Gene 536 (2014) 65–73

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Polymorphism p.Val231Ile alters substrate selectivity of drug-metabolizing arylamine *N*-acetyltransferase 2 (NAT2) isoenzyme of rhesus macaque and human

Theodora Tsirka [1], Sotiria Boukouvala [1], Bogos Agianian, Giannoulis Fakis *

Democritus University of Thrace, Department of Molecular Biology and Genetics, Alexandroupolis, Greece

ARTICLE INFO

Article history:
Accepted 30 November 2013
Available online 11 December 2013

Keywords:
NAT2 gene
Polymorphic gene variant
Xenobiotic metabolizing enzyme
Animal model
Primate
Drug metabolism

ABSTRACT

Arylamine *N*-acetyltransferases (NATs) are polymorphic enzymes mediating the biotransformation of arylamine/arylhydrazine xenobiotics, including pharmaceuticals and environmental carcinogens. The *NAT1* and *NAT2* genes, and their many polymorphic variants, have been thoroughly studied in humans by pharmacogeneticists and cancer epidemiologists. However, little is known about the function of *NAT* homologues in other primate species, including disease models. Here, we perform a comparative functional investigation of the NAT2 homologues of the rhesus macaque and human. We further dissect the functional impact of a previously described rhesus *NAT2* gene polymorphism, causing substitution of valine by isoleucine at amino acid position 231. Gene constructs of rhesus and human *NAT2*, bearing or lacking non-synonymous polymorphism c.691G>A (p.Val231Ile), were expressed in *Escherichia coli* for comparative enzymatic analysis against various NAT1- and NAT2-selective substrates. The results suggest that the p.Val231Ile polymorphism does not compromise the stability or overall enzymatic activity of NAT2. However, substitution of Val231 by the bulkier isoleucine appears to alter enzyme substrate selectivity by decreasing the affinity towards NAT2 substrates and increasing the affinity towards NAT1 substrates. The experimental observations are supported by in silico modelling localizing polymorphic residue 231 close to amino acid loop 125–129, which forms part of the substrate binding pocket wall and determines the substrate binding preferences of the NAT isoenzymes. The p.Val231Ile polymorphism is the first natural polymorphism demonstrated to affect NAT substrate selectivity via this particular mechanism. The study is also the first to thoroughly characterize the properties of a polymorphic NAT isoenzyme in a non-human primate model.

1. Introduction

Arylamine *N*-acetyltransferases (NATs; EC 2.3.1.5) are enzymes mediating the acetyl-coenzyme A (acetyl-CoA) dependent *N*-acetylation of arylamine and arylhydrazine xenobiotics in a range of phylogenetically divergent organisms (Glenn et al., 2010). NATs are active against numerous xenobiotic compounds of everyday life, including pharmaceuticals and environmental/dietary carcinogens. The competing activities of NATs and other xenobiotic metabolizing enzymes are important determinants of the effects of such chemicals on the human body. NAT biology thus becomes relevant when evaluating the safety and efficacy of drugs or the mutagenic potential of carcinogens, and NATs have been at the forefront of pharmacogenetic and toxicogenetic investigations for over 50 years (Cascorbi, 2006; Weinshilboum and Wang, 2004).

Early studies suggested the presence of two NAT isoenzymes in human liver, the first one active towards small arylamine substrates (e.g. *p*-aminosalicylate) and the second towards bulkier arylamines (e.g. sulphamethazine) and arylhydrazines (e.g. the anti-tubercular drug isoniazid) (Sim et al., 2012). Activity of the second NAT isoenzyme varied considerably within populations, prompting the trimodal classification of individuals as "rapid", "intermediate" or "slow" metabolizers (Weinshilboum and Wang, 2004). Cloning of the *NAT1* and *NAT2* genes verified the existence of two NAT isoenzymes with characteristic substrate selectivities, and revealed the molecular genetic nature of the acetylator polymorphism in humans (Grant et al., 1991). The ethnic/geographic distribution of variable *NAT* haplotypes has since been studied in the anthropological (Patin et al., 2006; Sabbagh et al., 2011), pharmacogenomic (Teixeira et al., 2011; Walker et al., 2009) and disease epidemiological (Ochs-Balcom et al., 2007; Sanderson et al., 2007) contexts.

Current understanding of the biochemistry, molecular biology, regulation and function of human NATs is to a significant extent the outcome of research with homologues of model species. The NAT protein of the enterobacterium *Salmonella typhimurium* was the first one to be

Abbreviations: Acetyl-CoA, acetyl-coenzyme A; 5-AS, 5-aminosalicylate; CoA, coenzyme A; DMAB, dimethylaminobenzaldehyde; DSF, differential scanning fluorimetry; DTNB, 5,5′-dithiobis-(2-nitrobenzoic acid); INH, isoniazid; NAT, arylamine *N*-acetyltransferase; PABA, *p*-aminobenzoate; PANS, *p*-anisidine; PA, procainamide; PAS, *p*-aminosalicylate; SMZ, sulphamethazine; SNP, single nucleotide polymorphism.

* Corresponding author at: Democritus University of Thrace, Department of Molecular Biology and Genetics, Building 10, University Campus (Dragana), Alexandroupolis 68100, Greece. Tel.: +30 25510 30628; fax: +30 25510 30613.
E-mail address: gfakis@mbg.duth.gr (G. Fakis).
[1] Equal contribution of authors.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.11.085

characterized at the structural level (Sinclair et al., 2000), opening the way towards similar studies with human NATs (Rajasekaran et al., 2011; Wu et al., 2007). Laboratory models have also been central to investigations of the enzyme properties and in vivo functions of human NATs (Sim et al., 2008). Mice, in particular, have been useful as pharmacological and toxicological models of human NATs, as various "rapid" and "slow" acetylator strains have been characterized (Boukouvala et al., 2002; Erickson et al., 2008; Martell et al., 1991). Moreover, various transgenic mouse models have been generated for NAT (Cornish et al., 2003; Sugamori et al., 2003, 2011), and these have been used to study the carcinogenicity of compounds (Sugamori et al., 2006, 2012). Such engineered strains have also been used to conduct studies on the epigenetic regulation (Wakefield et al., 2010) and developmental expression (Wakefield et al., 2008) of murine *Nat* genes, advancing our understanding of the endogenous roles of mammalian NATs (Butcher and Minchin, 2012). The mapping and functional characterization of basal transcriptional elements of human *NAT* genes were also facilitated by the high degree of synteny shared with the orthologous murine locus (Boukouvala et al., 2003; Fakis et al., 2000).

Little is known about NATs in non-human primates, although comparison of the sequenced genomes of human and other primate species suggests a highly conserved *NAT* locus (Sabbagh et al., 2013). Studies have also implicated *NAT* polymorphism as a possible toxicogenetic contributor to endometriosis (Baranova et al., 1999; Nakago et al., 2001), prompting us to investigate the *NAT2* gene in the rhesus monkey, a model animal for gynaecologic disorders (Fakis et al., 2007). In this previous study, we cloned two polymorphic variants of the *NAT2* gene (alleles *NAT2*1* and *NAT2*2*), which were different by one non-synonymous (c.691G>A) single nucleotide polymorphism (SNP) causing a p.Val231Ile amino acid substitution. Expression of the allozymes in mammalian cells suggested that the p.Val231Ile polymorphism might affect substrate selectivity of NAT2. In the present study, we have engineered the c.691G>A SNP into the human *NAT2*4* reference allele, and we report enzymatic analysis of the recombinant rhesus and human p.Val231Ile variants with NAT1- and NAT2-selective substrates, complemented by differential scanning fluorimetry and molecular modelling.

2. Methods

2.1. Generation and validation of gene constructs

The (MACMU)*NAT2*1* (GenBank: AJ504440) and (MACMU)*NAT2*2* (GenBank: AJ504439) alleles of the rhesus monkey (*Macaca mulatta*, Taxonomy: 9544) were previously cloned in pTargeT mammalian expression vector (Promega), as described (Fakis et al., 2007). The (HUMAN)*NAT2*4* reference allele (GenBank: X14672) was available in pET28b(+) *Escherichia coli* recombinant expression vector (Novagen). Primers were from VBC-Genomics. Sequencing was carried out by Macrogen. Gene, allelic and SNP nomenclature is according to the recommendations of the Human Gene Nomenclature Committee (HGNC: 7646), the Arylamine *N*-acetyltransferase Gene Nomenclature Committee (http://nat.mbg.duth.gr/background_2013.html) and the Human Genome Variation Society (http://www.hgvs.org/mutnomen/).

The inserts of pTargeT constructs [(MACMU)*NAT2** alleles] were amplified from the ends with primers incorporating restriction sites for in-frame transfer into the pET28b(+) vector [primer sequences: 5′-GCGG CAGCCATATGGACATTGAAGCATA-3′ (forward)/5′-GTGCTCGAGTGCGG CCGCCTAAATAGTGAAGGATC-3′ (reverse), with underlined *Nde*I and *Not*I restriction sites, respectively]. Amplifications were carried out using *Pfu*-DNA polymerase (Finnzymes), according to the manufacturer's instructions, and each PCR cycle included a primer annealing step at 56 °C. Gel-purified PCR products (QIAquick PCR Purification kit, QIAGEN) were subjected to digestion (37 °C) with *Nde*I and *Not*I (both from Takara) and purified again, prior to ligation (16 °C, 0.1 U/μl T4-DNA ligase by Promega) to 50 ng of pET28b(+) vector linearized with the same restriction enzymes. *E. coli* JM109 High Efficiency Competent Cells (Promega) were transformed by heat-shock, and recombinant clones were selected on standard LB-kanamycin (50 μg/ml) agar medium. Plasmid from sequence-validated clones was isolated (QIAprep Spin Miniprep Kit, QIAGEN) and transformed into *E. coli* BL21 (DE3) codon plus RIL competent cells (Promega).

Incorporation of the c.691G>A (p.Val231Ile) polymorphism into the open reading frame of the (HUMAN)*NAT2*4* reference allele in pET28b(+) vector was carried out by site-directed mutagenesis using the QuikChange II kit (Agilent), as instructed by the manufacturer. Two mutagenic primers were used [5′-GCAGACCCCAGAAGGGATTTA CTGTTTGGTGG-3′ (forward)/5′-CCACCAAACAGTAAATCCCTTCTGGGGT CTGC-3′ (reverse), with the mutation underlined], and the cycling conditions were 95 °C for 30 s, followed by 14 cycles for 30 s at 95 °C, 1 min at 55 °C and 6 min at 68 °C. The mutagenesis product was transformed into *E. coli* XL1-Blue Super-Competent cells (Agilent), grown on LB-kanamycin (50 μg/ml) agar medium. Mutagenized clones were selected by amplification with vector-specific primers (T7-promoter/T7-terminator), followed by digestion (37 °C) with *Hpy*8I (*Mja*IV) restriction endonuclease (Fermentas) and inspection of the change in digestion pattern (loss of *Hpy*8I restriction site at c.T694 of (HUMAN) *NAT2*4*, due to mutation at c.G691). Plasmid from sequence-validated clones was eventually transformed into *E. coli* BL21 (DE3) codon plus RIL competent cells (Promega), for recombinant protein expression.

The inserts of all pET28b(+) constructs used in subsequent experiments were verified by sequencing from both ends (vector-specific primers), and inspection of sequences with BioEdit Sequence Alignment Editor v.7.0.5.3.

2.2. Expression and purification of recombinant proteins

Frozen stocks (−80 °C, 10% v/v glycerol) of the pET28b(+) *NAT2* constructs in *E. coli* BL21 (DE3) codon plus RIL cells were used to initiate 10 ml cultures, grown overnight (37 °C, 180 rpm) in LB-kanamycin (50 μg/ml) medium. One milliliter of those cultures was used to inoculate 250 ml of fresh LB-kanamycin media, followed by incubation (37 °C, 120 rpm) until the OD (600 nm) was approximately 0.6. Induction of recombinant protein expression was carried out with isopropyl-1-thio-D-galactopyranoside (IPTG), added to final concentration of 0.5 mM. Following incubation for 4.5 h (25 °C, 120 rpm), the cells were harvested by centrifugation (3500 *g*, 20 min at 4 °C) and resuspended in 50 mM Tris–HCl (pH 7.5)/50 mM NaCl buffer, supplemented with a protease inhibitor cocktail (Thermo Scientific). Cell lysis was performed by sonication on ice (20 or more cycles, as necessary, of 45 s vibration, followed by 30 s of pause) and the soluble and insoluble fractions were separated by centrifugation at 12,000 *g* for 40 min (4 °C). Supernatants were either stored (−80 °C) in aliquots, or immediately processed. Lysates from uninduced *E. coli* BL21 (DE3) codon plus RIL cells were also prepared in a similar way, for use as controls in enzyme activity assays.

Purification of *N*-terminal hexa-histidine tagged recombinant NAT2 proteins was performed by affinity chromatography of recovered soluble fractions, through 0.5 ml of commercial nickel-charged Profinity IMAC resin (BioRad). The protein was allowed to bind to the resin by gentle orbital shaking for 3 h (4 °C), and the column was then washed with 2.5 ml of buffer (20 mM Tris–HCl pH 7.5, 50 mM NaH_2PO_4, 300 mM NaCl). The recombinant protein was subsequently eluted with 2.5 ml of the same buffer, supplemented with up to 400 mM imidazol. The eluate was stored (−80 °C) in aliquots that were thawed only once for immediate processing on ice.

Samples from each cell lysate (including uninduced controls) and from purified recombinant protein preparations were inspected for yield and purity by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Quantification of total recovered protein was performed using the Bradford's reagent (Sigma-Aldrich) against appropriate bovine serum albumin (BSA) standards, as instructed by the