Binding of the anti-tubercular drug isoniazid to the arylamine *N*-acetyltransferase protein from *Mycobacterium smegmatis*

JAMES SANDY,¹ SIMON HOLTON,² ELIZABETH FULLAM,¹ EDITH SIM,¹ AND MARTIN NOBLE²

¹Department of Pharmacology and ²Department of Molecular Biophysics, Oxford University, Oxford, OX1 3QT, United Kingdom

(RECEIVED October 6, 2004; FINAL REVISION November 17, 2004; ACCEPTED November 19, 2004)

Abstract

Isoniazid is a frontline drug used in the treatment of tuberculosis (TB). Isoniazid is a prodrug, requiring activation in the mycobacterial cell by the catalase/peroxidase activity of the *katG* gene product. TB kills two million people every year and the situation is getting worse due to the increase in prevalence of HIV/AIDS and emergence of multidrug-resistant strains of TB. Arylamine *N*-acetyltransferase (NAT) is a drug-metabolizing enzyme (E.C. 2.1.3.5). NAT can acetylate isoniazid, transferring an acetyl group from acetyl coenzyme A onto the terminal nitrogen of the drug, which in its *N*-acetylated form is therapeutically inactive. The bacterium responsible for TB, *Mycobacterium tuberculosis*, contains and expresses the gene encoding the NAT protein. Isoniazid binds to the NAT protein from *Salmonella typhimurium* and we report here the mode of binding of isoniazid in the NAT enzyme from *Mycobacterium smegmatis*, closely related to the *M. tuberculosis* and *S. typhimurium* NAT enzymes. The mode of binding of isoniazid to *M. smegmatis* NAT has been determined using data collected from two distinct crystal forms. We can say with confidence that the observed mode of binding of isoniazid is not an artifact of the crystallization conditions used. The NAT enzyme is active in mycobacterial cells and we propose that isoniazid binds to the NAT enzyme in these cells. NAT activity in *M. tuberculosis* is likely therefore to modulate the degree of activation of isoniazid by other enzymes within the mycobacterial cell. The structure of NAT with isoniazid bound will facilitate rational drug design for anti-tubercular therapy.

Keywords: arylamine *N*-acetyltransferase; isoniazid; tuberculosis; drug design

Tuberculosis (TB) is responsible for more deaths throughout the world than any other infectious disease; it is estimated that over three million people die each year of TB (World Health Organization, <http://www.who.int/mediacentre/factsheets/fs104/en/>). One-third of the world's population is infected with the *Mycobacterium tuberculosis* bacillus, resulting in the World Health Organization declaring TB a global emergency. Multidrug-resistant strains of TB are be-

coming more prevalent, which has led to the need for more effective treatments for this disease. Since its introduction in 1952, isoniazid (INH) has been one of the frontline treatments for TB; however, the drug must be taken for a minimum of 6 mo (<http://www.umdj.edu/~ntbcweb/history.htm>). Ensuring that TB-infected patients complete their treatment regimen is very difficult, particularly in the developing world where the disease is rife. Patients can remain infectious if the treatment regimen is not completed, the bacilli remaining in the lungs, and this in turn adds to the problem of multidrug-resistant strains of the bacteria developing. The bacterium has a unique cell wall made up from the mycolic acids (Barry et al. 1998) arabinogalactan-lipid

Reprint requests to: James Sandy, Department of Pharmacology, Oxford University, Mansfield Road, Oxford, OX1 3QT, United Kingdom; e-mail: james.sandy@pharm.ox.ac.uk; fax: +44-1865-271853.

Article and publication date are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.041163505>.

complexes and lipoarabinomannan (Brennan and Nikaido 1995). Together they form a waxy membrane around the bacterium that allows the bacterium to survive for long periods of time—hence the long treatment period—and also stops more traditional antibiotics from entering the cell due to decreased permeability at the cell surface. Isoniazid inhibits the formation of the mycolic acid cell wall (Slayden and Barry 2000). Isoniazid is a prodrug and requires activation before it becomes therapeutically active. This process is carried out by the catalase-peroxidase activity of the *katG* gene product (Heym et al. 1993; Bodiguel et al. 2001), and mutations in the *katG* gene contribute to resistance to isoniazid (Zhang et al. 1992; Slayden and Barry 2000). Once activated, isoniazid has a number of proposed targets within the mycobacterial cell (Slayden et al. 2000), including the enoyl acyl carrier protein (ACP) reductase *InhA*, and a β -ketoacyl ACP synthase, *KasA*. Modified isoniazid appears to become covalently attached to NAD^+ (Rozwarski et al. 1998) and causes inhibition of mycolic acid synthesis.

Both *M. tuberculosis* and *M. smegmatis* contain the *nat* gene (Payton et al. 1999) with the two coding sequences sharing 60% identity (Payton et al. 1999). It has been shown that the *nat* gene is expressed in *M. tuberculosis* (Upton et al. 2001). The NAT enzyme from *M. smegmatis* acetylates isoniazid (Sandy et al. 2002). When the *M. tuberculosis nat* gene was overexpressed in *M. smegmatis* the resultant bacteria showed increased resistance to isoniazid (Payton et al. 1999). In addition, when the gene was knocked out the bacteria exhibited increased sensitivity to isoniazid (Payton et al. 2001). It is known that NAT metabolizes isoniazid in growing organisms (Upton et al. 2001). It is likely therefore that NAT competes with *KatG* for isoniazid (Fig. 1). Further gene knockout experiments of *nat* have been carried out in *M. bovis BCG* and interestingly

the mycolic acid cell wall cannot be formed by the bacterium (Bhakta et al. 2004). This would suggest that NAT from *M. bovis BCG* has an endogenous role in cell wall formation. This enzyme therefore appears to be a possible novel target for anti-tubercular therapy.

To understand how small drug-like molecules interact with NAT, we have investigated the mode of binding of isoniazid to the NAT enzyme from *M. smegmatis*. While NAT from *M. tuberculosis* would clearly be the preferred subject for these studies, it has proved to be a difficult protein to work with because it is extremely insoluble when expressed recombinantly in *Escherichia coli* (Payton et al. 1999). *M. smegmatis* NAT was therefore chosen due to its high homology with *M. tuberculosis* NAT at the amino acid level (60% identity) and its solubility when produced in *E. coli*. We have used this structure, together with kinetic data determined for a range of NAT substrates, to begin to dissect the structure/activity relationship of the NAT active site. The substrates used were chosen as they have been used in many previous studies of NAT activity to probe the pharmacology and activity of the enzyme. In addition, they represent the two broad chemical classes (arylamines and arylhydrazines) of NAT substrates and span a wide range of affinities, so that they allow aspects of structure that impact on enzyme-binding to be identified. As well as highlighting properties that are likely to be present in the endogenous NAT substrate, these studies indicate the disposition of interaction sites that might be exploited in rationally designed NAT inhibitors.

Results and Discussion

Crystallization

The NAT protein from *M. smegmatis* has been crystallized in the presence of isoniazid (Fig. 2). Of the 96 conditions in our preliminary screen, 26 yielded crystals, so that no optimization of crystallization conditions was required. The first crystal to be analyzed was grown in 0.2 M ammonium sulphate, 0.1 M MES at pH 6.5, 30% PEG MME 5000, diffracted to 2.1 Å, and was in spacegroup $P2_12_12_1$. To confirm the mode of binding of isoniazid, a second crystal was analyzed. The second crystal was grown in 0.1 M MES at pH 6.5, 12% PEG 20,000 and was large enough for X-ray diffraction studies. This crystal, in spacegroup $P4_12_12$, diffracted to 1.91 Å. All crystals readily formed within 2–3 d, with most appearing overnight. Both of these crystal forms have also been observed for ligand-free *M. smegmatis* NAT, although the cell dimensions of crystals of the NAT isoniazid complex are consistently slightly different from those of apo-NAT. Crystallographic data are summarized in Table 1.

Data collection and structural analysis

The first data set was collected at the SRS Daresbury Laboratory, beamline 14.1. The crystal had four molecules of

Figure 1. Schematic to show activation and inactivation of INH. INH is a prodrug and requires activation by the catalase peroxidase protein (the *KatG* gene product). The NAT enzymes can *N*-acetylate INH, rendering the drug therapeutically inactive. The asterisk next to the terminal nitrogen in the active form of INH denotes a range of oxidized species (Bodiguel et al. 2001).

Figure 2. Crystals of arylamine *N*-acetyltransferase from *M. smegmatis* grown in the presence of INH. NAT protein was incubated for 30 min with 20 mM INH at 37°C prior to setting crystal trials. Crystal trials were incubated at 20°C and crystals appeared within 3 d. Two typical examples of the crystals produced are shown here. Conditions are (A) 0.1 M MES at pH 6.5, 12% PEG 20.00; (B) 0.2 M ammonium sulphate, 0.1 M MES at pH 6.5, 30% PEG MME 5000.

NAT in the asymmetric unit (Chains A–D). The three N-terminal residues introduced by the expression of the enzyme as a fusion protein (Glu⁻², Ser⁻¹, His⁰), together with the N-terminal residues of *M. smegmatis* NAT (Met¹, Ala²), were not resolved in the electron density. However, clear electron density was visible from residue A3–A275, B3–B275, C4–C275, and D8–D275, with the exception of one loop region (residues 99–104). This loop was not particularly clear in any of the molecules in the asymmetric unit, suggesting mobility of the loop within the crystal. The loop is away from contact regions between monomers and is very solvent exposed. There is a tryptophan residue that is adjacent to entrance to the active site cleft (Trp⁹⁹), and which has poor density in all monomers. This is part of a loop with high thermal factors as seen in the apo structure (Sandy et al. 2002).

The second data set was collected at the ESRF, Grenoble, beamline ID14eh3. The crystal had two molecules of NAT in the asymmetric unit. As before, there are loop regions in both molecules that are poorly defined in the electron density (residues A13–A17, A115–A122, B13–B17, and B115–B122). Again, these loop regions are away from intermolecular contacts and are solvent exposed.

The tertiary structure of the *M. smegmatis* NAT cocrystallized in the presence of isoniazid is identical to that of the apo structure (Fig. 3A). The electrostatic charge has been mapped to the surface and is shown in Figure 3B, with isoniazid clearly visible in the bottom of the active site cleft adjacent to the catalytic triad of residues (Cys⁷⁰-His¹¹⁰-Asp¹²⁷).

Orientation of isoniazid

Isoniazid binds in the active site cleft adjacent to the catalytically important cysteine residue (Cys⁷⁰), which is part of the catalytic triad of residues Cys⁷⁰-His¹¹⁰-Asp¹²⁷. Clear electron density was seen for the isoniazid moiety (Fig. 4). The terminal nitrogen atom of the hydrazyl group is positioned such that it is ~2.4 Å away from the Sγ of the

catalytic cysteine. Ligplot analysis (Wallace et al. 1995) of the bonding interactions is shown in Figure 5. This analysis identifies hydrophobic interactions within 3.9 Å of the isoniazid moiety with Phe³⁸, Val⁹⁵, Phe¹³⁰, and Phe²⁰⁷. In addition there are hydrogen bonded interactions with the backbone carbonyl group of Thr¹⁰⁹ and the hydrazyl group of isoniazid.

The NAT enzyme has been described as performing catalysis via a bi-bi ping-pong mechanism whereby an acetyl group is transferred from acetyl CoA onto the NAT enzyme, forming an acetylated intermediate, and then onto a substrate (Riddle and Jencks 1971). While it has been reported that human NAT1 can exist in a stable acetylated form prior to substrate binding (Butcher et al. 2004), and the significant level of sequence conservation across the NAT family supports a common mechanism for eukaryotic and prokaryotic homologs, there is no direct evidence for this in the prokaryotic NAT enzymes. The formation of a complex between isoniazid and NAT in the absence of either acetylation or acetyl CoA is, however, surprising, but is consistent with earlier NMR experiments using the enzyme from *S. typhimurium* (Delgoda et al. 2003) and may represent either a “substrate inhibition” complex or a complex that exists before initiation of the ping-pong mechanism. In either case, it is unlikely that the key determinants of INH binding observed in this structure are substantially different from those that are exploited in the catalytic cycle.

Isoniazid binding does not induce large changes in either the tertiary fold, or in the conformations of active-site residues

There are four well-defined water molecules within the active site that are conserved between the apo and cocrystallized structures of NAT. Five other well-conserved water molecules, consistently seen in the apo structure, are absent in the cocrystallized structure. Four of these would occupy the same space as the isoniazid molecule. The fifth is situated adjacent to the active site cysteine and has previously been proposed to occupy the oxy-anion hole (Sandy et al. 2002) found in other proteins that function through a catalytic triad. Hence substrate binding apparently vacates the oxyanion hole in preparation for necessary reconfigurations that accompany catalytic turnover.

It has been previously reported that a naturally occurring polymorphism (G207R) of *M. tuberculosis* NAT results in a lower apparent affinity of the *M. tuberculosis* enzyme for isoniazid (Upton et al. 2001). We have generated the equivalent mutation in *M. smegmatis* NAT and determined its structure by crystallography (Kawamura et al. 2003). This was shown to display only small conformational changes in amino acids within the active site. This is consistent with the observation that the corresponding mutation in the *M. smegmatis* enzyme has a much reduced influence

Table 1. Summary of crystallographic data for isoniazid and *M. smegmatis* NAT

Space group	P4 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2 ₁
Molecules/A.U.	2	4
Cell dimensions (Å)	99.41 99.41 130.29	101.36 106.13 140.39
Maximum resolution	1.91	2.10
Highest resolution shell	1.956–1.907	2.155–2.00
Observed reflections	1,211,593	1,352,137
Unique reflections	51,023	89,541
Completeness (%)	93.7 (65.4)	92.4 (90.4)
I/σI	9.4 (3.0)	5.7 (1.7)
Mosaicity (°)	0.3	0.4
Rmerge (%) ^a	9.1 (33.4)	10.9 (42.7)
VM (Å ³ /Da)	5.5	5.0
Refinement		
Protein atoms	4214	8508
Other atoms	2 Molecules isoniazid 350 Molecules water	4 Molecules isoniazid 740 Molecules water
Resolution range (Å)	32.704–1.907	31.62–2.00
R _{conv} (%) ^b	21.6	18.2
R _{free} (%) ^c	25.8	22.7
Rmsd bond lengths (Å)	0.013	0.019
Rmsd bond angles (°)	1.48	1.669
Crystallization conditions	0.1 M MES pH 6.5 12% PEG 20K	0.2 M Ammonium sulphate 0.1 M MES pH 6.5 30% PEG MME 5K

Values in parentheses indicate the specific values in the highest resolution shell.
^a $R_{sym} = \sum_h \sum_j |I_{h,j} - \langle I_h \rangle| / \sum_h \sum_j I_{h,j}$ where $I_{h,j}$ is the intensity of the j th observation of unique reflection h .
^b $R_{conv} = \sum_h ||F_o| - |F_c|| / \sum_h |F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h .
^c R_{free} is equivalent to R_{conv} but is calculated using a 5% disjoint set of reflections excluded from the maximum likelihood refinement stages.

on enzymic activity (Kawamura et al. 2003). How this mutation would affect binding of isoniazid remains unclear from our crystal structure, since the isoniazid moiety is bound >8.5 Å from the arginine residue (Arg²⁰⁷), suggesting that the observed perturbation results from a long-range phenomenon such as electrostatics or quantum tunneling. However, because the mutant structure was determined in the absence of a substrate, we cannot exclude the possibility that a conformational change in Arg²⁰⁷ accompanies substrate binding, to offer a more simple explanation.

Competition between NAT and the catalase/oxidase enzyme

We propose that there is competition between the catalase-oxidase protein (the *KatG* gene product) and the NAT enzyme for isoniazid (Fig. 1; Bhakta et al. 2004). The crystal structure of the catalase-oxidase (CP) enzyme from *M. tuberculosis* has recently been solved to a resolution of 2.4 Å (Bertrand et al. 2004). This is a big step forward in understanding the mechanism of INH activation. The mode of binding of INH to the CP enzyme from *M. tuberculosis* is postulated to involve an interaction between the amide

nitrogen of INH to the heme iron of the CP enzyme. Analysis of the chemical character of the INH binding site in CP using the program “GRID” suggests possible energetically

Figure 3. INH bound in the active site of the NAT enzyme. The INH moiety can be seen at the *bottom* of the active site cleft. (A) The backbone is shown in ribbon format, color ramped from red through white to blue over the length of the amino acid chain. The catalytic triad of residues (Cys⁷⁰-His¹¹⁰-Asp¹²⁷) and well-conserved Tyr¹⁷⁷ are also shown, with carbons colored magenta. (B) Electrostatic surface of NAT in complex with INH. The location of the active site at the *bottom* of a long deep cleft is apparent, as is the generally negative electrostatic potential of the catalytic site. The figure was generated using the programs ElectroSurface and AESOP (M. Noble and J. Gruber, unpubl.).

Figure 4. Orientation of INH in the active site of the NAT enzyme. Electron density can be clearly seen for INH in this omit map contoured at 1.5σ . The active site triad residues (Cys⁷⁰-His¹¹⁰-Asp¹²⁷) are shown in ball-and-stick format and the associated electron density is shaded in blue. The INH molecule is also shown in ball-and-stick format with the associated electron density shown colored in pink. The backbone is shown in ribbon format. This figure was produced using AESOP.

favorable interactions with Arg¹⁰⁴, Trp¹⁰⁷, and His¹⁰⁸ (Bertrand et al. 2004). The structure described here, in contrast, demonstrates how INH can be recognized in a cofactor-independent fashion. The molecular tools are now in place to measure directly competition between the CP enzyme and the NAT enzyme for INH.

Superposition of substrates: A structure–activity relationship investigation

Previous studies have looked at potential structure–activity relationships for the NAT enzyme from *M. smegmatis* (Brooke et al. 2003a). By analyzing the mode of binding of isoniazid to the NAT enzyme from *M. smegmatis*, we have a better idea of how substrates will bind and what interactions will take place within the active site. It should be noted that 10 of the 13 amino acids situated within 6 Å of the INH molecule in *M. smegmatis* NAT are identical in *M. tuberculosis* NAT, the differences being Tyr⁷¹/Phe⁷¹, Thr¹¹¹/Asn¹¹¹, and Ala¹⁹⁶/Val¹⁹⁶ (MSNAT/TBNAT). From our analysis none of these residues are implicated in substrate recognition, although it is possible that the Ala/Val substitution creates a slightly more hydrophobic environment within the active site.

K_m apparent values were determined for a series of different substrates of the NAT enzyme (Table 2). These values range over five orders of magnitude, indicating that, while tolerant of significant chemical diversity, the NAT active site shows a marked ability to discriminate. We have sought to rationalize this discrimination with reference to models of different substrate complexes derived from the NAT–isoniazid

complex structure. By aligning the terminal nitrogen of each substrate on top of the terminal nitrogen of the isoniazid molecule within the active site, we were able to build plausible models of each of the substrate complexes for which K_m apparent values have been determined (Fig. 6). While details of recognition may change in the context of the acetylated protein, the interactions observed remain as useful guides for the purposes of rational drug design of NAT inhibitors.

The substrates chosen can be divided into two distinct families: arylhydrazines INH and hydralazine (HDZ); and arylamines *p*-aminobenzoic acid (PABA), anisidine (ANS), 4-aminosalicylic acid (4-AS), and 5-aminosalicylic acid (5-AS). The arylhydrazines are the best substrates for *M. smegmatis* NAT with K_m values of <5 μ M for HDZ and 7.3 μ M for INH. The arylamines have K_m values of 586 μ M for 5-AS, 1460 μ M for ANS, 4500 μ M for 4-AS, and 56,000 μ M for PABA.

Collectively, the arylhydrazines exhibit much lower K_m values. This may in part be due to hydrogen bonding to the hydrazyl group from the carbonyl oxygen of Thr¹⁰⁹ and the S γ of Cys⁷⁰, which are conserved in the *M. tuberculosis* NAT enzyme. In addition HDZ (Fig. 6B) has further potential hydrogen bonding from heterocyclic nitrogen to the backbone oxygen of Gly¹²⁹. In the presence of an acetylated cysteine the substrate would have to be recessed away from the catalytic cysteine, possibly allowing hydrogen bonding

Figure 5. Ligplot analysis of the mode of binding of INH. Ligplot (Wallace et al. 1995) was used to analyze the mode of binding of INH to the NAT protein from *M. smegmatis*. Ligplot was run using the ‘A’ chain from the P2₁2₁2₁ structure. The output suggests hydrogen bonding to the amide nitrogen from Cys⁷⁰ to the carbonyl group from INH and also from the carbonyl group of the backbone from Thr¹⁰⁹ to both of the nitrogens in the hydrazyl group of INH. Hydrophobic interactions with the aromatic ring of INH are proposed from the residues Phe³⁸, Val⁹⁵, Phe¹³⁰, and Phe²⁰⁴.

Table 2. K_m values for *M. smegmatis* NAT with a range of substrates

Substrate	K_m (μ M)
Hydralazine (HDZ)	< 5
Isoniazid (INH)	7.3
5-Aminosalicylic acid (5-AS)	586
Anisidine (ANS)	1,460
4-Aminosalicylic acid (4-AS)	4,500
p-Aminobenzoic acid (PABA)	56,000

K_m values were calculated as described in Materials and Methods. All activity assays were carried out at 37°C in triplicate.

with the backbone oxygen of Phe¹³⁰, which has previously been implicated in docking studies (Mushtaq et al. 2002; Brooke et al. 2003b). This residue also plays another role in π -stacking interactions with the substrates. Although the phenylalanine at position 130 is conserved in most NAT enzymes, including the human isoforms, this study suggests that it is a useful handle that can be exploited to provide affinity in inhibitor design, where specificity might subsequently be achieved by interaction with less conserved amino acids.

Interestingly, the closely related Rifamycin amide synthase (RAS) enzyme (Pompeo et al. 2002) (which exhibits no acetylation activity) has a tyrosine residue in this position. The well-conserved Phe²⁰⁷ interacts with the ring system of the substrates by an edge-to-face contact (Mushtaq et al. 2004). Again, in RAS the corresponding amino acid is a threonine. These subtle differences suggest a more hydrophobic character to the endogenous substrate of NAT than that of RAS (Pompeo et al. 2002).

The arylamines 5-AS, 4-AS, ANS, and PABA all appear to interact slightly differently to the arylhydrazines. The affinity of this series appears to depend to some extent on the capacity to form interactions with Thr¹⁰⁹ and Phe¹³⁰. 5-AS is a good arylamine substrate and a model can readily be built in which 5-AS forms complementary hydrogen

bonds with both the side chain of Thr¹⁰⁹ and the main chain carbonyl of Phe¹³⁰ (Fig. 6C). This carbonyl is deprived of hydrogen bond donating partners in the equivalent complex with 4-AS, unless the binding mode is flipped through 180 degrees (Fig. 6D), which therefore leaves the side chain oxygen of Thr¹⁰⁹ partly occluded from solvent and with only poor options for forming hydrogen bonds. These differences go some way to explaining the lower apparent affinity of 4-AS for *M. smegmatis* NAT. ANS (Fig. 6E) has a lower K_m than 4-AS despite offering hydrogen bonds to neither Thr¹⁰⁹ nor Phe¹³⁰. In this case, the unfavorable desolvation of these groups upon ligand binding might be partly compensated by the generally lipophilic character of the methoxy group. Such groups are preferred in NAT substrates, presumably due to the formation of hydrophobic interactions in the active site, as has previously been discussed (Brooke et al. 2003b). PABA (Fig. 6F) has no additional chemical group on the aromatic ring facing the Thr¹⁰⁹ and has an even higher K_m than that of 4-AS. These models,

Figure 6. Modeling of substrates into the active site of NAT from *M. smegmatis*. Substrates were overlaid based upon the structure of INH using the terminal nitrogen as a reference point for alignment. Section A shows INH as seen in the crystal structure. Sections B–F represent the modeled mode of binding of substrates hydralazine, 5-aminosalicylate, 4-aminosalicylate, anisidine, and *p*-aminobenzoic acid, respectively. The secondary structure is colored blue with the molecular surface colored grey. Potential hydrogen bonds between the catalytic triad of residues and the Tyr¹⁷⁷ residue are indicated with dotted lines. In all frames the protein structure is exactly superimposable; the Cys⁷⁰ label has been placed to avoid occlusion of model substrates in frames B–F.

together with the kinetic data, suggest that the interaction with Thr¹⁰⁹ and Phe¹³⁰ is important in substrate binding.

This work has given an insight into the mode of substrate binding of the *M. smegmatis* NAT enzyme. By targeting the catalytically important residues in the active site, a range of new drugs can be produced that will specifically inhibit the activity of the NAT enzyme. In the *M. tuberculosis* bacterium, this could both improve the efficacy of INH, but also directly inhibit the formation of the mycolic acids (Bhakta et al. 2004). This in turn should lead to a more effective anti-tubercular therapy with a marked reduction in treatment time, improving patient compliance of the treatment regimen.

Materials and methods

Protein production and crystallization

M. smegmatis NAT was produced as a recombinant protein with an N-terminal hexa-histidine tag in *E. coli* as previously reported (Sandy et al. 2002). The protein was purified to homogeneity by binding to Ni-NTA agarose (Novagen), washing with increasing concentrations of imidazole, and eluting with 50 mM imidazole. The eluted protein was dialyzed against 100 volumes 20 mM Tris HCl at pH 8.0, 1 mM EDTA, 1 mM DTT (dialysis buffer) for 18 h at 4°C. The hexa-histidine tag was then excised using Thrombin (5 units/mg protein). Thrombin cleavage leaves three additional nonauthentic residues (a glutamic acid, a serine, and a histidine) at the amino terminus. The excised NAT protein was dialyzed against 100 volumes dialysis buffer for 18 h at 4°C. The NAT protein was finally concentrated using a spin column with a 10-kDa cut-off (Vivascience). Pure NAT protein at a concentration of 15 mg/mL was incubated with 20 mM isoniazid for 30 min at 37°C and then spun in a microcentrifuge at 14,000g prior to setting crystallization trials.

The NAT protein was incubated with Molecular Dimensions Structure Screen I and II and crystallization trials were set up using a Tecan Genesis ProTeam 150 liquid handling robot into 96 well plates. A drop size of 1 µL (50/50 protein: mother liquor) was used in all cases. Plates were sealed and incubated at 20°C. Crystals generally appeared between 1 and 3 d.

Data collection

X-ray diffraction data from the crystal produced in 0.1 M MES at pH 6.5, 12% PEG 20,000 were collected at CCLRC Daresbury Laboratory, beamline 14.1. All crystals were cryoprotected with 30% glycerol in mother liquor and data were collected at 100K. A second data set was collected at the ESRF synchrotron source, Grenoble, France, on beamline id14eh3 from a crystal grown in 0.2 M ammonium sulphate, 0.1 M MES at pH 6.5, 30% PEG MME 5000 with cryogenic protection as above.

Data processing

Diffraction images were analyzed and integrated using MOSFLM and were then scaled using SCALA (Collaborative Computational Project, Number 4, 1994). Molecular replacement was carried out using MOLREP using a modified 'A' chain from *M. smegmatis*

NAT (PDB accession number 1gx3). REFMAC5 (Collaborative Computational Project, Number 4, 1994) and O (Jones et al. 1991) were used for iterative rounds of refinement and model building respectively. The numbering of amino acids was the same as used in the apo structure (Sandy et al. 2002). A final refinement step was to locate the solvent molecules within the protein, using the program ARP (Collaborative Computational Project, Number 4, 1994). The stereochemical quality of the final model was verified using the program Procheck (Laskowski et al. 1993). Atomic coordinates have been deposited in the Protein Data Bank with the accession code 1w6f.

Activity assay

The rate of hydrolysis of acetyl CoA by arylamine *N*-acetyltransferase from *M. smegmatis* (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 (Brooke et al. 2003b).

The substrate and purified recombinant NAT enzyme were mixed and preincubated (37°C, 5 min) in a 250 µL 96-well plate. The substrate concentration was varied from 200 µM to 1 mM for isoniazid, 5-aminosalicylate, and anisidine from 20 µM to 1 mM for hydralazine, from 500 µM to 10 mM for 4-aminosalicylate and from 500 µM to 12.5 mM for *p*-aminobenzoic acid. It was confirmed that the substrates were following Michaelis-Menten kinetics and kinetic constants were determined from the Hanes plot for all substrates apart from isoniazid where the kinetic values were determined from the Lineweaver-Burke plot.

Acetyl CoA (1 mM) was added to start the reaction in a final volume of 100 µL. The reactions were started at $t = 0, 5, 10, 15$ min for isoniazid; $t = 0, 4, 8, 12$ min for 5-aminosalicylate, anisidine and hydralazine; $t = 0, 10, 20, 30$ min for *p*-aminobenzoic acid and 4-aminosalicylate, in order for a time course to be established. 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. The absorbance at 405 nm was measured on an Anthos 2020 plate-reader within 1 min of adding the quenching solution. The amount of CoA produced was determined from a standard curve. All reactions were carried out in triplicate.

Modeling of substrates into active site

Substrates were drawn using ChemDraw (CambridgeSoft) and energy minimized in Chem3D using the MM2 forcefield. Each substrate was superimposed upon isoniazid using O (Jones et al. 1991) with the terminal nitrogen from the hydrazyl group as a reference point. Potential modes of binding were investigated and checked against docking studies previously reported (Brooke et al. 2003a).

Acknowledgments

We thank the Wellcome Trust for continued financial support for this project. In addition, we also thank the staff of the ESRF and SRS synchrotrons for their assistance.

References

- Barry 3rd, C.E., Lee, R.E., Mdluli, K., Sampson, A.E., Schroeder, B.G., Slayden, R.A., and Yuan, Y. 1998. Mycolic acids: Structure, biosynthesis and physiological functions. *Prog. Lipid Res.* **37**: 143–179.

- Bertrand, T., Eady, N.A.J., Jones, J.N., Jesmin Nagy, J.M., Jamart-Gregoire, B., Raven, E.L., and Brown, K.A. 2004. Crystal structure of *Mycobacterium tuberculosis* catalase-peroxidase. *J. Biol. Chem.* **279**: 38991–38999.
- 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., et al. 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.
- Bodiguel, J., Nagy, J.M., Brown, K.A., and Jamart-Gregoire, B. 2001. Oxidation of isoniazid by manganese and *Mycobacterium tuberculosis* catalase-peroxidase yields a new mechanism of activation. *J. Am. Chem. Soc.* **123**: 3832–3833.
- Brennan, P. and Nikaido, H. 1995. The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**: 29–63.
- Brooke, E., Davies, S.G., Mulvaney, A.W., Pompeo, F., Sim, E., and Vickers, R.J. 2003a. An approach to identifying novel substrates of bacterial arylamine N-acetyltransferases. *Bioorg. Med. Chem.* **11**: 1227–1234.
- Brooke, E., Davies, S.G., Mulvaney, A.W., Okada, M., Pompeo, F., Sim, E., Vickers, R.J., and Westood, I.M. 2003b. Synthesis and in vitro evaluation of novel small molecule inhibitors of bacterial arylamine N-acetyltransferases (NATs). *Bioorg. Med. Chem.* **13**: 2527–2530.
- Butcher, N.J., Arulpragasam, A., and Minchin, R.F. 2004. Proteasomal degradation of N-acetyltransferase 1 is prevented by acetylation of the active site cysteine: A mechanism for the slow acetylator phenotype and substrate-dependent down-regulation. *J. Biol. Chem.* **279**: 22131–22137.
- Collaborative Computational Project, Number 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.* **D50**: 760–763.
- 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.
- Heym, B., Zhang, Y., Poulet, S., Young, D., and Cole, S. 1993. Characterisation of the *KatG* gene encoding catalase-peroxidase required for the isoniazid susceptibility of *Mycobacterium tuberculosis*. *J. Bacteriol.* **175**: 4255–4259.
- Jones, T.A., Cowan, S., Zou, J.-Y., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**: 110–119.
- 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 Expr. Purif.* **27**: 75–84.
- 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.
- 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.
- Mushtaq, A., Sandy, J., Holton, S.J., Shartau, P., Sim, E., and Noble, M. 2004. Investigation of the catalytic triad of arylamine N-acetyltransferases: Essential residues required for acetyl transfer to arylamines. Third International NAT Workshop, Vancouver, BC, Canada.
- 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.
- 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.
- 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 Expr. Purif.* **24**: 138–151.
- 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.
- Rozwarski, D.A., Grant, G.A., Barton, D.H.R., Jacobs, W.R., and Sacchettini, J.C. 1998. Modification of the NADH of the Isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* **279**: 98–102.
- 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.
- Slayden, R.A. and Barry 3rd, C.E. 2000. The genetics and biochemistry of isoniazid resistance in mycobacterium tuberculosis. *Microbes Infect.* **2**: 659–669.
- Slayden, R.A., Lee, R.E., and Barry 3rd, C.E. 2000. Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **38**: 514–525.
- Upton, A., 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**: 309–319.
- Wallace, A.C., Laskowski, R.A., and Thornton, J.M. 1995. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**: 127–134.
- Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. 1992. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* **358**: 591–593.