Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry 11 (2003) 1227–1234

BIOORGANIC & MEDICINAL CHEMISTRY

An Approach to Identifying Novel Substrates of Bacterial Arylamine *N*-Acetyltransferases

Edward W. Brooke,[a] Stephen G. Davies,[b,*] Andrew W. Mulvaney,[b] Frédérique Pompeo,[a] Edith Sim[a] and Richard J. Vickers[b]

[a]*Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK*
[b]*The Dyson Perrins Laboratory, University of Oxford, South Parks Rd, Oxford OX1 3QY, UK*

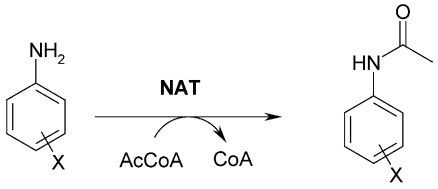
Received 1 October 2002; accepted 3 December 2002

Abstract—Arylamine *N*-acetyltransferases (NATs) catalyse the acetylation of arylamine, arylhydrazine and arylhydroxylamine substrates by acetyl Coenzyme A. NAT has been discovered in a wide range of eukaryotic and prokaryotic species. Although prokaryotic NATs have been implicated in xenobiotic metabolism, to date no endogenous role has been identified for the arylamine *N*-acetyl transfer reaction in prokaryotes. Investigating the substrate specificity of these enzymes is one approach to determining a possible endogenous role for prokaryotic NATs. We describe an accurate and efficient assay for NAT activity that is suitable for high-throughput screening of potential NAT ligands. This assay has been utilised to identify novel substrates for pure NAT from *Salmonella typhimurium* and *Mycobacterium smegmatis* which show a relationship between the lipophilicity of the arylamine and its activity as a substrate. The lipophilic structure/activity relationship observed is proposed to depend on the topology of the active site using docking studies of the crystal structures of these NAT isoenzymes. The evidence suggests an endogenous role of NAT in the protection of bacteria from aromatic and lipophilic toxins.

Introduction

Arylamine *N*-acetyltransferases (NATs) are polymorphic xenobiotic metabolizing enzymes first identified in humans as responsible for the inactivation of the front-line anti-tubercular drug isoniazid (INH).[1] NATs catalyze the transfer of an acetyl group from acetyl coenzyme A (AcCoA) to arylamine, hydrazine[2] and arylhydroxylamine substrates[3] (Scheme 1) including activated carcinogens. Interindividual variability in rates of drug metabolism[2,4,5] and susceptibility to arylamine carcinogens[6,7] have been linked to polymorphisms in the two active human isoenzymes NAT1 and NAT2.[8]

NAT has also been identified in the bacterium *Salmonella typhimurium* (STNAT), as a factor influencing carcinogen susceptibility in mutagenicity testing.[9] As bacterial genomes have become available, NAT has been discovered in more than 20 prokaryotes including *Mycobacterium tuberculosis*.[10] A NAT homologue, rifamycin amide synthetase,[11] has been found in *Amycolatopsis mediterranei* where it is involved in biosynthesis of the natural antibiotic rifamycin.[12] The *nat* gene in *M. tuberculosis* appears in a putative operon that encodes for enzymes capable of metabolizing aromatic compounds and suggests a role for mycobacterial NAT in xenobiotic metabolism.[10] The mycobacterial NATs acetylate and hence inactivate isoniazid which could have important implications for isoniazid resistance in *M. tuberculosis*.[13] Overexpression of NAT from *M. tuberculosis* in *M. smegmatis* has resulted in an increased resistance of the recombinant mycobacterium to isoniazid while NAT knockout strains have been shown to exhibit increased sensitivity to INH.[14]



Scheme 1.

*Corresponding author. Tel.: +44-1865-275680; fax: +44-1865-275674; e-mail: steve.davies@chem.ox.ac.uk

0968-0896/03/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0968-0896(02)00642-9

The three-dimensional crystal structure of NAT from *S. typhimurium* (STNAT) has been elucidated.[15] The catalytic mechanism has been previously determined as a 'Ping-Pong' bi-bi mechanism[16] and a cysteine residue was shown to be vital for catalysis.[17] The crystal structure reveals that the protein consists of three domains such that a cysteine, a histidine and an aspartate residue (Cys69, His107, Asp122) are juxtaposed to form a catalytic triad providing the potential mechanism to activate the cysteine thiol. A large cleft is believed to bind AcCoA and substrate molecules.[18] The crystal structure of *M. smegmatis* NAT has been determined to a resolution of 1.7 Å and shows an almost identical three-dimensional structure to STNAT.[19]

One approach to identify an endogenous substrate is to test the acetylation of a wide range of substrates in a high-throughput format to probe the substrate specificity of the protein. In order for this investigation to be undertaken, an efficient and accurate method of identifying NAT substrates was required. We demonstrate that NAT activity can be determined by spectrophotometric determination of free Coenzyme A (CoA) using 5,5′-dithio-bis(2-nitrobenzoic acid) (Ellman's reagent, DTNB).[20] We have used this method to elucidate the substrate specificity of two different recombinant prokaryotic NAT proteins and to identify novel small molecule ligands for bacterial NAT. The trends observed in the substrate specificity have been rationalised to the topology of the area immediately surrounding the active site using simulated annealing to dock the substrates into the available crystal structures.

Results and Discussion

Purification and stability of recombinant MSNAT

NATs from *M. smegmatis* (MSNAT) and *S. typhimurium* (STNAT) were produced in *E. coli* as recombinant proteins with an N-terminus hexahistidine tag and purified using Nickel-NTA affinity resin. The assay of MSNAT samples stored at +4, −20, −70 °C and frozen in liquid nitrogen over 18 days resulted in no less than 80% of the initial activity of the protein as determined by acetylation of anisidine (data not shown). The activity of freshly prepared MSNAT can be determined, as detected by acetylation of 100 μM anisidine, both with (161 ± 6 nmol min^{-1} mg^{-1}) and without (149 ± 5 nmol min^{-1} mg^{-1}) addition of 1 mM DTT to the assay buffer (20 mM Tris–HCl, pH 8.0).

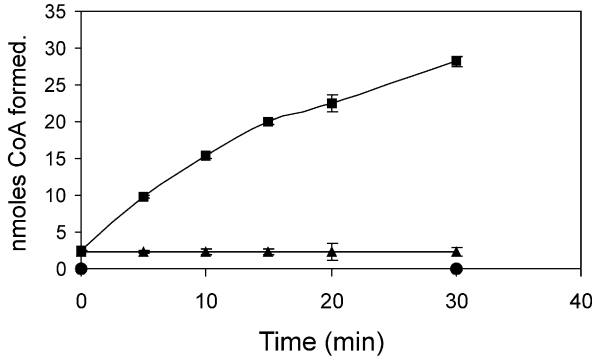


Figure 1. The use of DTNB to determine the rate of hydrolysis of acetyl Coenzyme A in the presence of isoniazid and MSNAT. Following incubation of MSNAT (1.2 μg per well), isoniazid (500 μM) and AcCoA (400 μM) for varying times (x-axis), the reaction was quenched and the quantity of CoA was determined with DTNB (squares). Results for identical experiments except without isoniazid (triangles) or without acetyl CoA (circles) are shown. Results shown are mean values ($n = 2$) with individual values shown as error bars.

Measurement of acetyl CoA hydrolysis

The NAT catalysed acetylation of arylamines results in hydrolysis of AcCoA to give free CoA. DTNB reacts with free thiol groups in solution to produce thionitrobenzoate (TNB) which has a maximum absorbance at $\lambda_{max} = 412$ nm.[20] DTNB can therefore be used to detect AcCoA hydrolysis by NAT as has been previously demonstrated with the carnitine acetyltransferases.[21] Preliminary experiments were carried out to determine whether this method could be used to detect NAT activity in a microtitre plate format. MSNAT, isoniazid (INH) and AcCoA were incubated for varying time periods in a 96-well plate and then DTNB was added. TNB was detected spectrophotometrically on a plate reader and the amount formed was found to be dependent on the time of incubation of the NAT/substrate mixture (Fig. 1). No hydrolysis of AcCoA was observed in the absence of INH or NAT.

In order to relate the amount of AcCoA consumed with the amount of acetylated substrate produced, *p*-anisidine was incubated with MSNAT and AcCoA, followed by spectrophotometric detection of CoA using DTNB with the simultaneous determination of consumption of *p*-anisidine (Fig. 2). The two products, CoA and *N*-acetylanisidine were formed in equal quantities over the time course of the reaction.

To determine whether this assay method had sufficient accuracy, the Z′ factor was calculated.[22] Standard assay conditions were set as 500 μM substrate, 400 μM AcCoA, sufficient NAT to maintain activity in the linear range and a 15 min timepoint. Using *p*-anisidine as a

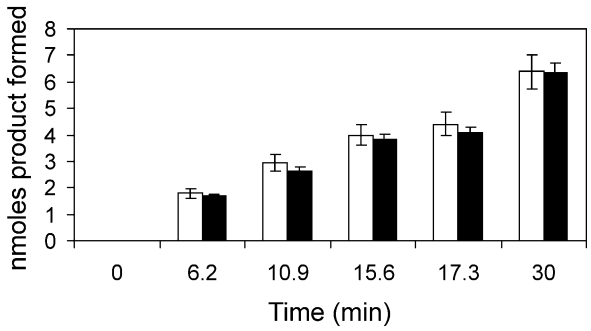


Figure 2. Comparison of the amounts of arylamine acetylated and acetyl coenzyme A hydrolysed by MSNAT. The rate of acetylation of anisidine (500 μM) in the presence of AcCoA (400 μM) and MSNAT over varying time periods (x-axis) was determined by simultaneously measuring loss of arylamine (filled bars) and detection of CoA (open bars).

positive control and $n=4$, the standard deviations of positive and negative controls were 0.0096 and 0.0065 AU for a change in absorbance of 0.198 AU. This gives a Z' factor of 0.756 which fits into the range for an acceptable assay ($0.5<Z'<1.0$). All substrates with specific activities greater than 200 nmol min^{-1} mg^{-1} under these conditions gave Z' factors well above 0.5. Poor substrates can be defined as having rates below 200 nmol min^{-1} mg^{-1} but the values in this region are inaccurate and cannot be used for further analysis.

These data demonstrate that the determination of NAT activity using DTNB gives an accurate and reproducible assay of the rate of AcCoA hydrolysis and is robust to most chemical functionalities, making it suitable for application to a high-throughput screening format. One potential drawback of this assay is that it cannot be used in the presence of dithiothreitol (DTT) or other thiol-based reducing agents. Although STNAT and MSNAT have an essential cysteine thiol, we have shown that the recombinant proteins can be assayed effectively without addition of a reducing agent. The NAT in the assay is present at sub-micromolar levels whereas the concentration of CoA formed is up to 100 μM. Hence, the contribution of the protein cysteine residues to the colour development is negligible (Fig. 1).

Previous assays for NAT activity have relied on determining the amount of acceptor substrate acetylated over time. One such method, the detection of arylamine with DMAB, relies on the nature of the acceptor substrate[23] although it has been adapted to 96-well format.[24] HPLC assay methods although very sensitive can be time-consuming and are also substrate specific.[25] The method described herein would also be potentially useful in the determination of rates of *O*-acetyl transfer from AcCoA to *N*-hydroxyarylamines where existing assays for this reaction are indirect, involving the quantification of DNA adducts formed after the reaction.[3]

Substrate screening

Twenty-one compounds were assayed as acceptor substrates of either MSNAT or STNAT using DTNB to detect hydrolysis of AcCoA (Table 1, Fig. 3). The compounds were selected to include both previously identified substrates of STNAT, human NAT1 and NAT2 and compounds that showed some structural relationship to known substrates (Scheme 2). Compounds were divided into four groups comprising arylamine drugs, other arylamines, alkoxyanilines and hydrazines. The rates of hydrolysis of AcCoA in the presence of these compounds and STNAT or MSNAT ranged from 10 to 20,000 nmol min^{-1} (mg protein)$^{-1}$. There was no observed production of CoA with any of the substrates in the absence of enzyme over the time periods used.

Hydralazine caused the most rapid rate of AcCoA hydrolysis with STNAT, and 4-hexyloxyaniline was the best substrate for MSNAT. 5-aminosalicylate and 2-aminofluorene were also effective substrates with these proteins in agreement with previous studies.[26] Previously identified STNAT substrates such as 4-anisidine

Table 1. Substrates of MSNAT

Compound (abbreviation)	Rate[a]	*c*log *P*[b]	Pop[n] of α site[c]	E_α[d]	E_β[e]
Arylamine drugs					
Sulfamethazine (SMZ)	7±7	1.62	nd	nd	nd
Procainamide (PRO)	54±46	1.05	nd	nd	nd
5-Aminosalicylate (5AS)	5670±110	0.02	nd	nd	nd
Other arylamines					
4-Aminobenzoic acid (pABA)	30±30	0.41	1	−6.69	−8.54±0.32
4-Aminopyridine (APY)	12±8	−0.06	nd	nd	nd
4-Trifluoromethylaniline (TFMA)	121±36	2.03	nd	nd	nd
4-Aminophenol (AMP)	122±24	0.66	0	—	−7.89±0.01
4-Chloroaniline (CLA)	430±22	1.67	1	−5.58	−7.56±0.28
4-Bromoaniline (BRA)	548±46	1.94	1	−5.27	−7.55±0.28
4-Iodoaniline (IOA)	956±61	2.35	nd	nd	nd
4-Phenoxyaniline (POA)	1620±46	3.13	10	−7.29±0.07	—
2-Aminofluorene* (2AF)	3610±80	2.51	10	−7.87±0.06	—
Alkoxyanilines					
4-Anisidine (ANS)	1970±100	1.18	4	−5.58±0.17	−7.42±0.31
4-Aminoveratrole (AMV)	3000±110	1.32	8	−6.09±0.36	−7.03±1.13
4-Ethoxyaniline (EOA)	3560±150	1.53	8	−5.95±0.09	−6.60±0.40
4-Butoxyaniline* (BOA)	5670±160	2.44	10	−6.96±0.18	—
4-Hexyloxyaniline* (HOA)	8510±110	3.35	10	−7.81±0.41	—
Hydrazines					
4-Methoxyphenylhydrazine (MPZ)	395±50	1.56	9	−5.99±0.23	−6.50
Isoniazid (INH)	1020±42	−0.64	6	−6.25±0.36	−7.69±0.13
4-Chlorobenzoic hydrazide (CBZ)	3380±70	1.25	10	−6.37±0.23	—
Hydralazine (HDZ)	7060±130	0.73	10	−7.58±0.42	—

[a]Rate of AcCoA hydrolysis in the presence of MSNAT, AcCoA (400 μM) and test compound (500 μM) (nmol min^{-1} (mg protein)$^{-1}$ mean shown±standard deviation, $n=4$). * In the presence of 5% DMSO, others 0.5% DMSO.
[b]Calculated partition coefficient between water and 1-octanol determined using ChemDraw Ultra v6.0 (CambridgeSoft).
[c]The number of docking solutions (out of 10) found in the α site. nd: Not determined.
[d]Average 'total binding energy' (kcal mol^{-1}, mean ± standard deviation) of solutions found in the α site. —: Not applicable.
[e]Average 'total binding energy' (kcal mol^{-1}, mean ± standard deviation) of solutions found in the β site. —: Not applicable.

(ANS), 4-aminoveratrole (AMV) and 4-iodoaniline (IOA) are also substrates for MSNAT.[27] The human NAT1 substrate 4-aminobenzoic acid (PABA) and the human NAT2 substrate sulfamethazine (SMZ) are both poor substrates for MSNAT and STNAT. Novel NAT substrates identified by this method include the alkoxyanilines (4-ethoxy-, 4-butoxy and 4-hexyloxyaniline) and the isoniazid analogue 4-chlorobenzoic acid hydrazide (Fig. 3). The substrate specificity appears to be virtually identical for both STNAT and MSNAT with substrates producing similar rates of reaction with either enzyme.

Kinetic analysis of MSNAT was performed with eight of the substrates chosen from each substrate class. Eadie-Hofstee plots showed that the reactions follow apparent Michaelis–Menten kinetics. Apparent kinetic constants (Table 2) were determined by non-linear optimisation of the curve to the data as shown for hydralazine and iodoaniline (Fig. 4). The $k_{cat,\ app}$ values

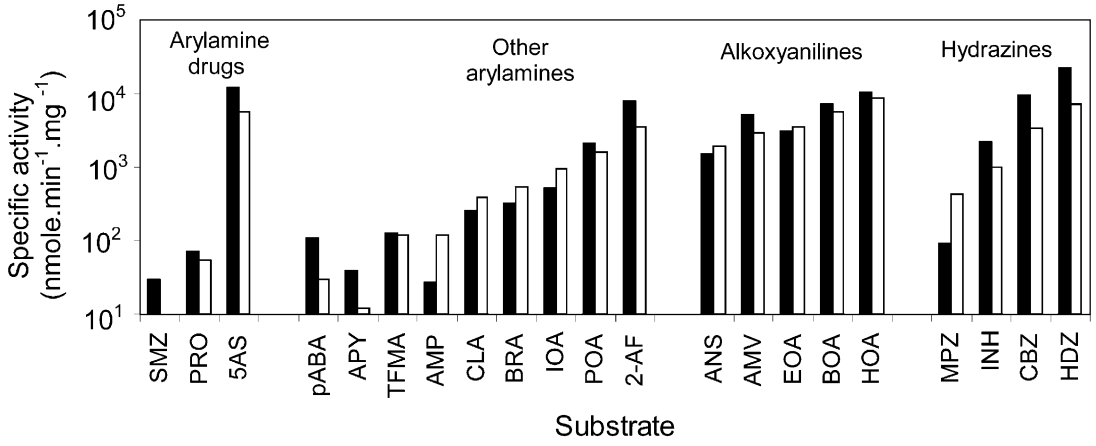
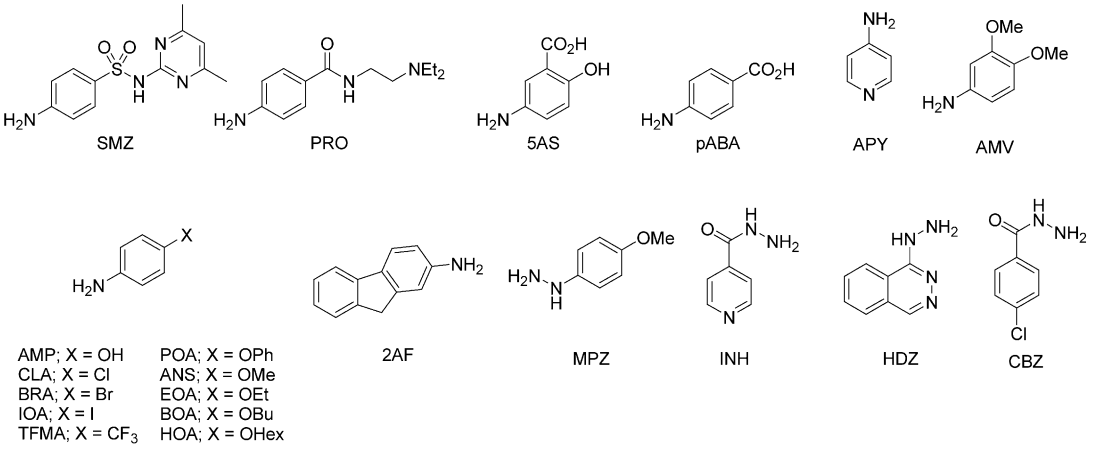


Figure 3. Substrate specificity of NATs from *Salmonella typhimurium* (STNAT) and *Mycobacterium smegmatis* (MSNAT). The activity of STNAT (dark shading) and MSNAT (light shading) is shown, determined by measuring the hydrolysis of acetyl CoA with DTNB with substrates shown in Table 1 (each at 500 μM) in assay buffer (20 mM Tris–HCl, pH 8.0) containing either 0.5% DMSO or 5% DMSO (2AF, BOA and HOA).



Scheme 2.

Table 2. Kinetic analysis of substrate acetylation

Substrate	$K_{m,\ app}$ (μM)	$V_{max,\ app}$ (μmol min^{-1} mg^{-1})	k_{cat} ($\times 10^3$ s^{-1})	k_{cat}/K_m (s^{-1} M^{-1})
Hydralazine	80±20	12±1.0	63	790
5-Aminosalicylate	250±40	8.6±0.6	44	180
4-Anisidine*	1800±150	5.7±0.6	30	17
4-Ethoxyaniline*	1010±90	6.9±0.9	36	36
4-Butoxyaniline*	330±50	7.5±0.2	39	120
4-Chloroaniline†	2000±100	1.3±0.1	6.9	3.5
4-Bromoaniline†	1800±200	1.7±0.2	8.6	4.8
4-Iodoaniline†	1000±100	1.9±0.1	9.8	9.7

The rate of acetyl CoA hydrolysis was determined as in Table 1 with varying concentrations of the substrates shown. Michaelis–Menten kinetics were confirmed using the Eadie–Hofstee plot (data not shown) and kinetic constants were determined by non-linear optimisation (Fig. 4). Mean±difference of two independent measurements are shown. The $K_{m,\ app}$ value for hydralazine is likely to be lower than that showed due to almost complete acetylation of substrate at the lower concentrations. Substrates were tested in the presence of 0%, 2.5%† or 5%* DMSO.

show a 10-fold variation across the series of substrates with turnover numbers for hydralazine of $0.063\,s^{-1}$ and for 4-chloroaniline of $0.0069\,s^{-1}$. The specificity constant, k_{cat}/K_m shows a 200-fold variation with hydralazine and 4-chloroaniline again representing the extremes.

Apparent differences were observed in the effect of the substituent on the aromatic ring of the substrates on the rate of the reaction. The data suggests that the electrostatic potential of the aromatic ring has little effect on the rate of acetylation as the electron-donating 4-aminophenol and the electron-withdrawing 4-trifluoromethylaniline both gave similar values. The series of alkoxyanilines suggested that the lipophilicity of the substrate may be a contributory factor to the rate of acetylation observed and so calculated partition coefficients (*c*log*P*) between water and 1-octanol of the substrate molecules were determined using ChemDraw v6.0 (CambridgeSoft) (Table 1). Double logarithmic plots of clogP against the MSNAT activity revealed a linear dependence in the rate of reaction with increasing lipophilicity for both the alkoxyanilines and also the other arylamines (Fig. 5).

Simulated annealing of substrates

The compounds tested were annealed to the MSNAT crystal structure using the Autodock suite of programs.[18] All solutions predict substrate binding into the active site with two specific loci: one immediately adjacent to the active cysteine residue (termed the α site)[18] and one at approximately 11 Å from the α site (termed the β site) (Fig. 6). Average final docked energy values for solutions in the β site across all the compounds were −7.68 kcal/mol and those for the α site were −6.65 kcal/mol. Where compounds bound at both sites, the calculated binding energy at the β site was in all cases more exothermic than at the α site. Six compounds gave solutions only at the α site with an average binding energy of −7.31 kcal/mol, including four of the five most effective substrates. 4-aminophenol was predicted

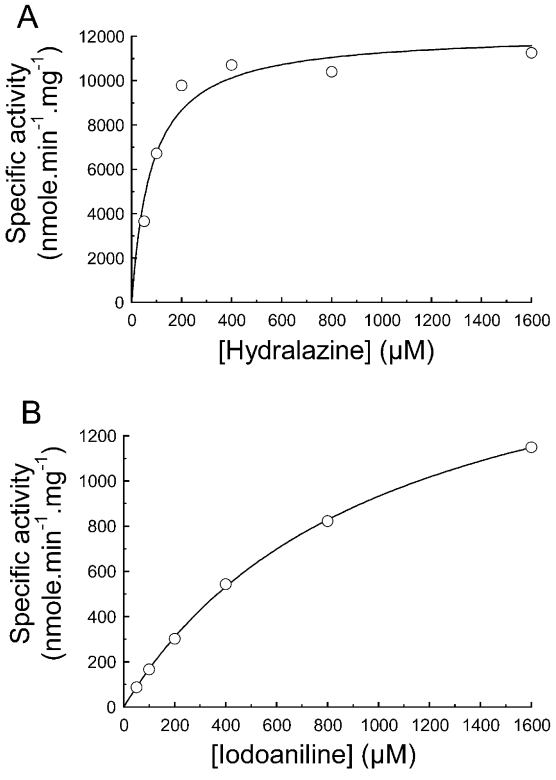


Figure 4. Kinetics of hydrolysis of acetyl CoA in the presence of hydralazine or iodoaniline. The specific activity of MSNAT is plotted against the substrate concentration for (A) hydralazine and (B) iodoaniline. 'Best-fit' Michaelis–Menten kinetics were plotted using non-linear optimization (KyPlot v6.0).

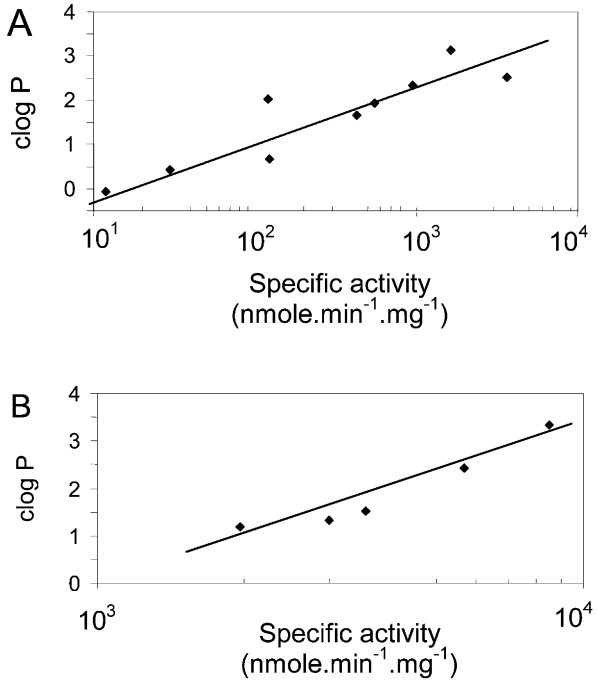


Figure 5. Effect of substrate lipophilicity on MSNAT activity. The MSNAT specific activity is plotted on a log scale against the substrate calculated partition coefficient between 1-octanol and water for the series of (A) arylamines and (B) alkoxyanilines as in Table 1. Logarithmic lines of regression are shown.

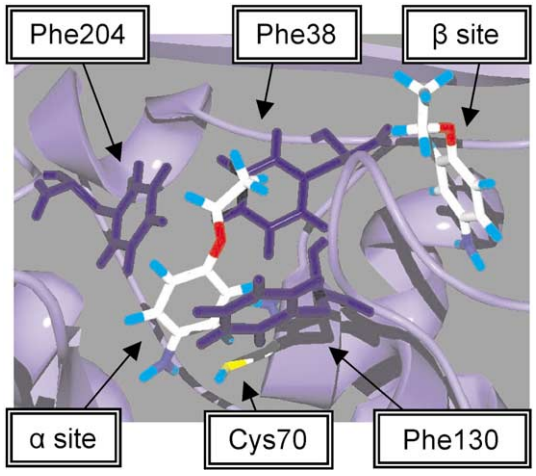


Figure 6. In silico docking of 4-ethoxyaniline to MSNAT. The 3-D structure of MSNAT at 1.7 Å[19] is depicted with two of the docking solutions for 4-ethoxyaniline showing the α and β sites of the enzyme. The EOA structures and the Cys70 sidechain are shown in standard CPK; Phe38, Phe130 and Phe204 are shown in blue.

to bind only at the β site and is a poor substrate for MSNAT. For the eight compounds that gave binding solutions at both sites, a general correlation was observed between the number of solutions in the alpha site and the specific activity. For alkoxyanilines, the binding energy at the α site was predicted to increase and the binding energy at the β site decrease with increasing specific activity.

Rationalising the structure–activity relationship of prokaryotic NAT substrates

The contact residues between the substrate and the protein in the active site can be used to rationalise the substrate specificity observed. The 3-D structures of MSNAT and STNAT are very similar, with a catalytic triad of Cys-His-Asp situated at the base of a cleft in the protein.[19] The sequence identity over the full length of STNAT and MSNAT is 34% and sequence identity between the first two domains of the two proteins is 52%.[10] The active site cysteine in both these proteins is in a cleft that is approximately 11 Å deep and 12 Å wide. In MSNAT, the sidechains of residues Phe38, Phe130 and Phe204 are each within 7 Å of the cysteine thiol group and appear to form a hydrophobic lid over the cysteine. In STNAT, Phe38, Phe125 and Phe199 occupy identical positions. Investigation of the docked substrates showed these residues to be in clear contact with the larger hydrophobic substrates (shown for 4-hexyloxyaniline and 2-aminofluorene, Fig. 7).

The lipophilic nature of these residues provides a good explanation for the increased activity observed with the more lipophilic substrates and substrates containing large planar ring systems such as 2-aminofluorene. The observed structure/activity relationship may be rationalised by lipophilic-lipophilic and pi-stacking interactions between the phenylalanine residues in the enzyme and the acceptor substrate. Accordingly, the specificity constants (k_{cat}/K_m) of the haloanilines and alkoxyanilines increase with the lipophilicity of the substrate yet the catalytic constants (k_{cat}) vary to a lesser extent. The chemical rate of reaction in the absence of enzyme catalysis would be expected to increase with the nucleophilicity of the arylamine nitrogen,[27] which can be affected by varying the substituents on the aromatic ring (OR > I > Br > Cl) and this explains the variation observed in k_{cat}. However, this influence is essentially constant for the series of alkoxyanilines thus protein-substrate interactions, shown by k_{cat}/K_m, must be involved in affecting the varying rates observed in this series. In contrast, hydralazine and 5-aminosalicylate have high specificity constants and low lipophilicity possibly indicating that they bind in an alternative fashion to the other arylamines. However, these observations do not take into account any conformational change that may occur on binding of the substrate into the active site and a crystal structure containing a substrate molecule would provide valuable insight on the mechanism of the reaction.

The phenylalanine residues that give rise to the observed substrate specificity are also highly conserved throughout the NAT family. Phe38 is located in the conserved PFENL region of NATs, Phe199 is conserved in all NATs with acetylating activity[10] and Phe125 is found in human NAT1 and highly conserved in prokaryotic NATs. In chimaeric enzyme studies of human NAT1 and NAT2, Phe125 has been implicated in causing the different acceptor substrate specificities between the two proteins[28] and the presence of Ser125 in human NAT2 may allow the acetylation of the more bulky and NAT2 specific substrates such as sulfamethazine and procainamide. Truncation mutants of STNAT with deletions in the third domain hydrolysed AcCoA in the absence of arylamine substrate and one of these mutants was lacking Phe199.[18] The phenylalanines may thus also serve to exclude water from the active site and prevent hydrolysis of the acetyl-cysteine intermediate by general base catalysis.

The elucidation of the endogenous role of NAT in bacteria will greatly assist in the understanding of this protein. Human NAT1 has been suggested to play a role in

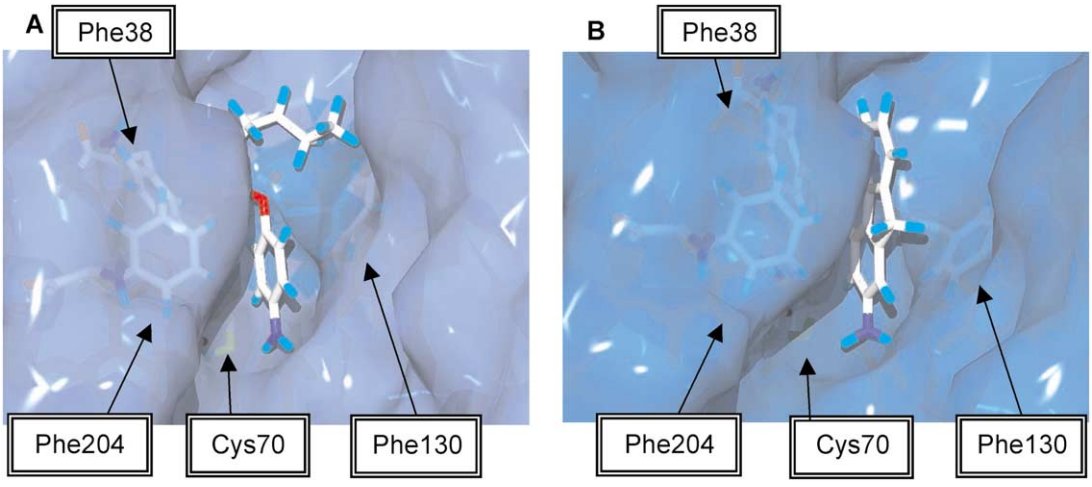


Figure 7. In silico docking of 4-hexyloxyaniline and 2-aminofluorene to MSNAT. The 3-D structure of MSNAT is shown with the lowest energy docking solution for (A) 4-hexyloxyaniline, and (B) 2-aminofluorene. The substrate structures and the enzyme sidechains are shown in standard CPK with the molecular surface in blue.

folate catabolism.[29,30] However, an endogenous substrate has not been identified for the bacterial enzymes. Analysis of the putative *nat* operon in *M. tuberculosis* identified several other predicted proteins that are transcribed with the *nat* gene and it has been suggested that these proteins are involved in aromatic and biphenyl metabolism.[10] The high rate of acetylation of lipophilic aromatic arylamines such as 2-aminofluorene by MSNAT provides further evidence for a role for NAT in the protection of these mycobacteria from environmental toxins of this type, which are exemplified by the cigarette smoke carcinogens 4-aminobiphenyl and β-naphthylamine. Interestingly, particularly broad-ranging NAT activity has been previously determined in the organism *Pseudomonas aeruginosa*[26] that exhibits considerable biodegradation capabilities for aromatic compounds. Further investigation of the NAT activity of this organism may provide increasing evidence as to whether this protective activity is a major role for NAT in bacteria.

Intriguingly, the series of alkoxyanilines that have been identified as novel NAT substrates have previously been shown to have anti-mycobacterial properties.[31] In view of the shared anti-mycobacterial properties and NAT activity of both isoniazid and these compounds, this may well represent a possible approach to novel therapies for tuberculosis.

Conclusions

DTNB (Ellman's reagent) can be used to assay NAT activity by detecting the amount of free CoA produced during the acetylation reaction. The assay has been modified to a high-throughput format and used to determine the substrate specificity of two recombinant bacterial NATs. The rate of acetylation and the specificity constants of simple arylamines increase with the lipophilicity of the compound. Docking studies showed that predicted binding into the α site correlates with a high rate of acetylation. The structure–activity relationship has been related to the location of three highly conserved phenylalanine residues that flank the active site cysteine. This suggests a role for bacterial NAT in protection of bacteria from toxic lipophilic arylamines.

Experimental

Materials

All chemicals were purchased from Sigma and used as received unless otherwise stated.

Expression of NATs

S. typhimurium NAT (STNAT) protein with an N-terminal His-tag was expressed in *E. coli* as described previously.[27] *M. smegmatis* NAT (MSNAT) with an N-terminal His-tag was cloned and expressed in *E. coli* BL21(DE3)pLysS.[14] Liquid cultures were grown as described previously.[19]

Purification of NAT proteins

Frozen cell pellets were thawed at 37 °C immediately before use and then sonicated (Soniprep 150) on wet ice using twenty bursts of 45 s on, 30 s off at 10 microns. Lysed cells were centrifuged (12,000*g*, 20 min, 4 °C) to pellet cell debris. Soluble lysate containing either MSNAT or STNAT was loaded onto 4.0 mL washed Ni-NTA agarose (Qiagen) and incubated at 4 °C for 2 h rotating at 15 rpm before centrifugation (3000*g*, 5 min, 4 °C) and removal of the supernatant. Protein was eluted from the resin by washing sequentially (2 × 10 mL each) with wash buffer (20 mM Tris–HCl, 300 mM NaCl, pH 8.0) containing imidazole (1, 10, 50 and 250 mM, pH 8.0). MSNAT was eluted in the 50 mM washes and STNAT was eluted in the 250 mM washes as determined by activity assays and SDS-PAGE. Fractions containing NAT were combined and dialysed against 100 volumes dialysis buffer (20 mM Tris–HCl, 1 mM EDTA, 1 mM DTT pH 8.0) at 4 °C. The N-terminal His-tag was cleaved with thrombin (Calbiochem) (5 U mg^{-1} protein, 22 °C, 2 h) and dialysed again. Protein was concentrated to approximately 5 mg mL^{-1} using a Centricon-10 concentrator (Millipore). Typically the yield of pure active protein was 20 mg (MSNAT) and 7 mg (STNAT) per litre of culture.

Enzymic assays

For assays described below, NAT enzyme samples were adjusted to ensure a linear initial rate of reaction. Kinetic analysis was performed by varying substrate concentration from 50–1600 μM and kinetic constants were determined by non-linear optimisation (KyPlot v6.0).

Acetylation of arylamines

Detection of acetylation of arylamines was performed as previously described[27] in a total volume of 100 μL. Samples of enzyme (0.1–2 μg) and arylamine (50–200 μM) in assay buffer (20 mM Tris–HCl, 1 mM DTT, pH 8.0) were pre-incubated at 37 °C for 5 min. AcCoA (400 μM) was added to start the reaction and the samples were incubated at 37 °C. The reaction was quenched with 100 μL of ice-cold aqueous TCA (20% w/v). 4-Dimethylaminobenzaldehyde (DMAB, 800 μL, 5% w/v in 9:1 acetonitrile/water) was added and the absorbance was measured in a 10 mm pathlength cuvette at $\lambda_{max} = 450$ nm (Cecil CE5502 Spectrophotometer). The amount of remaining arylamine was determined from a standard curve.

Hydrolysis of AcCoA

The rate of hydrolysis of AcCoA by arylamine *N*-acetyltransferase (NAT) was determined in 20 mM Tris–HCl, pH 8.0 using 5,5′-dithio-bis(2-nitrobenzoic acid) (DTNB) as a colorimetric developing agent.[20] TNB showed a molar extinction coefficient of 11.0 $mmol^{-1}$ dm^3 cm^{-1} at $\lambda_{max} = 412$ nm in buffer (100 mM Tris–HCl, 3.2 M guanidine.HCl pH 7.27) measured in a 10 mm pathlength cuvette. When a solution of CoA (20 mM

Tris–HCl, pH 8.0, 100 μL) is treated with DTNB solution (6.4 M guanidine–HCl, 0.1 M Tris–HCl, pH 7.3, 25 μL) in a flat-base polystyrene 250 μL 96-well plate the TNB has an extinction of $3.3 \pm 0.1\ mmol^{-1}\ dm^3$ at $\lambda_{max} = 405$ nm measured in a plate reader (Titertek Multiskan Plus MkII). The liquid depth in the centre of the well is approximately 3.9 mm and thus the molar extinction coefficient is $8.5\ mmol^{-1}\ dm^3\ cm^{-1}$ at this wavelength. A three-fold molar excess of DTNB over AcCoA was used to ensure saturation. The reducing agent dithiothreitol (DTT) releases two equivalents of TNB in the presence of DTNB with an effective molar extinction coefficient of $13.6\ mmol^{-1}\ dm^3\ cm^{-1}$ under these conditions.

The substrate (500 μM) and purified recombinant NAT were mixed and pre-incubated (37 °C, 5 min) in a 250 μL 96-well plate. AcCoA (400 μM) was added to start the reaction in a final volume of 100 μL. The reaction was quenched with guanidine hydrochloride solution (6.4 M guanidine–HCl, 0.1 M Tris–HCl, pH 7.3, 25 μL) containing 5 mM DTNB and the absorbance at 405 nm was measured on a plate-reader within 5 min. Reactions wherein substrate, AcCoA or NAT were omitted respectively were used as controls. The amount of CoA produced was determined from a standard curve.

Simulated annealing using autodock

The 3-D crystal structures of MSNAT[19] (accession number 1GX3) and STNAT[15] (accession number 1E2T) were downloaded from the Protein Data Bank (http://www.rcsb.org).[32] The Autodock v3.0 suite of programs[33,34] was used to perform the simulated annealing on a Silicon Graphics R12000 mips processor. Partial charges were added to the MSNAT crystal structure, the structure was solvated and saved in the 'pdbqs' format using AutoDock Tools. Substrate structures were created using the ChemOffice suite of programs (CambridgeSoft) and energy minimized in Chem3-D using the MM2 forcefield. Partial atomic charges, fixed rings and rotatable bonds were added using AutoDock Tools. Ten docking runs were performed and solutions were converted to 'pdb' (Brookhaven) format. Docked structures were observed in Swiss PDB Viewer (http://www.expasy.ch).

Acknowledgements

We are very grateful for financial support from the Oxford University Challenge Seedcorn Fund, the MRC (Discipline Hopper) and to the Wellcome Trust for continued support. We thank A. Mushtaq, J. Sandy and M. Noble for much helpful discussion. EB is funded by an MRC studentship and FP thanks the Blaschko Trustees for a scholarship.

References and Notes

1. Evans, D. A. P.; Manley, K. A.; McKuisick, V. A. *Br. Med. J.* **1960**, *2*, 485.
2. Weber, W. W.; Hein, D. W. *Pharmacol. Rev.* **1985**, *37*, 25.
3. Hein, D. W.; Doll, M. A.; Rustan, T. D.; Gray, K.; Feng, Y.; Ferguson, R. J.; Grant, D. M. *Carcinogenesis* **1993**, *14*, 1633.
4. Peters, J. H.; Miller, K. S.; Brown, P. *J. Pharmacol. Exp. Ther.* **1965**, *150*, 298.
5. Jenne, J. W. *J. Clin. Invest.* **1965**, *44*, 1992.
6. Hein, D. W.; Doll, M. A.; Fretland, A. J.; Leff, M. A.; Webb, S. J.; Xiao, G. H.; Devanaboyina, U. S.; Nangju, N. A.; Feng, Y. *Cancer Epidemiology, Biomarkers and Prevention* **2000**, *9*, 29.
7. Brockton, N.; Little, J.; Sharp, L.; Cotton, S. C. *Am. J. Epidemiol.* **2000**, *151*, 846.
8. Upton, A.; Johnson, N.; Sandy, J.; Sim, E. *Trends in Pharmacological Sciences* **2001**, *22*, 140.
9. Ames, B. N.; McCann, J.; Yamasaki, E. *Mutat. Res.* **1975**, *31*, 347.
10. Payton, M.; Mushtaq, A.; Yu, T.-W.; Wu, L.-J.; Sinclair, J.; Sim, E. *Microbiology* **2001**, *147*, 1137.
11. Pompeo, F.; Mushtaq, A.; Sim, E. *Prot. Expr. Purif.* **2002**, *24*, 138.
12. Floss, H.; Yu, T.-W. *Curr. Opinion Chem. Biol.* **1999**, *3*, 592.
13. Upton, A. M.; Mushtaq, A.; Victor, T. C.; Sampson, S. L.; Sandy, J.; Smith, D. M.; van Helden, P. V.; Sim, E. *Mol. Microbiol.* **2001**, *42*, 309.
14. Payton, M.; Auty, R.; Delgoda, R.; Everett, M.; Sim, E. *J. Bacteriol.* **1999**, *181*, 1343.
15. Sinclair, J. C.; Sandy, J.; Delgoda, R.; Sim, E.; Noble, M. E. *Nature Structural Biology* **2000**, *7*, 560.
16. Riddle, B.; Jencks, W. P. *J. Biol. Chem.* **1971**, *246*, 3250.
17. Watanabe, M.; Sofuni, T.; Nohmi, T. *J. Biol. Chem.* **1992**, *267*, 8429.
18. Mushtaq, A.; Payton, M.; Sim, E. *J. Biol. Chem.* **2002**, *277*, 12175.
19. Sandy, J.; Mushtaq, A.; Kawamura, A.; Sinclair, J.; Sim, E.; Noble, M. *J. Mol. Biol.* **2002**, *318*, 1071.
20. Riddles, P. W.; Blakeley, R. L.; Zerner, B. *Methods Enzymol.* **1983**, *91*, 49.
21. Cederblad, G.; Harper, P.; Lindgren, K. *Clin. Chem.* **1986**, *32*, 342.
22. Zhang, J. H.; Chung, T. D.; Oldenburg, K. R. *J. Biomol. Screen* **1999**, *4*, 67.
23. Andres, H. H.; Klem, A. J.; Szabo, S. M.; Weber, W. W. *Analytical Biochemistry* **1985**, *145*, 367.
24. Coroneos, E.; Gordon, J. W.; Kelly, S. L.; Wang, P. D.; Sim, E. *Biochim. Biophys. Acta* **1991**, *1073*, 593.
25. Delomenie, C.; Goodfellow, G. H.; Krishnamoorthy, R.; Grant, D. M.; Dupret, J. M. *Biochem. J.* **1997**, *323*, 207.
26. Delomenie, C.; Fouix, S.; Longuemaux, S.; Brahimi, N.; Bizet, C.; Picard, B.; Denamur, E.; Dupret, J.-M. *J. Bacteriol.* **2001**, *183*, 3417.
27. Sinclair, J.; Delgoda, R.; Noble, M.; Jarmin, S.; Goh, N.; Sim, E. *Prot. Exp. Purific.* **1998**, *12*, 371.
28. Goodfellow, G. H.; Dupret, J. M.; Grant, D. M. *Biochem. J.* **2000**, *348*, 159.
29. Minchin, R. F. *Biochem. J.* **1995**, *307*, 1.
30. Upton, A.; Smelt, V.; Mushtaq, A.; Aplin, R.; Johnson, N.; Mardon, H.; Sim, E. *Biochim. Biophys. Acta* **2000**, *1524*, 143.
31. Nodzu, R.; Watanabe, H.; Kuwata, S.; Nagaishi, C.; Teramatsu, T. *J. Pharm. Soc. Jpn.* **1954**, *74*, 872.
32. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
33. Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit.* **1996**, *9*, 1.
34. Olson, A. J.; Goodsell, D. S. *SAR QSAR Environ. Res.* **1998**, *8*, 273.