Expression, purification, characterization and structure of *Pseudomonas aeruginosa* arylamine N-acetyltransferase

Isaac M. WESTWOOD*, Simon J. HOLTON†, Fernando RODRIGUES-LIMA‡§, Jean-Marie DUPRET‡§, Sanjib BHAKTA*, Martin E. M. NOBLE† and Edith SIM*¹

*Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, U.K., †Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K., ‡CNRS-UMR 7000, Faculté de Médecine Pitié-Salpêtrière, 105 boulevard de l'Hôpital, 75013 Paris, France, and §UFR de Biochimie, Université Denis Diderot-Paris7, 75005 Paris, France

The gene for NAT (arylamine N-acetyltransferase) from *Pseudomonas aeruginosa* (*panat*) has been cloned from genomic DNA, and the gene product (PANAT) expressed as an N-terminal histidine-tagged protein in *Escherichia coli* and purified via nickel ion affinity chromatography. The specific activities of PANAT against a broad range of substrates have been investigated and compared with those of other prokaryotic NAT enzymes. For most arylamine substrates identified, PANAT exhibits *in vitro* specific activities typically one order of magnitude greater than those of recombinant NAT enzymes from *Mycobacterium smegmatis* or *Salmonella typhimurium*. Among the substrates of PANAT so far identified are the anti-tubercular drug isoniazid, 5-aminosalicylate (a drug used in the treatment of inflammatory bowel disease), as

well as important environmental pollutants such as 3,4-dichloroaniline and 2-aminofluorene. As well as acetylating common NAT substrates, PANAT is unique among the prokaryotic NATs so far studied in acetylating the folate precursor 4-aminobenzoic acid and the folate catabolite 4-aminobenzoylglutamate. The recombinant protein has been expressed in sufficient quantity to allow protein crystallization, and we have subsequently determined the 1.95 Å structure of PANAT by X-ray crystallography.

Key words: 5-aminosalicylic acid, arylamine N-acetyltransferase (NAT), *Pseudomonas aeruginosa*, xenobiotic metabolism, X-ray crystallography.

INTRODUCTION

NATs (arylamine N-acetyltransferases) constitute a family of drug-metabolizing enzymes of molecular mass 30–34 kDa that are found in a range of eukaryotes and prokaryotes [1]. NATs catalyse the transfer of an acetyl group from AcCoA (acetyl-CoA) to arylamines, arylhydrazines and hydroxyarylamines [2–5]. The NAT enzyme was first identified in humans in the 1960s due to its ability to inactivate the anti-tubercular drug isoniazid (INH) [6]. NATs have been identified more recently in a range of prokaryotes, including *Salmonella typhimurium*, *Mycobacterium smegmatis*, *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* [3,7–9].

According to current World Health Organization reports, approximately one-third of the world's population carry the *M. tuberculosis* bacterium, the causative agent of tuberculosis (www.who.int/gtb). The recent emergence of multidrug-resistant strains of *M. tuberculosis* has reduced the efficacy of the traditional three-drug anti-tubercular therapy of isoniazid, rifampicin and pyrazinamide [10,11], and so there is an urgent need to increase our understanding of tuberculosis biology.

Previous attempts to study TBNAT (NAT from *M. tuberculosis*; NAT104*1) have been unsuccessful due to low yields of recombinant protein [8]. *Escherichia coli*-based expression levels of recombinant NAT enzymes from both *M. smegmatis* (NAT103*1; MSNAT) and *S. typhimurium* (NAT101*1; STNAT) are significantly higher [8,12], and hence much of our biochemical and

structural knowledge of NATs is based on these proteins. The X-ray crystal structures of MSNAT and STNAT have revealed a conserved three-domain fold whose active site contains a triad of residues (Cys-His-Asp) that is proposed to catalyse substrate acetylation [13,14]. *In vitro*, both MSNAT and STNAT have been shown to acetylate a range of known NAT substrates, including 5-AS (5-aminosalicylate) and isoniazid [8,12,14–17]. Furthermore, it has been shown that overexpression of TBNAT in *M. smegmatis* decreases the susceptibility of the organism to isoniazid, while a genetic *nat* knockout in *M. smegmatis* exhibited heightened susceptibility to the drug [8,18]. More recently, characterization of a genetic *nat* deletion mutant of the *M. tuberculosis* model organism, *Mycobacterium bovis* BCG, has indicated that NAT is essential for mycolic acid biosynthesis in this organism [19].

Using the TBNAT sequence as a search template, we have identified a single putative NAT-encoding DNA sequence (*panat*) within the *P. aeruginosa* genome. Despite sharing only low sequence identity at the amino acid level (35%), the PANAT amino acid sequence contains many of the sequence motifs associated with NATs, for example the catalytic triad of Cys-His-Asp residues and the Phe-Glu-Asn-Leu (FENL) motif [20]. Interestingly, whereas STNAT and TBNAT share 43% sequence identity over the first and second domains, but only 16% identity in the third domain, the pattern of sequence identity between PANAT and TBNAT is distributed fairly homogeneously throughout the protein ($\approx 35\%$). The third domain of NATs confers substrate specificity

Abbreviations used: AcCoA, acetyl-CoA; 5-AS, 5-aminosalicylate; LB, Luria-Bertani; MALDI-TOF, matrix-assisted laser desorption/ionization-time-of-flight; MSNAT, NAT from *Mycobacterium smegmatis*; NAT, arylamine N-acetyltransferase; pABA, 4-aminobenzoic acid; pABA-Glu, 4-aminobenzoylglutamate; PANAT, NAT from *Pseudomonas aeruginosa*; RMSD, root mean square deviation; RT-PCR, reverse-transcription-PCR; STNAT, NAT from *Salmonella typhimurium*; TBNAT, NAT from *Mycobacterium tuberculosis*.

¹ To whom correspondence should be addressed (email edith.sim@pharm.ox.ac.uk).

The co-ordinates of *Pseudomonas aeruginosa* arylamine N-acetyltransferase have been submitted to the RSCB PDB under PDB identifier 1W4T.

[15], and hence in order to improve and complement our current enzymic and structural models of TBNAT in drug screening, we have recombinantly expressed and purified the NAT from *P. aeruginosa* (PANAT).

5-AS is commonly used in the treatment of inflammatory bowel diseases, such as Crohn's disease and ulcerative colitis [21]; however, its efficacy is reduced as a result of acetylation. Cultures containing *P. aeruginosa* have been demonstrated to acetylate and hence inactivate 5-AS [3,22]. *P. aeruginosa* is found in the gut, and so it is important to determine whether PANAT contributes to the acetylation of 5-AS.

A detailed characterization of PANAT is therefore pertinent to inflammatory bowel disease treatment and as a model system for the design of anti-tubercular agents. As a contribution to both of these research areas, in this paper we present the 1.95 Å structure of PANAT, determined by X-ray crystallography, and a detailed profile of PANAT activity against a broad range of substrates.

EXPERIMENTAL

Materials

All chemicals and reagents were purchased from Sigma Aldrich (Poole, Dorset, U.K.).

Protein and RNA isolation, and RT-PCR (reverse transcription-PCR)

Total RNA was isolated from a 4 ml culture of *Pseudomonas aeruginosa* (PA01 strain) using TRIzol (Gibco BRL, Cergy-Pontoise, France). Protein, produced as a by-product, was used for gel electrophoresis and Western blotting. cDNA was synthesized from 1 µg of total RNA template using the *C. therm.* Polymerase RT-PCR kit (Roche, Meylan, France). The reaction was performed in a total volume of 20 µl at 70°C using the antisense primer 5'-GGAATTCAGGCGGAGATCAGTCCCGCCA-3'.

For PCR analysis, 0.25 µg of cDNA was used as a template in a 50 µl amplification mixture containing 200 µM of each deoxynucleotide triphosphate, 10 mM Tris/HCl, pH 8.3, 1.5 mM MgCl₂, 50 mM KCl, 5% (v/v) DMSO, 0.3 µM of each primer and 1.5 units of DyNAzyme EXT DNA polymerase (Ozyme, Saint-Quentin en Yvelines, France). The primers used in the PCR were 5'-CGAGTGCCATATGACCCCGCTGACGCCCGA-3' (sense) and 5'-GGAATTCAGGCGGAGATCAGTCCCGCCA-3' (antisense). The PCR conditions were as follows: denaturation (94°C, 1 min), followed by 35 cycles of combined annealing and polymerization (72°C, 4 min) and denaturation (94°C, 1 min), followed by a final extension step (72°C, 10 min). To assess whether genomic DNA had contaminated the RNA samples, an additional PCR reaction was carried out under the same conditions with purified total RNA (1 µg). The PCR products were separated by gel electrophoresis in ethidium bromide-stained agarose (1.8%, w/v).

Cloning

The *panat* open reading frame was amplified from *P. aeruginosa* strain PA01 genomic DNA using PCR with the primers 5'-CGAGTGCCATATGACCCCGCTGACGCCCGA-3' (sense) and 5'-GGAATTCAGGCGGAGATCAGTCCCGCCA-3' (antisense). The start and stop codons for the *nat* gene are shown in bold type, and *Nde*I and *Eco*RI restriction sites are shown in italics. The PCR used *Pfu* DNA polymerase (Promega) and consisted of a denaturation step (95°C, 2 min) followed by 30 cycles of annealing (60°C, 30 s), extension (72°C, 2 min) and denaturation (95°C, 1 min), and a final extension step (72°C, 5 min). The PCR product was subcloned into the pET28b(+) vector (Novagen)

using the *Nde*I and *Eco*RI restriction sites. The pET28b(+) vector containing the *panat* insert was transformed initially into *Escherichia coli* strain JM109, which was grown in LB (Luria-Bertani) medium containing kanamycin (30 µg/ml). The identity of the insert was confirmed by DNA sequencing (University of Oxford Biochemistry Department Sequencing Service).

Expression and purification

The pET28b(+)-*panat* plasmid was transformed into the expression host *E. coli* BL21(DE3)pLysS as described previously [12]. An overnight culture was diluted into 100 vol. of LB broth containing kanamycin (30 µg/ml). Bacterial cultures of 1 litre were grown in a shaking incubator at 37°C to a *D*₆₀₀ of 0.4–0.6. Expression of PANAT was induced by addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 0.25 mM. The culture was maintained at 37°C until the *D*₆₀₀ had reached 1.8–2.0, at which point cells were harvested by centrifugation (5000 g, 4°C, 20 min). The cell pellet was resuspended in 15 ml of lysis buffer [300 mM NaCl, 20 mM Tris/HCl, pH 8.0, 1 × EDTA-free Complete Protease Inhibitor (Roche)] for each litre of original culture. The resuspended pellet was frozen and stored at –70°C. Pellets were thawed in a water bath (37°C) before sonication to complete the cell lysis. The lysate was clarified by centrifugation (12000 g, 4°C, 20 min), and then batch-bound to Ni²⁺-nitrilotriacetate resin (Qiagen) (pre-equilibrated in 300 mM NaCl, 20 mM Tris/HCl, pH 8.0) for 10 min at 4°C. The resin was packed into a PD-10 column (Amersham Biosciences, Chalfont St. Giles, Bucks., U.K.), and then washed with solutions (300 mM NaCl, 20 mM Tris/HCl, pH 8.0) containing increasing concentrations of imidazole (pH 8.0) at 4°C. The wash sequence was typically 10, 25, 125 and 250 mM imidazole (two 10 ml washes of each). Fractions containing pure PANAT were identified by SDS/PAGE [23], pooled and dialysed against 200 vol. of buffer (20 mM Tris/HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol) at 4°C for 16 h.

Western blot analysis

Protein samples were separated by SDS/PAGE and transferred to a nitrocellulose membrane (Hybond-C; Roche) for Western blot analysis as described previously [8]. An antibody raised against STNAT [8] was used to detect both PANAT in the *P. aeruginosa* lysate (antibody dilution 1:10000) and recombinant purified PANAT (antibody dilution 1:50000).

Enzymic assays

For the assays described below, the concentrations of NAT enzyme samples were adjusted to ensure a linear initial rate of reaction. Kinetic analyses were performed by varying the substrate concentration, and kinetic constants were determined from Hanes plots [24].

Acetylation of arylamines

Detection of the acetylation of arylamines was performed as described previously [25]. Samples of enzyme (0.3–1.5 pmol) and arylamine (50–200 µM) in assay buffer (20 mM Tris/HCl, 1 mM dithiothreitol, pH 8.0) were pre-incubated at 37°C for 5 min. Ac-CoA (400 µM final concentration) was added and the reaction was allowed to proceed at 37°C. The reaction (100 µl total volume) was quenched with 100 µl of aqueous trichloroacetic acid (20%, w/v) at 0°C. 4-Dimethylaminobenzaldehyde [800 µl; 5% (w/v) in 9:1 (v/v) acetonitrile/water] was then added and the absorbance was measured in a 10 mm path length cuvette at

$\lambda_{\max} = 450$ nm (Hitachi U-2001 spectrophotometer). The amount of arylamine remaining was determined by reference to a standard curve.

Hydrolysis of AcCoA

The rate of hydrolysis of AcCoA by NAT in the presence of an arylamine or hydrazine substrate was determined as described previously [17]. PANAT was diluted in assay buffer (20 mM Tris/HCl, pH 8.0) immediately prior to use to ensure that the final concentration of dithiothreitol was less than $1 \mu\text{M}$, which has been demonstrated previously not to interfere with the assay [17]. Samples of enzyme (0.3–1.5 pmol) and substrate ($500 \mu\text{M}$) in assay buffer were pre-incubated at 37°C for 5 min in a 96-well plate. AcCoA ($400 \mu\text{M}$ final concentration) was added to start the reaction and the plate was incubated at 37°C . The reaction ($100 \mu\text{l}$ total volume) was quenched with $25 \mu\text{l}$ of guanidine hydrochloride solution (6.4 M guanidine/HCl, 0.1 M Tris/HCl, pH 7.3) containing 5 mM 5,5'-dithiobis-(2-nitrobenzoic acid), and then the absorbance was measured at $\lambda_{\max} = 405$ nm (Anthos 2020 plate reader). The amount of CoA produced was determined by reference to a standard curve.

X-ray crystallography and data collection

PANAT was concentrated to 7–14 mg/ml using Amicon ultra-centrifugation concentrators (Millipore, Watford, Herts., U.K.). The crystals described were grown at 19°C using the sitting-drop vapour diffusion technique by mixing equal volumes ($1 \mu\text{l}$) of the concentrated PANAT solution with mother liquor {Molecular Dimensions Screen II; condition number 28 [100 mM Hepes, pH 7.5, 20% poly(ethylene glycol) 10K]}. These crystals are grown reproducibly by both hanging- and sitting-drop vapour diffusion methods [26], and both from manual 24-well plates and using a protein crystallization robot (Tecan Genesis 150 robot; Laboratory of Molecular Biophysics, University of Oxford). Crystals of dimensions of $15 \mu\text{m} \times 15 \mu\text{m} \times 100 \mu\text{m}$ typically grew after 2–5 days.

Prior to data collection, crystals were transferred briefly to a cryo-protection solution of 6 M sodium formate, before flash freezing in a nitrogen cryo-stream at 100 K (Oxford Cryosystems). Data were collected on beamline ID29 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) with an ADSC Quantum-4 CCD detector. The space group was determined, and data were integrated, scaled and merged using the programs MOSFLM and SCALA [27].

RESULTS

In order to determine whether the *panat* gene, identified from the complete genome sequence of *P. aeruginosa* strain PA01 [28], is expressed endogenously, RNA was extracted from *P. aeruginosa* cultures and RT-PCR was performed with NAT-specific primers. A PCR product of the expected size (855 bp) was observed (Figure 1A), indicating that the *panat* gene is expressed endogenously by *P. aeruginosa*. A Western blot using a monospecific antiserum against STNAT [8] indicated the presence of a cross-reacting protein with an approximate molecular mass of 31 kDa in lysates of *P. aeruginosa* (Figure 1B). The antibody has subsequently been shown to recognize pure recombinant PANAT (approx. 33.5 kDa with a hexa-histidine tag; Figure 1B).

The *panat* gene was then amplified by PCR, producing the expected product of size 870 bp, and cloned into the pET28b(+) expression vector to facilitate the overexpression of PANAT with an N-terminal hexa-histidine tag. The protein was purified to apparent homogeneity using Ni^{2+} -ion affinity chromatography

Figure 1 Endogenous expression of PANAT and purification of the recombinant enzyme

(A) Ethidium bromide-stained 1.8% (w/v) agarose gel showing the 855 bp fragment obtained after RT-PCR using total RNA isolated from *Pseudomonas aeruginosa* PA01 (lane 2). The absence of genomic DNA contamination was confirmed using a negative control with no reverse transcriptase in the RT-PCR experiment (lane 1). (B) Western blot showing PANAT detected in *P. aeruginosa* lysate using an antibody raised against STNAT at 1:10 000 dilution (lane 1). Recombinant pure PANAT with a hexa-histidine tag ($1.0 \mu\text{g}$) was detected using the anti-STNAT antibody at a 1:50 000 dilution (lane 2). (C) Purification of PANAT as described in the Experimental section. Lanes: M, markers; 1, insoluble lysate; 2, soluble lysate; 3, unbound lysate; 4, 10 mM imidazole wash; 5, 25 mM imidazole wash; 6, 125 mM imidazole wash; 7, 250 mM imidazole wash; 8, PANAT after 2 months at 4°C .

(Figure 1C). Typically, 30 mg of PANAT was purified per litre of bacterial culture. The identity of PANAT was confirmed by tryptic digestion and subsequent analysis of the fragments by MALDI-TOF (matrix-assisted laser desorption ionization–time-of-flight) MS analysis (see Supplementary Table S1 at <http://www.BiochemJ.org/bj/385/bj3850605add.htm>).

The apparent Michaelis constant and maximum reaction velocity for AcCoA were determined by following the acetylation of *p*-anisidine [K_m (app) = $136 \mu\text{M}$; $V_{\max} = 13280 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg PANAT}^{-1}$]. Substrates that have been shown to be acetylated by MSNAT, STNAT and TBNAT, as well as by bacterial cell lysates [3], were screened for substrate activity with PANAT (Table 1). Where it has been possible to compare these specific activities with literature values for the pure recombinant enzymes STNAT and MSNAT [17], the relative specific activities are shown in Figure 2. The activity of PANAT for the arylamine substrates, including 2-aminofluorene, the 4-haloanilines (4-chloroaniline, 4-bromoaniline, 4-iodoaniline) and 5-AS, is typically an order of magnitude greater compared with the recombinant MSNAT and STNAT enzymes. This trend does not extend to the hydrazine substrates, which constitute the most rapidly acetylated structural class of substrates for both STNAT and MSNAT. Additionally, PANAT acetylates the folate precursor pABA (4-aminobenzoic acid) and the folate catabolite pABA-Glu (4-aminobenzoylglutamate), both of which are substrates for the human NAT1 enzyme [29,30]. pABA acetylation activity is rare in prokaryotes, having been identified only in *P. aeruginosa* PA01, *Bacteroides* sp. and *S. typhimurium* YG1024 [3]. Acetylation of pABA-Glu by prokaryotic NATs has only been reported for STNAT [31].

The $1.95 \text{ \AA}$ structure of PANAT was solved by molecular replacement using the program MOLREP [27]. The search model

Table 1 Substrates of PANAT

The rate of hydrolysis of AcCoA ($\text{nmol} \cdot \text{min}^{-1} \cdot \text{mg}$ of PANAT $^{-1}$) was measured in the presence of PANAT, AcCoA ($400 \mu\text{M}$) and compound ($500 \mu\text{M}$) as described in the Experimental section. nd, no detectable activity.

Class/compound	Short name	Rate ($\text{nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$)
Arylamine drugs		
Sulphacetamide	SACa	nd
Sulphanilamide	SAnA	nd
Sulphadiazine	SDZ	nd
Sulphamethazine	SMZ	nd
Trimethoprim	TMP	nd
Sulphisoxazole	SIOZ	1000 ± 2
Sulphamethoxazole	SMOZ	1310 ± 5
5-Aminosalicylate	5-AS	73300 ± 3300
Other arylamines		
4-Aminobenzoic acid L-glutamate	pABA-Glu	532 ± 60
Aniline	AL	629 ± 40
4-Trifluoromethyl aniline	TFMA	2760 ± 233
4-Chloroaniline	4-CA	4140 ± 136
4-Bromoaniline	4-BA	5710 ± 622
3,4-Dichloroaniline	DCA	6760 ± 583
4-Aminobenzoic acid	pABA	8200 ± 78
4-Iodoaniline	4-IA	13800 ± 4860
4-Aminosalicylate	4-AS	18660 ± 467
4-Phenoxyaniline	POA	21970 ± 972
2-Aminofluorene*	2-AF	44710 ± 2720
Alkoxyanilines		
4-Hexyloxyaniline*	HOA	12050 ± 1170
4-Anisidine	ANS	13470 ± 0
4-Ethoxyaniline	EOA	19240 ± 1750
4-Butoxyaniline*	BOA	25660 ± 1170
4-Aminoveratrole	AMV	44510 ± 3690
Hydrazines		
Phenylhydrazine	PHZ	nd
4-Methoxyphenylhydrazine	MPZ	641 ± 603
Isoniazid	INH	2324 ± 0
4-Chlorobenzoic hydrazide	CBZ	14000 ± 1940
Hydralazine	HDZ	29550 ± 3110

* Assay performed in the presence of 5% (v/v) DMSO.

was created by modifying the high-resolution MSNAT structure (PDB identifier 1GX3), such that residues not conserved in PANAT were truncated to the C_β atom. Using this model, molecular replacement identified a single copy of PANAT within the asymmetric unit, consistent with a Matthew's coefficient of $2.5 \text{ \AA}^3 \cdot \text{Da}^{-1}$ and a solvent content of 49.7%. After manual rebuilding using program O [32], the PANAT structure was initially refined via a simulated-annealing step using the CNS suite of programs [33]. Subsequent refinement involved iterative manual rebuilding and maximum likelihood refinement using the program REFMAC5 [27]. A final refinement step was to solvate the protein, using the program ARP [27]. The stereochemical quality of the final model was verified using the program Procheck [34]. In the structure, 94.9% of residues were found in the most favoured regions, 4.7% were found in additionally allowed regions, and one residue (Asp²²⁵; 0.4%) was found in a generously allowed region. Clear electron density supports the conformation of Asp²²⁵. The final PANAT model has an R_{factor} of 19.1% and an R_{free} of 22.4%, consisting of a single molecule in the asymmetric unit, with an additional 133 water molecules. Electron density was attributable to residues 1–276, while the C-terminal three residues (Ile-Ser-Ala) could not be confidently located. Although the N-terminal hexa-histidine motif and linker residues (20 residues) could not be reliably traced, after 2 months at 4 °C there was no apparent degradation of the enzyme, as shown by SDS/PAGE (lane 8, Figure 1C) and no change in activity was observed (results not shown). This suggests that these residues are disordered and not degraded within the crystal. Co-ordinates have been submitted to the RSCB PDB under PDB identifier 1W4T (<http://www.rcsb.org/pdb/>). A summary of the final model statistics is presented in Supplementary Table S2 (<http://www.BiochemJ.org/bj/385/bj3850605add.htm>).

DISCUSSION

In this paper we have described the identification of the *panat* gene of *Pseudomonas aeruginosa* through its sequence identity with the NAT family of proteins. The *panat* open reading frame has

Figure 2 Comparison of the specific activities of PANAT with those of other bacterial NATs

Shown is the rate of hydrolysis of AcCoA (nmol of CoA $\cdot \text{min}^{-1} \cdot \text{mg}$ of protein $^{-1}$) in the presence of AcCoA ($400 \mu\text{M}$), substrate ($500 \mu\text{M}$) and NAT enzyme. PANAT from *P. aeruginosa* (black bars), MSNAT (grey bars) and STNAT (unshaded bars) were compared. MSNAT and STNAT data are literature values [17]. The same method was used for all three enzymes. Each bar represents the mean $\pm$ S.D. from triplicate determinations. Abbreviations are as described in Table 1.

Figure 3 X-ray crystal structure of PANAT

Orthogonal views of the X-ray crystal structures of PANAT (blue/purple), MSNAT (orange) and STNAT (green) superimposed using the program O [32]. All images were generated using Aesop [M. E. M. Noble, unpublished work; details available from M. E. M. N. on request (martin@biop.ox.ac.uk)]. Representative final $2F_{\text{obs}} - F_{\text{calc}}$ electron density for the active-site triad of PANAT is contoured at $0.2\text{e}^{-\text{\AA}^{-3}}$ in green.

been cloned and the gene has been expressed to generate sufficient levels of recombinant protein for both *in vitro* activity assays and X-ray crystallographic studies. Both our structural and biochemical data confirm the identification of the *panat* gene as encoding a NAT enzyme; PANAT is able to catalyse the *in vitro* acetylation of a broad range of arylamines and arylhydrazines (Table 1, Figure 2), and the X-ray structure of PANAT reveals a fold that is highly reminiscent of both MSNAT and STNAT structures (Figure 3). The anti-STNAT antibody cross-reacts with the PANAT recombinant protein (Figure 1B), as it does with MSNAT and, to a lesser extent, TBNAT [8].

There are now three model NAT proteins that readily crystallize and for which X-ray crystallographic structures have been determined [14,15]. Crystals from two of these proteins routinely diffract to better than $2\text{ \AA}$, and so provide good model systems with which to investigate protein–ligand interactions. The aims are twofold: first the improvement and validation of our current TBNAT model [18], for which no structural data are currently available; and secondly to facilitate the rational structure-led ligand design of small-molecule inhibitors with potential therapeutic value. Following on from studies of other bacterial NATs [20], in the present paper we have compared both the enzymic rates of acetylation and the structural properties of PANAT with both those of MSNAT and STNAT.

The X-ray crystallographic structure reveals that, despite low sequence similarity (Figure 4), PANAT shares an almost identical fold with that of both STNAT and MSNAT (Figure 3) [13,14]. PANAT has an RMSD (root mean square deviation) of $1.5\text{ \AA}$ with STNAT (over 248 equivalent residues) and of $1.3\text{ \AA}$ with MSNAT (over 240 equivalent atoms). The structure may be divided into three domains of approximately equal size. The first two domains, a helical bundle and a β -barrel, are aligned in such a way that three residues (Cys⁷³, His¹¹³ and Asp¹²⁸; *P. aeruginosa* numbering) form

a catalytic triad. The structural motif that these residues form is well documented in NATs and is identical with that of the cysteine protease superfamily [8,13,15,35–37]. The catalytic triad residues superpose with an RMSD of only $0.054\text{ \AA}$ (C_{α} residues), while a more general comparison of active-site residues in the different NAT structures reveals a significantly higher level of sequence identity than suggested by the global similarity, and further reinforces the similarity between these proteins (Figure 3). The third domain is linked to the second via an interdomain helix (approx. residue 180). The interface formed between domains II and III and the lid region together form a substantial active-site cleft.

Minor differences between PANAT and other NAT structures may be categorized as either (a) single point insertions/deletions, which are accommodated within loop regions of the structure and result in little perturbation of the protein fold, and (b) mobile loop regions between more stable secondary structural elements. These are located between $\beta 7/\beta 8$, $\beta 9/\beta 10$, $\beta 12/\beta 13$ and $\beta 13/\beta 14$. Similar conformations for these regions are seen in a second PANAT crystal form (results not shown), suggesting that the differences are a property of PANAT, and not due to an artefact of packing within the crystal lattice. Such differences are minor, since PANAT clearly shares the NAT fold. The invariant phenylalanine contained within the highly conserved Pro-Phe-Glu-Asn-Leu motif (residues 41–45; PANAT numbering) has been implicated in substrate binding [38], and so it is interesting to note that this sequence is Ala-Phe-Glu-Asn-Leu in PANAT. Surprisingly, the change from proline to alanine has no effect on the NAT fold, with both local secondary structural elements and individual side-chain conformations being identical in this region of MSNAT and PANAT. The change in residue leaves the main-chain nitrogen atom with an uncompensated hydrogen bond, which in the PANAT structure is compensated for by an additional water molecule

Figure 4 Multiple sequence alignment of selected bacterial NATs

The sequence alignment was constructed using ClustalX [51] and ESPrnt [52]. The secondary structural elements were identified from the PANAT structure using ESPrnt. α -Helices, η -helices, β -sheets and strict β -turns are denoted α , η , β and TT respectively. Grey stars indicate side chains for which multiple conformations were modelled. SwissProt accession numbers: *P. aeruginosa*, Q9HUY3; *M. smegmatis*, O86309; *S. typhimurium*, Q00267; *M. tuberculosis*, P96848. Similar amino acids are highlighted in boxes, and completely conserved residues are indicated by white lettering on a dark background.

forming hydrogen bonds with the carbonyl groups of both Arg³⁸ and Phe²⁵⁶.

This is now the third example of a crystallographic structure of a bacterial NAT, and suggests that other bacterial NAT proteins are likely to share an identical fold. Whether this conservation extends to more phylogenetically distant species remains to be seen, but the high level of bacterial structural conservation serves to support the validity of our previously published model structure of TBNAT [18].

To investigate whether the structural similarity is extended to the enzymic properties of NAT proteins, we measured the acetylation of a broad range of substrates, including those for which N-acetylation activity has been reported previously for *P. aeruginosa* cultures [3,9]. These included 2-aminofluorene, 5-AS, 4-aminosalicylic acid and pABA. Although the specific activity profile of PANAT substrates is broadly similar to those of MSNAT and STNAT, we have identified some key differences. Firstly, PANAT is significantly more active against arylamine substrates than is MSNAT or STNAT (Figure 2). Secondly, the relationship between the rate of acetylation and hydrophobicity of alkoxy-anilines by MSNAT {*p*-anisidine < 4-aminoveratrole < 4-ethoxy-aniline < 4-butoxyaniline < 4-hexyloxyaniline (Figure 2) [17]} is not conserved for PANAT, for which the rate of acetylation is in the order 4-hexyloxyaniline < *p*-anisidine < 4-ethoxy-aniline < 4-butoxyaniline < 4-aminoveratrole (Figure 2). Lastly, the arylhydrazine substrate hydralazine, which for both MSNAT and STNAT is the most rapidly acetylated substrate tested, is acetylated more slowly when compared with other PANAT substrates.

We propose that differences in the distribution of surface hydrophobicity in the MSNAT and PANAT active sites are at least partially responsible for the differing specific activity profiles (Figures 2 and 5). The cores of both NAT active sites do not possess significant hydrophobic character, in keeping with the presence

of several water molecules observed in this region of both crystal structures. However, regions of hydrophobicity are found on the edge of both active sites, albeit in differing locations. It is also interesting to note that, despite the overall fold conservation, subtle changes to side-chain conformations result in quite differing surface profiles. Co-crystallization experiments are currently under way in order to identify the binding site of these substrates, from which we hope to investigate these differences in more detail.

Normal faecal microflora, among which *P. aeruginosa* is found, have been shown to contribute to the acetylation of 5-AS [22]. Of the substrates measured in the present study, 5-AS was the most rapidly acetylated, suggesting that PANAT may contribute to the *in vivo* acetylation, and hence therapeutic inactivation, of 5-AS [39]. Recent studies with human NAT1 and NAT2 [40] have shown that PANAT is an order of magnitude more active in the acetylation of 5-AS than the human NAT enzymes. Hence PANAT in gut flora is likely to be an important contributor to the therapeutic inactivation of 5-AS. Inhibition of the N-acetylation of 5-AS by intestinal microflora using an exogenous PANAT-specific ligand may therefore have combinatorial therapeutic value [41]. Similarly, inhibition of PANAT-mediated acetylation of the anti-hypertensive drug hydralazine may help to reduce the occurrence of systemic lupus erythematosus, which this drug is known to cause [42].

The growing problem of *P. aeruginosa* multidrug resistance has been accounted for largely by its plethora of regulatory genes and efflux pumps [43–48]. Previous studies have shown that the *mexABoprM* multidrug efflux system is responsible for resistance of *P. aeruginosa* to the antibacterial agents trimethoprim and sulphamethoxazole [45]. Sulphamethoxazole and sulphisoxazole were poor substrates for PANAT, and no acetylation of trimethoprim or any of the other sulphonamides tested was observed.

pABA is required for the synthesis of folic acid in *P. aeruginosa*, as shown by hypersensitivity to sulphonamides and trimethoprim

Figure 5 Molecular surface and active-site architecture of PANAT and MSNAT

Molecular surfaces were calculated for both PANAT (A) and MSNAT (B) using the program Aesop [M. E. M. Noble, unpublished work; details available from M. E. M. N. on request (martin@biop.ox.ac.uk)]. Hydrophobicity was then calculated for each surface and underlying protein using the program GRID [50]. Surfaces are coloured on identical scales of increasing hydrophobicity from grey through green to yellow. Underlying secondary structural elements of PANAT (blue) and MSNAT (orange) and the active-site cysteine (ball and stick mode) are also shown.

when these agents are not removed by efflux, and so it was surprising to note that pABA is a substrate for PANAT. The literature apparent K_m for acetylation of pABA in *P. aeruginosa* lysate (1200 μM) is considerably higher than the reported K_m for the conversion of pABA into 7,8-dihydropteroate by dihydropteroate synthase from *Streptococcus faecalis* R (1.1 μM) [49]. Although the two enzymes are from different prokaryotes, we predict that the rate of removal of pABA by acetylation *in vivo* is insignificant compared with the rate of formation of 7,8-dihydropteroate in *P. aeruginosa*.

The data in the present paper present a picture of close structural similarity between bacterial NAT proteins. This is partially reflected in similar specific activity profiles, although we have identified a number of key differences that result from subtle structural differences. Current co-crystallization experiments are aimed at identifying the mode of substrate binding, from which we hope to understand further the nature of these differences. With its

ease of preparation and crystallization, PANAT is ideal for use in high-throughput screening assays for inhibitors of bacterial NATs, to complement data already obtained for the MSNAT enzyme. The PANAT structure is being used within our laboratory in an iterative, rational ligand-design process to identify both TBNAT and broad-range bacterial NAT antagonists for use in a range of clinical applications.

We recommend that PANAT is designated NAT105*1, in accordance with current NAT nomenclature, as agreed by the International NAT Nomenclature Committee (<http://www.louisville.edu/medschool/pharmacology/NAT.html>).

We are grateful to the Wellcome Trust for continued financial support. I. M. W. is in receipt of an MRC studentship. We thank Professor Stephen G. Davies (Chemistry Research Laboratory, University of Oxford) for his continued supervision of I. M. W., and Andrew W. Mulvaney, Richard J. Vickers and Angela J. Russell (Chemistry Research Laboratory, University of Oxford) for advice and helpful discussion. We also thank Professor A. M. Chakrabarty (University of Illinois College of Medicine, Chicago, IL, U.S.A.) for his kind donation of *Pseudomonas aeruginosa* genomic DNA, and Andrew Spiers (Plant Sciences, University of Oxford) for helpful discussion. MALDI-TOF analysis was carried out by Tony Willis (Department of Biochemistry, University of Oxford).

REFERENCES

- 1 Upton, A., Johnson, N., Sandy, J. and Sim, E. (2001) Arylamine N-acetyltransferases – of mice, men and microorganisms. *Trends Pharmacol. Sci.* **22**, 140–146
- 2 Evans, D. A. (1989) N-acetyltransferase. *Pharmacol. Ther.* **42**, 157–234
- 3 Delomnie, C., Fouix, S., Longuevaux, S., Brahimi, N., Bizet, C., Picard, B., Denamur, E. and Dupret, J. M. (2001) Identification and functional characterization of arylamine N-acetyltransferases in eubacteria: evidence for highly selective acetylation of 5-aminosalicylic acid. *J. Bacteriol.* **183**, 3417–3427
- 4 Weber, W. W. and Hein, D. W. (1985) N-acetylation pharmacogenetics. *Pharmacol. Rev.* **37**, 25–79
- 5 Riddle, B. and Jencks, W. P. (1971) Acetyl-coenzyme A:arylamine N-acetyltransferase. Role of the acetyl-enzyme intermediate and the effects of substituents on the rate. *J. Biol. Chem.* **246**, 3250–3258
- 6 Evans, D. A. P., Manley, K. A. and McKusick, V. A. (1960) Genetic control of isoniazid acetylation in man. *Br. Med. J.* **2**, 485–491
- 7 Watanabe, M., Sofuni, T. and Nohmi, T. (1992) Involvement of Cys69 residue in the catalytic mechanism of N-hydroxyarylamine O-acetyltransferase of *Salmonella typhimurium*. Sequence similarity at the amino acid level suggests a common catalytic mechanism of acetyltransferase for *S. typhimurium* and higher organisms. *J. Biol. Chem.* **267**, 8429–8436
- 8 Payton, M., Auty, R., Delgoda, R., Everett, M. and Sim, E. (1999) Cloning and characterization of arylamine N-acetyltransferase genes from *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: increased expression results in isoniazid resistance. *J. Bacteriol.* **181**, 1343–1347
- 9 Hsieh, S. E., Lo, H. H. and Chung, J. G. (1998) The characteristics of arylamine N-acetyltransferase in *Pseudomonas aeruginosa*. *Curr. Microbiol.* **36**, 353–360
- 10 Bass, Jr, J. B., Farer, L. S., Hopewell, P. C., O'Brien, R., Jacobs, R. F., Ruben, F., Snider, Jr, D. E. and Thornton, G. (1994) Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society and The Centers for Disease Control and Prevention. *Am. J. Respir. Crit. Care Med.* **149**, 1359–1374
- 11 Moore, M., Onorato, I. M., McCray, E. and Castro, K. G. (1997) Trends in drug-resistant tuberculosis in the United States, 1993–1996. *JAMA, J. Am. Med. Assoc.* **278**, 833–837
- 12 Sinclair, J. C., Delgoda, R., Noble, M. E., Jarman, S., Goh, N. K. and Sim, E. (1998) Purification, characterization, and crystallization of an N-hydroxyarylamine O-acetyltransferase from *Salmonella typhimurium*. *Protein Expression Purif.* **12**, 371–380
- 13 Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E. and Noble, M. E. (2000) Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat. Struct. Biol.* **7**, 560–564
- 14 Sandy, J., Mushtaq, A., Kawamura, A., Sinclair, J., Sim, E. and Noble, M. (2002) The structure of arylamine N-acetyltransferase from *Mycobacterium smegmatis* – an enzyme which inactivates the anti-tubercular drug, isoniazid. *J. Mol. Biol.* **318**, 1071–1083
- 15 Mushtaq, A., Payton, M. and Sim, E. (2002) The COOH terminus of arylamine N-acetyltransferase from *Salmonella typhimurium* controls enzymic activity. *J. Biol. Chem.* **277**, 12175–12181
- 16 Payton, M., Gifford, C., Schartau, P., Hagemeyer, C., Mushtaq, A., Lucas, S., Pinter, K. and Sim, E. (2001) Evidence towards the role of arylamine N-acetyltransferase in *Mycobacterium smegmatis* and development of a specific antiserum against the homologous enzyme of *Mycobacterium tuberculosis*. *Microbiology* **147**, 3295–3302

- 17 Brooke, E. W., Davies, S. G., Mulvaney, A. W., Pompeo, F., Sim, E. and Vickers, R. J. (2003) An approach to identifying novel substrates of bacterial arylamine N-acetyltransferases. *Bioorg. Med. Chem.* **11**, 1227–1234
- 18 Upton, A. M., Mushtaq, A., Victor, T. C., Sampson, S. L., Sandy, J., Smith, D. M., van Helden, P. V. and Sim, E. (2001) Arylamine N-acetyltransferase of *Mycobacterium tuberculosis* is a polymorphic enzyme and a site of isoniazid metabolism. *Mol. Microbiol.* **42**, 309–317
- 19 Bhakta, S., Besra, G. S., Upton, A. M., Parish, T., Sholto-Douglas-Vernon, C., Gibson, K. J. C., Knutton, S., Gordon, S., daSilva, R. P., Anderton, M. C. and Sim, E. (2004) Arylamine N-acetyltransferase is required for synthesis of mycolic acids and complex lipids in *Mycobacterium bovis* BCG and represents a novel drug target. *J. Exp. Med.* **199**, 1191–1199
- 20 Payton, M., Mushtaq, A., Yu, T. W., Wu, L. J., Sinclair, J. and Sim, E. (2001) Eubacterial arylamine N-acetyltransferases – identification and comparison of 18 members of the protein family with conserved active site cysteine, histidine and aspartate residues. *Microbiology* **147**, 1137–1147
- 21 Gisbert, J. P., Gomollon, F., Mate, J. and Pajares, J. M. (2002) Role of 5-aminosalicylic acid (5-ASA) in treatment of inflammatory bowel disease: a systematic review. *Dig. Dis. Sci.* **47**, 471–488
- 22 van Hogezaand, R. A., Kennis, H. M., van Schaik, A., Koopman, J. P., van Hees, P. A. and van Tongeren, J. H. (1992) Bacterial acetylation of 5-aminosalicylic acid in faecal suspensions cultured under aerobic and anaerobic conditions. *Eur. J. Clin. Pharmacol.* **43**, 189–192
- 23 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 24 Cornish-Bowden, A. (1995) *Fundamentals of Enzyme Kinetics*, Portland Press, London
- 25 Coroneos, E., Gordon, J. W., Kelly, S. L., Wang, P. D. and Sim, E. (1991) Drug metabolizing N-acetyltransferase activity in human cell lines. *Biochim. Biophys. Acta* **1073**, 593–599
- 26 McPherson, A. (1985) Crystallization of macromolecules: general principles. *Methods Enzymol.* **114**, 112–120
- 27 Collaborative Computational Project (1994) The CCP4 Suite: programs for protein crystallography (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
- 28 Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warren, P., Hickey, M. J., Brinkman, F. S., Huftnagle, W. O., Kowalik, D. J., Lagrou, M. et al. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature (London)* **406**, 959–964
- 29 Weber, W. W. and Vatsis, K. P. (1993) Individual variability in p-aminobenzoic acid N-acetylation by human N-acetyltransferase (NAT1) of peripheral blood. *Pharmacogenetics* **3**, 209–212
- 30 Minchin, R. F. (1995) Acetylation of p-aminobenzoylglutamate, a folic acid catabolite, by recombinant human arylamine N-acetyltransferase and U937 cells. *Biochem. J.* **307**, 1–3
- 31 Delgoda, R., Lian, L. Y., Sandy, J. and Sim, E. (2003) NMR investigation of the catalytic mechanism of arylamine N-acetyltransferase from *Salmonella typhimurium*. *Biochim. Biophys. Acta* **1620**, 8–14
- 32 Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119
- 33 Brunger, T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Panu, N. S. et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
- 34 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291
- 35 Pompeo, F., Mushtaq, A. and Sim, E. (2002) Expression and purification of the rifamycin amide synthase, Riff, an enzyme homologous to the prokaryotic arylamine N-acetyltransferases. *Protein Expression Purif.* **24**, 138–151
- 36 Kawamura, A., Sandy, J., Upton, A., Noble, M. and Sim, E. (2003) Structural investigation of mutant *Mycobacterium smegmatis* arylamine N-acetyltransferase: a model for a naturally occurring functional polymorphism in *Mycobacterium tuberculosis* arylamine N-acetyltransferase. *Protein Expression Purif.* **27**, 75–84
- 37 Butcher, N. J., Boukouvla, S., Sim, E. and Minchin, R. F. (2002) Pharmacogenetics of the arylamine N-acetyltransferases. *Pharmacogenomics J.* **2**, 30–42
- 38 Brooke, E. W., Davies, S. G., Mulvaney, A. W., Okada, M., Pompeo, F., Sim, E., Vickers, R. J. and Westwood, I. M. (2003) Synthesis and *in vitro* evaluation of novel small molecule inhibitors of bacterial arylamine N-acetyltransferases (NATs). *Bioorg. Med. Chem. Lett.* **13**, 2527–2530
- 39 Ireland, A., Priddle, J. and Jewell, D. (1992) Comparison of 5-aminosalicylic acid and N-acetylaminosalicylic acid uptake by the isolated human colonic epithelial cell. *Gut* **33**, 1343–1347
- 40 Kawamura, A. and Sim, E. (2004) Eukaryotic arylamine N-acetyltransferase: Investigation of substrate specificity by high-throughput screening. *Biochem. Pharmacol.*, in the press
- 41 Buchman, A. L. (2001) Side effects of corticosteroid therapy. *J. Clin. Gastroenterol.* **33**, 289–294
- 42 Sim, E., Gill, E. W. and Sim, R. B. (1984) Drugs that induce systemic lupus erythematosus inhibit complement component C4. *Lancet* **ii**, 422–424
- 43 Poole, K. (2004) Efflux-mediated multiresistance in Gram-negative bacteria. *Clin. Microbiol. Infect.* **10**, 12–26
- 44 Pirnay, J. P., De Vos, D., Cochez, C., Bilocq, F., Pirson, J., Struelens, M., Duinslaeger, L., Cornelis, P., Zizi, M. and Vanderkelen, A. (2003) Molecular epidemiology of *Pseudomonas aeruginosa* colonization in a burn unit: persistence of a multidrug-resistant clone and a silver sulfadiazine-resistant clone. *J. Clin. Microbiol.* **41**, 1192–1202
- 45 Khler, T., Kok, M., Michea-Hamzehpour, M., Plesiat, P., Gotoh, N., Nishino, T., Curty, L. K. and Pechere, J. C. (1996) Multidrug efflux in intrinsic resistance to trimethoprim and sulfamethoxazole in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **40**, 2288–2290
- 46 Carsenti-Etesses, H., Cavallo, J. D., Roger, P. M., Ziha-Zarifi, I., Plesiat, P., Garrabe, E. and Dellamonica, P. (2001) Effect of beta-lactam antibiotics on the *in vitro* development of resistance in *Pseudomonas aeruginosa*. *Clin. Microbiol. Infect.* **7**, 144–151
- 47 Greenberg, E. P. (2000) Bacterial genomics. Pump up the versatility. *Nature (London)* **406**, 947–948
- 48 Whiteley, M., Bangera, M. G., Bumgarner, R. E., Parsek, M. R., Teitzel, G. M., Lory, S. and Greenberg, E. P. (2001) Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature (London)* **413**, 860–864
- 49 Iwai, K. and Okinaka, O. (1980) Radioassay for dihydropteroate-synthesizing enzyme activity. *Methods Enzymol.* **66**, 560–564
- 50 Goodford, P. J. (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **28**, 849–857
- 51 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876–4882
- 52 Gouet, P., Courcelle, E., Stuart, D. I. and Metz, F. (1999) ESPript: multiple sequence alignment in PostScript. *Bioinformatics* **15**, 305–308

Received 4 August 2004/20 September 2004; accepted 24 September 2004

Published as BJ Immediate Publication 24 September 2004, DOI 10.1042/BJ20041330