation. The term monomorphic has been coined for genetically invariant N-acetylation activity (ref. 2, pp. 137–138). The molecular mechanisms for monomorphic and polymorphic N-acetylation in humans are not yet fully understood.

The amino acid composition reported for several tryptic peptides of electrophoretically homogeneous liver NAT from homozygous rapid acetylator rabbits (4) proved pivotal in the design of oligonucleotide screening probes and isolation of rabbit NAT cDNAs (6, 7), which eventually enabled identification of human NAT cDNA and genomic clones (7, 8). The libraries screened for human NAT cDNAs were made with mRNA from two livers of undetermined acetylator phenotype (7). It was inferred from the substrate selectivity profiles of the cDNA products expressed in Chinese hamster ovary cells that two of the isolated cDNA clones coded for the genetically variant NAT and one coded for the monomorphic protein (7). Three NAT gene loci, corresponding to *Eco*RI fragments of 1.3 kilobases (kb) (*NAT1*), 1.9 kb (*NAT2*), and 4.7 kb (*NATP*), were identified upon screening of a library constructed with leukocyte DNA from a heterozygous person. Indirect evidence from the deduced amino acid sequence of the gene products and from expression experiments suggested that *NAT1* and *NAT2* encoded the monomorphic and polymorphic NAT proteins, respectively, and that *NATP* was a pseudogene (8).

In the present study we have employed PCR to generate *NAT2* from human liver and leukocyte DNA from homozygous rapid, homozygous slow, and heterozygous acetylator phenotypes, ascertained by liver NAT activity determinations *in vitro* (9) and quantitation of urinary caffeine and serum dapsone metabolites after drug administration *in vivo* (10). Direct sequencing unveiled diverse point mutations in the protein-coding region of two distinct *NAT2* variants from slow acetylators: two nucleotide transitions leading to one amino acid change (Arg-197 → Gln) and elimination of a *Taq* I site were seen in *NAT2* from white blood cells (WBCs) of a slow acetylator individual, and three point mutations resulting in two amino acid substitutions (Ile-114 → Thr and Lys-268 → Arg) and obliteration of the sole *Kpn* I site were discernible in *NAT2* from a liver sample with low N-acetylation activity. In the former instance, the mutated guanine at position 590 was part of a CpG dinucleotide, the hypermutability of which has been attributed to the propensity of 5-methylcytosine to undergo deamination to thymine (11). Furthermore, duplicate comigrating nucleotide bands, indicative of heterozygosity, were exhibited by *NAT2* from one WBC and two liver samples.

EXPERIMENTAL PROCEDURES

Isolation of DNA from Liver and WBCs of Different Acetylator Phenotypes. Genomic DNA was isolated from 4 of 20 liver samples (12), the acetylator phenotype of which had been assessed by determination of cytosolic NAT activity with *p*-aminobenzoic acid and sulfamethazine as substrates (9). DNA was also extracted from WBCs (13) of three healthy

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: NAT, *N*-acetyltransferase; *NAT*, *N*-acetyltransferase gene; WBC, white blood cell; nt, nucleotide(s).

volunteers (subjects 8, 40, and 53) from among 75 persons comprising 16 families, whose acetylator status had been established by HPLC quantitation of urinary caffeine metabolites and by measurement of N-acetylated and unmetabolized dapsone in serum (10).

Synthesis and ^{32}P-Labeling of Oligonucleotides. Oligonucleotides for amplification and sequencing were synthesized at the University of Michigan DNA Synthesis Facility and purified by HPLC. Amplification primers in the forward (AP 1) and reverse (AP 2) directions corresponded to 28 bases at the 5′ and 3′ ends of *NAT2*, respectively (see Fig. 2), and in addition contained *Hin*dIII (AP 1) and *Sst* I (AP 2) linkers with a GC pair attached at the linker ends [total length, 36 nucleotides (nt)]. Sequencing primers were as follows: IP1, 5′-ACGTTATACCTATAATT-3′; IP2, 5′-TCTGACCACAATCGGTT-3′; IP3, 5′-TCTCCTGCCAAAGAAG-3′; and IP4, 5′-GCACATAAGTTGATAAT-3′.

Sequencing primers (20 pmol) were 5′ end-labeled with a 20% molar excess of [γ-^{32}P]ATP (120 μCi; 1 Ci = 37 GBq) and T4 polynucleotide kinase (10 units) in a total volume of 20 μl (13). 5′-^{32}P-labeled oligonucleotides were separated from unincorporated ATP on D-25 disposable columns (International Biotechnologies).

Amplification of *NAT2* by PCR. Typical reaction mixtures contained 1 μg of liver or WBC genomic DNA as template, forward (AP 1) and reverse (AP 2) amplification primers (each at 0.1–0.5 μM), the four deoxynucleoside triphosphates (each at 0.2 mM), 20 mM Tris·HCl buffer (pH 8.4 at room temperature), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% (wt/vol) gelatin, 2 mM dithiothreitol, and 5 units of *Taq* DNA polymerase (Beckman) in a total volume of 0.1 ml (14). Template DNA was omitted from control mixtures. After addition of a few drops of mineral oil, reaction vessels were placed in a thermal cycler (Ericomp TwinBlock, San Diego), and amplification was allowed to proceed for 30 cycles of denaturation (94°C, 1.5 min), annealing (55°C, 1 min), and extension (72°C, 3 min), followed by a 10-min extension step at 72°C after the last cycle. Ten-microliter aliquots of the reaction mixtures were subsequently analyzed by agarose gel electrophoresis.

Direct Sequencing of PCR-Generated *NAT2*. The composition of reaction mixtures and other conditions were as described by McGuire *et al.* (15) with the following modifications: 5′-^{32}P-labeled sequencing primers were included at 5 $\times$ 10^5 cpm, the temperature for the annealing and polymerization steps was 37°C, reactions were terminated on ice, and the DNA was precipitated and washed with ethanol. Heat-denatured samples in loading dye (15) were submitted to electrophoresis on 6% polyacrylamide/8 M urea gels, followed by autoradiography for 17–96 h.

RESULTS

Human *NAT2* consists of an intronless protein-coding region 870 base pairs (bp) in length and a 107-bp 5′ untranslated tract, which is interrupted by an intron 6 nt upstream from the initiator ATG codon (position +1). Two *Eco*RI sites, separated by 1.9 kb, are located within the intron (715 nt upstream from position +1) and in the 3′ untranslated segment (300 nt downstream from position 870), respectively (Fig. 1).

Amplification primers to the *Eco*RI sites of *NAT2* (Fig. 2 *Upper*) were mixed with WBC or liver DNA from homozygous (rapid and slow) and heterozygous acetylators and allowed to undergo 30 PCR cycles. An intense DNA band of the expected size (1.9 kb) was uniformly and consistently obtained upon agarose gel electrophoresis of the PCR mixtures, with very few nonspecific amplification products (Fig. 2 *Lower*, lanes 1–7). No bands were generated when genomic DNA was omitted from reaction mixtures (Fig. 2 *Lower*, lane 8). Both alleles of PCR-generated *NAT2* (Fig. 2 *Lower*, lanes 1–7) were sequenced directly with sequencing (internal)

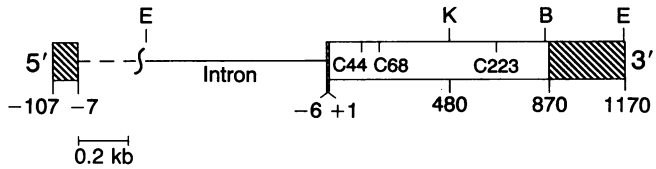
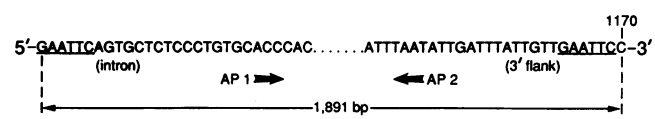


FIG. 1. Structure of human *NAT2*. The sequence of the entire 1.9-kb *Eco*RI fragment, amplified with genomic DNA from a liver sample with intermediate NAT activity (Fig. 2, lane 6), has been determined in this laboratory (K.P.V. and W.W.W., unpublished results) and found to be as reported by Blum *et al.* (8). The location of the intron has been verified by NAT cDNAs (7), but the actual intron size is not known. Cysteines 44, 68, and 223 (C44, C68, and C223) are conserved in the four species examined to date. A fourth cysteine (cysteine 233) is conserved in humans, rabbits, and mice, but not in chickens (6–8, 16, 17). Hatched bars denote 5′ and 3′ untranslated tracts. E, *Eco*RI; K, *Kpn* I; B, *Bam*HI.

primers IP1 (located 71–87 nt 5′ of position +1), IP2 (nt 234–250), IP3 (nt 537–552), and IP4 (complementary to nt 921–937). The length of the *NAT2* segments sequenced was ≈1 kb, encompassing 50 bases upstream from position +1, the entire coding region, and 31 bases in the 3′ flanking area.

The nucleotide alterations associated with one of the mutants are illustrated in Fig. 3. *NAT2* from WBCs of a putative homozygous slow acetylator (sample 2) exhibited two single-base changes in the coding region—namely, a C → T transition (nt 282; Fig. 3 *Upper*) and a G → A transition (nt 590; Fig. 3 *Lower*). Both bases (C/T and G/A) were present at the respective position in *NAT2* from liver L10 (sample 7), clearly demonstrating heterozygosity. No difference was detected in *NAT2* from samples 3–6 relative to that from WBC sample 1; the latter originated from a person phenotyped *in vivo* as a homozygous rapid acetylator.



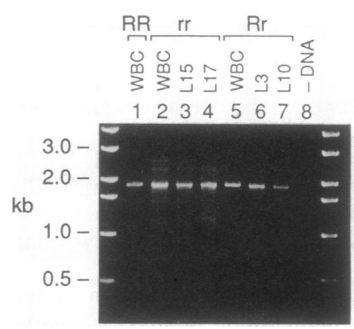


FIG. 2. Amplification and identification of *NAT2* from different acetylator phenotypes. (*Upper*) Forward (AP 1) and reverse (AP 2) amplification primers, made to the *Eco*RI sites (underlined) of *NAT2*, were mixed with WBC or liver DNA and *Taq* DNA polymerase, and 30 PCR cycles were carried out. (*Lower*) PCR products were separated on 1% agarose gels and visualized by staining with ethidium bromide. Mixtures with WBC DNA were loaded in lanes 1, 2, and 5, and those with liver DNA were loaded in lanes 3, 4, 6, and 7. DNA was omitted from the mixture in lane 8. The unidentified lanes at either end of the gel contain standard (1 μg of 1-kb ladder; BRL). A decrease in the amount of template DNA (from 1 to 0.5 μg) or in deoxynucleoside triphosphate concentration (from 0.2 to 0.05 mM) affected neither the yield nor the purity of the 1.9-kb DNA fragment. On the other hand, all nonspecific amplification products were eliminated when the concentration of the amplification primers was lowered from 0.5 to 0.1 μM (lane 5). Acetylator phenotypes: RR, homozygous rapid; rr, homozygous slow; and Rr, heterozygous.

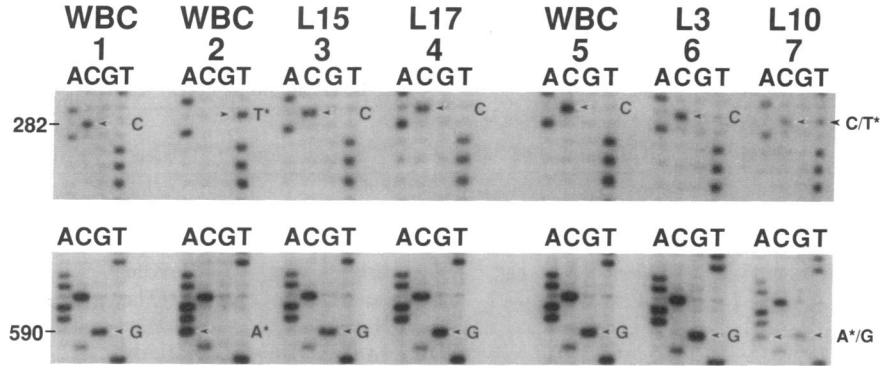


FIG. 3. Autoradiograms of sequencing gels with *NAT2* segments spanning nt 282 and 590. The sequences were determined with primers IP2 (*Upper*) and IP3 (*Lower*) and are in a 5′ to 3′ direction. The nucleotides at positions 282 (*Upper*) and 590 (*Lower*) are identified by arrowheads and letters. Asterisks indicate the nucleotides that differ from those of *NAT2* from the homozygous rapid acetylator (WBC sample 1). The heterozygote (liver L10, sample 7) is characterized by duplicate comigrating bands at both positions (282 and 590).

Multiple nucleotide substitutions were likewise observed in the coding region of a second *NAT2* mutant derived from liver tissue with low cytosolic NAT activity, as shown in Fig. 4 (liver L15; sample 3). These entailed T → C (nt 341; Fig. 4 *Top*), C → T (nt 481; Fig. 4 *Middle*), and T → C (nt 803; Fig. 4 *Bottom*) transitions. *NAT2* from liver samples L17 and L10 and WBC sample 5 displayed the typical heterozygote pattern of duplicate bands at all three positions. In contrast, *NAT2* from liver L3 (sample 6) was identical to that from WBC sample 1.

The amino acid changes arising from the aforementioned base alterations are summarized in Table 1, where the N-acetylation phenotypes are also compared to genotypes assigned on the basis of the direct sequence analysis. One of the *NAT2* structures identified originated from homozygous rapid acetylators (samples 1 and 6). Variant r_2r_2 was distinguished by a silent mutation and a G → A transition leading to replacement of Arg-197 by Gln (loss of a positive charge) and concomitant elimination of a *Taq* I site. Mutant r_3r_3 was characterized by a silent mutation resulting in obliteration of the sole *Kpn* I site and by two amino acid changes—i.e., replacement of Ile-114 by Thr and of Lys-286 by Arg. *NAT2* from liver L17 and WBC sample 5 exhibited heterozygosity for the rapid and one of the mutant alleles (r_3), but a heterozygote of the rapid and the other mutant allele (r_2) was not detected. *NAT2* from liver L10 had duplicate bands at all five nucleotide positions and was accordingly categorized as a compound heterozygote of both of the allelic variants (r_2r_3). As noted in Table 1, there was concordance between phenotype and genotype designations in four of seven cases.

The replacement of Ile-114 by Thr in mutant r_3r_3 represents a change from a nonpolar to an uncharged polar residue and is accompanied by a decrease in the hydropathy of NAT2 by 5.2 units (18). As a consequence, the hydropathy profile of the peptide segment flanking amino acid 114 markedly changed from hydrophobic to hydrophilic (Fig. 5). Neither the Arg for Lys-268 substitution nor that of Gln for Arg-197 is associated with a significant change in hydropathy.

DISCUSSION

The present study has shown that slow acetylator humans possess a gene for NAT2, which is in stark contrast to the absence of *NAT2* from slow acetylator rabbits (19). Secondly, two discrete allelic variants of human *NAT2* from homozygous slow acetylators were identified by comparison of the sequences to that of *NAT2* from a homozygous rapid acetylator. The 1-kb segments sequenced showed no differences at the 3′ splice acceptor site, but revealed that the last two-thirds of the protein-coding region of the two variants contained multiple nucleotide transitions, giving rise in each case to a silent mutation and one or two amino acid changes (Fig. 6, lines 1–3). These findings thus provide unambiguous evidence that human *NAT2* encodes the polymorphic NAT.

The G → A transition at position 857 of NAT cDNA, leading to replacement of Gly-286 by Glu and elimination of the sole *Bam*HI site (7) (Fig. 6, lines 1 and 4), was not seen in any of the *NAT2* samples examined. Although this does not rule out a *Bam*HI mutation among Caucasians, our results suggest that such a mutation should be infrequent in this ethnic group. Conversely, none of the mutations delineated herein was detected in Japanese (7). This and other ethnic differences (e.g., acetylator phenotype frequency) have prompted classification of NAT2 genes as "Caucasian" and "Oriental" (Fig. 6). Taking into account the relatively small sample size in the present investigation, the occurrence of additional mutations in *NAT2* from slow acetylators cannot be excluded.

Six genotype combinations are possible for one NAT2 gene with two allelic variants (Table 1, last column). Five of the six genotypes were identified in the present study despite the

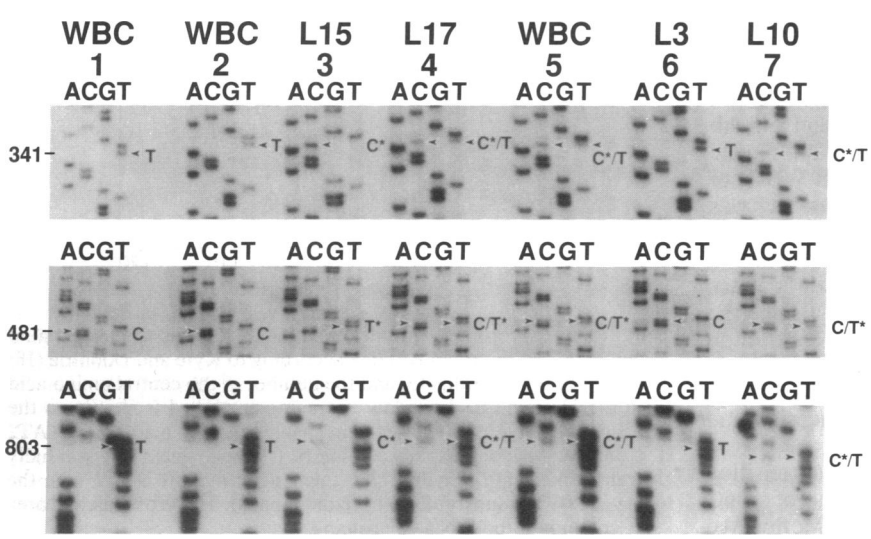


FIG. 4. Autoradiograms of sequencing gels with *NAT2* segments spanning nt 341 (*Top*), 481 (*Middle*), and 803 (*Bottom*). (*Top* and *Middle*) The sequences were determined with primer IP2 and proceed in a 5′ to 3′ direction. (*Bottom*) The segment was sequenced with primer IP4 and is in a 3′ to 5′ direction. Nucleotides differing from those of *NAT2* from the homozygous rapid acetylator (WBC sample 1) are marked with an asterisk. A pair of comigrating bands appears at all three nucleotide positions in *NAT2* from liver L17 (sample 4), WBC sample 5, and liver L10 (sample 7).

Table 1. Genotype designations based on direct sequencing of the gene for human polymorphic NAT2

Sample		N-Acetylation Activity			*NAT2* sequence*						
					Nucleotide			Deduced amino acid(s)			
No.	Tissue	Liver NAT†	*In vivo*‡	Phenotype	Position	Transition	Codon change	Position	Substitution	Present	Genotype designation§
1	WBC		8.0/0.95	RR							R_1R_1
6	Liver (L3)	17.0/0.78		Rr¶							
2	WBC		1.0/0.1	rr	282	C → T	TAC → TAT	94	None	Tyr/Tyr	r_2r_2
					590	G → A	CGA → CAA	197	Arg → Gln	Gln/Gln	
3	Liver (L15)	4.2/0.11		rr	341	T → C	ATT → ACT	114	Ile → Thr	Thr/Thr	r_3r_3
					481	C → T	CTG → TTG	161	None	Leu/Leu	
					803	A → G	AAA → AGA	268	Lys → Arg	Arg/Arg	
4	Liver (L17)	3.4/0.29		rr¶	341	T → T/C	ATT → ATT/ACT	114	Ile → Ile/Thr	Ile/Thr	R_1r_3
5	WBC		3.5/0.42	Rr	481	C → C/T	CTG → CTG/TTG	161	None	Leu/Leu	
					803	A → A/G	AAA → AAA/AGA	268	Lys → Lys/Arg	Lys/Arg	
7	Liver (L10)	16.4/0.76		Rr¶	282	C → C/T	TAC → TAC/TAT	94	None	Tyr/Tyr	r_2r_3
					341	T → T/C	ATT → ATT/ACT	114	Ile → Ile/Thr	Ile/Thr	
					481	C → C/T	CTG → CTG/TTG	161	None	Leu/Leu	
					590	G → G/A	CGA → CGA/CAA	197	Arg → Arg/Gln	Arg/Gln	
					803	A → A/G	AAA → AAA/AGA	268	Lys → Lys/Arg	Lys/Arg	

*Nucleotides, codons, and deduced amino acids preceding the arrows are those present in *NAT2* from a homozygous rapid acetylator (WBC sample 1). *NAT2* from liver L3 (sample 6) was identical to that from WBC sample 1 for the entire 1.0-kb segment sequenced. The mutated guanine at position 590 (variant r_2r_2) is part of a CpG dinucleotide and a *Taq* I restriction site. The sole *Kpn* I site in *NAT2* is obliterated as a result of the silent mutation at position 481 (variant r_3r_3).

†Shown are V_{max} values (in nmol per min per mg of protein) obtained with *p*-aminobenzoic acid and sulfamethazine as substrates, respectively. Liver samples from homozygous rapid acetylators gave average V_{max} values of 55.9 and 3.8 with the respective substrate (9).

‡The numbers shown are ratios calculated for urinary caffeine metabolites and for N-acetylated to unmetabolized dapsone in serum, respectively (10).

§These are informal designations. The letters refer to N-acetylation activity: R is for high activity, and r is for low activity. The numbers (1 to 3) were arbitrarily assigned to *NAT2* and the two allelic variants.

¶The phenotype ascribed to the indicated sample is in disagreement with the genotype ascertained by direct sequencing.

small sample size, lending credence to the concept that the mutations observed should be prevalent among Caucasians. On the other hand, a discrepancy was noted between acetylator phenotype of three of the samples and genotype ascertained by direct sequencing. Specifically, livers L3 and L10 had intermediate NAT activities that were virtually identical (Rr phenotype), yet showed genotypes R_1R_1 and r_2r_3, respectively. Such a disparity was also evident with liver L17 relative to L15 (Table 1). Inconsistencies of this nature are not surprising since classification of acetylator phenotype, determined *in vitro* or *in vivo*, relies on assays and drug tests with several variables (ref. 2, pp. 151–166) and, therefore, is not necessarily as accurate as genotype designation based on actual gene sequence. In this connection, Deguchi *et al.* (20) have also observed a similar discordance in two subjects with intermediate activity (phenotyped with isoniazid *in vivo*), whose genotype turned out to be homozygous rapid by restriction fragment length analysis.

Southern blot hybridization with a probe specific to the cDNAs thought to encode the genetically variant NAT (Fig. 6, lines 1 and 4) disclosed a restriction fragment length polymorphism involving two *Kpn* I sites 5.0 and 5.5 kb downstream from that in the coding region. The NAT gene with the 5.0-kb *Kpn* I fragment was derived from WBCs of slow acetylators phenotyped with isoniazid *in vivo* and was detected in 2 of 29 healthy Japanese (20). Caucasian *NAT2* variant r_3r_3 (Fig. 6, line 3) should likewise exhibit restriction fragment length polymorphism, but the banding pattern should differ from that seen by Deguchi *et al.* (20) since the coding region of mutant r_3r_3 lacks a *Kpn* I site. Restriction fragment length analysis would provide an additional reliable method for acetylator genotype assessment, as well as an expedient means of differentiating between Caucasian and Japanese slow NAT2 variants.

Of the three *NAT2* nucleotides that underwent a transition accompanied by an amino acid change, one was part of a CpG doublet and a *Taq* I site (nt 590; Fig. 6, line 1). Methylated cytosine in vertebrate DNA is present almost exclusively in CpG dinucleotides (21), which are at least 42 times more mutable within human coding sequences than is predicted by random mutation (11). In addition, restriction sites with the CpG dimer sequence (e.g., *Taq* I and *Msp* I) show a higher

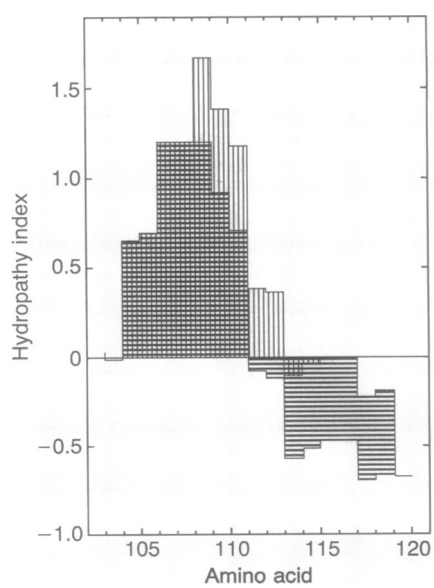


FIG. 5. Hydropathy profile of the NAT2 segment spanning amino acid 114. The hydropathy index according to Kyte and Doolittle (18) is plotted on the ordinate, and the number of the central amino acid in the 11-amino acid window averaging employed is shown on the abscissa. A hydrophobic segment spanning amino acid 114 in NAT2 from homozygous rapid acetylators (vertical lines) is essentially transformed into a hydrophilic segment in mutant NAT2 with the Ile-114 → Thr substitution (horizontal lines). The crosshatched area is common to both NAT2 proteins.

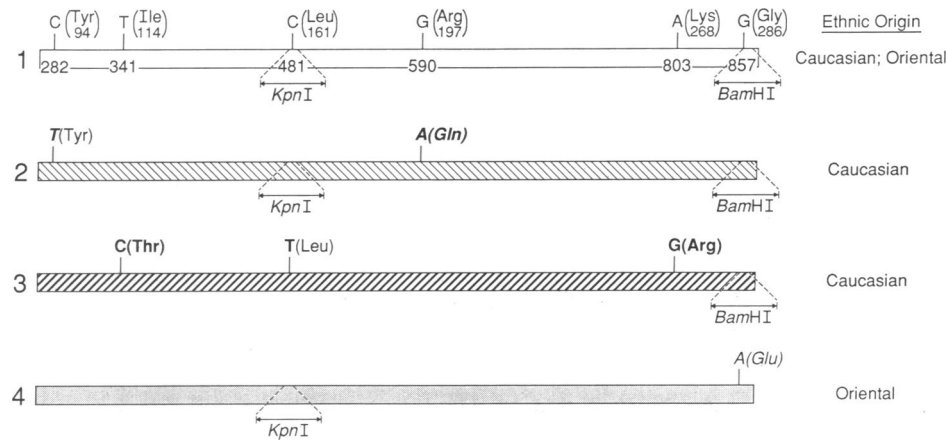


FIG. 6. Ethnic differences in the structure of genetically variant NATs. Lines 1–3 illustrate the 0.6-kb coding region segments (nt 270–870) of *NAT2* from homozygous rapid (line 1) and homozygous slow (lines 2 and 3) acetylators identified in this study by direct DNA sequencing. These are compared to the corresponding segments of NAT cDNAs 0-7 (line 1) and D-14 (line 4) reported by Ohsako and Deguchi (7). The positions of six mutable nucleotides and the identity of the deduced amino acid at the respective nucleotide position (in parentheses) are shown on line 1, as are the locations of *Kpn* I (nt 480–486) and *Bam*HI (nt 856–861) restriction sites (represented only once in *NAT2* and 0-7 cDNA). The guanine at position 590 is part of a CpG dinucleotide, and one of three coding region *Taq* I sites is also found at this location (nt 588–591; not shown). *NAT2* and 0-7 cDNA (line 1) are virtually identical over a total length of 1175 bp (nt −6 to 1170); they differ structurally only at three positions in the coding region, as follows (*NAT2* bases are listed before those in 0-7 cDNA): nt 438 (G instead of T), nt 450 (C instead of T), and nt 843 (C instead of A). None of these is associated with an amino acid change. Cytosol from Chinese hamster ovary cells transfected with 0-7 cDNA had a higher N-acetylation activity than did that from cells transfected with D-14 cDNA, suggesting that 0-7 (line 1) and D-14 (line 4) cDNAs may have originated from liver of rapid and slow acetylators, respectively.

frequency of restriction fragment length polymorphism (22). Collectively, these findings support the view that methylated cytosines are "hot spots" for mutation (11, 21, 22). The profound hypermutability of CpG dinucleotides encourages speculation that the G → A transition at position 590 may be the preponderant mutation among Caucasian *NAT2* variants.

One or both of the nucleotide transitions associated with the silent mutations (nt 282 and 481; Fig. 6, lines 1–3) and/or that which led to the conservative change (nt 803; Fig. 6, lines 1 and 3) may conceivably be unrelated to or play an auxiliary role in the human N-acetylation polymorphism. However, given the loss of a positive charge resulting from the Arg-197 → Gln substitution in variant r_2r_2 (Fig. 6, lines 1 and 2) and the large hydropathy decrease accompanying the Ile-114 → Thr change in variant r_3r_3 (Fig. 5), it seems reasonable to expect that either of these mutations could be deleterious to *NAT2* transcription or mRNA translation or lead to the production of NAT2 proteins with altered physicochemical properties. Perusal of the scant literature on this subject indicates that the single-nucleotide change in NAT cDNA (nt 857; Fig. 6, lines 1 and 4) may have caused an impairment in mRNA translation, or that the accompanying replacement of Gly-286 by Glu may have rendered the NAT protein more labile (7). In contrast, preliminary evidence suggests that the mutation associated with the 5.0-kb *Kpn* I fragment polymorphism may have resulted in defective transcription of the corresponding NAT gene (20). These possibilities will require evaluation by rigorous expression and site-directed mutagenesis studies.

This work was supported in part by Grants GM-27028, GM-44965, and 5T32 GM-07544 from the National Institutes of Health.

1. Weber, W. W. & Hein, D. W. (1985) *Pharmacol. Rev.* **37**, 25–79.
2. Weber, W. W. (1987) *The Acetylator Genes and Drug Response* (Oxford Univ. Press, New York).
3. Andres, H. H., Klem, A. J., Schopfer, L. M., Harrison, J. K. & Weber, W. W. (1988) *J. Biol. Chem.* **263**, 7521–7527.
4. Andres, H. H., Vogel, R. S., Tarr, G. E., Johnson, L. & Weber, W. W. (1987) *Mol. Pharmacol.* **31**, 446–456.
5. Evans, D. A. P. (1989) *Pharmacol. Ther.* **42**, 157–234.
6. Blum, M., Grant, D. M., Demierre, A. & Meyer, U. A. (1989) *Nucleic Acids Res.* **17**, 3589.
7. Ohsako, S. & Deguchi, T. (1990) *J. Biol. Chem.* **265**, 4630–4634.
8. Blum, M., Grant, D. M., McBride, W., Heim, M. & Meyer, U. A. (1990) *DNA* **9**, 193–203.
9. Kilbane, A. J., Petroff, T. & Weber, W. W. (1991) *Drug Metab. Dispos.* **19**, 503–507.
10. Kilbane, A. J., Silbart, L. K., Manis, M., Beitins, I. Z. & Weber, W. W. (1990) *Clin. Pharmacol. Ther.* **47**, 470–477.
11. Cooper, D. N. & Youssoufian, H. (1988) *Hum. Genet.* **78**, 151–155.
12. Levy, G. N. & Weber, W. W. (1988) *Anal. Biochem.* **174**, 381–392.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J., eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic, New York).
15. McGuire, M. C., Nogueira, C. P., Bartels, C. F., Lightstone, H., Hajra, A., Van Der Spek, A. F. L., Lockridge, O. & La Du, B. N. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 953–957.
16. Ohsako, S., Ohtomi, M., Sakamoto, Y., Uyemura, K. & Deguchi, T. (1988) *J. Biol. Chem.* **263**, 7534–7538.
17. Martell, K. J., Vatsis, K. P. & Weber, W. W. (1990) *FASEB J.* **4**, 753A (abstr.).
18. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
19. Blum, M., Grant, D. M., Demierre, A. & Meyer, U. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9554–9557.
20. Deguchi, T., Mashimo, M. & Suzuki, T. (1990) *J. Biol. Chem.* **265**, 12757–12760.
21. Bird, A. P. (1980) *Nucleic Acids Res.* **8**, 1499–1504.
22. Barker, D., Schafer, M. & White, R. (1984) *Cell* **36**, 131–138.