0026-895X/00/020288-12$3.00/0
MOLECULAR PHARMACOLOGY

MOL 58:288–299, 2000 /13105/828852



# Characterization of Naturally Occurring and Recombinant Human *N*-Acetyltransferase Variants Encoded by *NAT1**

JESÚS H. de LEÓN,[1] KOSTAS P. VATSIS,[2] and WENDELL W. WEBER

*Department of Pharmacology, Medical School, The University of Michigan, Ann Arbor, Michigan*

Received November 17, 1999; accepted March 8, 2000

This paper is available online at http://www.molpharm.org

## ABSTRACT

The genotype at the *NAT1** locus of an interethnic population of 38 unrelated subjects was determined by direct sequencing of 1.6-kb fragments amplified by PCR. The coding exon alone and together with the 3′ noncoding exon of the wild-type (*NAT1*4*) and the three mutant alleles (*NAT1*10*, **11*, and **16*) detected was expressed in *Escherichia coli* and COS-1 cells, respectively, and the cytosolic fraction of mononuclear leukocytes from *NAT1*4/*4* and *NAT1*10/*10* homozygotes was also isolated. Recombinant and leukocyte cytosolic preparations were thoroughly characterized by *N*-acetylation activity with several NAT1-specific and -selective substrates, as well as by steady-state kinetics with varying amounts of the substrate (fixed acetyl CoA) and acetyl CoA (fixed substrate), thermodynamics, stability, and protein immunoreactivity with a polyclonal human anti-NAT1. The polyadenylation signal mutation in the 3′ noncoding sequence of *NAT1*10* affected none of the aforementioned parameters evaluated both with recombinant *NAT1*10* and with the naturally occurring allele. Function was also unaffected by the coding and 3′ noncoding exon mutations in *NAT1*11*. In contrast, the three extra adenosines located immediately after the sixth position of the polyadenylation signal in the 3′ untranslated region of *NAT1*16* ostensibly caused disruption of the predicted secondary structure of the pre-mRNA for NAT1 16, culminating in parallel 2-fold decreases in the amount and catalytic activity of NAT1 16 in COS-1 cell cytosol. This novel finding in *N*-acetylation pharmacogenetics clearly demonstrates a direct link between reduced catalytic activity and structural alteration in the 3′ untranslated region of an *NAT* variant (*NAT1*16*) brought about by mutation.

The isolation of two functional *NAT** loci in humans a decade ago spurred a litany of molecular genetic and biochemical studies on NAT polymorphisms. The two loci, *NAT1** and *NAT2**, are independently regulated and encode proteins with 81% deduced amino acid sequence similarity. Both NAT1 and NAT2 catalyze *N*-acetylation (detoxification) and *O*-acetylation (activation) reactions, and possess selectivity with various therapeutic agents and with carcinogenic and mutagenic compounds. In *N*-acetylation reactions, NAT1 has specificity for the drugs 4(*p*)-aminosalicylate and 4-aminobenzoate and selectivity for several carbocyclic arylamine carginogens occurring as environmental/occupational pollutants (e.g., 2-aminofluorene, 4-aminobiphenyl, and benzidine), whereas NAT2 has a preference for sulfamethazine. For *O*-acetylation and *N*,*O*-transacetylation reactions, NAT1 is selective with carbocyclic arylamines and NAT2 with dietary heterocyclic amine mutagens. Overall, the data indicate that NAT1 is a much more efficient and effective catalyst than NAT2 (cf. Vatsis et al., 1995; Grant et al., 1997; Vatsis and Weber, 1997).

The initial focus on *NAT2** soon generated unequivocal evidence for association of this locus with the human genetic variability that had been amply documented with isoniazid and sulfamethazine, and typically illustrated as a bimodality, and often trimodality, in the distribution of populations for N-acetylation of these compounds. The pronounced variation and ethnic distribution of *NAT2** alleles have also been thoroughly investigated, and, in general, *NAT2** genotypic-NAT2 phenotypic indices are reasonably well correlated. Of the 26 human *NAT2** allelic variants known to date, five to

This work was supported by Grants GM44965 and CA39018 from the National Institutes of Health. The data are taken from a thesis submitted by Jesús H. de León to the Rackham School of Graduate Studies, The University of Michigan, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Pharmacology. Preliminary reports of this study have been presented at the annual meetings of the American Society of Human Genetics [*Am J Hum Genet* **55(suppl):**A340, 1994] and of Experimental Biology (*FASEB J* **10**:A456, 1996), as well as at the North American meeting of the International Society for the Study of Xenobiotics (*ISSX Proc* **12**:97, 1997).

[1] Present address: Health Effects Laboratory Division, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Morgantown, WV 26505-2845.

[2] Present address: Department of Biological Chemistry, The University of Michigan Medical School, M5440 Medical Science Building I, Ann Arbor, MI 48109-0606.

**ABBREVIATIONS:** NAT(s), *N*-acetyltransferase(s). NAT genes (*NAT1**; *NAT2**), allelic variants (e.g., *NAT1*4*; *NAT2*4*), and proteins (NAT1 4; NAT2 4) are designated in accordance with published nomenclature guidelines (Vatsis et al., 1995). The substituent at the *para* position of the substrates is interchangeably denoted by number (4-) or letter (*p*-). The designations T (thymidine) and U (uridine) also appear interchangeably. PCR, polymerase chain reaction; 3′ UTR, 3′ untranslated region; nt, nucleotide(s); bp, base-pair(s); kb, kilobase(s).

six remarkably account for almost all of the variation seen in diverse ethnogeographic populations [reviewed in Vatsis and Weber (1997) and in Grant et al. (1997)].

Although a wealth of publications had shown appreciable differences in *N*-acetylation of 4-aminobenzoate and 4-aminosalicylate [up to 90-fold in vitro and more than 100-fold in vivo (cf. Vatsis and Weber, 1994)], the human acetylase with activity toward these compounds was long regarded as genetically invariant (Weber, 1987) because the existence of two independently regulated NATs with discrete substrate selectivities had not yet been discovered. Strong indirect evidence for an acetylation polymorphism at the human *NAT1** locus was provided by disclosure of *NAT1** structural heterogeneity (Vatsis and Weber, 1993), and demonstration of a tendency toward bimodality in the distribution in vivo (Grant et al., 1992) and in vitro (Weber and Vatsis, 1993) of established NAT1-specific substrates (Grant et al., 1991). The evidence was indirect in that *NAT1** allelic variation was observed with persons of unknown NAT1 phenotype (Vatsis and Weber, 1993; Vatsis et al., 1994), and, conversely, NAT1 phenotypic variation was determined with subjects whose *NAT1** genotype had not been ascertained (Grant et al., 1992; Weber and Vatsis, 1993; Vatsis and Weber, 1994). These concerns have begun to be addressed, as evidenced by three very recent studies on direct phenotypic-genotypic correlations: elimination of 4-aminosalicylate in vivo and *N*-acetylation by whole blood and cytosol of heterologously expressed alleles from individuals of defined *NAT1** genotype (Hughes et al., 1998); 4-aminosalicylate *N*-acetylation by several structurally characterized recombinant NAT1 variants (Lin et al., 1998); and 4-aminobenzoate *N*-acetylation by whole blood from 85 unrelated subjects of established *NAT1** genotype, with 8% of the bimodally distributed population displaying the slow NAT1 phenotype (Butcher et al., 1998).

In the present study, we have found a prevalence of wild-type *NAT1*4* and mutant *NAT1*10*, and a considerably lower distribution frequency for mutant *NAT1*11* and *NAT1*16* alleles, identified by direct sequencing of 1.6-kb fragments of PCR-generated *NAT1** from an interethnic group of 38 unrelated subjects. Heterologous expression in bacterial and mammalian cells and rigorous biochemical and immunochemical characterization of the recombinant alleles, as well as of cytosol from mononuclear leukocytes of *NAT1*4/*4* and *NAT1*10/*10* homozygotes, revealed no major dissimilarities between wild-type *NAT1*4* and mutants *NAT1*10* (polyadenylation signal mutation) and *NAT1*11* (coding and 3′ noncoding exon mutations). In *NAT1*16*, an AAA insertion on the 3′ side of the AATAAA signal resulted in a 2-fold decrease in immunoreactivity and a commensurate decline in *N*-acetylation activity of NAT1 16. This finding constitutes the first example of a causative relationship between NAT1 phenotype and a mutationally modified sequence in the 3′ untranslated region of an *NAT** variant.

## Experimental Procedures

### Materials

Restriction endonucleases and other DNA-modifying enzymes were obtained from GibcoBRL/Life Technologies (Bethesda, MD), New England Biolabs (Beverly, MA), Promega (Madison, WI), and Boehringer Mannheim (Indianapolis, IN). Wizard PCR Preps DNA purification system was obtained from Promega, SeaPlaque GTG agarose from FMC Corporation (Rockland, ME), [$\alpha$-$^{32}$P]dATP and the Sequenase version II kit from Amersham (Arlington Heights, IL), and Immobilon-P transfer membranes from Millipore (Bedford, MA). USB Specialty Biochemicals was the source of inorganic pyrophosphatase, and of 7-deaza-dGTP and dITP termination nucleotide mixes. The pET expression system, which includes the *pET-17b* vector and *Escherichia coli* strains Novablue and BL21(DE3)*pLysS*, was purchased from Novagen (Madison, WI). Prokaryotic and eukaryotic cell culture media were from Difco (Detroit, MI) and GibcoBRL, respectively, the eukaryotic TA cloning kit from Invitrogen (San Diego, CA), fetal bovine serum from Hyclone (Logan, UT), the $\beta$-galactosidase enzyme assay system from Promega, $^{125}$I-conjugated goat anti-rabbit secondary antibody from DuPont NEN (Boston, MA), Biomax MR film from Eastman Kodak Company (Rochester, NY), and the protein assay and silver stain kits from Bio-Rad (Hercules, CA). Antibodies and molecular biology grade chemicals and buffers for recombinant DNA experiments were from Sigma (St. Louis, MO). Oligonucleotide primers were made at the University of Michigan DNA Synthesis Facility and purified by HPLC. The *pGEX-KN* vector (Hakes and Dixon, 1992) and *E. coli* strain BL21 were gifts from Dr. Jack E. Dixon of the Department of Biological Chemistry at this institution. *E. coli* strain JM105 harboring the plasmid *pKK2233-NAT2*4* was kindly donated by Dr. David W. Hein, Department of Pharmacology and Toxicology, University of Louisville School of Medicine.

Vacutainer tubes were purchased from Becton-Dickinson (Rutherford, NJ). The red blood cells lysis buffer and phosphate-buffered saline, as well as the Polymorphprep density gradient medium, a sterile solution of 13.8% (w/v) sodium metrizoate and 8.0% (w/v) dextran 500, were obtained from GibcoBRL.

Leupeptin, phenylmethylsulfonylfluoride, acetyl-DL-carnitine, carnitine acetyltransferase, and acetyl CoA were purchased from Sigma. Arylamine substrates, two arylamide derivatives [phenacetin (*p*-ethoxyacetanilide) and 2-acetamidofluorene], and (+)amethopterin were from Aldrich (Milwaukee, WI). The acetylated derivative of *p*-anisidine (*p*-methoxyacetanilide) was obtained from Pfaltz and Bauer (Waterbury, CT), 4-acetamidobenzoate was from Eastman Kodak, and 5-acetamidosalicylate was a generous gift from Dr. Marshall Montrose (The Johns Hopkins University School of Medicine, Baltimore, MD). Acetylated 4-aminosalicylate was synthesized by one of us (K. P. Vatsis) with the help of Dr. Alfin D. N. Vaz of the Department of Biological Chemistry at this institution, and the purity of the compound was ascertained by gas chromatography-mass spectrometry and HPLC.

### Subjects

A total of 26 healthy and unrelated volunteers (19 males; 7 females) were recruited from the graduate student body, postdoctoral trainees, faculty, and staff at the University of Michigan Medical School. To avoid genetic "ad-mixtures" as much as possible (Lin et al., 1993), the ethnically diverse population studied consisted of individuals born in Japan ($n$ = 5), Korea ($n$ = 5), mainland China ($n$ = 6), India ($n$ = 5), and Latin America (Mexico and South America; $n$ = 5). Written consent was obtained from all volunteers, and the experimental procedures were approved by the University Human Subjects Committee.

Blood samples (30–40 ml) were collected in yellow-capped vacutainer tubes at the outpatient clinics of the University of Michigan Hospital. The blood was submitted to density gradient centrifugation for isolation of mononuclear and polymorphonuclear leukocytes, from which genomic DNA and the cytosolic fraction were prepared for *NAT1** amplification by PCR and for catalytic activity determinations, respectively.

### Methods

**Amplification and Direct Sequencing of *NAT1**.** The separation of mononuclear and polymorphonuclear leukocytes from whole

human blood is outlined toward the end of the *Experimental Procedures* section. Separated white cells corresponding approximately to 40% of the blood sample were mixed well with 4 volumes of phosphate-buffered saline, and centrifuged at 12,000*g* for 15 min (4°C) in a fixed angle rotor. The pellets were resuspended in an equal volume of 10 mM Tris-HCl buffer (pH 7.5) containing 1 mM EDTA, and combined. Genomic DNA template was extracted from the pooled white blood cells (Vatsis and Weber, 1993), and the size and intactness of the preparation verified by electrophoresis on 1% agarose gels.

Amplification of *NAT1** from human DNA and direct sequencing of the entire 1.6-kb amplified fragment (nt −440 to 1175), as well as of 0.87-kb (nt 1–870) and 1.2-kb (nt 1–1175) fragments of recombinant *NAT1**, were carried out by previously published procedures (Vatsis et al., 1991; Vatsis and Weber, 1993).

To eliminate the marked gel compression originally observed within a palindromic sequence (nt 446–464) of *NAT1*11* from persons heterozygous for this allele (Vatsis and Weber, 1993), two strategies were adopted as recommended in the "Nucleic Acid Sequencing" manual from USB Specialty Biochemicals. 1) The segment encompassing the palindrome was sequenced in both directions with three different sets of deoxy/dideoxy nucleotide solutions containing dGTP or one of its two analogs, namely, 7-deaza-dGTP or dITP. Reaction mixtures consisted of experimentally determined optimal concentrations of the deoxynucleotide (40 $\mu$m) and dideoxynucleotide (4 $\mu$m) triphosphates, pyrophosphatase in a 1:3000 molar ratio with sequenase, and the additional components described previously (Vatsis et al., 1991). 2) Electrophoresis of the samples was done on regular 7% polyacrylamide-7 M urea gels, as well as under stronger denaturing conditions effected by addition of 40% formamide to the sequencing gel.

**Heterologous Expression of Human *NAT1** Alleles: Transformation of *E. coli*.** *pCR3-NAT1*4* and *pCR3-NAT1*11* recombinant plasmids (see COS-1 cells below) were digested with *Bam*HI and *Eco*RI to release the *NAT1*4* and **11* inserts, which were ligated with similarly cleaved and dephosphorylated *pET-17b* prokaryotic expression vector with a T7 promoter, and electroporated into *E. coli* cloning host Novablue. The recombinant plasmids *pET-17b-NAT1*4* and *pET-17b-NAT1*11* were mapped with *Ssp*I to confirm *NAT1** insert orientation, and sequenced to establish unambiguously the identity of the *NAT1** alleles. The T7 tag of the *pET-17b* vector was excised by digestion with *Nde*I and *Bam*HI, blunt ends were generated by treatment with Mung bean nuclease, and the plasmids were electroporated into *E. coli* expression host BL21(DE3)*pLysS* as described in the Bio-Rad instruction manual.

*E. coli* BL21(DE3)*pLysS* cells harboring recombinant plasmids *pET-17b-NAT1*4* and *pET-17b -NAT1*11* were grown at 37°C (to an $A_{600}$ of 0.6) in Luria-Bertani medium containing ampicillin (100 $\mu$g/ml) and chloramphenicol (25 $\mu$g/ml), treated with 0.4 mM isopropylthio-$\beta$-galactoside and 11 mM glucose, and incubated at 37°C until maximal induction of NAT1 proteins was achieved (experimentally determined to be 7 h). Cultures were centrifuged at 4,000*g* for 5 min (4°C), and the pellets were resuspended in a solution containing 3 mM Tris-HCl buffer (pH 7.8 at room temperature), 2 mM EDTA, 2 mM dithiothreitol, 0.02 mM leupeptin, and 0.1 mM phenylmethylsulfonylfluoride. The bacterial cells were disrupted by sonication on ice, and the homogenate was submitted to differential centrifugation (17,000*g*, 30 min; 100,000*g*, 60 min) for isolation of the 100,000*g* supernatant (cytosolic) fraction.

**Heterologous Expression of Human *NAT1** Alleles: Transient Transfection of COS-1 Cells.** The coding exon and 3′ UTR of *NAT1** (nt 1–1175) were amplified with template DNA from *NAT1*4/*11* and *NAT1*10/*16* heterozygotes; the sense and antisense amplification primers contained *Bam*HI and *Eco*RI restriction sites, respectively, to facilitate transfer of *NAT1** inserts to other vectors. The PCR products were ligated directly into eukaryotic expression vector *pCR3* with the TA cloning system, yielding recombinant plasmids *pCR3-NAT1*4*, *pCR3-NAT1*10*, *pCR3-NAT1*11*, and *pCR3-NAT*16*, which were screened for the presence of insert by the procedure of Akada (1994). Positive colonies on 0.8% agarose gels were identified by size, insert orientation in positive clones was verified by mapping with *Bsa*I, and the identity of the *NAT1** alleles was conclusively affirmed by sequencing of the 1.2-kb fragments. Since each *NAT1** insert contained the 3′ UTR with its own polyadenylation signal, the bovine growth hormone polyadenylation signal of the *pCR3* vector was removed by cleavage with *Not*I and *Afl*III. Blunt ends were generated with T4 DNA polymerase, the recombinant vectors were submitted to electrophoresis on 1% SeaPlaque GTG agarose gels, and were then ligated and electroporated in *E. coli* cloning host Top10 *F* as described in the instruction booklet from FMC Corporation. The reporter vector *pCR3βGal* was constructed by ligation of the *E. coli* $\beta$-galactosidase gene with the *Not*I restriction site in the polylinker region of previously dephosphorylated *pCR3* vector, and was mapped with *Eco*RV to confirm orientation. All recombinant plasmids were purified by double-banding CsCl gradient centrifugation (Sambrook et al., 1989).

COS-1 cells were acutely transfected by the calcium phosphate precipitation technique (Chen and Okayama, 1987). Subconfluent cells were cotransfected (10-cm plates) with 16 $\mu$g of the respective *NAT1** recombinant plasmid and 4 $\mu$g of *pCR3βGal* in 1 ml of a solution containing 25 mM *N*,*N*-*bis*(2-hydroxyethyl)-2-aminoethanesulfonic acid buffer, 140 mM NaCl, 1 mM sodium phosphate, and 250 mM $CaCl_2$ (pH 7.0); cells were harvested 2 to 3 days after transfection. Subsequent disruption of the cells by sonication, as well as resuspension in dilute Tris-HCl buffer and differential centrifugation for isolation of the cytosolic fraction, were exactly as described above for transformed *E. coli*.

**Preparation of Polyclonal Antibody to Human NAT1.** PCR-generated *NAT1*4* (nt 1–1175) with *Not*I and *Eco*RI restriction sites at the 5′ and 3′ ends, respectively, was extracted with phenol/chloroform, digested with *Not*I and *Eco*RI for 20 h, and purified by agarose gel electrophoresis in conjunction with the Wizard PCR Preps DNA Purification System. Digested and gel-purified *NAT1*4* was ligated in the *Not*I and *Eco*RI sites of previously dephosphorylated *pGEX-KN* prokaryotic expression vector, which contains the glutathione *S*-transferase gene from *Schistosoma japonicum* as a fusion partner. The recombinant plasmid (*pGEX-KN-NAT1*4*) was mapped with *Pst*I to confirm insert orientation, and the identity of *NAT1*4* was established unequivocally by sequencing.

Host *E. coli* BL21 cells were made competent by the $CaCl_2$ method, transformed with *pGEX-KN-NAT1*4* by standard procedures (Sambrook et al., 1989), and grown at 37°C in 2× YT medium containing ampicillin (100 $\mu$g/ml) to an $A_{600}$ of 1.2. After cooling to room temperature, cultures were treated with 0.3 mM isopropylthio-$\beta$-galactoside, incubated at room temperature for 17 h, and centifuged; the pellets were resuspended in phosphate-buffered saline containing 1% (v/v) Triton X-100 and lysozyme (1 mg/ml buffer), placed on ice for 30 min, and stored frozen overnight. After thawing in cool water, the cells were disrupted by sonication on ice, and the homogenate was centrifuged at 8,000*g* for 10 min (4°C). The supernatant fraction was then applied to a glutathione Sepharose affinity column, and the column was eluted with glutathione. The eluate was assayed for *N*-acetylation activity with 4-aminobenzoate as substrate, and for NAT1 protein content by electrophoresis and silver staining of the gel. The purified glutathione *S*-transferase-NAT1 4 fusion protein was sent to Organon Teknika Corporation (West Chester, PA) for production of anti-NAT1 serum in rabbits.

The antiserum was found to recognize both human NAT1 4 and NAT2 4 expressed in *E. coli*, as expected from the high extent of deduced amino acid sequence similarity (81%) between these proteins. Despite an almost identical molecular mass, NAT1 4 and NAT2 4 migrated with an electrophoretic mobility corresponding to 34 and 31 kDa, respectively, as previously observed by others (Grant et al., 1991; Dupret and Grant, 1992; Dupret et al., 1994).

**Preadsorption of anti-NAT1 Serum to *E. coli* Lysates.** Nonspecific immunoreactive components were eliminated by adsorption

of human NAT1 antiserum to bacterial lysates before immunoblotting. To this end, *E. coli* strain BL21(DE3)*pLysS* bearing expression vector *pET-17b* without NAT insert was grown overnight in Luria-Bertani medium in the presence of ampicillin (100 $\mu$g/ml). The culture was centrifuged at 4,000*g* for 15 min (4°C), and pellets were resuspended in 5 mM Tris-HCl buffer (pH 7.5 at 37°C) containing 3 mM EDTA, 2 mM dithiothreitol, 0.02 mM leupeptin, and 0.1 mM phenylmethylsulfonylfluoride. Native and heat-denatured *E. coli* suspensions (1 mg of protein/ml) were added to a solution of 0.2% Tween 20 in 0.15 M NaCl-0.02 M Tris-HCl buffer (pH 7.4) containing appropriately diluted antiserum, the mixtures were rotated overnight at room temperature, centrifuged at 8,000*g* for 10 min (4°C), and the supernatant fraction submitted to immunoblotting.

**Immunoblotting.** Cytosol from transformed bacteria and transfected COS-1 cells was submitted to electrophoresis on 12% polyacrylamide gels in the presence of sodium dodecyl sulfate, and electrophoretically transferred to Immobilon-P membranes. The membranes were incubated for 1 h at room temperature in 0.15 M NaCl-0.02 M Tris-HCl buffer (pH 7.4) containing 2% bovine serum albumin and 1 mM EDTA, the solution was supplemented with appropriately diluted NAT1 antiserum that had been previously adsorbed, and the incubation was continued overnight at 4°C. After washing with 0.15 M NaCl-0.02 M Tris-HCl buffer (pH 7.4), the filters were incubated for 1 h with goat anti-rabbit $^{125}$I-conjugated secondary antibody, and exposed to film for 18 to 20 h at −70°C. Quantitation of NAT1 bands on autoradiograms was achieved by densitometric scanning (Bio-Rad model GS-670 Imaging Densitometer and accompanying software). Without exception, the amount of protein expressed was obtained from the slope of the line for the relative band density (in arbitrary units) plotted against different amounts of cytosolic protein to ensure the direct proportionality of these two parameters.

**Isolation of Human White Blood Cell Types and Preparation of Cytosol.** Anticoagulated whole human blood (3–5 ml) was layered at room temperature over an equal volume of sodium metrizoate-dextran density gradient medium (1:1.7; Polymorphprep solution). The samples were centrifuged at 500*g* for 20 to 30 min (20°C) in the swinging bucket rotor of a tabletop centrifuge, resulting in separation of the leukocytes into two clearly demarcated bands: an upper band consisting of mononuclear leukocytes and a lower band of polymorphonuclear leukocytes; the erythrocytes were pelleted. Mononuclear and polymorphonuclear white cells from approximately 40% of the blood sample were processed for template DNA isolation (described above). The remainder was transferred to separate sterile tubes, contaminating erythrocytes were removed by treatment with an equal volume of red blood cell lysing buffer for 5 min at room temperature (hypotonic lysis), and the white cells were washed twice with an equal volume of phosphate-buffered saline. The leukocytes were pelleted each time by centrifugation at 5,000*g* for 10 min (4°C) in a fixed angle rotor, and were finally resuspended in a solution of 3 mM Tris-HCl buffer (pH 7.78 at room temperature), 2 mM EDTA, 2 mM dithiothreitol, 0.02 mM leupeptin, and 0.1 mM phenylmethylsulfonylfluoride.

Leukocytes were disrupted by sonication on ice, the clear homogenate was centrifuged at 12,000*g* for 30 min (4°C), and the supernatant fraction from this step was centrifuged again at 100,000*g* for 1 h (4°C). The final supernatant fraction (cytosol) was the source of enzyme for catalytic, steady-state kinetic, and thermostability determinations.

**Catalytic Properties and Thermostability of Human White Blood Cell and Recombinant NAT1 Variants.** *N*-acetylation reactions were initially optimized for buffer type (Tris versus triethanolamine) and concentration, and the effect on catalytic activity of various agents such as KCl, dimethyl sulfoxide, bovine serum albumin, and butylated hydroxytoluene (Cribb et al., 1991; Grant et al., 1991; Dupret and Grant, 1992; Dupret et al., 1994; Hughes et al., 1998) was assessed. Since, in accordance with the ping-pong reaction mechanism of NAT, initial rates and steady-state kinetic constants will vary with the acetyl CoA concentration, catalytic experiments were done with the cofactor at 0.1 mM to enable comparison of the results with those in the literature (Grant et al., 1991; Dupret and Grant, 1992; Dupret et al., 1994; Hughes et al., 1998). Saturating amounts of the substrates were determined from kinetic experiments, and were 0.05 mM for 5-aminosalicylate; 0.1 mM for 4-aminosalicylate, 4-aminobenzoate, and 2-aminofluorene; 0.25 mM for *p*-phenetidine and *p*-anisidine; and 0.5 mM for sulfamethazine. 2-Aminofluorene was prepared in 1% methanol, with the solvent at a final concentration of 0.1% in the reaction mixtures (0.1% methanol was found to have no effect on 4-aminobenzoate and 4-aminosalicylate *N*-acetylation activities); all other substrates were aqueous solutions. The linearity of the reactions with cytosolic protein and time was established in comprehensive pilot experiments with each of the substrates and variants of heterologously expressed and human leukocyte NAT1.

Reaction mixtures (in triplicate) for determination of initial rates consisted of the cytosolic fraction from transformed *E. coli* (1–2 $\mu$g of protein), transfected COS-1 cells (1 $\mu$g of protein), or human leukocytes (10 $\mu$g of protein), saturating substrate (see preceding paragraph), an acetyl CoA regenerating system (containing 4.5 mM acetyl-DL-carnitine and 0.02 unit of carnitine acetyltransferase), 0.1 mM acetyl CoA, 2 mM EDTA, 2 mM dithiothreitol, and 20 mM Tris-HCl buffer (pH 7.5 at 37°C) in a total volume of 0.1 ml. Reactions were initiated with the substrate, carried out at 37°C for 5 to 10 min, and terminated with 0.01 ml of 15% perchloric acid. Zero-time controls contained all the components, except that the mixtures were quenched with perchloric acid before addition of substrate. After precipitation of the denatured protein, aliquots of the supernatant fraction were assayed for *N*-acetylated product formation by HPLC, as detailed below.

In steady-state kinetic experiments, the substrate was varied from about 0.1 to 10 × $K_m$, and acetyl CoA was held constant at 0.1 mM. Kinetic parameters were also determined with varying amounts of the cofactor in the presence of saturating (0.1 mM) 4-aminosalicylate. The concentration range for the variable component, reaction time, and amount of cytosolic protein are given in the appropriate table. Kinetic constants were derived from Lineweaver-Burk double reciprocal plots and also computed with the Cleland Hyper program (Cleland, 1967).

The *N*-acetylation of 4-aminosalicylate by NAT1 variants expressed in bacterial cells was investigated at temperatures ranging from 4.6 to 39.4°C, and the energy of activation for this reaction was derived from the slope of Arrhenius plots of the data (Segel, 1975).

The catalytic activity of NAT1 isozymes in cytosol from *E. coli* and from human mononuclear leukocytes was assessed with subsaturating and saturating concentrations of 4-aminosalicylate, respectively, in the absence and presence of varying amounts of amethopterin, a competitive inhibitor of NAT1-catalyzed *N*-acetylation of 4-aminobenzoate (Ward et al., 1995). The $K_i$ values were obtained directly from Dixon plots of the data (Segel, 1975).

In thermostability experiments, cytosol from transformed *E. coli* and from human mononuclear leukocytes was incubated for varying time periods at 37°C. Aliquots were removed at specified intervals and assayed simultaneously for NAT1 immunoreactivity and catalytic competence; *N*-acetylation activity was determined with 4-aminosalicylate as substrate, and reactions were carried out at 30°C to slow down the inactivation process.

**Quantitation of *N*-Acetylated Products by HPLC.** Parent compounds and arylamide metabolites were separated on a reversed phase $C_{18}$ column (Waters Nova-Pak $C_{18}$: 3.9 × 150 mm; 5-$\mu$m particle size), eluted at a flow rate of 1.0 ml/min under isocratic conditions. Elution was effected with a mobile phase of 0.1% aqueous trifluoroacetic acid/methanol (for 4-acetamidobenzoate, 4-acetamidosalicylate, and 5-acetamidosalicylate), 0.5% aqueous trifluoroacetic acid/methanol (for phenacetin and *p*-methoxyacetanilide), or 0.5% aqueous trifluoroacetic acid/methanol/acetonitrile (for 2-acetamidofluorene). The composition of the mobile phase was adjusted so that

the retention time of the *N*-acetylated products was between 5.5 and 7 min. The UV/visible detector was set at the wavelength maximum of the arylamide metabolite, determined by spectral scanning with a Cary model 3E double beam spectrophotometer.

The *N*-acetylated products were quantitated by comparison of the integrated area of the elution peak with that of known amounts of authentic arylamide standard. All analyses were performed with a Waters HPLC system consisting of a model 600E solvent delivery system, a model 486 absorbance detector, a model 717plus autosampler, and the Millennium 2010 chromatography manager computer software.

**Analytical and Other Methods.** Protein was determined by a dye-binding method with bovine serum albumin as the standard (Bradford, 1976). Statistical analysis was done by the unpaired Student's *t* test with the aid of the InStat program from GraphPad (San Diego, CA).

## Results

### Clarification of the Mutations in *NAT1*11*

**Coding Region.** A pronounced band compression resulting in the appearance of comigrating bands in adjacent lanes had been observed for palindromic nt 446–464 in *NAT1*11* from subjects heterozygous for this allele (e.g., *NAT1*4/*11*), making it impossible to decide whether these were true heterozygosities or DNA secondary structure artifacts [see Table 1, footnote c, in Vatsis and Weber (1993)]. The compressed area in one of the original *NAT1*11* samples (Vatsis and Weber, 1993) was fully disrupted upon sequencing in either direction with 7-deaza-dGTP or dITP, two dGTP analogs which form weaker base pairs that are readily denatured for polyacrylamide gel electrophoresis under regular conditions. Band compression was likewise eliminated when the same *NAT1*11* sample was sequenced with dGTP, but submitted to electrophoresis on a polyacrylamide gel prepared with 40% formamide to effect stronger denaturing conditions (not shown). Since preparation of formamide gels is not always an easy task, the remainder of the original *NAT1*11* samples and all new samples were routinely sequenced with dITP. Unlike the double allelic constitution of PCR-amplified genes that are sequenced directly (no subcloning) (Vatsis et al., 1991; Vatsis and Weber, 1993), only a single allele is introduced when a gene is cloned into a prokaryotic or eukaryotic host. It is noteworthy in this connection that no band compression was seen in the corresponding segment (nt 446–464) of *NAT1*11* expressed in *E. coli* or COS-1 cells (de León, 1996), indicating that the structural abnormality at the palindrome was exacerbated by the presence of a second allele.

Restoration of normal spacing for the palindromic nucleotides revealed that, in addition to the point mutations delineated several years ago (Vatsis and Weber, 1993), the coding region of *NAT1*11* also contains a $G^{445}A$ transition involving replacement of Val-149 by Ile (not shown). The coding exon of *NAT1*11* is, therefore, distinguished by one silent (T153T) and two missense mutations (V149I; S214A) (Fig. 1), in accordance with the very recent findings of Lin et al. (1998). Another recent report had assigned a new allelic designation to these three point mutations (Doll et al., 1997), but there is now agreement that they are indeed associated with *NAT1*11* (D. W. Hein, personal communication).

**3′ UTR.** Deletion of 9 bp from an AT-rich segment in the 3′ UTR of PCR-generated *NAT1*11* was also reported some time ago, but the actual location of the deletion site could not be ascertained because eight repetitive triplets (AAT, ATA, or TAA) within this segment gave 18 nonanucleotide combinations as potential deletion sites; elimination of any of these nonanucleotides resulted in the same sequence for *NAT1*11* (Vatsis and Weber, 1993). Sequencing of the single allele of recombinant *NAT1*11* expressed in *E. coli* or COS-1 cells (de León, 1996) has allowed a precise definition of the deletion end points as nt 1077 and 1088 (Fig. 1).

### Determination of *NAT1** Genotype

The *NAT1** genotype of 26 unrelated individuals from five non-Caucasian groups was ascertained unambiguously by direct sequencing of both strands of 1.6-kb fragments (nt

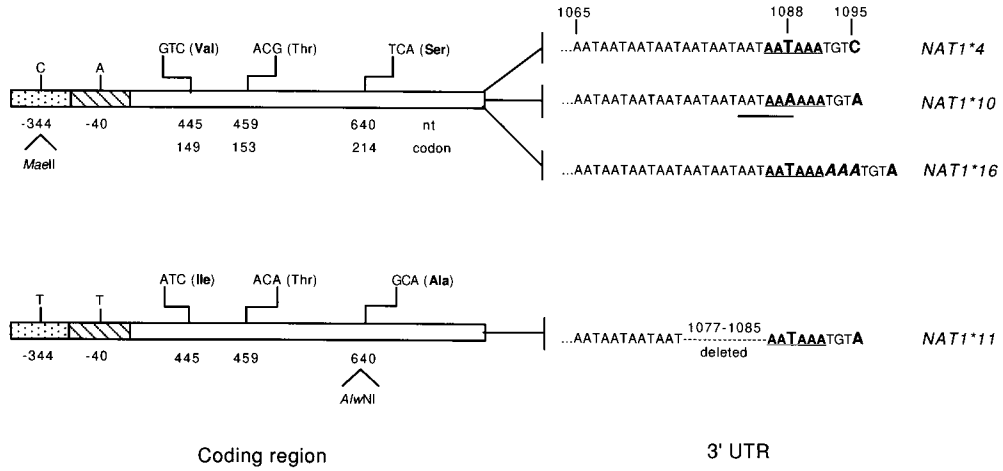


**Fig. 1.** Structural variation in four human *NAT1** alleles. The PCR-generated fragments of the variant *NAT1** alleles (1.6 kb, nt −440 to 1175) were unambiguously identified by direct sequencing (no subcloning). The gene regions are marked as follows: stippled bar, 5′ flank; hatched bar, 5′ UTR; open bar, coding exon. Nucleotide and codon numbers are depicted below the bars. An AT-rich segment in the 3′ UTR (nt 1065–1091) contains eight identical copies of AAT, ATA, or TAA motifs, as well as the consensus polyadenylation signal (AATAAA; nt 1086–1091, underlined). The highly conserved AATAAA hexanucleotide is obliterated in *NAT1*10* by virtue of the $T^{1088}A$ transversion. The double underlining of this segment of *NAT1*10* emphasizes that the AAT triplet (nt 1083–1085) immediately 5′ to the mutated (AAAAAA) signal serves to generate a new and, according to the results of the present study, completely functional polyadenylation signal (AATAAA; nt 1083–1088). The end points of the nonanucleotide deleted from the 3′ UTR of *NAT1*11* have been identified by sequencing of the recombinant allele as nt 1077 and 1085. *NAT1*16* has an insertion of three adenosines (***AAA***) immediately 5′ to the AATAAA signal. Alleles *NAT1*10*, **11*, and **16* share the $C^{1095}A$ transversion (position 1098 in *NAT1*16*).

−440 to 1170) of PCR-generated *NAT1**. The three structural variants (*NAT1*4*, **10*, and **11*) originally identified in 12 unrelated Caucasians (Vatsis and Weber, 1993) were also detected in the non-Caucasian population in this study together with a new allele (Vatsis et al., 1994), hereby designated *NAT1*16* in compliance with the NAT nomenclature system (Vatsis et al., 1995). Mutations in *NAT1*16* were confined to the 3′ UTR, and included an insertion of three adenosines immediately 3′ to the sixth position of the AATAAA signal (AATAAA***AAA***), as well as a C to A transversion four nucleotides downstream from the trinucleotide insertion ($C^{1095}A$, also displayed by *NAT1*10* and **11*) (Fig. 1). No other variant was found, and none could have gone undetected because allelic structure was determined by sequencing.

There are ten possible allelic combinations (genotypes) for four variant *NAT1** alleles, and seven of these were detected despite the comparatively small size of the population (Table 1). By far the most prevalent genotypes were those of homozygotes and heterozygotes of *NAT1*4*, found in 33 of the 38 subjects. Genotypes with *NAT1*10* were the second most abundant, accounting for 20 of the individuals (two homozygotes, 15 heterozygotes, and three compound heterozygotes), whereas *NAT1*11* and **16* were present in only six subjects (three heterozygotes and three compound heterozygotes). Not surprisingly, the three genotypes not represented in the population were homozygotes and compound heterozygotes of the relatively infrequent *NAT1*11* and **16* alleles.

*NAT1*4* and **10* were widely distributed in all six ethnic groups, whereas *NAT1*11* was present only in Caucasians and Asian Indians, and *NAT1*16* was seen only in two Far Eastern groups (Japanese and mainland Chinese) (not shown). Given the limited number of subjects we have examined, it cannot be stated with certainty that *NAT1*11* and **16* occur exclusively in Caucasians/Asian Indians and Orientals, respectively, nor can the allelic incidence in each of the six ethnic groups be accurately determined for the same reason. Preliminary estimates of the *combined* interethnic frequencies ($n$ = 76 alleles) are possible, however, and show a distribution of 63% for *NAT1*4*, 29% for *NAT1*10*, 5% for *NAT1*11*, and 3% for *NAT1*16* (see Table 1 for the allelic frequency computations). These frequency distribution estimates are quite close to those obtained in two other recent studies with subjects of mixed ethnic origin (Hughes et al., 1998; Lin et al., 1998).

TABLE 1
Incidence of *NAT1** genotypes in racially diverse groups[a]

| *NAT1** Genotype[b] | Number Observed | Frequency |
|---|---|---|
| | | *% total* |
| Homozygotes | | |
| *NAT1*4/*4* | 15 | 39.5 |
| *NAT1*10/*10* | 2 | 5.3 |
| Heterozygotes | | |
| *NAT1*4/*10* | 15 | 39.5 |
| *NAT1*4/*11* | 2 | 5.3 |
| *NAT1*4/*16* | 1 | 2.6 |
| Compound heterozygotes | | |
| *NAT1*10/*11* | 2 | 5.3 |
| *NAT1*10/*16* | 1 | 2.6 |

[a] 38 subjects: Caucasians, 12; Japanese, 5; mainland Chinese, 6; Koreans, 5; Asian Indians, 5; Hispanics, 5. The *NAT1** genotype of the Caucasian subjects is from a previous study (Vatsis and Weber, 1993).
[b] Determined by direct (no subcloning) sequencing of 1.6-kb fragments (nt −440 to 1175) of *NAT1** amplified with leukocyte DNA as template (Vatsis and Weber, 1993).

### Characterization of Recombinant Human *NAT1** Variants

To enable an unobscured interpretation of the data, the effects of the coding region point mutations on NAT1 structure and function were evaluated separately from those of the mutations in the 3′ UTR. This was accomplished by expression of the *NAT1** coding region (nt 1–870) in *E. coli,* and of the coding region together with the 3′ UTR (nt 1–1170) in COS-1 cells.

**Expression of the Coding Region of *NAT1*4* and **11* in *E. coli*.** Since *NAT1*4*, **10*, and **16* have identical coding sequences (Fig. 1), *NAT1*4* was selected as representative of these alleles for transformation, along with the coding exon of *NAT1*11*, into *E. coli*. The expression level varied slightly (4–7%) for each of the expressed proteins in three separate bacterial inductions, and was approximately 25% lower for NAT1 11 [2.69 ± 0.09 units (NAT1 4) and 2.12 ± 0.05 units (NAT1 11) of immunoreactive NAT1/mg of *E. coli* cytosolic protein].

5-Aminosalicylate metabolically generated in the colonic lumen is thought to be the therapeutically active agent against ulcerative colitis treated with drugs such as sulfasalazine, and several clinical studies have demonstrated a large variation in *N*-acetylation of 5-aminosalicylate among human subjects (reviewed in de León, 1996). The catalytic efficiency ($V_{max}/K_m$ ratio) with 5-aminosalicylate was found in preliminary experiments to be 125 times greater for wild-type NAT1 4 than for wild-type NAT2 4 (not shown), indicating for the first time that the specificity for 5-aminosalicylate resides in NAT1.

Steady-state *N*-acetylation kinetics with 5-aminosalicylate and two additional NAT1-specific substrates, as well as with 4-aminosalicylate in the presence of varying amounts of acetyl CoA, are summarized for NAT1 4 and NAT1 11 in Table 2. The kinetic constants were first determined with 4-aminosalicylate in Tris-buffered mixtures that did not contain salt or organic solvent, as has routinely been practiced in this laboratory (Andres et al., 1985), and in triethanolamine-buffered mixtures containing KCl and dimethyl sulfoxide [in an amount equal to that contributed when the substrate is dissolved in the organic solvent (Cribb et al., 1991; Grant et al., 1991; Dupret and Grant 1992; Dupret et al., 1994; Hughes et al., 1998)]. Significantly higher $K_m$ and lower $V_{max}$ values were obtained with the triethanolamine-buffered mixtures and either NAT1 isozyme; the $K_m$ with NAT1 4 is comparable to that reported for this isozyme expressed in *E. coli* (Dupret and Grant, 1992; Hughes et al., 1998; Lin et al., 1998). A comparison of the kinetic constants with Tris-buffered mixtures shows NAT1 11 with a 2-fold higher affinity for all three substrates, a very significant difference statistically, and a somewhat lower affinity for the cofactor relative to wild-type NAT1 4. There were no statistically significant differences in $V_{max}$ with the two isozymes. As a consequence of the uniformly lower $K_m$ of NAT1 11 for the substrate, the $V_{max}/K_m$ ratio with this isozyme was 60 to 70% greater than that with NAT1 4. The invariably higher affinity of NAT1 11 for arylamines (Table 2) was not seen in another laboratory, but only one experiment was performed with 4-aminosalicylate as substrate (Hughes et al., 1998).

As with 4-aminobenzoate *N*-acetylation by purified human

recombinant NAT1 of unspecified isozymic identity (Ward et al., 1995), the *N*-acetylation of 4-aminosalicylate by cytosolic recombinant NAT1 4 and NAT1 11 was also inhibited competitively by amethopterin. The reaction was blocked 50% by about 30 $\mu$M amethopterin (Table 3), and more than 90% by 150 $\mu$M amethopterin. On the other hand, the *N*-acetylation of sulfamethazine by recombinant wild-type NAT2 4 was found to proceed unimpeded in the presence of 200 $\mu$M amethopterin (not shown). These findings illustrate the potency and specificity of amethopterin as an inhibitor of NAT1-catalyzed *N*-acetylation. The specificity is further evidenced by the similarity in $K_i$ (Table 3) and $K_m$ (Table 2) values with 4-aminosalicylate as substrate. No differences were seen in the degree of inhibition or magnitude of the inhibitory constants with NAT1 4 and NAT1 11 (Table 3).

The *N*-acetylation of 4-aminosalicylate by NAT1 4 and NAT1 11, respectively, was monitored as a function of temperature, and Arrhenius plots of the data with either isozyme exhibited a transition temperature at about 27°C, indicative of a change in the catalytic mechanism (Segel, 1975). The energy of activation below the transition temperature was identical with NAT1 4 and NAT1 11, whereas that above the transition temperature was slightly but significantly higher with NAT1 11 (Table 3).

The catalytic activity and fate of NAT1 4 and NAT1 11 in *E. coli* cytosol were simultaneously assessed in time course experiments at 37°C. The *N*-acetylation of 4-aminosalicylate gradually decreased with time, reaching values at 15 h that were 2.5% of those recorded at zero time with either isozyme (Table 3), and the activity was abolished after incubation for 20 h. Semilog plots of the data revealed a monophasic reaction with a first order inactivation rate constant of the same magnitude for NAT1 4 and NAT1 11 (Table 3), and similar to that reported for recombinant NAT1 3 in COS-1 cells (Grant et al., 1991) and for NAT1 3, NAT1 4, and NAT1 11 in *E. coli* (Dupret and Grant, 1992; Dupret et al., 1994; Hughes et al., 1998; Lin et al., 1998).

NAT1 protein content was also evaluated in aliquots of cytosolic fractions from the zero time, 3-, 6-, and 15-h points of the time course study. There was little change in the extent of immunoreactivity of NAT1 4 (~10% decrease) and NAT1 11 (20% decrease) (Table 3) after 15 h of exposure of bacterial cytosol to 37°C, which resulted in cessation of the catalytic activity of either isozyme. Proteolytic processing of these variants thus appears to have been minimal, but it should be noted that *E. coli* may not possess specific proteases that degrade human NATs (Hein et al., 1994). By far the bulk of the catalytic activity loss most likely reflects the intrinsic instability of human NAT1 proteins, or at least of the four variants examined in this study, which is much more pronounced than that of wild-type and some of the mutant NAT2 molecules (Dupret and Grant, 1992; Dupret et al., 1994).

**Expression of the Coding Region and 3′ UTR of *NAT1*4*, **10*, **11*, and **16* in COS-1 Cells.** The role of the 3′ UTR in NAT1 expression and catalytic competence was investigated upon transient transfection of COS-1 cells with the coding exon and 290 bp of the 3′ UTR of *NAT1*4*, **10*, **11*, and **16*, respectively; all four alleles have dissimilar 3′ UTR sequences (Fig. 1).

NAT1 4, NAT1 10, and NAT1 11 had comparable transfection efficiency, immunoreactivity, and specific *N*-acetylation activities (Table 4 and Fig. 2). Even though *NAT1*16* displayed a considerably greater transfection efficiency, the cytosolic content of NAT1 16 was 2-fold lower than that of the other isozymes (Fig. 2 and Table 4). The decrease in NAT1 16 protein could not have been the result of proteolysis, given the fact that the coding region of this variant is identical to that of NAT1 4 and NAT1 10. Furthermore, the 50% decline in immunoreactivity of NAT1 16 was accompanied by a corresponding decrease in the specific *N*-acetylation activity of this isozyme with all three substrates (Table 4). As expected from these data, correction of the specific activities by the amount of NAT1 protein expressed in COS-1 cell cytosol gave normalized values, expressed per unit of immunoreactive NAT1, that were similar for the four isozymes (not shown), indicating that NAT1 16 was catalytically unaffected. No activity was detected with the *pCR3* vector devoid of *NAT1** insert (Table 4), and the normalized activity of the NAT1 variants with sulfamethazine, an NAT2-selective substrate (cf. Vatsis and Weber, 1997), was less than 1% of those

TABLE 2

Steady-state kinetics of *N*-acetylation by NAT1 4 and NAT1 11 isozymes expressed in *E. coli* strain BL21(DE3)*pLysS*

Reaction mixtures with cytosol from transformed bacteria (1–2 $\mu$g of protein) were prepared in 20 mM Tris-HCl buffer (pH 7.5 at 37°C) containing an acetyl CoA regenerating system and the additional components described under *Experimental Procedures*. Mixtures with varying amounts of the substrate (2–100 $\mu$M) contained a fixed amount of acetyl CoA (0.1 mM); conversely, 4-aminosalicylate was held constant (0.1 mM) when the cofactor was varied from 0.005 to 0.5 mM. Kinetic constants with 4-aminosalicylate as the variable component were also determined with reaction mixtures prepared in 55 mM triethanolamine-HCl buffer (pH 7.5 at 37°C) containing 25 mM KCl and 2.5% (v/v) dimethyl sulfoxide; the acetyl CoA regenerating system and other components were as for the Tris-HCl buffer system. Reactions were at 37°C for 2 to 5 min, depending on the particular substrate or cofactor serving as the variable component, and arylamide products were quantitated by HPLC. NAT1 protein content was ascertained by densitometric scanning of 34-kDa bands on immunoblots, generated with a polyclonal human NAT1 antibody; 1 unit (U) of immunoreactive NAT1 was arbitrarily equated to 1000 relative band density units. Apparent $K_m$ and $V_{max}$ constants were calculated with the Cleland Hyper program (Cleland, 1967). The values (mean ± S.D.) represent data from three experiments (three separate bacterial inductions), with triplicate determinations in each experiment. Statistical analyses were performed with the unpaired Student's *t* test.

| Buffer | Component Varied | Apparent $K_m$ | | Normalized Apparent $V_{max}$ | | Ratio, $V_{max}/K_m$ | |
|---|---|---|---|---|---|---|---|
| | | NAT1 4 | NAT1 11 | NAT1 4 | NAT1 11 | NAT1 4 | NAT1 11 |
| | | *μM* | | *nmol of arylamide formed / min / U of NAT1* | | | |
| Triethanolamine | Substrate | | | | | | |
| | 4-aminosalicylate | 13.6 ± 1.4 | 9.3 ± 0.2[a] | 22.2 ± 1.9 | 18.4 ± 0.7[b] | 1.6 | 2.0 |
| Tris | 4-aminosalicylate | 8.9 ± 0.3 | 4.3 ± 0.2[c] | 34.6 ± 6.9 | 28.9 ± 4.4[b] | 3.9 | 6.7 |
| | 5-aminosalicylate | 14.5 ± 1.5 | 7.8 ± 0.5[d] | 44.8 ± 3.5 | 39.5 ± 5.4[b] | 3.1 | 5.1 |
| | 4-aminobenzoate | 20.8 ± 1.2 | 9.8 ± 0.2[d] | 39.2 ± 6.3 | 30.1 ± 7.6[b] | 1.9 | 3.1 |
| | Acetyl CoA | 186 ± 8.6 | 258 ± 16[a] | 76.7 ± 7.8 | 81.0 ± 24[b] | 0.4 | 0.3 |

[a] Significantly different from NAT1 4 ($P < .05$).
[b] Not significantly different from NAT1 4 ($P > .1$).
[c] Significantly different from NAT1 4 ($P < .0001$).
[d] Significantly different from NAT1 4 ($P < .01$).

obtained with the NAT1-specific and -selective substrates listed in Table 4 (not shown).

Steady-state kinetics with 4-aminosalicylate as the variable component gave a similar $K_m$ for NAT1 4, NAT1 10, and NAT1 16 (10–13 $\mu$M), an expected finding in view of the identical coding sequence of these variants; as with the recombinant proteins in *E. coli*, a small but statistically significant decrease was seen for the $K_m$ of NAT1 11 (7 $\mu$M). The normalized $V_{max}$ values were the same for NAT1 4 and NAT1 10, and about 20% lower for NAT1 11 and NAT1 16 (not shown).

TABLE 3

Inhibition, thermodynamics, and stability of NAT1 4 and NAT1 11 isozymes expressed in *E. coli* strain BL21(DE3)*pLysS*

NAT1-specific 4-aminosalicylate served as substrate for activity determinations in all experiments summarized in this table. The reaction mixtures, in 20 mM Tris-HCl buffer (pH 7.5), also contained bacterial cytosol, acetyl CoA (0.1 mM), an acetyl CoA regenerating system, and the additional components detailed under *Experimental Procedures*; 4-acetamidosalicylate was quantitated by HPLC. Statistical analysis was by the unpaired Student's *t* test.

| Parameter | Observation | |
|---|---|---|
| | NAT1 4 | NAT1 11 |
| Inhibition by amethopterin[a] | | |
| $K_i$ ($\mu$M) | 16.2 | 19.3 |
| $I_{50}$ ($\mu$M) | 27.7 | 29.7 |
| Energy of activation ($E_a$; kcal mol$^{-1}$)[b] | | |
| $E_{a_1}$ (below ~27°C) | 17.1 ± 0.4 | 17.5 ± 0.1[c] |
| $E_{a_2}$ (above ~27°C) | 7.8 ± 1.5 | 11.4 ± 1.1[d] |
| Thermostability and immunoreactivity[e] (37°C; 0–15 h) | | |
| Normalized catalytic activity (nmol of 4-acetamidosalicylate formed/min/U of NAT1)[f] | | |
| zero time | 30.2 ± 3.6 | 25.3 ± 3.4[c] |
| 15 h | 0.8 ± 0.10 | 0.6 ± 0.15[c] |
| $k_{inact}$ (h$^{-1}$) | 0.12 ± 0.008 | 0.1 ± 0.001[c] |
| Immunodetectable NAT1 (U of immunoreactive NAT1/mg of *E. coli* cytosolic protein)[g] | | |
| zero time | 2.8 ± 0.17 | 2.0 ± 0.14[h] |
| 15 h | 2.6 ± 0.11[i] | 1.6 ± 0.06[h,j] |

[a] The respective NAT1 in bacterial cytosol (1 $\mu$g of protein) was incubated at 37°C with 4-aminosalicylate at 8 and 30 $\mu$M, respectively, in the absence and presence of varying amounts of amethopterin (10–150 $\mu$M); the inhibitor was delivered in dimethyl sulfoxide (up to 1% final concentration), which was found in control experiments to have a negligible effect on the activity of either isozyme. The $K_i$ constants were obtained directly from Dixon plots of the data (Segel, 1975), and $IC_{50}$ values were derived from plots of residual (uninhibited) activity as a function of time. The results are the average of two experiments, with triplicate determinations in each experiment.

[b] Reactions with bacterial cytosol (2 $\mu$g of protein) and the additional components described above were carried out at temperatures ranging from 4.6 to 39.4°C, and linearity with time was established at each temperature. The Arrhenius plots with either isozyme displayed a discontinuity (transition temperature) between 25.5 and 30°C, and the $E_a$ constants were derived from the slope of the corresponding line (Segel, 1975). The results (mean ± S.D.) are from three experiments (three separate bacterial inductions), with quadruplicate determinations (duplicate reaction mixtures for each of two time points) per experiment.

[c] Not significantly different from NAT1 4 ($P > .1$).

[d] Significantly different from NAT1 4 ($P < .05$).

[e] *E. coli* cytosol (2 $\mu$g of protein) containing NAT1 4 or NAT1 11 was incubated at 37°C for varying time periods up to 15 h, and aliquots were simultaneously assayed for 4-aminosalicylate *N*-acetylation activity at 30°C and for NAT1 protein content by immunoreactivity.

[f] The values represent the mean ± S.D. from three experiments (three separate bacterial inductions), with triplicate determinations in each experiment. A unit (U) of immunoreactive NAT1 is defined in Table 2.

[g] Bacterial cytosol from the time course at 37°C was loaded on polyacrylamide gels in triplicate per time point assessed in each of three experiments (three separate bacterial inductions); the amount of cytosol loaded (10 $\mu$g) was from the linear segment of the relative band density versus protein concentration plot. Immunoreaction of the electrophoretically separated proteins was as described under *Experimental Procedures*; the definition of a unit (U) of immunoreactive NAT1 is given in Table 2. Values are the mean ± S.D.

[h] Significantly different from NAT1 4 ($P < .0005$).

[i] Significantly different from zero time ($P < .05$).

[j] Significantly different from zero time ($P < .001$).

The above findings with the eukaryotic expression system support the conclusion that the parallel decreases in the cytosolic content and specific *N*-acetylation activity of NAT1 16 are a consequence of the structural alteration in the 3′ UTR of *NAT1*16*, entailing insertion of three extra adenosines immediately 3′ to the sixth position of the AATAAA signal (Fig. 1). In sharp contrast, the changes in the 3′ UTR of *NAT1*10* and **11* (Fig. 1) affect neither the expression nor the catalytic function of the corresponding protein.

## Functional Correlates of NAT1 4 and NAT1 10 in Human White Blood Cells

The distribution of NATs in human leukocytes has been studied indirectly by steady-state kinetics with NAT1- and NAT2-specific substrates, and the findings suggested that NAT1 may be the sole acetylase in cytosol from mononuclear leukocytes and the principal acetylase in cytosol from neutrophils; the *NAT1** genotype of the donors of the white blood cells was not evaluated in that study (Cribb et al., 1991).

Catalytic parameters and thermostability were investigated with the cytosolic fraction of mononuclear and polymorphonuclear leukocytes from two subjects homozygous for wild-type *NAT1*4*, and from one of the two *NAT1*10* homozygotes detected in our population study (see Table 1). As demonstrated in Table 5, rates of *N*-acetylation of two NAT1-specific, two NAT1-selective, and one as yet uncategorized compound (*p*-anisidine; cf. Vatsis and Weber, 1997) were essentially identical for NAT1 4 and NAT1 10 in cytosol from either white blood cell type. By the same token, no differences were discernible in the apparent $K_m$ [9–15 $\mu$M (NAT1 4); 9.5 $\mu$M (NAT1 10)] and $V_{max}$ [7–10 nmol/min/mg of protein (NAT1 4); 8 nmol/min/mg of protein (NAT1 10)] determined with 4-aminosalicylate as the substrate. The $K_i$ for the competitive inhibition exerted by amethopterin was the same for NAT1 4 (13 $\mu$M) and NAT1 10 (16.5 $\mu$M) in mononuclear leukocyte cytosol, and comparable to that obtained with the coding region of NAT1 4 (which is the same as that of NAT1 10 and NAT 1 16) expressed in *E. coli* cytosol (see Table 3).

Both NAT1 4 and NAT1 10 underwent rapid thermal inactivation, as evidenced by a 90–95% loss of 4-aminosalicylate *N*-acetylation activity after only 8 h of exposure of mononuclear leukocyte cytosol to 37°C. Unlike recombinant NAT1 in *E. coli*, thermal inactivation of naturally occurring NAT1 4 and NAT1 10 was a biphasic process. More importantly, the extent and rate constant associated with each phase of inactivation were the same for NAT1 4 and NAT1 10: about 35% of the activity catalyzed by either isozyme was lost in the first (more rapid) phase with a first order rate constant of 1 h$^{-1}$, and the remaining two-thirds went in the second (slower) phase with a rate constant of 0.2 h$^{-1}$. Inactivation of naturally occurring NAT1 4 and NAT1 10 was, therefore, two to ten times faster than that of the enzymes expressed in *E. coli* (~0.1 h$^{-1}$; Table 3), quite possibly reflecting less than optimal folding or configuration of the NAT1 proteins in the artificial expression system.

## Secondary Structure of the 3′ UTR in pre-mRNAs for the Human NAT1 Variants

The secondary structure of the 3′ UTR in the pre-mRNA for each of the NAT1 variants is characterized by a series of stems and loops (Fig. 3). Three structural elements are emphasized: the polyadenylation signal (AAUAAA; nt 1087–

1096); the predicted cleavage site (GA ↓ A) 3′ to the AAUAAA hexamer (between nt 1124 and 1125); and a U-rich sequence farther downstream (nt 1149–1152). The AAUAAA signal is present in bulges of the stem in NAT1 4, NAT1 10, and NAT1 16 pre-mRNAs, but in loop 3 of NAT1 11 pre-mRNA. The predicted cleavage site (GA ↓ A) is also located in a bulge in NAT1 4, NAT1 10, and NAT1 11, but found on the double-stranded stem of NAT1 16 pre-mRNA. The U-rich sequence is present on the single-stranded loop 4 in NAT1 4, NAT1 10, and NAT1 11, but is notably absent from NAT1 16 pre-mRNA. Both the predicted cleavage site and the U-rich downstream element appear to be sequestered in NAT1 16. Moreover, as a consequence of the altered secondary structure, the distance between the AAUAAA hexamer and the U-rich sequence is four nucleotides longer in NAT1 16 pre-mRNA (61 nt total) than in the pre-mRNA for the other three NAT1 variants (57 nt total).

## Discussion

### Compensatory Elements in the 3′ UTR of Human *NAT1** Give Rise to a Fully Operational Polyadenylation Signal in *NAT1*10*

A $T^{1088}$A transversion in the third position of the polyadenylation signal (AATAAA; nt 1086–1091) of *NAT1*10* obliterates this highly conserved element, which is situated at the tail end of an AT-rich segment of iterative AAT triplets (Fig. 1). Point mutations within the AATAAA hexamer are known to interfere markedly with the maturation and, consequently, promote the degradation of a host of functionally distinct mRNAs (Sheets et al., 1990), inferring that less NAT1 10 protein may be produced because of impaired 3′ end processing of the corresponding transcript (Vatsis and Weber, 1993). We also considered that the AAT triplet (nt 1083–1085) immediately 5′ to the first adenosine (nt 1086) of the mutated (AAAAAA) hexanucleotide could possibly serve to generate a new AATAAA signal (nt 1083–1088) in *NAT1*10*, but this would change the optimal distance between AATAAA and the cleavage/poly(A) addition site. Since there appears to be some spatial requirement for efficient mRNA production (Nevins, 1983), the functionality of a new polyadenylation signal with a 5′ shift of three nucleotides could not have been accurately predicted for *NAT1*10* (Vatsis and Weber, 1993), and, hence, had to be established experimentally.

The experiments in the present study clearly show that recruitment of the AAT triplet at positions 1083–1085 toward formation of a new AATAAA signal does not disrupt the predicted secondary structure of the pre-mRNA for NAT1 10, which is indistinguishable from that of wild-type NAT1 4 (Fig. 3, A and B). Consistent with this observation was the finding that the content of NAT1 10 protein in COS-1 cell cytosol was identical with that of wild-type NAT1 4. Furthermore, no differences existed in any of the biochemical parameters assessed with recombinant NAT1 4 and NAT1 10 in the eukaryotic host, and, more importantly, none was discernible with cytosol from mononuclear leukocytes of *NAT1*4/*4* and *NAT1*10/10* homozygotes. It may be concluded that the $T^{1088}$A transversion in *NAT1*10* is well tolerated because the AAT triplet on the 5′ side of the mutated polyadenylation signal does indeed generate a perfectly functional AATAAA hexamer. Our findings do not support the suggestion that NAT1 10 has greater *N*-acetylation activity than NAT1 4 (Bell et al., 1995), a conclusion also reached from NAT1 phenotypic assessment upon administration of 4-aminosalicylate to 148 unrelated subjects (Hughes et al., 1998).

The greater instability of NAT1 4 and NAT1 10 in their natural milieu (mononuclear leukocytes) as compared to the prokaryotic expression system appears to be the explanation for the substantially larger decrease in $V_{max}$ seen with blood lysates from a heterozygote of *NAT1*14* than with *E. coli* cytosol harboring the coding region of *NAT1*14* (Hughes et al., 1998). Negative modulation of the catalytic activity by the 3′ UTR mutations in *NAT1*14* cannot explain the discrepant $V_{max}$ results of Hughes et al. (1998), since the data presented above (Tables 4 and 5) amply demonstrate that *NAT1*10*, which has the same polyadenylation signal mutation as

TABLE 4

Efficiency of transfection, NAT1 protein content, and specific *N*-acetylation activities in cytosol from COS-1 cells cotransfected with β-galactosidase and *NAT1** variants

COS-1 cells (21 dishes per *NAT1** allele) were transiently cotransfected with *NAT1** recombinant construct (*NAT1** coding and 3′ noncoding exons) and the reporter vector *pCR3βGal*. The cells were pooled, disrupted by sonication, and the cytosolic fraction isolated by differential centrifugation of the homogenate. Catalytic activity of the cytosol was determined at 37°C with reaction mixtures in 20 mM Tris-HCl buffer, pH 7.5, containing 4-aminosalicylate (0.1 mM), 5-aminosalicylate (0.05 mM), or 2-aminofluorene (0.1 mM), acetyl CoA (0.1 mM), and the additional components detailed under *Experimental Procedures*. The activities are the mean ± S.D. of a total of 24 values obtained as follows: a) the average of the slopes of the lines constructed from data on protein and time dependence of the reactions, respectively [18 determinations: triplicates for each of three time points (5, 10, and 15 min with mixtures containing 1 μg of cytosol), and three amounts of cytosolic protein (0.5, 1, and 3 μg in mixtures incubated for 5 min)]; and b) the average of triplicates from each of two additional experiments conducted under the previously established initial rate conditions with respect to amount of cytosolic protein and reaction time (1 μg of protein irrespective of substrate; reactions carried out for 5 min with 2-aminofluorene and for 10 min with 4- and 5-aminosalicylate).

| Recombinant NAT1 | Transfection Efficiency[a] | NAT1 Content[b] | Specific Activity | | |
|---|---|---|---|---|---|
| | | | 4-Aminosalicylate | 5-Aminosalicylate | 2-Aminofluorene |
| | *U of β-gal/mg of cytosolic protein* | *U of NAT1/**mg** of cytosolic protein* | *nmol of arylamide formed/min/**mg** of cytosolic protein* | | |
| NAT1 4 | 0.57 (100)[c] | 0.87 (100) | 139 ± 35 (100) | 112 ± 23 (100) | 149 ± 16 (100) |
| NAT1 10 | 0.59 (104) | 0.82 (94) | 120 ± 28 (86) | 103 ± 21 (92) | 128 ± 13 (86) |
| NAT1 11 | 0.80 (140) | 0.86 (99) | 111 ± 22 (80) | 96 ± 18 (86) | 129 ± 18 (87) |
| NAT1 16 | 1.40 (246) | 0.44 (51) | 53 ± 9 (38) | 46 ± 6 (41) | 58 ± 10 (39) |
| None (*pCR3* vector) | | 0 | ND[d] | ND | 3 (2) |

[a] Transfection efficiency was ascertained by the catalytic activity of β-galactosidase at 37°C, determined from the absorbance of the reaction product (*o*-nitrophenol) at 420 nm. Reaction mixtures were in triplicate, and initial experiments were done to ensure the linearity of the reaction with COS-1 cell cytosolic protein and time; standard curves were generated with varying amounts of β-galactosidase (0.05–3 units).

[b] The numbers represent the slope of the linear segment of the curves depicted in Fig. 2. One unit (U) of immunoreactive NAT1 was artibitrarily equated to 1000 relative band density units.

[c] Values in parentheses are the percentage of those obtained with NAT1 4.

[d] ND, not detected.

*NAT1*14*, encodes a protein with catalytic and other properties indistinguishable from those of wild-type NAT1 4.

### Coding and 3′ UTR Mutations in *NAT1*11* Do Not Significantly Affect Function

Maximal rates of substrate *N*-acetylation were comparable for recombinant wild-type NAT1 4, with a coding sequence superimposable to that of NAT1 10 and NAT1 16, and for mutant NAT1 11, the only variant with missense mutations among the four identified in this study. NAT1 4 and NAT1 11 also exhibited the same sensitivity to the competitive inhibitor amethopterin, and possessed similar energies of activation and thermal stability assessed with 4-aminosalicylate as the substrate. On the other hand, NAT1 11 was invariably found to have a greater catalytic efficiency by virtue of a 1.5- to 2-fold higher affinity for all arylamines examined, a statistically significant effect observed irrespective of reaction buffer or expression host.

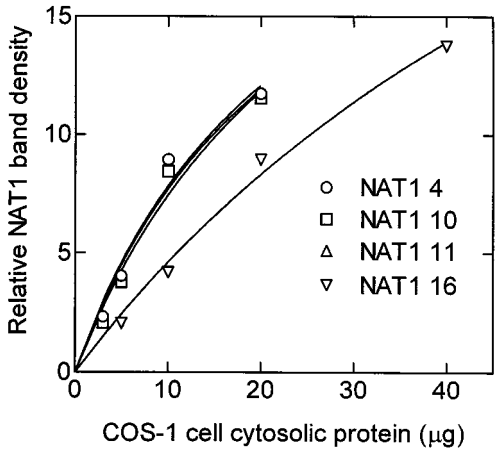


**Fig. 2.** Quantitation of NAT1 protein content in COS-1 cell cytosol. Autoradiograms with varying amounts of COS-1 cell cytosolic protein (2.5–40 μg, each in triplicate) were generated with human anti-NAT1 as described under *Experimental Procedures*, scanned with an imaging densitomer, and the relative density of 34-kDa (NAT1) bands plotted against protein concentration. A direct proportionality is seen between NAT1 band density and protein up to 10 μg of cytosol containing NAT1 4, NAT1 10, or NAT1 11, and up to 20 μg of cytosol containing NAT1 16. Protein content is obtained from the slope of the linear segment of the curves.

TABLE 5

*N*-Acetylation of drugs and toxicants by cytosol from mononuclear and polymorphonuclear leukocytes of *NAT1*4/*4* and *NAT1*10/*10* homozygotes

Mononuclear and polymorphonuclear leukocytes in whole blood from *NAT1*4/*4* and *NAT1*10/*10* homozygotes were separated by density gradient centrifugation, sonicated, and the cytosolic fraction isolated by differential centrifugation of the homogenate. Cytosol (5–10 μg of protein) from the respective white blood cell type was incubated in 20 mM Tris-HCl buffer (pH 7.5) containing acetyl CoA (0.1 mM) in the presence of an acetyl CoA regenerating system, saturating substrate (0.1 mM 4-aminobenzoate, 4-aminosalicylate, or 2-aminofluorene; 0.25 mM *p*-phenetidine or *p*-anisidine), and the additional components described under *Experimental Procedures*. Reactions were at 37°C for 5 and 10 min to ensure linearity of *N*-acetylated product formation with time. The values are the mean of quadruplicate determinations.

| Substrate | Specific Activity | | | |
|---|---|---|---|---|
| | Mononuclear Leukocytes | | Polymorphonuclear Leukocytes | |
| | NAT1 4 | NAT1 10 | NAT1 4 | NAT1 10 |
| | *nmol arylamide formed/min/mg protein* | | | |
| 4-Aminobenzoate | 5.6;6.7[a] | 6.8 | 5.1;5.4[a] | 4.7 |
| 4-Aminosalicylate | 4.6;5.3 | 5.1 | 4.0;4.0 | 3.5 |
| 2-Aminofluorene | 6.3;6.2 | 6.5 | 5.6;5.7 | 4.2 |
| *p*-Phenetidine | 4.7;4.6 | 5.1 | 4.3;4.2 | 3.5 |
| *p*-Anisidine | 5.0;5.6 | 5.7 | 4.3;4.4 | 3.7 |

[a] Values with mononuclear and polymorphonuclear leukocytes from two subjects with *NAT1*4/*4* genotype.

One of the mutational events in *NAT1*11* concerns a sizeable 9-bp deletion (nt 1077–1088) from the same AT-rich segment containing the polyadenylation signal (nt 1064–1091; Fig. 1). AT-rich stretches have been found in the 3′ UTR of many transcripts with high turnover rates, and are referred to as destabilizing motifs because they facilitate mRNA degradation and also act as negative modifiers of mRNA translation. The corollary is that elimination of nine bases from a destabilizing sequence should result in higher steady-state levels of NAT1 11 transcript and protein (Vatsis and Weber, 1993). More recent studies have shown, however, that only a specific AT-rich sequence, namely, TTATTTAT, can act as a destabilizing unit to bring about rapid mRNA decay (Zubiaga et al., 1995). Human *NAT1** alleles do not contain the specific TTATTTAT element, and the identical immunoreactivity of NAT1 4, NAT1 10, and NAT1 11 in COS-1 cell cytosol is consistent with the absence of such a motif, as is the virtually identical secondary structure of the pre-mRNAs for these isozymes, as well as the similarity in catalytic activities and $V_{max}$ constants of NAT1 4 and NAT1 11.

Collectively, the prokaryotic and eukaryotic expression studies have shown no major differences in catalytic or other properties of wild-type NAT1 4 and mutant NAT1 11, and others have obtained overlapping values for *NAT1*4/*11* heterozygotes and *NAT1*4/*4* homozygotes phenotypically evaluated with 4-aminosalicylate in vivo (Hughes et al., 1998). It should be pointed out, however, that the observed similarity in *N*-acetylation activities of NAT1 4 and NAT1 11 with the substrates examined in this study, including three simple amine drug molecules and one carbocyclic arylamine carcinogen (2-aminofluorene), does not in any way preclude differences in catalytic activity with other environmental/occupational carbocyclic carcinogens and dietary heterocyclic mutagens, which are of particular importance to chemical carcinogenesis and cancer epidemiology (Lin et al., 1998).

### Formation of NAT1 16 Protein is Obstructed by Structural Modification in the 3′ UTR

Numerous publications have documented the pivotal role of RNA secondary structure in the assembly of a stable formation complex, a stepwise process involving the cooperative binding of six nuclear factors to a pre-mRNA molecule, which is absolutely essential for efficient 3′ end processing of the pre-mRNA. A mutationally modified pre-mRNA may promote destabilization of the formation complex, thereby adversely affecting maturation. This topic is reviewed extensively by de León (1996), and a synopsis is presented in the ensuing paragraph.

Consensus sequences in target elements of RNA-binding proteins are frequently found in loops, bulges, or interior loops of the RNA, representing single-stranded regions for component binding (McCarthy and Kollmus, 1995). Additionally, the AATAAA signal in many viruses and certain mammalian mRNAs is flanked by nucleotide segments that form base-pairs, giving rise to a hairpin structure with the AATAAA signal located in the single-stranded loop. Cleavage efficiency was diminished 50% by disruption of the stem structure, and 90% by a decrease in the size of the loop

(Berkhout et al., 1995). Cleavage and polyadenylation efficiencies were likewise severely curtailed when a sequence insertion between the AATAAA and GT/T elements, as seen in *NAT1*16*, failed to form a stem-loop structurally similar to these two 3′ UTR elements. Poly(A) sites can, therefore, tolerate sequence insertions only if the inserted sequence forms RNA secondary structures that resemble the AATAAA and GT/T elements (Brown et al., 1991). Nuclease accessibility/sensitivity to the cleavage site is also affected by alterations in stem-loop structures brought about by mutation (cf. de León, 1996).

The insertion of three adenosines on the 3′ side of the AATAAA signal in *NAT1*16* resulted in parallel (2-fold) decreases in COS-1 cell cytosolic content of NAT1 16 protein and specific *N*-acetylation activity with three substrates, indicating that the structural alteration in the 3′ UTR was deleterious solely to the synthesis, not the catalytic competence, of NAT1 16 protein. The structural alterations resulting from the AAA insertion in the 3′ UTR of *NAT1*16* (Fig. 1) entail obliteration or drastic reduction in the size of loop 4, sequestration of the predicted cleavage site, and a longer distance between the polyadenylation signal and the downstream T-rich element (compare Fig. 3A and 3D). It follows from the discussion elaborated in the preceding paragraph that these structural modifications may have several untoward consequences, such as decreased binding affinity of cleavage stimulation factor for the T-rich element because of the apparent sequestration or marked reduction of this element; less accessibility/sensitivity of the predicted cleavage site to endonucleases; diminished cooperativity among the various nuclear factors, resulting from a lower binding affinity of cleavage stimulation factor for the T-rich motif, as well as from the increased distance between AATAAA and the T-rich motif; and/or assembly of an unstable formation complex that is detrimental to transcript maturation. Suffice it to say that, whatever the precise mechanism, structural abnormalities of this sort cause disruption of the secondary structure of the pre-mRNA for NAT1 16, culminating in the observed decrease in the amount and corresponding decrease in the catalytic activity of NAT1 16 protein in COS-1 cell cy-

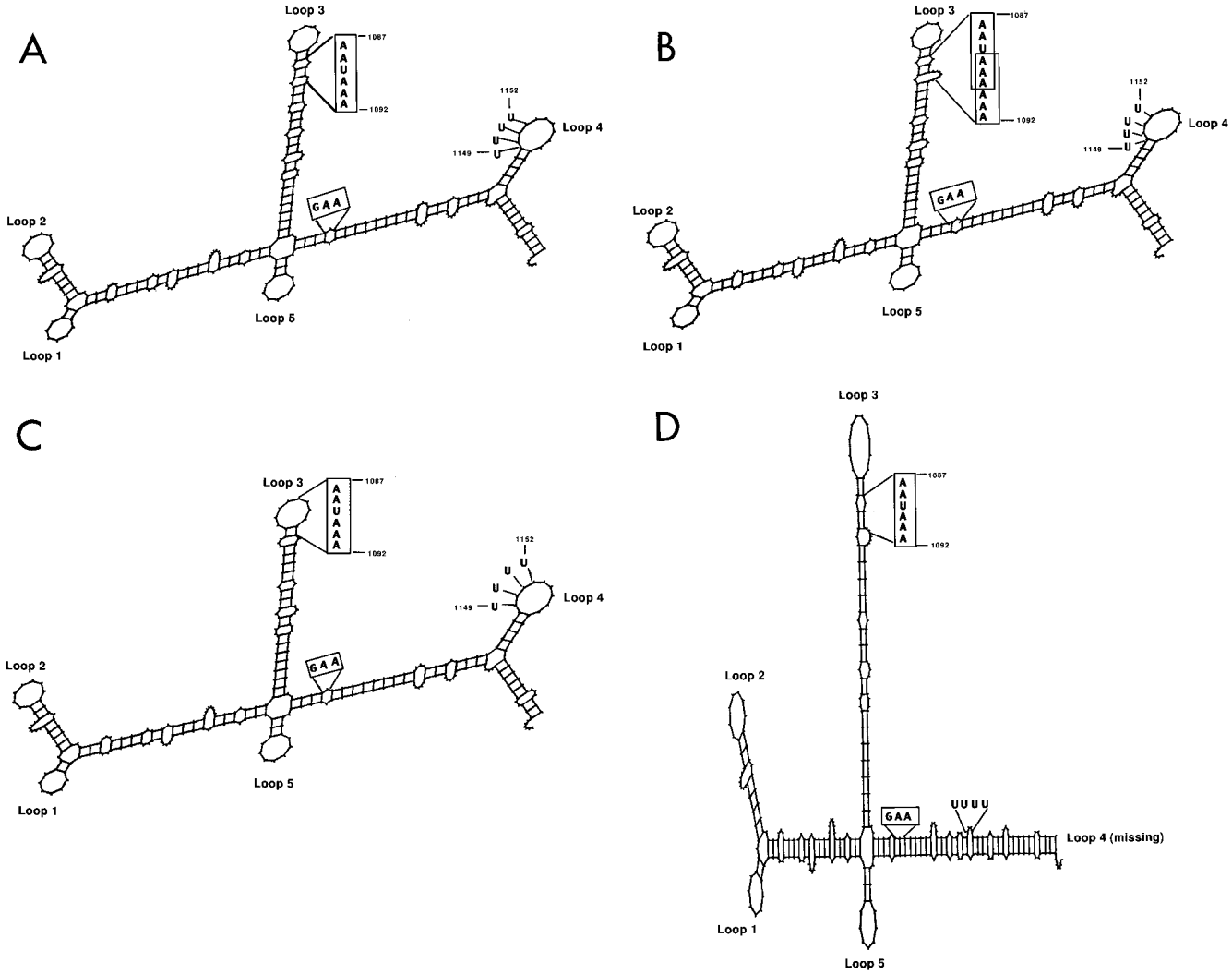


**Fig. 3.** Predicted secondary structures of 3′ UTR in the pre-mRNA for NAT1 4 (A), NAT1 10 (B), NAT1 11 (C), and NAT1 16 (D). The pre-mRNA secondary structures were predicted by the "Fold" program (Zuker and Stiegler, 1981) and "Squiggles" program (Osterburg and Sommer, 1981) provided by GenBank. "Fold" finds an optimal secondary structure of minimum free energy for an RNA molecule based on published values of stacking and loop destabilizing energies (Zuker and Stiegler, 1981). The three elements of interest, depicted in larger font and/or framed, include the polyadenylation signal (AAUAAA), the predicted cleavage site (GA ↓ A), and the U-rich sequence on loop 4.

tosol. It should also be mentioned at this juncture that, since mRNA levels were not measured, other possibilities for the decreased immunoreactivity of NAT1 16, such as impaired translation efficiency, cannot be ruled out at this time.

## References


Akada R (1994) Quick-check method to test the size of *Escherichia coli* plasmids. *Biotechniques* **17:**58.

Andres HH, Klem AJ, Szabo SM and Weber WW (1985) New spectrophotometric and radiochemical assays for acetyl CoA:arylamine *N*-acetyltransferase applicable to a variety of arylamines. *Anal Biochem* **145:**367–375.

Bell DA, Badawi AF, Lang NP, Ilett KF, Kadlubar FF and Hirvonen A (1995) Polymorphism in the *N*-acetyltransferase 1 (*NAT1*) polyadenylation signal: Association of *NAT1*10* allele with higher N-acetylation activity in bladder and colon tissue. *Cancer Res* **55:**5226–5229.

Berkhout B, Klaver B and Das AT (1995) A conserved hairpin structure predicted for the poly(A) signal of human and Simian immunodeficiency viruses. *Virology* **207:** 276–281.

Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:**248–254.

Brown PH, Tiley LS and Cullen BR (1991) Effect of RNA secondary structure on polyadenylation site selection. *Genes Dev* **5:**1277–1284.

Butcher NJ, Ilett KF and Minchin RF (1998) Functional polymorphism of the human arylamine *N*-acetyltransferase type 1 gene caused by $C^{190}T$ and $G^{560}A$ mutations. *Pharmacogenetics* **8:**67–72.

Chen C and Okayama H (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* **7:**2745–2752.

Cleland WW (1967) The statistical analysis of enzyme kinetic data. *Adv Enzymol* **29:**1–32.

Cribb AE, Grant DM, Miller MA and Spielberg SP (1991) Expression of monomorphic arylamine *N*-acetyltransferase (NAT1) in human leukocytes. *J Pharmacol Exp Ther* **259:**1241–1246.

de León JH (1996) Characterization of congenic mouse (*NAT2**) and human (*NAT1**) *N*-acetyltransferases. Ph.D. thesis, The University of Michigan.

Doll MA, Jiang W, Deitz AC, Rustan TD and Hein DW (1997) Identification of a novel allele at the human *NAT1* acetyltransferase locus. *Biochem Biophys Res Commun* **233:**584–591.

Dupret J-M, Goodfellow GH, Janezic SA and Grant DM (1994) Structure-function studies of human arylamine N-acetyltransferases NAT1 and NAT2: Functional analysis of recombinant NAT1/NAT2 chimeras expressed in *Escherichia coli*. *J Biol Chem* **269:**26830–26835.

Dupret J-M and Grant DM (1992) Site-directed mutagenesis of recombinant human arylamine *N*-acetyltransferase expressed in *Escherichia coli*: Evidence for direct involvement of $Cys^{68}$ in the catalytic mechanism of polymorphic human NAT2. *J Biol Chem* **267:**7381–7385.

Grant DM, Blum M, Beer M and Meyer UA (1991) Monomorphic and polymorphic human arylamine *N*-acetyltransferases: A comparison of liver isozymes and expressed products of two cloned genes. *Mol Pharmacol* **39:**184–191.

Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A, Yu VL and Grewal R (1997) Human acetyltransferase polymorphisms. *Mutat Res* **376:**61–70.

Grant DM, Vohra P, Avis Y and Ima A (1992) Detection of a new polymorphism of human arylamine *N*-acetyltransferase NAT1 using *p*-aminosalicylic acid as an *in vivo* probe. *J Basic Clin Physiol Pharmacol* **3 (Suppl):**244.

Hakes DJ and Dixon JE (1992) New vectors for high level expression of recombinant proteins in bacteria. *Anal Biochem* **202:**293–298.

Hein DW, Ferguson RJ, Doll MA, Rustan TD and Gray K (1994) Molecular genetics of human polymorphic *N*-acetyltransferase: enzymatic analysis of 15 recombinant wild-type, mutant, and chimeric NAT2 allozymes. *Hum Mol Genet* **3:**729–734.

Hughes NC, Janezic SA, McQueen KL, Jewett MAS, Castranio T, Bell DA and Grant DM (1998) Identification and characterization of variant alleles of human acetyltransferase NAT1 with defective function using *p*-aminosalicylate as an *in-vivo* and *in-vitro* probe. *Pharmacogenetics* **8:**55–66.

Lin HJ, Han C-Y, Lin BK and Hardy S (1993) Slow acetylator mutations in the human polymorphic *N*-acetyltransferase gene in 786 Asians, Blacks, Hispanics, and Whites: Application to metabolic epidemiology. *Am J Hum Genet* **52:**827–834.

Lin HJ, Probst-Hensch NM, Hughes NC, Sakamoto GT, Louie AD, Kau IH, Lin BK, Lee DB, Lin J, Frankl HD, Lee ER, Hardy S, Grant DM and Haile RW (1998) Variants of *N*-acetyltransferase NAT1 and a case-control study of colorectal adenomas. *Pharmacogenetics* **8:**269–281.

McCarthy JEG and Kollmus H (1995) Cytoplasmic mRNA-protein interactions in eukaryotic gene expression. *Trends Biochem Sci* **20:**191–197.

Nevins JR (1983) The pathway of eukaryotic mRNA formation. *Annu Rev Biochem* **52:**441–466.

Osterburg G and Sommer R (1981) Computer support of DNA sequence analysis. *Comput Programs Biomed* **13:**101–109.

Sambrook J, Fritsch EF and Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Segel IH (1975) *Enzyme Kinetics*. John Wiley & Sons, New York.

Sheets MD, Ogg SC and Wickens MP (1990) Point mutations in AAUAAA and the poly(A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucleic Acids Res* **18:**5799–5805.

Vatsis KP, Martell KJ and Weber WW (1991) Diverse point mutations in the human gene for polymorphic *N*-acetyltransferase. *Proc Natl Acad Sci USA* **88:**6333–6337.

Vatsis KP and Weber WW (1993) Structural heterogeneity of Caucasian *N*-acetyltransferase at the NAT1 gene locus. *Arch Biochem Biophys* **301:**71–76.

Vatsis KP and Weber WW (1994) Human *N*-acetyltransferases, in *Handbook of Experimental Pharmacology: Conjugation-Deconjugation Reactions in Drug Metabolism and Toxicity* (Kauffman FC ed) vol 122, pp 109–130, Springer-Verlag, Berlin.

Vatsis KP and Weber WW (1997) Acetyltransferases, in *Comprehensive Toxicology: Biotransformation* (Guengerich FP ed) vol 3, pp 385–399, Elsevier Science, Oxford.

Vatsis KP, Weber WW, Bell DA, Dupret J-M, Price Evans DA, Grant DM, Hein DW, Lin HJ, Meyer UA, Relling MV, Sim E, Suzuki T and Yamazoe Y (1995) Nomenclature for *N*-acetyltransferases. *Pharmacogenetics* **5:**1–17.

Vatsis KP, Yao Y and Weber WW (1994) Allelic variation at the *NAT1* locus of *N*-acetyltransferase among Orientals. *Am J Hum Genet* **55(suppl):**A340.

Ward A, Summers MJ and Sim E (1995) Purification of recombinant human *N*-acetyltransferase type 1 (NAT1) expressed in *E. coli* and characterization of its potential role in folate metabolism. *Biochem Pharmacol* **49:**1759–1767.

Weber WW (1987) *The Acetylator Genes and Drug Response*. Oxford University Press, New York.

Weber WW and Vatsis KP (1993) Individual variability in *p*-aminobenzoic acid *N*-acetylation by human *N*-acetyltransferase (NATI) of peripheral blood. *Pharmacogenetics* **3:**209–212.

Zubiaga AM, Belasco JG and Greenberg ME (1995) The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. *Mol Cell Biol* **15:**2219–2230.

Zuker M and Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res* **9:**133–148.


**Send reprint requests to:** Dr. Kostas P. Vatsis, Department of Biological Chemistry, The University of Michigan Medical School, M5440 Medical Science Building I, Ann Arbor, MI 48109-0606. E-mail: kvatsis@umich.edu